Review

The evolving story of human leukocyte antigen and the immunogenetics of peanut allergy

Jonathan A. Hemler, MD*; Elizabeth J. Phillips, MD†,‡,§; Simon A. Mallal, MBBS†,‡,§; and Peggy L. Kendall, MD*,*

ABSTRACT

Objective: Peanut allergy (PA) clearly has a heritable component. Specific genetic contributions are unknown, but human leukocyte antigen (HLA) loci are obvious candidates. This review focuses on emerging studies of HLA associations with PA.

Data Sources: PubMed was searched with no time limitations using key terms human leukocyte antigen, HLA, MHC, peanut, peanut hypersensitivity, and peanut allergy.

Study Selections: Qualifying studies were English-language reports of genetic analyses examining PA and HLA associations.

Results: Seven relevant citations were identified, which were published from 1996 to 2015. Early studies using candidate gene approaches found associations between PA and HLA-DR and -DQ alleles (HLA-DRB1*08 and DQB1*06:03P) when comparing subjects with peanut allergy with nonallergic unrelated control groups. No significant associations were found between siblings with and without peanut allergy. However, a recent large genomewide association study of patients with peanut allergy and their family members found 2 PA-associated single-nucleotide polymorphisms (rs9275596 and rs7192) mapping to regions involving the HLA-DR and HLA-DQ genes. Associations with differential DNA methylation partly mediated the associations between PA and single-nucleotide polymorphisms.

Conclusion: Early studies using candidate gene approaches identified HLA associations with PA compared with the general population, suggesting a link with atopy but failing to identify a PA-specific association. These studies had various limitations that included small samples. The most compelling evidence for a PA-specific HLA association comes from a genomewide association study, which examined the entire genome in large, well-defined, related cohorts. More research is needed to validate and replicate these findings, to perform fine genetic mapping of specific HLA loci, and to demonstrate underlying mechanisms of HLA contributions to PA.

Introduction

Food allergy is defined as an immune response that occurs reproducibly to a given food, typically an immunoglobulin E (IgE)-mediated clinical reaction to specific protein epitopes.1,−3 During the past 20 to 30 years, food allergy has grown into a major public health problem, especially in developed countries, including the United States.4−6 Some recent studies have estimated the prevalence at 2% to 8% of the general population.4 Peanut allergy is a common type of food allergy that accounts for a disproportionate number of fatal and near-fatal anaphylactic events among all common food allergens.7−9 Some studies have estimated that this condition affects 1.5% to 3% of all children, making it one of the most common chronic conditions of childhood.4,10−12 Therefore, identifying additional prevention and treatment strategies for this disease is of major clinical importance.

One hindrance to developing such strategies is lack of understanding of the mechanisms governing the biological development...
of peanut allergy in vivo. There are data to suggest that genetic factors play a significant role in this disease based on familial aggregation\textsuperscript{13,14} and twin\textsuperscript{15} studies. Therefore, studying the interplay of certain genes in the development of peanut allergy is warranted. One particular area of the genome that has garnered interest in regard to atopic disease, including food allergy, is the human leukocyte antigen (HLA) locus. Some recent studies have suggested that HLA plays a role in these diseases\textsuperscript{16,17} including a recent large-scale genomewide association study (GWAS) that identified the HLA-DQB1 locus to be associated with asthma.\textsuperscript{18} There also is a significant body of literature implicating the major histocompatibility complex (MHC) in the immunopathogenesis of drug hypersensitivity reactions, suggesting that HLA class I and/or class II alleles might be necessary but not sufficient for the development of severe immunologically mediated drug reactions.\textsuperscript{19–21} In humans, genes of the HLA family are part of the MHC locus located on chromosome 6p21.\textsuperscript{22,23} These genes are crucial to the functioning of the immune system and to the development of allergy (Fig 1). Protein antigens, including those found in foods, are internalized by antigen-presenting cells such as dendritic cells, macrophages, and B cells. These antigens are processed into peptides (12–25 amino acids long) that bind to a limited number of MHC class II molecules (Fig 2).\textsuperscript{24} Antigen-specific T-cell receptors (TCRs) on CD4\textsuperscript{+} T cells recognize their peptide in complex with MHC class II. This, in combination with the presence of various cytokines and different cell–cell interactions, leads to activation of the T cells and their development into T-helper type 2 (Th2) cells. B cells also are antigen-specific, recognizing proteins by membrane-bound antibodies on their surfaces that bind and internalize proteins, allowing concentration and processing of specific antigens for loading into MHC class II. Thus, TCRs on allergen-specific CD4\textsuperscript{+} Th2 cells interact with MHC peptide on B cells that recognize the same antigen (Fig 1). This cognate T–B interaction, in concert with the production of Th2 cytokines, induces IgE class switch and further

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Diagram illustrating the role of major histocompatibility complex (MHC) class II in antigen presentation of allergens to T cells and antigen-specific T–B interactions that give rise to allergen-specific immunoglobulin E (IgE). APC, antigen presenting cell; IL, interleukin; TH, T-helper cell. Image reprinted with permission from Medscape Drugs & Diseases (http://emedicine.medscape.com/), 2015, available at: http://emedicine.medscape.com/article/136217-overview.}
\end{figure}
The study's aim was to determine whether peanut-specific allergy or more general allergy involved shared TCR-α, TCR-β, or HLA haplotypes. They found that peanut allergy segregated with the paternal HLA-DR4 haplotype and not with the TCR-α or TCR-β alleles. This study is obviously limited by its small sample; however, it did offer the first evidence that HLA genes might play a role in peanut allergy.

Howell et al.29 performed the first larger study evaluating HLA class II loci and peanut allergy. They examined 50 subjects with peanut allergy, 34 siblings without peanut allergy, parents without peanut allergy, and 293 unrelated controls. Genotype frequencies were compared between the overall study group (patients with peanut allergy plus family members without peanut allergy) and the control group, between patients with peanut allergy and the control group, and between patients with peanut allergy and their siblings without peanut allergy (with and without atopy). HLA-DRB1*08 and DQB1*04 were increased in frequency in the overall study group and in patients with peanut allergy compared with controls, but no significant HLA class II associations were found when comparing siblings with vs without peanut allergy. These findings suggested that the identified HLA alleles are important for atopy overall rather than just for peanut allergy.

Another interesting study by Boehncke et al.30 examined cohorts of patients with pollen allergy and pollen-associated food allergy and analyzed the HLA class II genotypes of these patients. With respect to peanut allergy, patients who were allergic to grass pollen and peanut were analyzed, given the association between those allergens that the investigators previously reported.31 Patients with peanut allergy showed strong association with HLA-DRB1*08 compared with the control population, the same allele found in the previous study by Howell et al.29 However, this association lost significance when compared with patients who were allergic to grass pollen, again suggesting that this allele might be a marker for atopy rather than being specific for peanut allergy. Major limitations of this study included the small sample (9 patients with peanut allergy) and the fact that it was not designed specifically to identify associations between HLA and peanut allergy.

The final study using a candidate gene approach that identified an HLA association with peanut allergy was performed in 2013 by Madore et al.31 They examined the HLA-DQB1 locus owing to its association with asthma found in a large GWAS.18 The study

**Figure 2. Major histocompatibility complex (MHC) class II structure and function.** (A) Schematic of MHC class II consisting of α and β chains, with extracellular domains, transmembrane region, and cytoplasmic domains. (B) Ribbon diagram of a typical MHC class II protein, with β-pleated sheet flanked by α helices, forming a groove for a peptide antigen (red), aa, amino acid. Figure adapted from White K, Gaudieri S, Phillips E. HLA and the pharmacogenomics of drug hypersensitivity. In: Padmanabhan S, ed. Handbook of Pharmacogenomics and Stratified Medicine. Vol 1. Philadelphia, PA: Elsevier; 2014:437–459.

**Table 1.** What is contained within the table?

<table>
<thead>
<tr>
<th>MHC CLASS II</th>
<th>STRUCTURE</th>
<th>Alpha and beta chain heterodimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOURCE</td>
<td>Antigen presenting cells</td>
<td></td>
</tr>
<tr>
<td>PEPTIDE</td>
<td>Endosomal (extracellular)</td>
<td></td>
</tr>
<tr>
<td>LOCUS</td>
<td>DR, DQ and DP</td>
<td></td>
</tr>
<tr>
<td>FUNCTION</td>
<td>Presentation of antigenic peptides to CD4+ T-cells</td>
<td></td>
</tr>
</tbody>
</table>

**Tables and figures: How are these represented in the text?**

**References:**


**Methods**

The literature search was conducted using PubMed, with no time limitation and with results limited to English-language human data. The Medical Subject Heading key term human leukocyte antigen was searched and combined with the following terms: peanut hypersensitivity and peanut allergy. Additional searches using the terms HLA, MHC, and peanut were performed to ensure capture of all relevant literature. Bibliographies of identified citations also were searched for relevant references. The search criteria identified 7 articles of interest that specifically dealt with studies evaluating genetic associations of HLA alleles with peanut allergy.

**Studies Showing Association between HLA and Peanut Allergy**

The first study of an association between HLA alleles and peanut allergy was performed in 1996 by Donovan et al.28 They studied a single family in which 4 of the 5 siblings had strong allergy to peanut and several other allergens.
Studies Showing No Association between HLA and Peanut Allergy

In 2006, Shreffler et al identified a marker of families with a propensity toward peanut allergy. This study differed from the study by Howell et al, in which DQB1*04, rather than DQB1*06, was associated with peanut allergy. Strengths of this study include its well-defined cohorts of relatively large size. A weakness is the large number of ambiguous genotypes at the allele type level indicating a high failure rate in genotyping calls. Again, the significant differences found in this study were between patients with peanut allergy and normal, unrelated controls, and the possibility that these genes are associated primarily with an overarching atopic phenotype cannot be ruled out.

A study by Dreskin et al in 2010 studied peanut-specific IgG levels in subjects with vs without peanut allergy related to HLA class II alleles. The investigators hypothesized that differences in peanut-specific IgG between subjects with peanut allergy and those with peanut tolerance might be related to differences in the ability to present these allergens by HLA class II molecules. In this study, they determined anti-peanut IgG (total IgG and IgG4), peanut specific IgG levels, and HLA class II alleles by high-resolution typing in a population of subjects with peanut allergy and their siblings with peanut tolerance. No significant differences in HLA class II between these discordant sibling pairs were found. Furthermore, there were highly significant differences in peanut-specific IgG and IgE between the siblings with and without peanut allergy with identical HLA class II alleles. Notably, there was an increased frequency of a rare allele, HLA-DRB1*08:03, in the siblings with and without peanut allergy compared with a large control group of bone marrow donors of European descent, similar to the lower-resolution finding for HLA-DRB1*08 by Howell et al. Dreskin et al concluded that DRB1*08 and particularly DRB1*08:03 appear to be a marker of families with a propensity toward peanut allergy.
allergy and not a marker of peanut allergy itself or it could be a statistical aberration of no biological consequence.

Genomewide Association Study

The reports discussed earlier were candidate gene studies, which limits analysis to 1 small area of the genome. Large-scale, adequately powered GWASs have become the method of choice for identifying genes that influence complex disease. GWASs investigate associations between genetic variables and a specified condition and test hundreds of thousands of genes simultaneously with the benefit of microarray technology. The variable most commonly studied is the single-nucleotide polymorphism (SNP), which is a change in a base pair of DNA that exists in at least 1% of the population. Studies using this method are typically case–control studies, with the cases being patients with the disease of interest and controls selected from the general population. The advantages are evident in that the entire genome can be examined for genetic associations with a particular disease.

The first GWAS of a well-defined food allergy cohort of US children and their biological parents was published in February 2015. This impressive study was performed in 3 stages. The first stage (“gene discovery”) used the modified quasi-likelihood score (MQLS) test to detect genetic associations with any food allergy (including peanut, egg white, cow’s milk, soy, wheat, walnut, fish, shellfish, and sesame seed) and the 3 most common types of food allergy independently (peanut, egg white, and cow’s milk). The MQLS test is a statistical tool used in genetic studies designed to maximally use available information in complex family datasets.

For analysis, 2,759 subjects (853 families) from the Chicago Food Allergy Study (1,315 children and 1,444 parents) were included. These families included 1 or 2 parents with at least 1 biological child with or without food allergy verified by clinical history questionnaires, allergy skin prick testing, and food-specific IgE levels. Normal controls were defined as children with neither clinical allergic reaction nor evidence of sensitization to any of the 9 foods. Other children were categorized as having an “uncertain phenotype” if these criteria could not be met. All parents were defined as having uncertain food allergy phenotypes. The final sample included 2,197 individuals of European ancestry (671 children with food allergy) and 497 individuals of non-European ancestry (155 children with food allergy) for genome discovery analysis using the MQLS test after quality control exclusions. In the MQLS analysis for any food allergy, no SNP reached genomewide significance. Subtype analyses specifically examining peanut allergy, egg allergy, or milk allergy identified genomewide significant associations for peanut, but no significant associations for milk or egg. In 319 patients with peanut allergy, 144 nonallergic controls, and 1,737 controls of uncertain type from the families of European descent, 40 SNPs spanning the HLA class II DQ genes at the 6p21.32 region were identified as significant. The SNP rs9275596, which is intergenic between the HLA-DQB1 and HLA-DQA2 genes, showed the most significant association with peanut allergy. The other 39 were found to be in linkage disequilibrium with rs9275596 and were not significant when conditioned on that SNP. A second SNP, rs7192, which leads to a Leu242Val change in the HLA-DRA gene product, also was found to be significant, and the association between rs9275596 and peanut allergy was decreased when conditioned on rs7192, suggesting that these 2 SNPs might represent a single risk factor. Of note, these associations were not borne out in studies of the smaller cohorts of non-European ancestry.

In the second stage, the investigators narrowed the scope of the study to focus only on peanut allergy, because it was the only significant association found. In this stage, a replication study was performed for the 2 peanut allergy SNPs using an independent sample of 86 peanut allergy cases and 127 controls from the same Chicago Food Allergy cohort. All these samples were independent of those used in stage 1. They found that rs9275596 (the most significant SNP) and rs7192 (potential functional SNP) were significantly associated with peanut allergy in children of European ancestry, and that the 2 SNPs had a similar effect size as seen in the stage 1 GWAS results.

The third stage examined relations between the 2 SNPs associated with peanut allergy and a regulator of gene expression, DNA methylation. They found that the 2 SNPs were significantly associated with differential DNA methylation at multiple sites and that differential DNA methylation of the HLA-DQB1 and HLA-DRB1 genes partly mediated the identified association between SNP and peanut allergy.

This study presented the first GWAS of food allergy, identifying 2 SNPs associated with HLA class II genes, specifically HLA-DR and HLA-DQ. Then, it replicated these findings with an independent sample and discovered that the 2 SNPs significantly affected DNA methylation in several nearby genes. This study is the first to provide powerful evidence that the HLA-DR and -DQ gene region harbors significant genetic risk for peanut allergy in subjects of European ancestry.

Conclusion

The HLA region of the genome presents a complex area of study for identifying genes that might contribute to peanut allergy. This is the most highly polymorphic component of the genome, and all genes are codominantly expressed. This emerging field tracks with technical advances in cellular and molecular biology. Interestingly, the very first study, using only a single family at a time when polymerase chain reaction was laborious, found an association with HLA, but not with TCR, genes. Of note, these early findings linked peanut allergy and HLA with atopy (dust mite allergy) and thus did not suggest any independent relation between peanut allergy and HLA. Subsequent studies, using candidate gene approaches and increasingly larger cohorts, produced data with similar overall outcomes: significant associations emerged when patients were compared with normal controls, but not when compared within families, again suggesting that HLA might govern atopy, without confirming any specific link to peanut allergy. These studies were intriguing, especially because some found associations with the same alleles. However, they had inherent limitations that might have masked true associations between HLA and peanut allergy, including small samples, inadequately or imprecisely defined phenotypes, inadequate racial stratification, and the use of low-resolution DNA typing.

Technology has advanced another step forward, and the recent GWAS took advantage of higher-powered approaches and larger GWASs to clearly identify a relation between HLA and peanut allergy for the first time, even among family members, that is independent of atopy. This study found a strong association between specific SNPs that tag the HLA class II DR and DQ region and peanut allergy that appears to explain approximately 20% of children with peanut allergy. These 2 SNPs also are associated with differential DNA methylation of the HLA-DRB1 and HLA-DQB1 genes, suggesting that alteration of the expression of HLA genes through epigenetic mechanisms might be a potential explanation for why not all children who carry specific HLA class II risk alleles develop peanut allergy.

These studies highlight the complex mechanisms underlying development of tolerance to foods. The roadmap of immunogenetic factors necessary but not sufficient for the development of peanut allergy is likely to be equally complex. As with any immunologically mediated response, antigen presentation by HLA-encoded MHC molecules should be a primary factor governing the peptides available to generate food-specific responses. However, only with
the advent of large databases and refined molecular techniques do patterns implicating HLA in this disease begin to emerge. This indicates that the relation between HLA and food allergy is not simple. Interestingly, HLA also might confer risk or protection in a less direct manner, for example, by helping to shape the microbiome. Therefore, future study of HLA in oral tolerance studies might help explain why early feeding of peanut antigen induces tolerance in some, but not all, patients at risk. In addition, because studies to date have focused mainly on European populations, studies across diverse populations, including African and African American populations in which the burden of peanut allergy is increasing, are warranted.

References


