

Received Date : 27-Jun-2016
Revised Date : 06-Nov-2016
Accepted Date : 07-Nov-2016
Article type : Original Article-Basic Mechanisms in Allergic Disease

Clinical and Experimental Allergy

Genome-wide association study of peanut allergy reproduces association with amino acid polymorphisms in *HLA-DRB1*

David J Martino ^{1 *}, Sarah Ashley ^{1,4, *}, Jennifer Koplin ^{1,2}, Justine Ellis ¹, Richard Saffery ¹, Shyamali C Dharmage ², Lyle Gurrin ², Melanie C Matheson ², Birgit Kalb ^{5,6,7}, Ingo Marenholz ^{6,7}, Kirsten Beyer ⁵, Young-Ae Lee ^{6,7}, Xiumei Hong ⁸, Xiaobin Wang ⁸, Damjan Vukevic ^{1,3}, Allan Motyer ^{1,3}, Stephen Leslie ^{1,3}, Katrina J. Allen ¹ and Manuel A.R. Ferreira ⁹.

¹*Murdoch Childrens Research Institute, The University of Melbourne Department of Paediatrics, The Royal Children's Hospital, Melbourne, Australia.*

²*The University of Melbourne, School of Population and Global Health, Australia*

³*The University of Melbourne, Centre for Systems Genomics, Schools of Mathematics and Statistics and Biosciences, Australia*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cea.12863

This article is protected by copyright. All rights reserved.

⁴ *Hudson Institute of Medical Research, Australia*

⁵ *Pediatric Pneumology and Immunology, Charité Universitätsmedizin Berlin, Germany*

⁶ *Clinic for Pediatric Allergy, Experimental and Clinical Research Center of MDC and Charité, Germany*

⁷ *Max-Delbrück-Center for Molecular Medicine (MDC), Germany*

⁸ *Center on the Early Life Origins of Disease, Department of Population, Family and Reproductive Health, Johns Hopkins University Bloomberg School of Public Health*

⁹ *QIMR Berghofer Medical Research Institute, Brisbane, Australia*

^{*} *Denotes equal contribution*

Key words: Food Allergy, genetics, GWAS, hypersensitivity, peanut allergy,

Corresponding author: Professor Katrina Allen

Gastro and Food Allergy

Murdoch Childrens Research Institute

Flemington Road, Parkville

Melbourne, Australia

Phone: 61 3 9936 6585 Fax: 61 3 9348 1391

Email: Katie.Allen@rch.org.au

Abstract

Background: Genetic variants for IgE-mediated peanut allergy are yet to be fully characterized and to date only one genome-wide association study (GWAS) has been published.

Objective: To identify genetic variants associated with challenge proven peanut allergy.

Methods: We carried out a GWAS comparing 73 infants with challenge-proven IgE-mediated peanut allergy against 148 non-allergic infants (all ~ 1 year old). We tested a total of 3.8 million single nucleotide polymorphism (SNPs), as well as imputed HLA alleles and amino acids. Replication was assessed by *de novo* genotyping in a panel of additional 117 cases and 380 controls, and *in silico* testing in two independent GWAS cohorts.

Results: We identified 21 independent associations at $P \leq 5 \times 10^{-5}$ but were unable to replicate these. The most significant HLA association was the previously reported amino acid variant located at position 71, within the peptide-binding groove of HLA-DRB1 ($P = 2 \times 10^{-4}$). Our study therefore reproduced previous findings for the association between peanut allergy and HLA-DRB1 in this Australian population.

Conclusions & Clinical Relevance: Genetic determinants for challenge proven peanut allergy include alleles at the HLA-DRB1 locus.

Introduction

IgE-mediated peanut allergy is a common food allergy that frequently manifests in early life (1). It affects 1-3% of infants world-wide (2), is often a persistent type of food allergy, more likely to be associated with severe reactions and potentially fatal anaphylaxis (3). Twin and family studies provide compelling evidence that genetic factors contribute to peanut allergy, with estimated heritability of around 80% (4)(5) for peanut allergy and 15 – 35% for IgE sensitization to different foods (6). Despite a significant heritability, only a small number of genetic studies of food allergy have been previously published. Candidate gene studies have identified a number of genetic associations in genes involved in immune function, or skin barrier integrity. For the latter, the filaggrin loss-of-function mutations have been consistently reported as a risk factor for peanut sensitization (7), and more recently peanut allergy (8,9), potentially by increasing trans-cutaneous sensitization to peanut in childhood (10-12). Variants in the Serine protease inhibitor Kazal type 5 (SPINK5) (13) gene have been associated with atopic eczema in Japanese children, and may increase the risk for food allergy. Polymorphisms in immunological genes associated with food allergy include Interleukin-10 (IL-10) (14), interleukin-13 (IL-13) (15), Cluster of differentiation 14 (CD14) (16), Forkhead box P3 (FOXP3) (17) and Signal transducer and activator 6 (STAT6) (18). To date, only one genome-wide association study (GWAS) has been conducted for peanut allergy (19), although case-control status in that study was not defined using challenge-proven outcomes. In that GWAS, the single nucleotide polymorphisms (SNP) rs7192 and rs9275596, located in the major histocompatibility locus (MHC), and the amino acid polymorphism at position 71 of *HLA-DRB1* were associated with peanut allergy. HLA class II variants have also been associated with peanut allergy in candidate gene studies (20-22),

although this has not always been a consistent finding (23). Thus, most genetic determinants of peanut allergy remain to be identified. The aim of this study was to carry out a GWAS on a well-characterized population of challenge-proven paediatric peanut allergic cases and controls to identify SNPs with a large effect on disease risk using an agnostic genome-wide approach, and to perform a comprehensive assessment of the MHC.

Methods

Statement of ethics

This study was approved by the Office for Children HREC (ref. no. CDF/07/492), Department of Human Services HREC (ref. no. 10/07), Royal Children's Hospital HREC (ref. no.27047) and the QIMR Berghofer Medical Research Institute (ref. no. P710).

GWAS study population

All participants included in the GWAS are part of the HealthNuts study, which utilized a population-based sampling frame and diagnostic food challenges. All participants included in the GWAS are part of the HealthNuts population-based cohort (24). Briefly, 5,300 infants aged 11-15 months inclusive were recruited from council-run immunization sessions across Melbourne, Australia, between June 2008 and August 2011. Skin prick testing to egg and peanut was carried out at immunization sessions and those testing positive (~1000) were invited for allergy assessment at the Royal Children's Hospital Melbourne (24). Clinical allergy assessment was carried out by a registered nurse, under hospital supervision using

skin prick testing and open peanut challenges using predetermined objective criteria (25)).

Non-allergic controls were skin prick test negative to a panel of common food allergens (egg white, peanut, sesame, shrimp or cow's milk, cashew, almond, hazelnut, soy and wheat) and safely tolerated peanut during oral peanut challenge. Blood samples were collected from participants after peanut challenge with guardian's written consent. After quality control (summarized below and described in detail in the **Supplementary Methods**), genotyping data were available for 73 cases and 148 controls. Cases were defined as having a skin prick test greater or equal to 2 mm above the negative control, or specific IgE > 0.35 kU/L to peanut on the day of challenge, with evidence of clinical reactivity during oral peanut challenge. Clinical characteristics of the 221 genotyped infants are given in **Supplementary Table 1**.

Peanut food challenges

Peanut oral food challenges (OFC) were conducted at the Royal Children's Hospital under medical supervision. Positive oral food challenge was defined as 1 or more than 1 of the following: 3 or more concurrent noncontact urticaria lasting 5 minutes or more; perioral/periorbital angioedema; vomiting; or circulatory or respiratory compromise within 2 hours of ingestion of a challenge dose. Infants underwent OFC irrespective of their history of ingestion or SPT wheal size, unless there was a parent report of a clear history of an immediate reaction to peanut within the past 2 months (as per the HealthNuts challenge criteria (Koplin 2012)). Three infants were deemed peanut allergic based on the latter definition.

Genotyping and quality control

DNA was extracted from peripheral blood using the Qiagen FlexiGene protocol and stored at -20 degrees prior to genotyping. Genotyping was carried out by the Australian Genome Research Facility using the Illumina HumanOmni 2.5-8 SNP array. DNA was submitted as a single batch comprising 9 plates with subjects randomised across these, and raw data were processed and analysed together. Only unrelated individuals with call rate > 95% were included in the analysis. SNPs were excluded if MAF <1%, call rate <95% and/or with Hardy Weinberg Equilibrium (HWE) test P-value < 10^{-6} , and/or had significant ($P < 0.001$) MAF differences when compared against the 1000 Genomes Project samples (Europeans only), leaving 389,427 high quality directly genotyped variants. Ancestry was inferred from visual inspection of results of a multidimensional scaling (MDS) analysis of identity-by-state (IBS) distance between all individuals and samples of known ancestry from the 1000 Genomes project (26), using PLINK (v1.07) (27). Full details of quality control are provided in **Supplementary Methods.**

Imputation of variants from the 1000 Genomes Project and association analyses

Genotype data for 389,427 directly genotyped variants from 364 individuals was used to impute unmeasured variants, using Impute2 (28) with default options and the 1000 Genomes Project March 2012 release of reference haplotypes available through the Impute2 website (files ALL_1000G_phase1integrated_v3_chr*_impute.*). Variants in the X-chromosome were imputed using the same approach. The association between case-control status and allelic dosage was tested using SNPTTEST (29) with sex and ancestry strata

included as covariates. A P-value of 5×10^{-8} was considered the threshold for genome-wide significance, and variants achieving an a priori determined P-value $\leq 10^{-5}$ were selected for subsequent replication studies. The options used for the association test of autosomes were `--frequentist 1 --method expected`. For the X-chromosome, options were `--method newml` and `--assume_chromosome X`. After excluding variants with imputation info < 0.95 , HWE test P-value $< 10^{-6}$, outlier beta estimates ($|\text{beta}| > 3$ and/or $\text{SE} > 1.5$), or MAF $< 1\%$, results were available for 3,814,967 variants.

Imputation of HLA alleles and amino acid variants

HLA alleles for 8 loci (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1*) were imputed from genotyped SNPs using the HLA*IMP:02 method (30). These imputed alleles were then mapped to amino acid alleles and SNP alleles within the HLA genes, using the same method as was done to build the reference panel for SNP2HLA (31). Association testing for the HLA variants was carried out using logistic regression adjusted for sex and ancestry strata using PLINK (v1.07) (27). Conditional analysis was carried out using forward step-wise logistic regression adjusted for sex and ancestry strata.

Replication study – de novo genotyping

The replication study was conducted in two phases: *de novo* genotyping in a replication sample, and in silico replication in two independent GWAS studies of paediatric food allergy (**Supplementary Figure 1**). A replication panel was assembled to include 117 peanut allergic cases and 380 non-allergic controls from HealthNuts, the Barwon Infant Study (BIS) (32), the

This article is protected by copyright. All rights reserved.

Peanut Allergen Threshold study (PAT) (33) and the Peanut and Probiotic Oral Immunotherapy Trial (PPOIT) (34), all Melbourne-based paediatric allergy studies with identical phenotyping criteria including diagnostic peanut challenges using the same protocol. Clinical characteristics are reported in **Supplementary Table 2**. Case and control definitions were synonymous with the HealthNuts discovery samples. DNA was extracted from cryopreserved peripheral blood using Qiagen AllPrep DNA/RNA mini method at the Murdoch Childrens Research Institute laboratories and stored at -30 degrees prior to genotyping. Targeted genotyping was carried out as a single batch using the MassARRAY platform and IPLEX chemistry (Agena Bioscience) for 19 selected SNPs and a panel of ancestry informative markers (35). Full details are provided in the **Supplementary Methods**.

In silico replication population (American and German studies)

In silico replication was carried out using data from 316 peanut allergy cases, 144 non-allergic controls, and 1,737 controls of uncertain phenotypes (589 children and 1,148 parents) from the Chicago Food Allergy study (19), tested with MQLS (36). Cases were defined as having a history of a clinical allergic reaction on ingestion of peanut and evidence of sensitization to peanut defined as having detectable peanut-specific IgE ($\geq 0.1 \text{ kUI-1}^{-1}$) and/or positive skin prick test (wheal diameter $>3\text{mm}$) to peanut. Non-allergic non-sensitized controls had no evidence of clinical reactivity by food challenge or sensitization to a panel of allergens tested.

In silico replication was also assessed using data from a Caucasian population of 205 peanut allergy cases and 2,387 controls from the German Understanding Food Allergy (UFA) study. Peanut allergic cases were defined by double-blind placebo-controlled food challenge or a history of a severe allergic reaction to peanuts plus specific sensitization to peanut protein ($>0.35 \text{ kUI-1}^{-1}$). Controls were unrelated individuals from the German population based Heinz Nixdorf Recall Study (37). Full details in **Supplementary Methods**.

RESULTS

GWAS of challenge-proven peanut allergy

Demographics and clinical characteristics of the 221 participants included in the discovery GWAS are summarized in **Supplementary Table 1**. After quality control and imputation of unmeasured variants, a total of 3,814,967 SNPs were tested for association with case-control status. The genomic inflation factor of this analysis was 0.998, consistent with no significant effects of technical artefacts or unaccounted population substructure on the association results. A single variant (rs10018666, OR = 5.9, $P = 4 \times 10^{-8}$) located in the *SLC2A9* gene reached genome-wide significance for association with peanut allergy (**Figure 1**). Multiple variants in linkage disequilibrium (LD) with rs10018666 supported this association, including two variants (rs13129697, imputed, $P = 10^{-5}$; rs10939650, genotyped, $P = 2 \times 10^{-6}$) previously reported to associate with variation in the expression of *SLC2A9* in blood cells (**Supplementary Table 3**). This variant, and an additional 20 that were associated with peanut allergy at $P \leq 5 \times 10^{-5}$ (**Supplementary Table 4**), were selected for a replication analysis and 19 of the 21 SNPs (designated in **Supplementary Table 4**) were incorporated

into a multiplexed genotyping assay for the replication stage. For most of these variants, the direction of effect was the same across the three ancestry strata analysed (**Supplementary Table 5**).

Replication of top associations in three independent studies

We directly genotyped the 19 SNPs and 41 ancestry markers in an additional 117 peanut allergy cases and 380 non-allergic non-sensitized controls from Melbourne (HealthNuts, Barwon, Peanut Allergen Threshold, PPOIT studies), who were unrelated to the participants included in the discovery GWAS. Four variants were found to have suggestive associations with peanut allergy (unadjusted $P < 0.05$), with similar effects between the two ancestry populations (**Supplementary Table 6**), but these associations did not survive a correction for multiple testing. To increase power of the replication study, we extracted results for the top 21 SNPs from two additional studies of food allergy – the German UFA (205 peanut allergy cases and 2,387 controls, **Supplementary Table 7**) and Chicago (316 peanut allergy cases, 144 controls, **Supplementary Table 8**) studies – but found no overall consistent evidence (same direction of effect) for association with any of the 21 variants. Therefore, we conclude that the top 21 SNPs identified in the discovery GWAS likely represent false-positive associations.

Analysis of HLA polymorphisms

We studied in greater detail the association between HLA polymorphisms and food allergy by imputing classical 102 2-digit alleles, 187 4-digit alleles, as well as 2,205 polymorphic amino acid loci, in addition to 2,412 SNPs across the MHC. The most significant associations were observed with amino acid polymorphisms at position 37 (OR = 0.3, $P = 9.8 \times 10^{-5}$) and 71 (OR = 0.34, $P = 1.5 \times 10^{-4}$) of the *HLA-DRB1* gene (**Table 1**). The direction of effect for these associations was the same as that reported for position 71 previously in the Hong et al GWAS (OR= 0.38, $P = 9.8 \times 10^{-11}$) (19). Conditioning on position 37 in our study substantially reduced the effect of position 71 (OR = 0.4, $P = 0.03$), whilst condition on position 71 also reduced the effect of 37 (OR = 0.4, $P = 0.02$; **Supplementary Figure 2**). These two variants are in moderate linkage disequilibrium ($r^2=0.5$) and we could not statistically distinguish the effect of one over the other, suggesting that they are unlikely to be independent. All of the remaining associations were greatly reduced after conditioning on amino acid positions 37 or 71, indicating that the associations at HLA were primarily driven by these polymorphisms.

Discussion

In this study, we report the second GWAS of IgE-mediated peanut allergy, but the first using challenge-proven outcomes. With the sample size for this study, only variants with a very large effect on disease risk would be detected at the genome-wide significance level. In total, we identified 21 SNPs with a suggestive association with peanut allergy in the discovery GWAS but were unable to replicate these in three replication studies, suggesting that they were likely to be false-positive associations. On the other hand, a comprehensive

analysis of HLA polymorphism identified an association with amino acid positions 37 and 71 of the *HLA-DRB1* gene which supports the association originally reported with this polymorphism in a US population of children with peanut allergy (19). Although the p-value did not reach genome-wide significance, the association was clearly reproduced in our data set as evidenced by the consistent direction and effect size, and the fact that it was the strongest signal at the MHC. We were unable to assess the HLA associations for rs7192 and rs9275596 also reported by Hong (19) since these SNPs did not pass QC in our data set. Collectively, our data support a role for polymorphisms in the HLA region on the risk of peanut allergy, specifically those relating to *HLA-DRB1* amino acids. *HLA-DR* molecules comprise the major HLA class II isotypes present on the surface of antigen presenting cells, and *HLA-DRB1* is a highly polymorphic locus. The binding groove of *DRB1* consists of pockets that interact with peptide side chain residues. Polymorphisms in the binding groove can alter the overall binding affinity and specificity of *HLA-DRB1* for peptide ligands. The outcome of these effects can play a central role in T-helper cell activation and influence the quality of the immune response. As outlined in Hong et al, and originally reported by Sturniolo et al (38), polymorphic amino acid residue 71 affects the binding specificity of pocket 4 of the peptide binding groove. It is therefore biologically plausible that this amino acid variant plays an important role in the development of antigen-specific immune responses and the development of allergy. Future studies employing targeted genotyping approaches are now needed to properly resolve the *DRB1* gene association and complex haplotype structure in order to clarify functional and clinical significance. Given that SNP arrays are limited by LD structure they indicate a region of interest, and targeted genotyping with functional studies would be required in future to assess causality.

The strengths of this study include the well-characterized cohort in which clinical phenotypes were determined by oral food challenge, and clinical assessment protocols were harmonised across the discovery and replication populations. Despite this we cannot exclude some heterogeneity between our discovery and replication populations as a potential caveat. The discovery population were young infants (age 12 months) whilst the replication studies (*de novo* and *in silico*) consisted of infants from a range of ages (mean age of peanut allergics in UFA was 3.5 years and 7 years in Chicago study). This was reflected to some extent as variations in peanut specific IgE levels, with the older children exhibiting higher antibody titres (HealthNuts: mean=13.5, SD=19.6 kU/L, UFA: mean=57.5, SD=95.9 kU/L and Chicago: mean 65.6, SD 87.4 kU/L), and mean wheal diameters (HealthNuts: mean=8.7 mm, *de novo* replication children ranged from 5.71 to 18.8mm across the studies **Supplementary Table 2**). Subjects from the *in silico* analysis were entirely Caucasian, whilst the discovery and *de novo* genotyping populations consisted of Caucasian and Asian ancestries. Other caveats include the limited sample size, which precluded the ability to detect all but substantial effect sizes. Moreover, our analysis was restricted to common variants, which is typical for most GWAS, and alternative study designs will be required to address the role of rare variants.

In summary, this study supports a role for *HLA-DRB1* alleles as genetic risk factors of IgE-mediated peanut allergy.

Acknowledgements: The genome wide association in this study was funded by the US Department of Defense (grant # W81XWH-10-1-0487). The HealthNuts cohort is funded by the Australian National Health and Medical Research Council. This work was supported by

This article is protected by copyright. All rights reserved.

the Australian National Health and Medical Research Council (NHMRC), Career Development Fellowship ID 1053756 (S.L.) Australian Research Council Future Fellowship ID FT120100253 (JE), Early Career Fellowships 1072752, 1054592 (DM, JE); and by the Victorian Life Sciences Computation Initiative (VLSCI) grant number VR0240 on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia (S.L.). Research at the Murdoch Childrens Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. S.L. is a partner in Peptide Groove LLP. The UFA study was supported by a Clinical Exchange Program fellowship of the Experimental and Clinical Research Center of Max-Delbrück-Center and Charité Medical School. We thank Markus M Noethen (Institute of Human Genetics, University of Bonn, Germany) and Sven Cichon (Division of Medical Genetics, University Hospital Basel and Department of Biomedicine, University of Basel) for providing genotype data of the Heinz Nixdorf RECALL (HNR) study.

References

1. Sicherer SH. Clinical update on peanut allergy. *ANAI*. 2002 Apr;88(4):350–61–quiz361–2–394.
2. Osborne NJ, Koplin JJ, Martin PE, Gurrin LC, Lowe AJ, Matheson MC, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *J Allergy Clin Immunol*. 2011 Mar;127(3):668–76.e1–2.
3. Finkelman FD. Peanut allergy and anaphylaxis. *Curr Opin Immunol*. 2010 Dec;22(6):783–8.
4. Sicherer SH, Furlong TJ, Maes HH, Desnick RJ, Sampson HA, Gelb BD. Genetics of peanut allergy: a twin study. *Journal of allergy and clinical immunology*. 2000 Jul;106(1 Pt 1):53–6.

- Accepted Article
5. Hourihane JO, Dean TP, Warner JO. Peanut allergy in relation to heredity, maternal diet, and other atopic diseases: results of a questionnaire survey, skin prick testing, and food challenges. *BMJ*. BMJ Publishing Group Ltd; 1996 Aug 31;313(7056):518–21.
 6. Tsai HJ, Kumar R, Pongracic J, Liu X, Story R, Yu Y, et al. Familial aggregation of food allergy and sensitization to food allergens: a family-based study. *Clin Exp Allergy*. 2009 Jan;39(1):101–9.
 7. Tan H-TT, Ellis JA, Koplin JJ, Matheson MC, Gurrin LC, Lowe AJ, et al. Filaggrin loss-of-function mutations do not predict food allergy over and above the risk of food sensitization among infants. *J Allergy Clin Immunol*. 2012 Nov;130(5):1211–3.
 8. Brown SJ, Asai Y, Cordell HJ, Campbell LE, Zhao Y, Liao H, et al. Loss-of-function variants in the filaggrin gene are a significant risk factor for peanut allergy. *Journal of allergy and clinical immunology*. 2011;127(3):661–7.
 9. Venkataraman D, Soto-Ramírez N, Kurukulaaratchy RJ, Holloway JW, Karmaus W, Ewart SL, et al. Filaggrin loss-of-function mutations are associated with food allergy in childhood and adolescence. *J Allergy Clin Immunol*. 2014 Oct;134(4):876–882.e4.
 10. Thyssen JP, Tang L, Husemoen LLN, Stender S, Szecsi PB, Menné T, et al. Filaggrin gene mutations are not associated with food and aeroallergen sensitization without concomitant atopic dermatitis in adults. *J Allergy Clin Immunol*. 2015 Feb 10;135(5):1375–1378.e1.
 11. Brough HA, Liu AH, Sicherer S, Makinson K, Douiri A, Brown SJ, et al. Atopic dermatitis increases the effect of exposure to peanut antigen in dust on peanut sensitization and likely peanut allergy. *J Allergy Clin Immunol*. 2014 Nov 18.
 12. Brough HA, Simpson A, Makinson K, Hankinson J, Brown S, Douiri A, et al. Peanut allergy: effect of environmental peanut exposure in children with filaggrin loss-of-function mutations. *J Allergy Clin Immunol*. 2014 Oct;134(4):867–875.e1.
 13. Kusunoki T, Okafuji I, Yoshioka T, Saito M, Nishikomori R, Heike T, et al. SPINK5 polymorphism is associated with disease severity and food allergy in children with atopic dermatitis. *Journal of allergy and clinical immunology*. 2005 Mar;115(3):636–8.
 14. Campos Alberto EJ, Shimojo N, Suzuki Y, Mashimo Y, Arima T, Matsuura T, et al. IL-10 gene polymorphism, but not TGF-beta1 gene polymorphisms, is associated with food allergy in a Japanese population. *Pediatr Allergy Immunol*. 2008 Dec;19(8):716–21.
 15. Liu X, Beaty TH, Deindl P, Huang S-K, Lau S, Sommerfeld C, et al. Associations between specific serum IgE response and 6 variants within the genes IL4, IL13, and IL4RA in

German children: the German Multicenter Atopy Study. *Journal of allergy and clinical immunology*. 2004 Mar;113(3):489–95.

16. Woo JG, Assa'ad A, Heizer AB, Bernstein JA, Hershey GKK. The -159 C-->T polymorphism of CD14 is associated with nonatopic asthma and food allergy. *Journal of allergy and clinical immunology*. 2003 Aug;112(2):438–44.
17. Torgerson TR, Linane A, Moes N, Anover S, Mateo V, Rieux-Laucat F, et al. Severe food allergy as a variant of IPEX syndrome caused by a deletion in a noncoding region of the FOXP3 gene. *Gastroenterology*. 2007 May;132(5):1705–17.
18. Amoli MM, Hand S, Hajeer AH, Jones KP, Rolf S, Sting C, et al. Polymorphism in the STAT6 gene encodes risk for nut allergy. *Genes Immun*. 2002 Jun;3(4):220–4.
19. Hong X, Hao K, Ladd-Acosta C, Hansen KD, Tsai H-J, Liu X, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nat Commun*. 2015;6:6304.
20. Howell WM, Turner SJ, Hourihane JO, Dean TP, Warner JO. HLA class II DRB1, DQB1 and DPB1 genotypic associations with peanut allergy: evidence from a family-based and case-control study. *Clin Exp Allergy*. 1998 Feb;28(2):156–62.
21. Hand S, Darke C, Thompson J, Stingl C, Rolf S, Jones KP, et al. Human leucocyte antigen polymorphisms in nut-allergic patients in South Wales. *Clin Exp Allergy*. Blackwell Science Ltd; 2004 May;34(5):720–4.
22. Madore A-M, Vaillancourt VT, Asai Y, Alizadehfar R, Ben-Shoshan M, Michel DL, et al. HLA-DQB1*02 and DQB1*06:03P are associated with peanut allergy. *Eur J Hum Genet*. 2013 Feb 26.
23. Shreffler WG, Wanich N, Moloney M, Nowak-Wegrzyn A, Sampson HA. Association of allergen-specific regulatory T cells with the onset of clinical tolerance to milk protein. *Journal of allergy and clinical immunology*. 2009 Jan;123(1):43–7.
24. Koplin JJ, Wake M, Dharmage SC, Matheson M, Tang ML, Gurrin LC, et al. Cohort Profile of the HealthNuts study: Population prevalence and environmental/genetic predictors of food allergy. *Int J Epidemiol*. 2015 Jan 21.
25. Koplin JJ, Tang MLK, Martin PE, Osborne NJ, Lowe AJ, Ponsonby A-L, et al. Predetermined challenge eligibility and cessation criteria for oral food challenges in the HealthNuts population-based study of infants. *J Allergy Clin Immunol*. 2012 Apr;129(4):1145–7.

26. Consortium T1GP. An integrated map of genetic variation from 1,092 human genomes. *Nature*. Nature Publishing Group; 2012 Nov 1;491(7422):56–65.
27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007 Sep;81(3):559–75.
28. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. Schork NJ, editor. *PLoS Genet*. Public Library of Science; 2009 Jun;5(6):e1000529.
29. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet*. Nature Publishing Group; 2010 Jul 1;11(7):499–511.
30. Dilthey A, Leslie S, Moutsianas L, Shen J, Cox C, Nelson MR, et al. Multi-population classical HLA type imputation. Browning S, editor. *PLoS Comput Biol*. 2013;9(2):e1002877.
31. Jia X, Han B, Onengut-Gumuscu S, Chen W-M, Concannon PJ, Rich SS, et al. Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. Tang J, editor. *PLoS ONE*. Public Library of Science; 2013 Jun 6;8(6):e64683.
32. Vuillermin P, Saffery R, Allen KJ, Carlin JB, Tang ML, Ranganathan S, et al. Cohort Profile: The Barwon Infant Study. *Int J Epidemiol*. 2015 Aug;44(4):1148–60.
33. Zurzolo GA, Allen KJ, Taylor SL. Peanut Allergen Threshold Study (PATS): validation of eliciting doses using a novel single-dose challenge protocol. *Allergy Asthma Clin* 2013.
34. Tang MLK, Ponsonby A-L, Orsini F, Tey D, Robinson M, Su EL, et al. Administration of a probiotic with peanut oral immunotherapy: A randomized trial. *J Allergy Clin Immunol*. 2015 Mar;135(3):737–8.
35. Bousman CA, Yung AR, Pantelis C, Ellis JA. Effects of NRG1 and DAOA genetic variation on transition to psychosis in individuals at ultra-high risk for psychosis. *Translational* 2013.
36. Thornton T, McPeck MS. Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *The American Journal of Human Genetics*. 2007.
37. Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged

subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. Am Heart J. 2002 Aug;144(2):212–8.

38. Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nat Biotechnol. 1999 Jun;17(6):555–61.

Figures and Tables

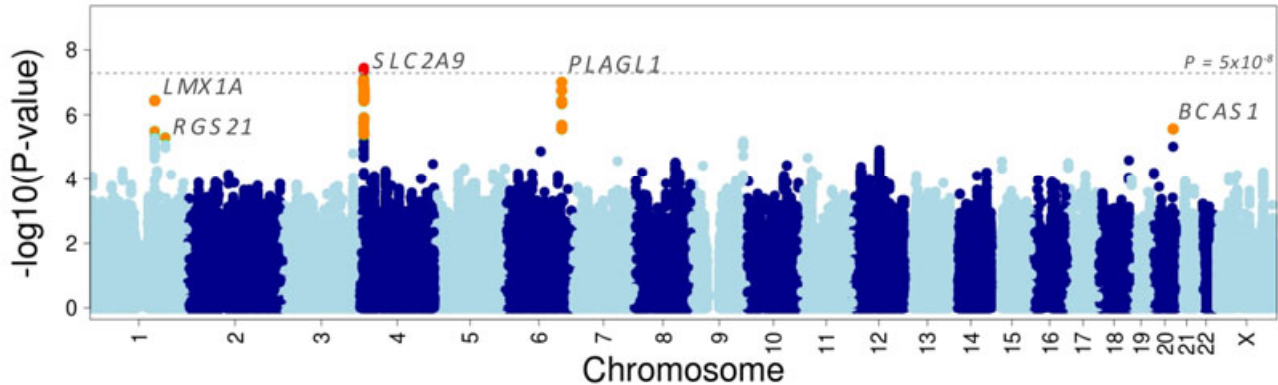


Figure 1. Summary of association results for 3,814,967 SNPs tested in 73 peanut food allergy cases and 148 controls. Suggestive associations ($P \leq 10^{-5}$) are shown in orange and genome-wide significant associations ($P = 5 \times 10^{-8}$) in red.

Table 1 Top 5 HLA associations stratified by variant class ranked according to P-value

Amino acids and imputed SNP	CHR	BP	OR	SE	L95	U95	STAT	P
AA_DRB1_37_32660037_FY	6	32552059	0.3	0.31	0.16	0.55	-3.9	9.80E-05
AA_DRB1_71_32659935_R	6	32551957	0.34	0.29	0.19	0.59	-3.79	1.50E-04
SNP_DRB1_32659935_C	6	32551957	0.34	0.29	0.19	0.59	-3.79	1.50E-04
AA_DRB1_37_32660037_NS	6	32552059	0.34	0.29	0.19	0.6	-3.7	2.16E-04
AA_DRB1_47_32660007	6	32552029	0.35	0.29	0.2	0.62	-3.64	2.76E-04
Genotyped SNPs								
rs154975	6	32900601	2.6	0.25	1.59	4.24	3.83	1.30E-04
rs2858305	6	32670464	2.22	0.22	1.44	3.4	3.64	2.72E-04

rs2856717	6	32670308	2.19	0.22	1.43	3.36	3.59	3.28E-04
rs2858332	6	32681161	2.28	0.23	1.45	3.58	3.58	3.48E-04
rs2858309	6	32668713	2.17	0.22	1.42	3.33	3.57	3.61E-04

2-digit HLA alleles

HLA_DQB1_06	6	32631061	2.26	0.25	1.38	3.69	3.24	1.20E-03
HLA_DPB1_03	6	33049368	3.23	0.41	1.45	7.21	2.86	4.25E-03
HLA_DQA1_02	6	32608306	0.36	0.42	0.16	0.82	-2.44	1.48E-02
HLA_DRB1_15	6	32552064	2.15	0.32	1.15	3.99	2.41	1.60E-02
HLA_DRB1_07	6	32552064	0.34	0.46	0.14	0.82	-2.39	1.68E-02

4-digit HLA alleles

HLA_DPB1_0301	6	33049368	3.23	0.41	1.45	7.21	2.86	4.25E-03
HLA_DQA1_0201	6	32608306	0.36	0.42	0.16	0.82	-2.44	1.48E-02
HLA_DRB1_0701	6	32552064	0.34	0.46	0.14	0.82	-2.39	1.68E-02
HLA_DQB1_0602	6	32631061	2.11	0.31	1.14	3.91	2.38	1.72E-02
HLA_DQA1_0102	6	32608306	1.81	0.25	1.11	2.95	2.38	1.75E-02

CHR = Chromosome, BP = base position, OR= odds ratio, SE = standard error, L95= lower 95% confidence interval, U95 = upper 95% confidence interval, STAT = value of the test statistic, P= P-value.