

# Frequency of single nucleotide polymorphisms of some immune response genes in a population sample from São Paulo, Brazil

Frequência de polimorfismo de nucleotídeo único de alguns genes da resposta imune em amostra populacional da cidade de São Paulo, Brasil

Léa Campos de Oliveira<sup>1,2</sup>, Rajendranath Ramasawmy<sup>1</sup>, Jaila Dias Borges<sup>3</sup>, Maria Lucia Carnevale Marin<sup>1</sup>, Natalie Guida Muller<sup>1</sup>, Jorge Kalil<sup>1,4,5</sup>, Anna Carla Goldberg<sup>1,3,5</sup>

## ABSTRACT

**Objective:** To present the frequency of single nucleotide polymorphisms of a few immune response genes in a population sample from São Paulo City (SP), Brazil. **Methods:** Data on allele frequencies of known polymorphisms of innate and acquired immunity genes were presented, the majority with proven impact on gene function. Data were gathered from a sample of healthy individuals, non-HLA identical siblings of bone marrow transplant recipients from the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, obtained between 1998 and 2005. The number of samples varied for each single nucleotide polymorphism analyzed by polymerase chain reaction followed by restriction enzyme cleavage. **Results:** Allele and genotype distribution of 41 different gene polymorphisms, mostly cytokines, but also including other immune response genes, were presented. **Conclusion:** We believe that the data presented here can be of great value for case-control studies, to define which polymorphisms are present in biologically relevant frequencies and to assess targets for therapeutic intervention in polygenic diseases with a component of immune and inflammatory responses.

**Keywords:** Polymorphism, genetic; Immunity, innate; Cytokines

## RESUMO

**Objetivo:** Apresentar a frequência de polimorfismo de nucleotídeo único de alguns genes da resposta imune em amostra populacional da cidade de São Paulo (SP). **Métodos:** Foram apresentadas as frequências de alelos de conhecidos polimorfismos de genes de imunidade inata e adquirida, a maioria com impacto funcional comprovado. Os dados foram coletados a partir de amostras de indivíduos saudáveis, irmãos não-HLA idênticos, de receptores de

transplante de medula óssea do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, obtidos entre 1998 e 2005. O número de amostras variou para cada polimorfismo de nucleotídeo único analisado por reação em cadeia pela polimerase seguida de clivagem com enzimas de restrição. **Resultados:** Apresentou-se a distribuição de alelos e genótipos de 41 polimorfismos genéticos, a maioria de genes para citocinas, mas também incluindo outros genes de resposta imune. **Conclusão:** Acreditamos que os dados apresentados aqui possam ser de grande valor para definir quais os polimorfismos presentes em frequências relevantes, para estudos caso-controle e para avaliar alvos de intervenção terapêutica nas doenças poligênicas com componente de resposta imune ou inflamatória.

**Descritores:** Polimorfismo genético; Imunidade inata; Citocinas

## INTRODUCTION

Over the last years, it has become increasingly clear that individual genetic variation is an essential component of overall immune responses that contributes to susceptibility, progression, and outcome of infectious and autoimmune diseases, and cancer. Innumerable studies have uncovered the extent of human genetic variation while trying to map their role in multifactorial and polygenic diseases. These studies showed that there are different types of variants, ranging from single nucleotide polymorphisms (SNPs) with one sole base change to repetitive short- and medium-sized sequences, as well as copy number variations, which may extend

*Study carried out at Instituto do Coração da Faculdade de Medicina, Universidade de São Paulo – USP, São Paulo (SP), Brazil; Instituto Israelita de Ensino e Pesquisa Albert Einstein – IIEPAE, São Paulo (SP), Brazil; Instituto de Investigação em Imunologia of Institutos Nacionais de Ciência e Tecnologia – INCT, São Paulo (SP), Brazil.*

<sup>1</sup> Instituto do Coração, Faculdade de Medicina, Universidade de São Paulo – USP, São Paulo (SP), Brazil.

<sup>2</sup> Laboratório de Pediatria Clínica (LIM-36), Instituto da Criança, Hospital das Clínicas, Universidade de São Paulo – USP, São Paulo (SP), Brazil.

<sup>3</sup> Instituto Israelita de Ensino e Pesquisa Albert Einstein – IIEPAE, São Paulo (SP), Brazil.

<sup>4</sup> Divisão de Imunologia Clínica e Alergia, Faculdade de Medicina, Universidade de São Paulo – USP, São Paulo (SP), Brazil.

<sup>5</sup> Instituto de Investigação em Imunologia of Institutos Nacionais de Ciência e Tecnologia – INCT, São Paulo (SP), Brazil.

Corresponding author: Anna Carla Goldberg – Avenida Albert Einstein, 627, 2SS, Bloco A – Morumbi – CEP 05652-000 – São Paulo (SP), Brasil – Tel.: 11 2151-0941 – E-mail: goldberg@einstein.br

Received on: Aug 26, 2010 – Accepted on: Jun 22, 2011

Conflict of interest: None

throughout large segments of chromosomes. Individual genomes are currently being analyzed in the 1000 Genomes Project to catalogue the full extent of human genetic variation. Previous efforts like HapMap focused on SNPs aiming to map gene variants that affect health, disease, and individual responses to medications and environmental factors. Today, the number of known SNPs is in the range of 10 million and HapMap is estimated to contain 80% of all SNPs with frequencies of >10%<sup>(1)</sup> (useful web links are listed in Appendix 1). These projects have a common characteristic, namely, the mapping of hundreds of thousands of allelic variants. Useful variants for these genome-wide association studies (GWAS) typically exhibit frequencies between 0.5 and 5% and a great number of cases and controls is needed to obtain statistical power. For example, a 90% power to detect an allele with 1% frequency, with a risk factor of 2 GWAS, demands approximately 20,000 individual samples<sup>(2)</sup>. This number can be considerably lower (~300) for a rare allele frequency of roughly 10%. A good example are the studies conducted by the Wellcome Trust Case Control Consortium, which analyzed several thousand individuals in genome-wide scans in order to map the genetic risk for type 2 diabetes, rheumatoid arthritis, and Crohn's disease, to name a few of the target diseases studied<sup>(3)</sup>.

On the other hand, pooling samples can also pose a problem. Disease heterogeneity and population differences act as confounding factors, hindering identification of relevant genes due to the small effect many of the variants have on the disease itself.

Studying the genetic background of Brazilian populations has always been a challenging issue. Not only there is a high degree of miscegenation, but the input from different migratory currents differs from one region to another. Amerindian, Caucasian, and African ethnicities contributed to this true "melting pot" since the very beginning of colonization five centuries ago by the Portuguese. Studies aiming to assess the relative contribution of the three races forming the gene pool of the Brazilian population through matrilinear<sup>(4)</sup> and patrilinear<sup>(5)</sup> descent, using mtDNA and chromosome Y-based methods, confirm historical data of crossbreeding between European men and Amerindian and African women<sup>(6,7)</sup>.

In the Southeast region of Brazil where São Paulo is located, the composition of the mixture includes a majority of Italians, Spaniards, and Germans. However, this region also received a significant input from Africans and Indigenous peoples. In addition, as the most populated city in South America, São Paulo has always been an important destination of migrants from other places in Brazil and from neighboring countries. To account for the genetic admixture, some surrogate

markers, such as skin color, are employed in many studies. However, Parra et al.<sup>(6)</sup> showed that in Brazilians this is a poor substitute for the assessment of individual ancestry. We chose to analyze healthy individuals coming from the same socioeconomic strata as the patients seen at the largest hospital in São Paulo. These individuals mirror the great diversity present in the population of the city of São Paulo, and though ancestry markers were not analyzed, we feel the information should be reported in order to be available for fellow researchers in the field when searching for candidate genes for case-control studies of the many different types of diseases that affect 19 million inhabitants of the extended metropolitan region of São Paulo.

The data shown in this paper include allele frequencies of known polymorphic genes in innate and acquired immunity, the majority with proven impact on gene function. Most of the polymorphisms shown here have an impact upon transcription levels. Others lead to changes in the binding strength to their corresponding ligands and/or in intracellular signal transduction, and even in the half-life of messenger RNAs.

The scope of this paper does not include detailed information on the genes or their polymorphisms, which can be assessed in good textbooks<sup>(8)</sup>.

### Innate immunity genes

The innate immune system relies on the presence of pathogen-associated molecular patterns (PAMPs) on microbial surfaces resulting in the activation of effector cells capable of clearing infection and inducing inflammation in this process. These pattern recognition molecules can be cell-associated, as is the case for Toll-like receptors (TLRs)<sup>(9,10)</sup>, or soluble as in the case of the mannose-binding lectin family of proteins<sup>(11)</sup>.

TLRs are predominantly expressed in antigen-presenting cells, either on the surface (TLR 1, 2, 4, 6) or in the cell (TLR 3, 7, 9). Each recognizes a specific PAMP-like *Gram*-negative bacterial lipopolysaccharide (TLR-4), flagellin (TLR-5), single-stranded (TLR-7) or double-stranded (TLR-3) viral RNA, bacterium-derived CpG DNA (TLR-9). TLR-4 is the major representative of the family, and is known to respond to exogenous and endogenous ligands, participating in inflammation and local tissue responses to wounds, hypoxia, or other forms of stress. Diminished response to lipopolysaccharides was mapped to amino acid substitutions in the TLR4 gene<sup>(12)</sup>.

The complement system can be activated in different ways, one of which is the lectin pathway, initiated by mannose-binding lectin (MBL), reviewed in detail by Dommett et al.<sup>(13)</sup>. MBL is an acute phase reactant that binds to mannose, sugars, and other microbial

compounds by way of the lectin domain. Once activated, the cascade begins with the cleavage of C4 and C2, by MBL-activated MASP-1 and -2. It is intriguing that MBL-2, the functional form of MBL in humans, harbors polymorphisms in coding regions leading to non-preserved amino acid substitutions. In addition, there are promoter polymorphisms in strong linkage disequilibrium, which result in extended haplotypes. Of notice, the frequency of these alleles varies greatly worldwide and many different diseases, infectious or otherwise were studied<sup>(13)</sup>.

### Interleukin 10 and receptor genes

IL-10 is the prototypical down-regulating effector cytokine in immune response. Its action is crucial for control of inflammation accompanying immune responses. Absence of IL-10 in animal models of disease leads to tissue damage, autoimmunity, or chronic infection. IL-10 is produced by many different immune cells, such as macrophages and dendritic cells, TH1, TH2, and TH17 lymphocytes. It counteracts the production of pro-inflammatory interferon- $\gamma$  and increases the differentiation of regulatory T cells<sup>(14)</sup>. The gene encoding IL-10 is highly polymorphic, presenting several variants in the 5' and promoter regions with impact on *in vitro* and *in vivo* production<sup>(15,16)</sup>. IL-10 exerts its actions through the specific heterodimeric IL-10 receptor, where chain 1 is responsible for high affinity binding<sup>(17)</sup>.

### Immunomodulatory genes in the MHC class III region

Several immunomodulatory genes are located in the MHC class III region. In addition to the well-known inflammatory cytokines tumor necrosis factor alpha (TNF- $\alpha$ ) and lymphotoxin alpha (LT- $\alpha$ ), functional polymorphisms are being studied in recently described genes, like HLA-B associated transcript 1 (BAT1), nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1 (NFKBIL1), and leukocyte-specific transcript 1 (LST1)<sup>(18)</sup>. The exact functional role of these three genes is not yet clear.

BAT1, the most telomeric of the three genes, is a DEAD-box RNA helicase, one of several genes in this region involved in RNA processing<sup>(19)</sup> and, importantly, a target for immune evasion by cytomegalovirus<sup>(20)</sup>. NFKBIL1, a protein with some homology to the I $\kappa$ B family, was suggested to be involved in the control of cytosolic nuclear transcription factor-kappa B (NF- $\kappa$ B), a major molecule governing the transcription of over 200 different immune response genes<sup>(21)</sup>. A detailed paper<sup>(22)</sup> recently showed that NFKBIL1 is expressed in all tissues, and that the protein is present in macrophages and T cells in the synovium of rheumatoid arthritis patients. Furthermore, in spite of the apparent homology, this

product binds mRNA and seems to function in RNA processing. Finally, LST1 is translated into multiple isoforms with varying expression according to cell type and induction form. Expression in immune cells is widespread and, as a rule, leads to diminished lymphocyte proliferation.

### IL-4, -5, -13, and receptor genes

IL-4 is a pleiotropic cytokine essential for IgE synthesis in B cells and for T cell differentiation into the TH2 phenotype<sup>(23)</sup>. Th2 cells secrete IL-4, IL-13, and IL-5. Mast cells also secrete IL-5 and this interleukin activates eosinophils. IL-5, together with IL-4 and IL-13, is deeply involved in the induction and maintenance of allergic processes. The functions of IL-13 in immune surveillance and in TH2 type immune response partially overlap with IL-4, but IL-13 also has an impact on tissue eosinophilia, as well as tissue remodeling and development of fibrosis<sup>(24)</sup>. Both IL-4 and IL-13 genes harbor functionally relevant polymorphisms. The biological activity of these two cytokines occurs through binding on target cells to their specific receptor. IL-13, which shares several biological functions with IL-4, operates through the IL-13 receptor, a heterodimer formed by the shared IL-4R alpha chain and IL-13R alpha chain.

### Other cytokine and chemokine genes

A great variety of cytokines, chemokines, growth factors, and other molecules are produced in the beginning or during expansion of innate and acquired immune responses. The type of trigger, the site of the tissue, and the combination of the input of different molecules will shape the nascent immune response, defining the predominant type of cell and portfolio of effector molecules produced, according to Amsen et al.<sup>(23)</sup> and Pulendran et al.<sup>(25)</sup>. The pro-inflammatory cascade of events is well-known, and two of the classical effector molecules involved are IL-12 and interferon gamma (IFN- $\gamma$ ). In addition, monocyte chemoattractant protein 1 (MCP-1) and chemokine C-C motif receptor 5 (CCR5) are, respectively, chemokine and chemokine receptor playing major roles in proinflammatory events<sup>(8)</sup>. Cytotoxic T lymphocyte-associated protein 4 (CTLA-4) is an essential ligand governing activity of T lymphocytes. Thus, functional polymorphisms which modify levels or efficiency of any of these molecules may impact the outcome of the immune response due to the shift in balance of inflammatory and regulatory responses. The polymorphisms presented here have been widely studied, and have been found to play a role in several infectious and autoimmune diseases.

## OBJECTIVE

To present the frequency of SNPs of some immune response genes in a population sample from São Paulo city.

## METHODS

### Subjects

Data in this study were gathered from a sample of healthy individuals, non-HLA identical siblings of bone marrow transplant recipients from the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, obtained between 1998 and 2005. These individuals underwent clinical and laboratory screening as possible donors, and were deemed fit for donation, but were dismissed due to MHC incompatibility with respective recipients. Samples were obtained after informed consent and permission from the Ethics Committee of the hospital. Since the sample number varied for each SNP analyzed, numbers (n) are shown in tables 1-5.

### DNA extraction and genotyping

Blood samples were drawn and DNA was extracted by dodecyltrimethylammonium bromide/cetyltrimethylammonium bromide (DTAB/CTAB)<sup>(26)</sup> or alternatively, by salting-out methods<sup>(27)</sup>.

### RFLP-PCR genotyping

There was no deviation from expected Hardy-Weinberg proportions in any of the genes analyzed. For genotyping of all SNPs, 100 ng genomic DNA was used. Polymorphisms were typed by PCR-RFLP as described elsewhere (additional information on the SNPs presented is available in Appendix 2 and upon request). Briefly, PCRs were performed in a final volume of 25  $\mu$ L containing 100 ng genomic DNA, 40  $\mu$ M of dNTP and 0.2 U of Taq polymerase, 1.5 mM of MgCl<sub>2</sub>, 0.25  $\mu$ M of each primer. In some cases, protocols employed 2.0 mM of MgCl<sub>2</sub> and 0.5  $\mu$ M of each primer. PCR was usually carried out with an initial 5-minute denaturation step at 95°C followed by 35 cycles at 95°C for 20 seconds, annealing for 30 seconds, followed by an extension at 72°C for 20 seconds and a final extension step of 5 to 7 minutes at 72°C. An aliquot of 10  $\mu$ L of the PCR product was digested for 3 hours with the specified restriction enzyme (New England Biolabs), in a total volume of 20  $\mu$ L at the temperature specified by the manufacturer. Digested products were separated by electrophoresis on 2 to 4% agarose gel, stained with ethidium bromide, and visualized under ultraviolet (UV) light.

## RESULTS

Allele and genotype distribution of 41 different gene polymorphisms, mostly cytokines, but also including other immune response genes, are shown in tables 1-5.

**Table 1.** Genotype and allele frequencies of innate immunity gene polymorphisms in a sample of healthy individuals

SNP	Alleles	Genotypes (%)	AF
TLR4 +896, n = 255	AA	223 (87.4)	0.94
	AG	31 (12.2)	
	GG	1 (0.4)	
TLR5 +1174, n = 279	CC	261 (93.5)	0.97
	CT	17 (6.1)	
	TT	1 (0.4)	
MBL-550, n = 281	CC	117 (41.5)	0.65
	CG	134 (48)	
	GG	30 (10.5)	
MBL-221, n = 281	CC	10 (4.0)	0.16
	CG	69 (24)	
	GG	202 (72)	
MBL +4, n = 281	CC	158 (56)	0.74
	CT	101 (36)	
	TT	22 (8.0)	
MBL 52*, n = 281	CC	259 (92.1)	0.96
	CT	21 (7.5)	
	TT	1 (0.4)	
MBL 54*, n = 281	AA	6 (2)	0.14
	GA	64 (23)	
	GG	210 (75)	
MBL 57*, n = 281	GG	262 (93)	0.96
	GA	19 (7)	
	AA	0	
MBL A/O**, n = 281	A/A	176 (63)	0.80
	A/O	91 (32)	
	O/O	14 (5)	
MASP2 120, n = 281	GG	269 (96)	0.98
	GA	12 (4)	
	AA	0	

SNP: single nucleotide polymorphism; AF: allele frequencies.

\* codon; \*\* C+D+E, which corresponds to variants in positions 52,54,57.

**Table 2.** Genotype and allele frequencies of IL-10 and IL-10 receptor gene polymorphisms in a sample of healthy individuals

SNP	Alleles	Genotypes (%)	AF
IL-10 -592, n = 278	AA	38 (14)	0.35
	AC	119 (43)	
	CC	121 (43)	
IL-10 -819, n = 278	CC	120 (43)	0.65
	TC	120 (43)	
	TT	38 (14)	
IL-10 -1082, n = 278	AA	137 (49.3)	0.7
	AC	115 (41.4)	
	CC	26 (9.3)	
IL10 -2763, n = 278	AA	29 (10)	0.25
	AC	83 (30)	
	CC	166 (60)	
IL-10 -2849, n = 278	AA	15 (5.4)	0.18
	AG	68 (24.5)	
	GG	195 (70.1)	
IL-10 -3575, n = 278	TT	163 (59)	0.76
	TA	98 (35)	
	AA	17 (6)	
IL10R 138*, n = 265	AA	211 (80)	0.9
	AG	49 (18)	
	GG	5 (2)	
IL10R 330*, n = 259	GG	151 (58.3)	0.76
	AG	94 (36.3)	
	AA	14 (5.4)	

SNP: single nucleotide polymorphism; AF: allele frequencies. \* codon.

**Table 3.** Genotypes and allele frequencies (AF) of MHC III gene polymorphisms in a sample of healthy individuals

SNP	Alleles	Genotypes (%)	AF
TNFA -238, n = 281	AA	1 (0.4)	0.04
	AG	18 (6.4)	
	GG	262 (93.2)	
TNFA -308, n = 281	AA	4 (1)	0.1
	AG	45 (16)	
	GG	232 (83)	
LTA +80, n = 273	CC	116 (42)	0.6
	AC	114 (42)	
	AA	43 (16)	
LTA +252, n = 279	AA	132 (47.3)	0.7
	AG	118 (42.3)	
	GG	29 (10.4)	
NFKBIL1 -63, n = 276	AA	38 (14)	0.35
	AT	118 (43)	
	TT	120 (43)	
BAT1 -22, n = 265	GG	127 (48)	0.7
	CG	110 (41.5)	
	CC	28 (10.5)	
BAT1 -348, n = 266	CC	192 (72)	0.85
	CT	68 (26)	
	TT	6 (2)	
LST1 +290, n = 141	AA	7 (5)	0.09
	AG	10 (7)	
	GG	124 (88)	

SNP: single nucleotide polymorphism; AF: allele frequencies.

**Table 4.** Genotypes and allele frequencies of IL-4, IL-5, IL-13, and receptor gene polymorphisms in a sample of healthy individuals

SNP	Alleles	Genotypes (%)	AF
IL5 - 746, n = 128	CC	20 (16)	0.43
	CT	71 (55)	
	TT	37 (29)	
IL4 -589, n = 180	TT	21 (11)	0.31
	CT	68 (38)	
	CC	91 (51)	
IL4 +33, n = 220	TT	30 (14)	0.31
	CT	77 (35)	
	CC	113 (51)	
IL4 +3017, n = 86	GG	15 (17.4)	0.51
	GT	58 (67.4)	
	TT	13 (15.2)	
IL13 +2044, n = 160	AA	6 (4.0)	0.2
	AG	60 (37.0)	
	GG	94 (59.0)	
IL4R +223, n = 212	A/A	49 (19)	0.50
	A/G	129 (61)	
	G/G	43 (20)	
IL13RA1 +1398, n = 85	AA	57 (67)	0.76
	AG	16 (19)	
	GG	12 (14)	

SNP: single nucleotide polymorphism; AF: allele frequencies.

**Table 5.** Genotypes and allele frequencies of several immune response gene polymorphisms in a sample of healthy individuals

SNP	Alleles	Genotypes (%)	AF
IL6 -174, n = 264	GG	168 (64)	0.8
	CG	83 (31)	
	CC	13 (5)	
INFG +874, n = 273	AA	95 (35)	0.6
	AT	129 (47)	
	TT	49 (18)	
IL12B +1188, n = 266	AA	134 (50)	0.7
	AC	109 (41)	
	CC	23 (9)	
CCR5Δ32, n = 278	WT**/WT	252 (90)	0.95
	DEL**/WT	24 (9)	
	DEL/DEL	2 (1)	
MCP1 - 2518, n = 256	AA	122 (48)	0.7
	AG	103 (40)	
	GG	31 (12)	
CTLA4-318, n = 217	CC	191 (88)	0.93
	CT	23 (11)	
	TT	3 (1)	
CTLA4 +49, n = 228	AA	94 (41)	0.66
	AG	113 (50)	
	GG	21 (9)	
CTLA4 CT60, n = 181	AA	42 (23)	0.47
	AG	87 (48)	
	GG	52 (29)	

SNP: single nucleotide polymorphism; AF: allele frequencies; \* wild type; \*\* deletion.

## DISCUSSION

In this article we present a series of allele and genotype frequencies of known and novel immune response genes. Though these genes show a modest contribution to the overall phenotype, it is important to detail the effects each gene has in the development and progression of a given disease. The sum of multiple genetic and environmental factors leads to different clinical presentations and therapeutic responses in each patient<sup>(28)</sup>. Thus, the study of significant numbers of patients carrying the same disease, as well as the comparison between similar diseases (for example, autoimmune diseases), opens the way to identify relevant mechanisms in their pathophysiology. Genetic polymorphisms, such as those shown in this article, were associated with a variety of autoimmune, inflammatory, and infectious diseases, ranging from celiac disease and rheumatoid arthritis to acute myocardial infarction, Chagas' disease, and viral hepatitis.

The majority of polymorphic sites in the genome is common in populations worldwide, and variants exhibit moderate frequency<sup>(29)</sup>, implying that most have gone through a balancing selection, that is, they have been preserved because in addition to imparting susceptibility to certain diseases, they also have a beneficial role according to the environmental background of the populations. Two important points should be made concerning some of the polymorphisms we have studied that showed very low frequencies (for example TNF- $\alpha$ 238 and TLR5 +1174). Low frequencies impact the statistical power and a greatly increased number of samples need to be examined to reach significance in association studies. When the impact of the variant upon a phenotype is low, the issue is further complicated. In candidate gene studies, in which cases and controls typically number only in the hundreds, this is an important issue and should be taken into account when choosing target genes. On the other hand, genes with greater effects can be reliably analyzed.

Some considerations may help circumvent or lessen the impact of the issues pointed out here before: first and foremost, a robust hypothesis based on clinical evidence is needed. It should be noted that while genome-wide screening does not employ *a priori* hypotheses, case-control studies will benefit from correlation with data obtained through careful clinical follow-up and detailed laboratory data. The choices of additional markers in the same gene or chromosome region, the analysis in subsequent independent samples, the use of two to four times more controls than patient samples, and care to avoid hidden population structure, which can result in false differences, are additional points to be taken into account. Although many claims of associations have been published, few are subsequently replicated, a problem affecting GWAS studies as well<sup>(2,30)</sup>.

However, as pointed out by Eric Lander et al.<sup>(2)</sup>, there is still a role for association studies. The primary value of genetic mapping is not risk prediction, but providing novel insights about mechanisms of disease. Knowledge of disease pathways can suggest strategies for prevention, diagnosis, and therapy.

## CONCLUSION

Finally, though ancestry was not defined in our study population, we believe that the data presented here can be of great value for case-control studies, to define which polymorphisms are present in biologically relevant frequencies, and to assess targets for therapeutic intervention in polygenic diseases with a component of immune and inflammatory responses.

## ACKNOWLEDGMENTS

This project was supported by The State of Sao Paulo Research Foundation (FAPESP), grants 2001/09850-0 and 2005/00553-3) and National Council for Scientific and Technological Development (CNPq), grant 0062/2001-0. ACG and JK are recipients of personal grants from CNPq.

## REFERENCES

- International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005;437(7063):1299-320.
- Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008;322(5903):881-8.
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-78.
- Alves-Silva J, da Silva Santos M, Guimarães PE, Ferreira AC, Bandelt HJ, Pena SD, et al. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 2000;67(2):444-61.
- Carvalho-Silva DR, Santos FR, Rocha J, Pena SD. The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet*. 2001;68(1):281-6.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD, Parra, F.C., et al., Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci U S A*. 2003;100(1):177-82.
- Suarez-Kurtz G, Vargens DD, Struchiner CJ, Bastos-Rodrigues L, Pena SD. Self-reported skin color, genomic ancestry and the distribution of GST polymorphisms. *Pharmacogenet Genomics*. 2007;17(9):765-71.
- Kindt TJG, Goldsby RA, Osborne BA, Kuby. *Immunology*. 6th ed. New York: W.H. Freeman; 2007.
- Ishii KJ, Coban C, Kato H, Takahashi K, Torii Y, Takeshita F et al. A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nat Immunol*. 2006;7(1):40-8.
- Iwasaki A, Medzhitov R. Toll-like receptor control of the adaptive immune responses. *Nat Immunol*. 2004;5(10):987-95.
- Garred P, Honoré C, Ma YJ, Munthe-Fog L, Hummelshøj T. MBL2, FCN1, FCN2 and FCN3-The genes behind the initiation of the lectin pathway of complement. *Mol Immunol*. 2009;46(14):2737-44.
- Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M et al., TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet*. 2000;25(2):187-91.
- Dommett RM, Klein N, Turner MW. Mannose-binding lectin in innate immunity: past, present and future. *Tissue Antigens*. 2006;68(3):193-209.
- Saraiva M, O'Garra A. The regulation of IL-10 production by immune cells. *Nat Rev Immunol*. 2010;10(3):170-81.
- Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ, Hutchinson IV. An investigation of polymorphism in the interleukin-10 gene promoter. *Eur J Immunogenet*. 1997;24(1):1-8.
- Shin HD, Park BL, Kim LH, Kim JS, Kim JW. Interleukin-10 haplotype associated with total serum IgE in atopic dermatitis patients. *Allergy*. 2005;60(9):1146-51.
- Gasche C, Grundtner P, Zwirn P, Reinisch W, Shaw SH, Zdanov A. Novel variants of the IL-10 receptor 1 affect inhibition of monocyte TNF-alpha production. *J Immunol*. 2003;170(11):5578-82.
- Koch W, Hoppmann P, Michou E, Jung V, Pfeufer A, Mueller JC et al. Association of variants in the BAT1-NFKBIL1-LTA genomic region with protection against myocardial infarction in Europeans. *Hum Mol Genet*. 2007;16(15):1821-7.
- Lehner B, Semple JI, Brown SE, Counsell D, Campbell RD, Sanderson CM. Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics*. 2004;83(1):153-67.
- Lischka P, Toth Z, Thomas M, Mueller R, Stamminger T. The UL69 transactivator protein of human cytomegalovirus interacts with DEXD/H-Box RNA helicase UAP56 to promote cytoplasmic accumulation of unspliced RNA. *Mol Cell Biol*. 2006;26(5):1631-43.
- Carlsen H, Alexander G, Austenaa LM, Ebihara K, Blomhoff R. Molecular imaging of the transcription factor NF-kappaB, a primary regulator of stress response. *Mutat Res*. 2004;551(1-2):199-211.
- Greetham D, Ellis CD, Mewar D, Fearon U, an Ultaigh SN, Veale DJ et al., Functional characterization of NF-kappaB inhibitor-like protein 1 (NFKBIL1), a candidate susceptibility gene for rheumatoid arthritis. *Hum Mol Genet*. 2007;16(24):3027-36.
- Amsen D, Spiliarakis CG, Flavell RA. How are T(H)1 and T(H)2 effector cells made? *Curr Opin Immunol*. 2009;21(2):153-60.
- Wynn TA. IL-13 effector functions. *Annu Rev Immunol*. 2003;21:425-56.
- Pulendran B, Tang H, Manicassamy S. Programming dendritic cells to induce T(H)2 and tolerogenic responses. *Nat Immunol*. 2010;11(8):647-55.
- Gustincich S, Manfioletti G, Del Sal G, Schneider C, Carninci P. A fast method for high-quality genomic DNA extraction from whole human blood. *Biotechniques*. 1991;11(3):298-300, 302.
- Fernandez-Vina M, Moraes JR, Moraes ME, Miller S, Stastny P. HLA class II haplotypes in Amerindians and in black North and South Americans. *Tissue Antigens*. 1991;38(5):235-7.
- Loscalzo J, Kohane I, Barabasi AL. Barabasi, Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Mol Syst Biol*. 2007;3:124
- Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet*. 2001;17(9):502-10.
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*. 2003;33(2):177-82.

### Appendix 1. Useful web links

<p>International Hapmap Project (<a href="http://snp.cshl.org/">http://snp.cshl.org/</a>)</p> <p>1000 Genomes Project (<a href="http://www.1000genomes.org/page.php">http://www.1000genomes.org/page.php</a>)</p> <p>Database of Genomic Variants, a database of human copy number variations (<a href="http://projects.tcag.ca/variation/">http://projects.tcag.ca/variation/</a>)</p> <p>Entrez SNP (<a href="http://www.ncbi.nlm.nih.gov/snp">http://www.ncbi.nlm.nih.gov/snp</a>)</p> <p>Wellcome Trust Case Control Consortium (<a href="http://www.wtccc.org.uk/">http://www.wtccc.org.uk/</a>)</p>
---

## Appendix 2. Single nucleotide polymorphism identification and references

Gene	rs	SNP	References
IL-6 -174	rs1800795	G/C	Sainz J, Pérez E, Gómez-Lopera S, López-Fernández E, Moratalla L, Oyonarte S, Jurado M. Genetic variants of IL6 gene promoter influence on C-reactive protein levels but are not associated with susceptibility to invasive pulmonary aspergillosis in haematological patients. <i>Cytokine</i> . 2008;(3):268-78.
IL-10R 138	rs 4252272	A/G	Gasche C, Grundtner P, Zwirn P, Reinisch W, Shaw SH, Zdanov A, et al. Novel variants of the IL-10 receptor 1 affect inhibition of monocyte TNF-alpha production. <i>J Immunol</i> . 2003;170(11):5578-82
IL-10R 330	rs2229113	G/A	Gasche C, Grundtner P, Zwirn P, Reinisch W, Shaw SH, Zdanov A, et al. Novel variants of the IL-10 receptor 1 affect inhibition of monocyte TNF-alpha production. <i>J Immunol</i> . 2003;170(11):5578-82
IL-10 -592	rs1800872	C/A	Lech-Maranda E, Baseggio L, Bienvenu J, Charlot C, Berger F, Rigal D, et al. Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. <i>Blood</i> . 2004;103(9):3529-34.
IL-10 -819	rs1800871	C/T	Lech-Maranda E, Baseggio L, Bienvenu J, Charlot C, Berger F, Rigal D, et al. Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. <i>Blood</i> . 2004;103(9):3529-34.
IL-10 1082	rs1800896	A/G	Lech-Maranda E, Baseggio L, Bienvenu J, Charlot C, Berger F, Rigal D, et al. Interleukin-10 gene promoter polymorphisms influence the clinical outcome of diffuse large B-cell lymphoma. <i>Blood</i> . 2004;103(9):3529-34.
IL-10 -2849	rs6703630	G/A	Moraes MO, Santos AR, Schonkeren JJ, Vanderborght PR, Ottenhoff TH, Moraes ME, et al. Interleukin-10 promoter haplotypes are differently distributed in the Brazilian versus the Dutch population. <i>Immunogenetics</i> . 2003;54(12):896-9.
IL-10 -2763	rs6693899	C/A	Moraes MO, Santos AR, Schonkeren JJ, Vanderborght PR, Ottenhoff TH, Moraes ME, et al. Interleukin-10 promoter haplotypes are differently distributed in the Brazilian versus the Dutch population. <i>Immunogenetics</i> . 2003;54(12):896-9.
IL-10 -3575	rs1800890	T/A	Moraes MO, Santos AR, Schonkeren JJ, Vanderborght PR, Ottenhoff TH, Moraes ME, et al. Interleukin-10 promoter haplotypes are differently distributed in the Brazilian versus the Dutch population. <i>Immunogenetics</i> . 2003;54(12):896-9.
INFG +874	rs2430561	A/T	Pravica V, Perrey C, Stevens A, Lee JH, Hutchinson IV. A single nucleotide polymorphism in the first intron of the human IFN-gamma gene: absolute correlation with a polymorphic CA microsatellite marker of high IFN-gamma production. <i>Hum Immunol</i> . 2000;61(9):863-6.
IL-12B 1188	rs3212227	A/C	van Veen et al., <i>Genes and Immunity</i> . 2000;1:219-24
CCR5[delta]32	rs333	del	Shahbazi M, Ebadi H, Fathi D, Roshandel D, Mahamadhoseeni M, Rashidbaghan A, et al. CCR5-delta 32 allele is associated with the risk of developing multiple sclerosis in the Iranian population. <i>Cell Mol Neurobiol</i> . 2009;29(8):1205-9.
TLR4 299D/G 896	rs4986790	A/G	Ramasawmy R, Cunha-Neto E, Fae KC, Borba SC, Teixeira PC, Ferreira SC, et al. Heterozygosity for the S180L variant of MAL/TIRAP, a gene expressing an adaptor protein in the Toll-like receptor pathway, is associated with lower risk of developing chronic Chagas cardiomyopathy. <i>J Infect Dis</i> . 2009;199(12):1838-45.
MBL57 Gly/Asp	rs1800451	G/A	Ramasawmy R, Spina GS, Fae KC, Pereira AC, Nisihara R, Messias Reason IJ, et al. Association of mannose-binding lectin gene polymorphism but not of mannose-binding serine protease 2 with chronic severe aortic regurgitation of rheumatic etiology. <i>Clin Vaccine Immunol</i> . 2008;15(6):932-6 / Lipscombe RJ, Sumiya M, Hill AV, Lau YL, Levinsky RJ, Summerfield JA, et al. High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. <i>Hum Mol Genet</i> . 1992;1(9):709-15.
MBL52 Arg/Cys	rs5030737	C/T	Ramasawmy R, Spina GS, Fae KC, Pereira AC, Nisihara R, Messias Reason IJ, et al. Association of mannose-binding lectin gene polymorphism but not of mannose-binding serine protease 2 with chronic severe aortic regurgitation of rheumatic etiology. <i>Clin Vaccine Immunol</i> . 2008; 15(6):932-936/ Garred P. Mannose-binding lectin genetics: from A to Z. <i>Biochem Soc Trans</i> . 2008;36:1461-6.
MBL54 Gly/Asp	rs1800450	G/A	Ramasawmy R, Spina GS, Fae KC, Pereira AC, Nisihara R, Messias Reason IJ, et al. Association of mannose-binding lectin gene polymorphism but not of mannose-binding serine protease 2 with chronic severe aortic regurgitation of rheumatic etiology. <i>Clin Vaccine Immunol</i> . 2008;15(6):932-6.
MBL-221 X/Y	rs7096206	G/C	Laisk T, Peters M, Saare M, Haller-Kikkatalo K, Karro H, Salumets A. Association of CCR5, TLR2, TLR4 and MBL genetic variations with genital tract infections and tubal factor infertility. <i>J Reprod Immunol</i> . 2010; Jul 1. [Epub ahead of print]
MBL-550 H/L	rs11003125	C/G	Laisk T, Peters M, Saare M, Haller-Kikkatalo K, Karro H, Salumets A. Association of CCR5, TLR2, TLR4 and MBL genetic variations with genital tract infections and tubal factor infertility. <i>J Reprod Immunol</i> . 2010; Jul 1. [Epub ahead of print]
MBL +4 P/Q	rs7095891	C/T	Laisk T, Peters M, Saare M, Haller-Kikkatalo K, Karro H, Salumets A. Association of CCR5, TLR2, TLR4 and MBL genetic variations with genital tract infections and tubal factor infertility. <i>J Reprod Immunol</i> . 2010; Jul 1. [Epub ahead of print]
MASP2 120D/G	rs72550870	A/G	Ramasawmy R, Spina GS, Fae KC, Pereira AC, Nisihara R, Messias Reason IJ, et al. Association of mannose-binding lectin gene polymorphism but not of mannose-binding serine protease 2 with chronic severe aortic regurgitation of rheumatic etiology. <i>Clin Vaccine Immunol</i> . 2008;15(6):932-6.
NOD2 3020insC 1007fs	rs2066847	insC	Ghandil P, Chelala C, Dubois-Laforgue D, Senée V, Caillat-Zucman S, Kockum I, et al. Crohn's disease associated CARD15 (NOD2) variants are not involved in the susceptibility to type 1 diabetes. <i>Mol Genet Metab</i> . 2005;86(3):379-83.
MCP1 -2518	rs1024611	A/G	Rovin BH, Lu L, Saxena R. A novel polymorphism in the MCP-1 gene regulatory region that influences MCP-1 expression. <i>Biochem Biophys Res Commun</i> . 1999;259(2):344-8.
TNFA -238	rs361525	G/A	Migot-Nabias F, Mombo LE, Luty AJ, Dubois B, Nabias R, Bisseye C, et al. Human genetic factors related to susceptibility to mild malaria in Gabon. <i>Genes Immun</i> . 2000;1(7):435-41.
TNFA -308	rs1800629	G/A	Bittencourt PL, Palácios SA, Caçado EL, Porta G, Drigo S, Carrilho FJ, et al. Autoimmune hepatitis in Brazilian patients is not linked to tumor necrosis factor alpha polymorphisms at position -308. <i>J Hepatol</i> . 2001;35(1):24-8.
LTA +80	rs2239704	A/C	Ramasawmy R, Fae KC, Cunha-Neto E, Müller NG, Cavalcanti VL, Ferreira RC, et al. Polymorphisms in the gene for lymphotoxin-alpha predispose to chronic Chagas cardiomyopathy. <i>J Infect Dis</i> . 2007;196(12):1836-43
LTA +252	rs909253	A/G	Ramasawmy R, Fae KC, Cunha-Neto E, Müller NG, Cavalcanti VL, Ferreira RC, et al. Polymorphisms in the gene for lymphotoxin-alpha predispose to chronic Chagas cardiomyopathy. <i>J Infect Dis</i> . 2007;196(12):1836-43

Continue...

...Continuation

**Appendix 2. Single nucleotide polymorphism identification and references**

NFkB1L1 -63	rs2071592	T/A	Ramasawmy R, Faé KC, Cunha-Neto E, Borba SC, Ianni B, Mady C, et al. Variants in the promoter region of IKBL/NFKBIL1 gene may mark susceptibility to the development of chronic Chagas' cardiomyopathy among Trypanosoma cruzi-infected individuals. <i>Mol Immunol.</i> 2008;45(1):283-8
BAT1 -22	rs2239527	C/G	Ramasawmy R, Cunha-Neto E, Faé KC, Müller NG, Cavalcanti VL, Drigo SA, et al. BAT1, a putative anti-inflammatory gene, is associated with chronic Chagas cardiomyopathy. <i>J Infect Dis.</i> 2006;193(10):1394-9.
BAT1 -348		C/T	Ramasawmy R, Cunha-Neto E, Faé KC, Müller NG, Cavalcanti VL, Drigo SA, et al. BAT1, a putative anti-inflammatory gene, is associated with chronic Chagas cardiomyopathy. <i>J Infect Dis.</i> 2006;193(10):1394-9.
LST1 +290		A/G	Rollinger-Holzinger I, Eibl B, Pauly M, Griesser U, Hentges F, Auer B, et al. LST1: a gene with extensive alternative splicing and immunomodulatory function. <i>J Immunol.</i> 2000;164(6):3169-76.
IL-4 -589	rs2243250	C/T	Verra F, Luoni G, Calissano C, Troye-Blomberg M, Perlmann P, Perlmann H, et al. IL4-589C/T polymorphism and IgE levels in severe malaria. <i>Acta Trop.</i> 2004;90(2):205-9.
IL-4 +33	rs2070874	C/T	Shibata N, Ohnuma T, Takahashi T, Baba H, Ishizuka T, Ohtsuka M, et al. The effect of IL4 +33C/T polymorphism on risk of Japanese sporadic Alzheimer's disease. <i>Neurosci Lett.</i> 2002;323(2):161-3.
IL-4 +3017	rs2227284	G/T	Basehore MJ, Howard TD, Lange LA, Moore WC, Hawkins GA, Marshik PL, et al. A comprehensive evaluation of IL4 variants in ethnically diverse populations: association of total serum IgE levels and asthma in white subjects. <i>J Allergy Clin Immunol.</i> 2004;114(1):80-7.
IL-13 Q110R	rs20541	G/A	Heinzmann A, Mao XQ, Akaiwa M, Kreomer RT, Gao PS, Ohshima K, et al. Genetic variants of IL-13 signalling and human asthma and atopy. <i>Hum Mol Genet.</i> 2000;9(4):549-59.
IL-4R I50V	rs1805011	A/G	Mitsuyasu H, Yanagihara Y, Mao XQ, Gao PS, Arinobu Y, Ihara K, et al. Cutting edge: dominant effect of Ile50Val variant of the human IL-4 receptor alpha-chain in IgE synthesis. <i>J Immunol.</i> 1999;162(3):1227-31.
IL-5 -746	rs2069812	C/T	Hong SJ, Lee SY, Kim HB, Kim JH, Kim BS, Choi SO, et al. IL-5 and thromboxane A2 receptor gene polymorphisms are associated with decreased pulmonary function in Korean children with atopic asthma. <i>J Allergy Clin Immunol.</i> 2005;115(4):758-63.
IL-13RA1 +1398		A/G	Heinzmann A, Mao XQ, Akaiwa M, Kreomer RT, Gao PS, Ohshima K, et al. Genetic variants of IL-13 signalling and human asthma and atopy. <i>Hum Mol Genet.</i> 2000;9(4):549-59.
CTLA4 -49	rs231775	G/A	Fan LY, Tu XQ, Cheng QB, Zhu Y, Feltens R, Pfeiffer T, et al. Cytotoxic T lymphocyte associated antigen-4 gene polymorphisms confer susceptibility to primary biliary cirrhosis and autoimmune hepatitis in Chinese population. <i>World J Gastroenterol.</i> 2004;10(20):3056-9.
CTLA4 -60	rs3087243	A/G	Brophy K, Ryan AW, Thornton JM, Abuzakouk M, Fitzgerald AP, McLoughlin RM, O'morain C, Kennedy NP, Stevens FM, Feighery C, Kelleher D, McManus R. Haplotypes in the CTLA4 region are associated with coeliac disease in the Irish population. <i>Genes Immun.</i> 2006;7(1):19-26 / Hradsky O, Dusatkova P, Lenicek M, Bronsky J, Nevoral J, Vitek L, Lukas M, Zeniskova I, Cinek O. The CTLA4 variants may interact with the IL23R- and NOD2-conferred risk in development of Crohn's disease. <i>BMC Med Genet.</i> 2010;11:91.
CTLA -318	rs5742909	C/T	Fan LY, Tu XQ, Cheng QB, Zhu Y, Feltens R, Pfeiffer T, et al. Cytotoxic T lymphocyte associated antigen-4 gene polymorphisms confer susceptibility to primary biliary cirrhosis and autoimmune hepatitis in Chinese population. <i>World J Gastroenterol.</i> 2004;10(20):3056-9.

SNP: single nucleotide polymorphism.