Genetic haplotypes of Th-2 immune signalling link allergy to enhanced protection to parasitic worms

Maria Moller1, Mike B. Gravenor1, Stephen E. Roberts1, Dejan Sun2, Peisong Gao1,† and Julian M. Hopkin1,*

1School of Medicine, Institute of Life Science, Swansea University, Singleton Park, Swansea SA2 8PP, UK and 2Institute of Parasitology, Shanghai, China

Received March 2, 2007; Revised and Accepted May 10, 2007

Parasitic worm infection, allergy and asthma involve increased IgE production, eosinophil activity, mucus secretion and smooth muscle reactivity, effected through Th-2 immune signalling. These pathological features of allergic disorder, common in developed countries, appear to be protective features in resistance to parasitic worm infections prevalent in many developing countries. We investigated how genetic variation in the Th-2 signalling transduction molecule STAT6 relates to these clinical disorders, using immune phenotyping by serum IgE levels and haplotyping nine STAT6 genetic variants in a rural Chinese population, where Ascaris infection is prevalent, and an urban UK population where Ascaris is largely unknown but asthma and allergy are prevalent. We show for the first time that STAT6 haplotypes relate clearly to IgE levels, allergy and worm burden. The haplotypes segregated into two groups: those with raised IgE/low worm burden tended to have increased risk of allergic disorder, whereas low IgE/high worm burden tended to have a reduced risk of allergies. By estimating the mean worm burden for each haplotype in China and the relative risk of asthma for the matching haplotype in the UK, we draw a cross-population comparison and show a negative correlation between worm burden and expected risk of asthma. These data imply that the origin of common up-regulating variants of Th-2 signalling, involving STAT6, promotes asthma and allergy in developed countries, whereas in developing countries it protects against parasitic worm infections. Selective evolutionary mechanisms, driven by parasitic worm infection, may underlie the genetic contribution to risk of allergy and asthma in humans.

INTRODUCTION

The substantial genetic variation evident in the human population, which impacts on almost all physiological functions including the immune system, is moulded by well-recognized mechanisms including mutation, genetic drift and selective advantage. In relation to selection, the constant threat to humans from parasites plays a vital role. A long-established exemplar is the sickle variant in beta-globin in humans which when present in homozygous form results in severe anaemia and when present in heterozygous state confers increased fitness through increased resistance to cerebral malaria. High prevalence of the variant is confined to regions where malaria is or has recently been endemic. Parasitic challenge also influences genetic variation within the human immune system itself, the dominant system shaping resistance to parasitic infections, together with other important factors such as nutritional status. Recent comprehensive linkage studies identify several susceptibility loci, emphasizing the importance of human genetic variation on protection against parasitic worms (1).

We and others have studied genetic variation in one major component of immune function, Th-2 immune signalling (2–5). Th-2 immune signalling operates principally through the cytokines interleukin (IL)-13 and IL-4, both of which utilize the transduction and transactivating transcription factor STAT6, to orchestrate transcriptional activation of a number of loci and hence upregulate functional activity leading to increased synthesis of immunoglobulin E (IgE) and a characteristic set of increased mucosal protective mechanisms.

*To whom correspondence should be addressed at: School of Medicine, University of Wales, Swansea SA2 8PP, UK. Tel: +44 1792513062; Fax: +44 1792513054; Email: j.m.hopkin@swansea.ac.uk
†Present address: Division of Allergy and Clinical Immunology, The Johns Hopkins University, Baltimore, MD 21224, USA.
infection, a parasitic worm infection endemic to China, Ascaris lumbricoides. We have tested for the possibility of a balancing advantage for these variants against a protective activity. The Th-2 response is shown moderate up-regulating effects for certain variants within the UK and Chinese populations.

In addition to these protective activities, the Th-2 response is also now recognized to play a key role in promoting the bronchial inflammation that is an important pathological element of the common chronic disorder of asthma, in which many children and young adults suffer from chronic chest tightness, wheeziness and cough with sputum. Asthma has been long recognized to show familial clustering suggestive of the influence of genetic factors, but also shows striking world-wide geographical variation with a consistent emphasis towards increased prevalence in highly developed countries, the latter suggesting the impact of environmental influences. Many genetic studies have been conducted on asthma and its associated phenotypes of atopic allergic disorder, with high IgE levels often directed to common environmental antigens such as house dust mite proteins and pollens and the clinical syndromes of allergic rhinitis (hay fever) and allergic dermatitis (eczema). A number of reproducible chromosomal linkages have been observed for asthma, some population specific and others general. These indicate that the genetic origins of asthma are polygenic and comprise the actions, in combination, of genetic variants with small to moderate functional effects at a number of distinct loci. Among these prevalent variants, some have been identified in Th-2 immune signalling, within the immune cytokine IL-13 itself, one of its receptor subunits IL-4 receptor alpha (IL-4Rα), and STAT6 (11–13). Reproducible associations have been observed between these variants and asthma and/or allergy and in vitro functional investigations have shown moderate up-regulating effects for certain variants within each of these three loci.

Because the asthma and allergy-promoting variants in Th-2 immune signalling occur at high frequencies and because of the role of Th-2 immune mechanisms in both allergy and in protection to parasitic worms, we have tested for the possibility of a balancing advantage for these variants against a parasitic worm infection. We conducted a genetic association study in a population of rural Chinese children, testing for association between the variants and the burden of Ascaris lumbricoides infection, a parasitic worm infection endemic in that region; we found that a 3′-UTR variant of STAT6 associated strongly with diminished worm burden (14). We also found a weaker association for the degree of worm burden to a genetic variant in the IL-13 gene and a significant interaction of STAT6 with IFN-γ (14).

Out of the 15 genes involved in either the Th-1 or Th-2 signalling pathways that we previously tested for worm burden associations, STAT6 was clearly the strongest candidate to analyse further. In order to test this finding more rigorously, and to clarify the genetic association, we have now conducted detailed haplotype analysis of nine genetic variants spanning all of STAT6, including four microsatellites and five single nucleotide polymorphisms (SNPs), in the same Chinese population and in comparison with a British population. In both populations, we included serum IgE levels, with the phenotypes of asthma, hay fever or eczema (UK), and the burden of Ascaris infection (China).

Results

Genotypes and haplotype reconstruction

The location of the markers across the STAT6 gene is shown in Figure 1, and the genotype frequency distributions for all markers in both populations are shown in Figure 2. The dinucleotide frequencies show a similar pattern in both UK and China, whereas the tetranucleotide in intron 13 has a completely different distribution pattern between the populations; this greatly reduced the opportunities for cross-population statistical comparison of haplotype effects that included this polymorphism. The majority of the UK tetranucleotide alleles are much shorter than those in the Chinese and moreover there is a more even distribution among alleles in the Chinese. The tetranucleotides display high diversity with 36 alleles identified by sequence length. We sequenced a limited set of these alleles and discovered that some with identical length had different repeat pattern, i.e. a different origin so the true number of alleles exceeds 36. The microsatellite denoted GT3 has also been genotyped in an Indian population (15) where they showed that the repeat length of 15 and 17 GTs were the major alleles compared with the repeat lengths of 13 and 15 GTs in both the UK and Chinese with 13 repeats most common in UK and 15 repeats most common in China. Two different studies in Japanese populations showed the 15-repeat allele in highest frequency and the 17-repeat allele not present at all (16,17).

SNPs 1, 2 and 4 have a very dissimilar distribution between the UK and the Chinese, and the SNPs 3 and 5 have more or less identical distribution (Fig. 2). All five SNPs show, as expected, the same frequencies in the British as a previous study of a German population (18) and a Finnish population (19). Pairwise linkage disequilibrium (LED) tests showed that all dinucleotide and SNP markers are in strong LED in both populations (data not shown).

Three validated non-synonymous SNPs have been reported for STAT6 in the NCBI SNP database (rs11771202, rs3024952 and rs2626577), and we screened part of the British population for variation at these sites (between 30 and 100 individuals). No polymorphism was detected. It is quite striking that no common variant in the coding region of STAT6 has as yet been detected, although several screenings for polymorphisms across the gene have been carried out (18,20,21). This is a remarkable contrast to the eight common amino acid changes detected in the IL4Rα gene which is a molecule that forms part of the STAT6 signalling pathway.

Figure 1. Schematic representation of the location of the STAT6 polymorphisms genotyped in the UK and Chinese population.
The estimated haplotype frequencies are summarized in Supplementary Material, Wright's $F$-statistics revealing a significant difference between the population frequencies ($P < 0.001$).

**Genotype/haplotype and phenotype associations**

Using score tests, we first identified those haplotypes at either the 5'- or 3'-UTR that were significantly associated with either IgE levels or worm burden. For each of these selected haplotypes, we then tested for potential association with allergic disorder (asthma, rhinitis or eczema). The results are shown in Table 1. With few exceptions, there was a remarkably clear pattern: those haplotypes showing significantly raised IgE/low worm burden also showed increased risk of asthma (three significant at $P < 0.05$), increased risk of rhinitis (six significant) and increased risk of eczema (three significant) (upper half of Table 1). In contrast, those haplotypes identified as significantly low IgE/high worm burden showed reduced risk of asthma, rhinitis (two significant) and eczema (one significant at 10% level) (lower half of Table 1).

Using regression analyses (haplo.glm), we estimated mean phenotypic outcomes for all STAT6 haplotypes (defined by the three 5'-UTR and the five 3'-UTR loci) present in the populations at a frequency $> 1\%$. Trends in the relationship between serum IgE levels, burden of Ascaris infection and allergic disorder are plotted in Figure 3. In the UK population we identified 15 haplotypes, showing a positive correlation between risk of asthma and IgE level (Fig. 3A; $P < 0.05$ for logarithmic trend). While in the Chinese population, the 14 most common haplotypes demonstrated a negative correlation between worm burden and IgE (Fig. 3B; $P < 0.05$ for linear trend).

By combining the phenotype estimates for those haplotypes present in both the UK and Chinese populations, we can generate a novel cross-population comparison that predicts the haplotype-specific asthma risk, based on worm burden. This predicts an inverse relationship between worm burden and asthma risk (Fig. 3C). Note that since asthma is confined to the UK population and Ascaris to the Chinese population,
Table 1. Score tests for STAT6 haplotype associations to IgE, worm burden and allergic disease in a British (Brit) and Chinese (Chin) population

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>IgE level</th>
<th>Asthma</th>
<th>Rhinitis</th>
<th>Eczema</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-UTR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>153-123-212</td>
<td>0.71/0.043</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>151-123-212</td>
<td>0.007/0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>153-123-N</td>
<td>0.81/0.032</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>151-123-N</td>
<td>0.013/0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-123-212</td>
<td>0.042/0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-115-216</td>
<td>0.014/0.043</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Haplotypes of the upper half of the table (A) are associated with raised IgE levels, low egg counts and susceptibility to allergic disease (exceptions marker $\gamma$); haplotypes of the lower half of the table (B) are associated with low IgE levels, high egg counts and protection against allergic disease. Mean IgE and egg counts are presented as a % above or below the value for the most common (base) haplotype in the population as defined by the specific set of haplotypes in the analysis.

*Haplotype frequency below 1%.

Human Molecular Genetics, 2007, Vol. 16, No. 15
then this is illustrative of the likely relationship. The prediction reflects the real relationships within each population (Fig. 3A and B) and a correlation between relative IgE levels for given haplotypes common to both populations. Very similar patterns were obtained using the combined phenotype ‘allergic disease’ (asthma/hay-fever/eczema) in the UK population.

**DISCUSSION**

These associations imply that certain variants in the transduction molecule STAT6 upregulate STAT6 function to cause high IgE levels, and the different clinical phenotypes in the UK and Chinese populations. Relatively little is known about the functional aspects of any variant of STAT6. Only three rare amino acid-changing variants are known, and no information is available on any functional result. Genetic variation is prevalent in the promoter and UTR regions as well as the intronic regions and includes both SNPs and microsatellites. In the NCBI SNP database as of May 2006, 125 polymorphisms had been reported in the STAT6 region.

For one of these variants, a microsatellite in exon 1, we and our colleagues have previously been able to show that shorter repeats (13 repeat units) of this dinucleotide structure enhance transcription of STAT6 in vitro, and hence promote upregulated Th-2 immune signalling and hence high IgE levels, prominent eosinophil activity, enhanced mucus secretion and increased muscle contractivity (5). Such activities are typical pathological features in asthma and underlie the observed fluctuation of lung function in asthmatics (as in this group with peak flow variability or >30%), and the increased air flow obstruction in asthma in response to a range of inhaled challenges. Here, we show that this shorter allele is present in the haplotypes that are associated with high IgE levels in both British and Chinese, with allergy in the British, and low levels of worm infection in the Chinese (Table 1). In experimental mice, genetic knock-out of STAT6 fully inhibits the production of experimental eosinophilic bronchitis and bronchial reactivity (22). Hence the common genetic variation we observe in STAT6 can increase the risk of asthma through upregulated function acting with other common variants at other loci and environmental factors to cause the disease.

This Chinese rural area shows endemic infection with *Ascaris*, most notable in children, and promoted by social and environmental factors that promote transmission of *Ascaris* by faecal–oral spread, such as the fertilization of home-grown crops by ‘night soil’. For ascariasis and other parasitic worm infections in endemic areas, it has been repeatedly observed that the intensity of infection varies significantly among children, and that some 20% of any human population are heavily infected, suffer clinical consequences and are the principal sources of cycles of infection in the community [‘20% rule’ (23)]. Genetic linkage studies (e.g. 24,25) suggest that human genetic variation contributes significantly to this variation in the intensity of individual infections with parasitic worms, for instance reproducible linkage between intensity of schistosomiasis infection and a chromosome 5 locus. Strong association has since been found on chromosome 5, between a promoter region variant (−1055C/T) of the major Th-2 cytokine IL-13 and increased resistance to schistosomiasis (26). In this context it is interesting to note that in our previous study, when screening for genetic association in *Ascaris* burden in the Chinese population and single common variants at a number of immune cytokine pathways, we observed that the −1055 variant of IL-13 contributed with a common 3′-region variant of STAT6 to the association with diminished *Ascaris* burden in the regression model (14). The results point to the likely importance of genetic variation of Th-2 immune signalling in influencing the intensity of parasitic worm infection and the likelihood that the genetic influences on burden are polygenic and comprise actions of moderate effect to common variants at a number of loci. This accords with the polygenic influences on the development of
allergy and asthma which include variation of TH-2 immune signalling, and other genes including the epithelium-associated protein, ADAM33, the β2 adrenergic receptor on muscle and the beta-subunit of the high affinity IgE receptor (FCER1B) on mast cells (27–29). Of interest, the variant of FCER1B associates with high IgE levels in parasitized aboriginal children, but no data were available to test directly the association with worm burden.

Our findings here specifically suggest that genetic variants of STAT6 can confer enhanced resistance to ascariasis, and imply that through this they confer increased biological fitness in environments where Ascaris worm is prevalent. However, since there are currently no proper estimates of the impact of ascariasis on biological fitness (fertility), this needs formal investigation. Nevertheless, there is substantial information available on the ill-health due to ascariasis. Ascariasis is estimated to occur in 1.3–1.5 billion of the world’s population, involving at least 150 of the world’s 240 states, and with especially high prevalence in East Asia, Africa and South America (30,31). Infected individuals with the highest burden of worm carriage suffer the most severe consequences (32,33). Some 200 million (4% of the world’s population) have some serious morbidity due to ascariasis, and this includes growth retardation (34) which may be permanent (35); impaired physical fitness (35), anaemia (36); cognitive impairment (37); various intestinal complications such as obstruction and perforation which are an important source of surgical admissions and procedures in affected countries (38,39); and increased predisposition to other diseases such as severe malaria (40). Although ascariasis is recognized as a disease of high morbidity, there are additionally some 20 000 deaths directly attributable to ascariasis per annum (41). Overall, parasitic worms with their evasive/subversive actions against human immunity (42) provide a serious challenge to human health, and this implies an important impact on biological fitness and strong impact on human evolution. We suggest here that one such effect may be the selection for up-regulating variants of human TH-2 immune signalling. We propose that such variants have been disseminated across many human populations in the face of widespread challenge from parasitic worms. Some of these populations are known to have only recently achieved control of ascariasis and other geo-helminths, such as Japan (43), whereas others, such as Europe including Britain, appear to have had more remote but still significant exposures to parasitic worms (44). In the British population, where parasitic worm infection including ascariasis are absent, we observe strong associations between the up-regulating Th-2 variants and a range of manifestations including raised IgE levels, and the clinical allergic disorders of rhinitis (hay fever), dermatitis (eczema) and asthma.

Of further note in our study is that the prevalence of these clinical allergic syndromes was extremely low in the Chinese children. This observation matches those from other recent studies in which individuals with helminth infection show less clinical allergic disorder, even though they show similarly positive skin prick tests to common allergens (45–47). There is now increasing evidence that the mechanism behind the suppression of clinical allergic disorder by helminths is mediated by the parasite’s induction of CD4(+)CD25(+) regulatory T-cells (Treg) which downregulate both Th-2 and Th-1 immune phenomena at mucosal surfaces, for instance in the respiratory tract (48,49). Other potential mechanisms, including the saturation of IgE receptors on mast cells by helminth-induced polyclonal IgE or the induction of the regulatory cytokine IL-10 (50), seem less important. These findings have led to the consideration of worm products as preventive or therapeutic agents in asthma (51). Thus insights from comparative studies from different countries in which human genetic variation in Th-2 immune signalling provides a common link may provide for novel strategies towards improved treatment and prevention of allergic clinical disorder which is rising to epidemic proportions in certain countries. Parasitic worm infections in developing countries and allergic and asthmatic disorder in developed countries, both intimately involve Th-2 immune signalling mechanisms. Our data show a clear relationship between up-regulating genetic variants of this pathway and increased risk of allergy, and increased resistance to infection. This points to a selective evolutionary process, the clinical consequences of which are dependent on the prevailing environmental conditions.

MATERIALS AND METHODS

Subjects

Caucasian population from Oxfordshire, UK. Three hundred subjects were studied, 150 suffering from asthma and requiring on-going treatment at a specialist respiratory unit, and 150 un-selected who were then attending hospital for routine obstetric care. The asthmatic subjects had specialist physician-diagnosed asthma with recurrent breathlessness and chest tightness requiring continuous treatment, physician documented wheeze and documented labile airflow obstruction with serial peak expiratory rate variability of >30%. The diagnosis of rhinitis was based on chronic nasal obstruction or congestion, rhinorhoea and sneezing. The obstetric care population was un-selected and thus included, by chance, some subjects with family diagnosed and treated asthma and rhinitis because of the prevalence of these disorders. All subjects had their skin examined for the presence of eczema applying the diagnostic criteria of Hanafin and Rajka and had total serum IgE assayed by solid phase immunoassay (Pharmacia CAP system). Ascaris infection does not occur at this location and stool was not examined for worm eggs. The mean age of asthmatics was 26 and of the obstetric care patients, 22.

Xing-Chang Province population, People’s Republic of China. A total of 612 children, aged 11–15 years, were recruited as an un-selected population from four Schools in Xing-Chang district of Shanghai as part of an Ascaris testing and eradication programme (14). Infection of the roundworm Ascaris lumbricoides is prevalent in this area. The children had fresh stool examined for Ascaris eggs by the Kato-Katz method and egg counts quantified per gram of stool (14). If the stool had eggs from other parasitic worms it was also recorded.

The children responded to the ISAAC (The International Study of Asthma and Allergies in Childhood) video questionnaire in respiratory allergy (asthma and hay fever) and were
examined for eczema. Total serum IgE was assayed as for the UK population.

**Genotyping of microsatellites and single nucleotide polymorphisms**

Both the UK and Xing-Chang populations provided circulating white cells for isolation of genomic DNA. The microsatellites, three with dinucleotide GT-repeats denoted GT1, GT2 and GT3, located in the 5′-region and one tetranucleotide microsatellite, (CTTT), in intron 13 (Fig. 1) were amplified with a PCR profile as follows: denaturing 6 min at 94°C; 37 cycles of 94°C, 30s; 62°C for 30 s; 72°C for 60 s followed by an additional extension of 10 min at 72°C. For GT2 and the tetranucleotide the annealing temperature was 59°C. Each forward primer was end-labelled using WellRED dyes (Proligo). The fragments were subsequently analysed with capillary electrophoresis using the Beckman Coulter CEQ8000 machine together with the CEQ8000 software (Genetic Analysis System, version 5.0.360, Beckmann Coulter). The five SNPs in the 3′-UTR were genotyped by sequencing in the UK population and by the TaqMan (ABI) assay on the Chinese population. Sequencing was carried out by cycle sequencing (BigDye, Perkin-Elmer) and analysed on the ABI310 capillary electrophoresis sequence analyzer (Perkin-Elmer). The TaqMan analysis was carried out by MRC geneservice (Cambridge, UK). All primer sequences, marker description and fragment details are presented in Supplementary Material.

**Statistical analysis and haplotype reconstruction**

We used haplotype reconstruction (for the case of unknown linkage phase) and analysis tools provided by the Haplo.Stats suite of software (52) run on the R statistical environment (53). Our first aim was to identify whether any individual haplotypes were significantly associated with IgE/worm burden using the haplo.score function. The score tests were then grouped according to IgE and worm burden level (high/low), and a further set of score tests performed to test whether those haplotypes identified as having IgE/worm effects were significantly associated with any of the allergic phenotypes (asthma/eczema/rhinitis). Covariates (age and sex) did not confound the haplotype associations.

To guard against false positives due to multiple testing, we report *P*-values for the IgE/worm associations with individual haplotypes only if the global *P*-value for those under consideration was significant at the 5% level. We did not make further adjustment of *P*-values due to the increased chance of type II errors, in particular for haplotypes present at low frequency. Although construction of the haplotype across the whole dataset was possible (see below), due to the very low frequency of many of the resulting haplotypes, the analysis of phenotype effects was performed separately for the dinucleotide, tetranucleotide and SNP regions in order to increase the power for those haplotypes defined in each region.

In the above score tests, our aim was to first identify those specific haplotypes with significant phenotype associations and secondly, to group them according to effect on allergy. In order to inspect the allergy trends in more detail, we considered the ‘full’ haplotype defined by the three 5′- and five 3′-loci and estimated, for all haplotypes, mean values for asthma risk, IgE UK, IgE China and worm burden. This was done by fitting a set of four regression models using the haplo.glm function. This approach is illustrative, allowing us to demonstrate overall trends for all haplotypes and phenotypes. It is also useful for making a hypothetical quantitative comparison of phenotypic response across the populations, and inferring (indirectly) the relationship between asthma risk and worm burden. To illustrate this we plot the asthma risk in the UK population according to those haplotypes for which we have measured the worm burden in the Chinese population. We note that in this final comparison, only seven common haplotypes (for which accurate mean values could be calculated) matched exactly at all loci in the Chinese and UK populations. To expand the sample size, the worm burden/asthma risk comparison was also inferred for an additional 12 haplotypes that differed at one locus only.

In constructing the haplotypes for the regression analysis we were restricted to using a reduced version of the dinucleotide GT1, due to its considerable variability (and reduced overlap between populations), and coded individuals according to the tertiles (the median, ‘low’ and ‘high’ numbers of repeats). In each case, worm burden and IgE levels were log-transformed before analysis in order to improve the fit of the regression model, and age and sex covariates checked for confounding effects. *F*-statistics for measurement of genetic distances between populations were analysed using ARLEQUIN version 3.0 (54).

**Conflict of Interest statement.** None declared.

**REFERENCES**


