Genetic variation at the Th2 immune gene IL13 is associated with IgE-mediated paediatric food allergy

DOI:
10.1111/cea.12942

Document Version
Accepted author manuscript

Genetic Variation at the Th2 Immune Gene *IL13* is Associated with IgE-mediated Paediatric Food Allergy

Sarah Ashley[1,2], Hern-Tze Tina Tan[1,3], Rachel Peters[1,4], Katrina Allen[1,4,9,10], Peter Vuillermin[1,6,7], Shyamali C Dharmage[1,8], Mimi LK Tang[1,4,9], Jennifer Koplin[1,4], Adrian Lowe[1,8], Anne-Louise Ponsonby[1,4], John Molloy[1,7], Melanie C Matheson[1,4], Richard Saffery[1,2,4], Justine A Ellis[1,4,11]*, David Martino[1,4,5]* and the HealthNuts team.

†On behalf of the Barwon Infant Study, the Melbourne Atopy Cohort study, the Peanut Allergen Threshold Study and the Probiotics and Peanut Oral ImmunoTherapy study.

*Denotes equal contribution

[1] Murdoch Childrens Research Institute, Royal Children's Hospital, Australia

[2] The Hudson Institute, Monash Translational Health Precinct (MTHP), Monash University, Clayton, Australia

[3] Department of Immunology, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Malaysia

[4] University of Melbourne, Department of Paediatrics, Parkville, Australia

[5] University of Western Australia, Department of Paediatrics, Australia

[6] Child Health Research Unit, Barwon Health, Geelong, Australia

[7] Deakin University, Waurn Ponds, Australia

[8] The University of Melbourne, Centre for Epidemiology and Biostatistics, School of Population Health, Melbourne, Australia

[9] Department of Allergy and Immunology, Royal Children's Hospital, Parkville, Australia

[10] Institute of Inflammation and Repair, University of Manchester, UK

[11]Centre for Social and Early Emotional Development, Faculty of Health, Deakin University

Abbreviations

SNP: Single nucleotide polymorphism

OFC: Oral food challenge

SPT: Skin prick test

IL13: *Interleukin 13*

AIMs: Ancestry informative markers

IgE: Immunoglobulin E

Background: Food allergies pose a considerable worldwide public health burden with incidence as high as one in ten in 12-month old infants. Few food allergy genetic risk variants have yet been identified. The Th2 immune gene *IL13* is a highly plausible genetic candidate as it is central to the initiation of IgE-class switching in B cells.

Objective: Here we sought to investigate whether genetic polymorphisms at *IL13* are associated with the development of challenge-proven IgE-mediated food allergy.

Method: We genotyped nine *IL13* 'tag' single nucleotide polymorphisms (tag-SNPs) in 367 challenge-proven food allergic cases, 199 food sensitised-tolerant cases and 156 non-food allergic controls from the HealthNuts study. 12-month old infants were phenotyped using open oral food challenges. SNPs were tested using Cochran-Mantel-Haenszel test adjusted for ancestry strata. A replication study was conducted in an independent, co-located sample of four paediatric cohorts consisting of 203 food allergic cases and 330 non-food allergic controls. Replication sample phenotypes were defined by clinical history of reactivity, 95% PPV or challenge and *IL13* genotyping was performed.

Results: *IL13* rs1295686 was associated with challenge-proven food allergy in the discovery sample (P=0.003; OR=1.75; CI=1.20-2.53). This association was also detected in the replication sample (P=0.03, OR=1.37, CI=1.03-1.82) and further supported by a meta-analysis (P=0.0006, OR=1.50). However, we cannot rule out an association with food sensitisation. Carriage of the rs1295686 variant A allele was also associated with elevated total plasma IgE.

Conclusions: We show for the first time, in two independent cohorts, that *IL13* polymorphism rs1295686 (in complete linkage disequilibrium with functional variant rs20541) is associated with challenge-proven food allergy.

Key words: food allergy, food sensitisation, *IL13*, interleukin-13, IgE, single nucleotide polymorphism

Introduction

Interleukin-13 has a well-described role in mediating Immunoglobulin E (IgE) class switching, and is necessary for the development of T-helper type 2 (Th2) mediated allergic immune responses (1). IL-13 deficient mice exhibit impaired Th2 development and reduced Th2 cytokine production and IgE levels (1). In humans, polymorphisms at *IL13* have been studied in the context of asthma, eczema and atopy (2-10). *IL13* is one of the most replicated asthma and atopy genetic associations. Variants at *IL13* have been linked to elevated cord blood IgE (cbIgE) (11, 12) and elevated IgE at 12-months, 2, 4 and 8 years of age (13, 14). Of the 21 SNPs associated with cbIgE in the Yang *et al.* study an *IL13* polymorphism was one of the two SNPs with highest predictive accuracy for cbIgE (12). One well-studied variant, rs20541 (*IL13*+2044GA), produces an IL-13 gain of function amino acid substitution (Arg130Gln) variant with notably higher induction of downstream processes including CD23 monocyte expression and IgE class switching in B cells (15, 16). Another variant rs1800925 enhances IL-13 production in Th2 polarised cells and has been associated with asthma, eczema and sensitisation to foods (2, 17, 18). Liu *et al.*, identified epistatic effects of genetic variants of *IL13* and *IL4RA* on atopy and specific IgE production. These findings demonstrated that carriage of the TT genotype at rs1800925 of *IL13* was associated with sensitisation to foods, dependent on variants of *IL4RA* (18). Overall there is compelling evidence that *IL13* polymorphisms are associated with allergic disease. There is however a paucity of data regarding polymorphisms at *IL13* and the development of challenge-proven food allergy. We sought to comprehensively investigate a potential association between *IL13* genetic polymorphisms and the development of challenge-proven IgE-mediated food allergy in childhood from an Australian community-based population.

Methodology

Discovery Phase Cohort

This study included 722 Australian infants from the HealthNuts study, methodology of which is published elsewhere (19). Briefly, recruitment for HealthNuts took place between 2007 and 2011 for 5,276 12 month-old infants presenting for scheduled immunisations at council run clinics (19). All infants underwent a skin prick test (SPT) to peanut, egg white, sesame, and either cow's milk or shrimp (or both) for food sensitisation status during attendance at immunisation clinics. Those who had a wheal size of ≥ 1mm were invited to the Royal Children's Hospital for an open oral food challenge (OFC) using pre-determined objective diagnostic criteria (20). A random selection of 200 individuals with a negative SPT (no detectable wheal) to the food allergens underwent food challenges to confirm their status as negative controls. No individuals who originally had negative SPTs had positive OFCs. The main phenotypes of interest were any *food allergy* defined as having a positive SPT result (wheal ≥2mm) to peanut, egg white or sesame and clinical reactivity by OFC. Those individuals with a positive SPT (wheal ≥2mm) to peanut, egg white, cow's milk (or shrimp) or sesame but were asymptomatic by OFC were deemed *food sensitised tolerant*. *Non-food allergic* controls were those without a detectable SPT wheal to any foods at recruitment and had a subsequent negative OFC result in the clinic. Eczema was defined by parent reported doctor diagnosis or nurse observation on the day of OFC. Parents of study participants completed a questionnaire that included questions to parental country of birth, which were used to define ancestry strata. Individuals were classified as Caucasian if both parents were born in Australia, Europe, UK, Northern America or New Zealand (n=503). Individuals were classified as Asian if both parents were born in South East Asia (n=74). Those with one parent in each category were classified as mixed Asian-Caucasian (n=145). We validated these ancestry strata using a random sample of 344 individuals for which ancestry was also

determined by genome-wide SNP typing and identity-by-descent cluster analysis, which revealed excellent (>90%) agreement between the parent-reported and genetically determined categories (**Supplemental methods**). Ten mL of blood was collected from 836 infants who attended clinics for IgE titers and genetic studies. Ethical approval was obtained from the Office for Children HREC (Human Research Ethics Committee) (CDF/07/492), the Department of Human Services HREC (10/07) and the Royal Children's Hospital (RCH) HREC (27047).

Plasma IgE measures

Total IgE was quantified by ImmunoCAP System FEIA (Phadia AB, Uppsala, Sweden) using plasma aliquots derived from blood samples.

Tag-SNP selection

To comprehensively measure polymorphisms at *IL13*, nine tag SNPs within a region of 91.5kb incorporating *IL13* and sequence ~5kb upstream and ~11kb downstream of the gene were incorporated into a multiplex genotyping assay (**Supplementary figure 1**). Tag SNPs were selected using HapMap data (HapMap Genome Browser Phase 1, 2 & 3 - data source: HapMap Data Rel 27 (Feb 09)) and the tagger function in Haploview to broadly cover 22 variants with linkage disequilibrium (LD) of $r^2 \geq 0.8$ (21). Primers for the multiplex assay were designed using Agena Bioscience MassArray Design 3.1 software (sequences available from the authors).

Genotyping and quality control

Genomic DNA derived from peripheral blood samples (n=836) were genotyped using Agena Bioscience iPLEX Gold chemistry and the MassARRAY mass spectrometer system

according to manufacturer's instructions. Quality control of genotyping data was conducted in PLINK (22). Samples with a genotyping success rate of less than 95% were excluded (n=59), resulting in a final post-QC sample size of 722 in the HealthNuts discovery cohort of 12-month old infants; 367 food allergy cases, 199 food sensitised but tolerant cases and 156 non-food allergic controls (**Supplementary Table 1**). All nine *IL13* tag SNPs passed QC with a genotype call rate of ≥ 95%, there was no evidence of significant deviation from the Hardy Weinberg Equilibrium (HWE) ($p<0.01$).

Discovery phase statistical analysis

We compared allele frequencies between cases and controls using the Cochran-Mantel-Haenszel (CMH) test in the PLINK software (22), which controls for heterogeneity arising from population stratification using an "average" odds ratio. The total plasma IgE data were positively skewed and thus log transformed prior to analysis. The linear regression test in PLINK was used to test for an association between *IL13* variants and log transformed total plasma IgE, adjusted for food allergy as a covariate.

Replication phase

Candidate gene associations were replicated in an independent paediatric sample from multiple Melbourne based studies with comparable food allergy measures; 132 non-allergic controls and 36 food allergic cases from the Barwon Infant Study (BIS) (23), 198 non-allergic controls and 57 food allergic cases from the Melbourne Atopic Cohort (MACs) (24), 72 food allergic cases from the Peanut Allergen Threshold Study (PATS) (25) and 38 food allergic cases from the Probiotic and Peanut Oral Immuno-Therapy study (PPOIT) (26). Demographics and clinical characteristics are presented in **Supplementary Table 2**). Ethical approval was provided by Office for Children HREC for PATS (HRECApp32166A and

2012P002475). The RCH Human Research and Ethics Committee HREC 27086Q (PPOIT). Mercy Maternity Hospital Ethics Committee (R88/06) for MACS with 18-year follow-up, including collection of DNA, was approved by the Royal Children's Hospital (HREC 28035); BIS: Barwon Health Human Research Ethics Committee HREC (10/24). The definition of phenotypes for the majority of the sample was similar to the discovery study, OFC (BIS, PPOIT, PATs) or clear history of reactivity within 1-2 hours in addition to SPT (>2mm) or sIgE (>0.35 kUA/L) sensitisation (PPOIT, PATs). The widely accepted 95% PPV (SPT > 95% PPV) (27) for diagnosing clinical allergy was utilised for MACs samples. Non-food allergic cases were defined with negative SPT to a panel of foods in the MACs study and with OFC in addition to negative SPT wheal in the Barwon Infant Study. Eczema was defined by history of doctor-diagnosis or nurse-observation at clinic. Individuals with an unclear phenotype below the 95% PPV (n=60) were removed from analysis. Assay design and genotyping were carried out as for the discovery phase. Ancestry strata were genetically determined (described below) using a panel of 49 ancestry informative markers (AIMs) derived from a panel described in Bousman *et al.*, 2013 (28). Ancestral population clusters were determined in PLINK using Identity-by-state (IBS) distance clustering resulting in 463 individuals of European descent, 51 of mixed European-Asian descent and 15 of Asian descent (post-QC). Consistent with the discovery phase, individuals with a genotyping success rate of less than 95% were excluded (n=71). All SNPs pass the 95% genotyping call rate and none significantly deviated from HWE ($P<0.01$). After QC there were 533 phenotyped and genotyped individuals (203 food allergic cases and 330 non-food allergic controls).

The replication cohort genotyping data were analysed in PLINK using logistic regression models adjusted for population stratification using the first and second principal components (PCs). Meta-analysis of discovery and replication samples was conducted in

PLINK with random and fixed effects models to assess any heterogeneity of effect sizes between studies (570 food allergic cases and 486 non-food allergic controls).

Results

Following genotyping and quality control, 367 out of 722 (50.8%) participating infants with blood samples in the discovery phase of the study had clinical food allergy. Egg allergy was the most common type of food allergy (89.3%), followed by peanut allergy (36.2%) whilst allergy to sesame was less common (6.7%). Food sensitised or food allergic infants were more likely to be male with one or more Asian parents. Eczema rates were higher in the food allergy group (56.4%) (**Supplementary Table 1**).

IL13 variant rs1295686 is associated with challenge-proven food allergy

Allele and genotype frequencies between discovery phase food allergy cases (n=367) and non-allergic controls (n=156) were tested using the CMH clustering test adjusted for ancestry strata. *IL13* variant rs1295686, which is in complete LD (r^2 =1) with a previously described functional variant rs20541 (Gln144Arg), was associated (nominal P<0.05) with challenge proven food allergy (P=0.003; OR=1.75; CI=1.20-2.53, **Table 1. A.**). A sensitivity analysis of infants without eczema (174 food allergic cases and 90 non-food allergic controls) reproduced this association suggesting it was independent of eczema co-morbidities (P=0.008; OR=1.82; CI=1.17-2.83, **Supplementary Table 3**). We also adjusted for atopic morbidities using logistic regression, which did not materially alter the pattern of association (**Supplementary Table 4**). There was weak evidence of an association between rs1295686 (P=0.06; OR=1.48; CI=0.98-2.23) and rs1295687 (P=0.05; OR=2.19; CI=1.01-4.76) with the food sensitised tolerant phenotype (n=199) (**Table 1. B.**). There was no evidence that any of the variants tested increased the risk for food allergy when comparing food allergic cases (n=367) to food sensitised tolerant children (n=199) (**Table 1.C**)

SNPs in *IL13* are associated with variations in total plasma IgE

Total plasma IgE measures were available for 423 individuals from the discovery study (69 non-food allergic, 106 food sensitised but tolerant and 248 food allergic). We tested log IgE titers as the dependent variable in a regression analysis using genotyped allele frequencies as predictors. The strongest evidence for an association with IgE titres was with the minor allele of rs1295686, which was associated with elevated log transformed total plasma IgE levels (P=0.0003, SE=0.10, Beta=0.38). There was also evidence that three additional variants (rs3091307, rs1800925 and rs1295683) were associated with elevated IgE levels (**Supplementary table 5 & Supplementary figure 2**).

Replication confirms *IL13* variant rs1295686 is associated with food allergy

To support our findings from the discovery phase, we genotyped rs1295686 and rs1295687 in an additional 533 independently selected children. In the replication sample, egg allergy was again the most common food allergy (72.5%), followed by peanut (62.1%). Eczema was more common among food allergic children (41.9%) and there was a higher proportion of Europeans (78.5%) in the control group.

In a logistic regression analysis adjusted for ancestry, we replicated the association between food allergy and rs1295686 in the independent sample group (P=0.03, OR=1.37, CI=1.03-1.82) (**Table 2. A.**). A sensitivity analysis of infants without eczema (96 food allergic cases and 201 non-food allergic controls) again did not suggest eczema was a significant confounder. The rs12954687 SNP did not show evidence of an association with phenotype in this population (**Table 2.A**). To maximize the study power, we performed a meta-analysis of both discovery and replication samples. Genotyped SNP rs1295686 remained the top associated variant (P=0.0006, OR=1.50, Q=0.31, I=3.32) while rs1295687 was not associated (**Table 2. B.**). In this analysis the p-value for Cochrane's Q statistic (Q) and I^2 heterogeneity

index (I) (0-100) was indicative of low heterogeneity of effect sizes between study populations.

Discussion

Genetic variants at *IL13* have been reported to increase the risk of a range of atopic conditions including asthma (3, 6, 7, 9) and eczema (5, 8). To our knowledge this is the first demonstration of an association of *IL13* variants with clinical food allergy in an ethnically diverse population. Our data provide evidence that the minor allele of variant rs1295686 (r^2=1 with functional variant rs20541, see **Supplementary figure 1**) predisposes to an increased risk of food sensitisation and food allergy, possibly through enhanced IgE production. These conclusions are based on the following lines of evidence: Association testing of allele frequencies between food allergy cases and controls in two independent samples identified rs1295686 as the strongest association with food allergy phenotype, independently of eczema, exhibiting a consistent direction of effect between discovery and replication analyses. Despite there being noted differences in some clinical characteristics between the discovery and replication populations (eczema rates, ethnicity strata, age of diagnosis), and criteria used for phenotyping (OFC, 95%PPV) a meta-analysis across both cohorts suggested limited heterogeneity of observed effect sizes between studies, indicative of a robust association. We also found evidence that rs1295686 was associated with elevated total plasma IgE levels supporting a functional role in disease etiology.

The rs1295686 variant is a T -> C polymorphism located in the intronic region of *IL13* that has previously been associated with asthma in a consortium-based meta-analysis (6), with dysregulated total plasma IgE in a US population (4), and with elevated cord blood plasma IgE levels (11, 14). This particular functional variant is in complete linkage disequilibrium with rs20541 (**Supplementary figure 1**), an A -> G missense mutation that

has previously been linked to allergic rhinitis in a meta-analysis of eight studies (10), asthma in a meta-analysis of 34 studies (7) and elevated IgE levels (4). In our study we found evidence that rs1295686 was associated with elevated plasma IgE levels in addition to food sensitisation and food allergy, however we are unable to distinguish which variant is driving these associations, and future fine mapping studies would be needed to resolve this.

A strength of this study was the robust phenotyping measures in the discovery cohort determined by open food challenge with predetermined, objective diagnostic criteria, standard for infants as per PRACTALL food challenge guidelines (29). These measures provide the study with the rare opportunity to distinguish between the mechanisms of allergic sensitisation and clinical food allergy. We found only weak evidence ($p = 0.06$) that rs1295686 was associated with food sensitisation although the direction and size of the effect was consistent with that observed for the association between rs1295686 and clinical food allergy. These observations suggest that the lack of association was most likely a function of reduced sample size. There was no evidence of an effect when clinically allergic cases were compared to sensitised tolerant controls suggesting that variant rs1295686 increases the risk of food allergy via sensitisation to dietary antigen.

We carefully addressed the issue of population structure in the discovery phase by CMH analyses with clusters defined by parental country of birth and validated our categories using genome-wide SNP data. In the replication sample population structure was adjusted for with principal components. One limitation of the study was the absence of challenge-proven food allergy phenotyping measures in some of the studies utilised for replication, instead relying upon 95% PPV definitions. This may have led to inclusion of some food sensitised but tolerant individuals in the groups and/or the exclusion of some true food allergy cases. We addressed this by comparing effect sizes across the different study populations within the

replication sample, which did not suggest phenotype definition or age at diagnosis between studies as a significant source of heterogeneity.

Our study establishes association between food allergy and *IL13* genetic variation in two Australian food allergy samples. Further independent studies to confirm and extend these findings are now warranted. Understanding the genetic factors and molecular pathways that contribute to perturbation of Th2 immune responses will assist in the development of novel therapeutic approaches to immune-mediated diseases.

References

1. McKenzie GJ, Emson CL, Bell SE, Anderson S, Fallon P, Zurawski G, et al. Impaired development of Th2 cells in IL-13-deficient mice. Immunity. 1998;9(3):423-32.

2. Cameron L, Webster RB, Strempel JM, Kiesler P, Kabesch M, Ramachandran H, et al. Th2 cell-selective enhancement of human IL13 transcription by IL13-1112C> T, a polymorphism associated with allergic inflammation. The Journal of Immunology. 2006;177(12):8633-42.

3. DeMeo D, Silverman E, Senter J, Drazen J, Barth M, Laird N, et al. Univariate and multivariate family‐based association analysis of the IL‐13 ARG130GLN polymorphism in the Childhood Asthma Management program. Genetic epidemiology. 2002;23(4):335-48.

4. Granada M, Wilk JB, Tuzova M, Strachan DP, Weidinger S, Albrecht E, et al. A genome-wide association study of plasma total IgE concentrations in the Framingham Heart Study. Journal of Allergy and Clinical Immunology. 2012;129(3):840-5. e21.

5. Liu X, Nickel R, Beyer K, Wahn U, Ehrlich E, Freidhoff LR, et al. An IL13 coding region variant is associated with a high total serum IgE level and atopic dermatitis in the German multicenter atopy study (MAS-90). Journal of Allergy and Clinical Immunology. 2000;106(1):167-70.

6. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. New England Journal of Medicine. 2010;363(13):1211-21.

7. Nie W, Liu Y, Bian J, Li B, Xiu Q. Effects of polymorphisms-1112C/T and+ 2044A/G in interleukin-13 gene on asthma risk: a meta-analysis. PloS one. 2013;8(2):e56065.

8. Tsunemi Y, Saeki H, Nakamura K, Sekiya T, Hirai K, Kakinuma T, et al. Interleukin-13 gene polymorphism G4257A is associated with atopic dermatitis in Japanese patients. Journal of dermatological science. 2002;30(2):100-7.

9. van der Pouw Kraan Tv, Van Veen A, Boeije L, Van Tuyl S, De Groot E, Stapel S, et al. An IL-13 promoter polymorphism associated with increased risk of allergic asthma. Genes and immunity. 1999;1(1):61-5.

10. Ying X-J, Zhao S-W, Wang G-L, Xie J, Xu H-M, Dong P. Association of interleukin-13 SNP rs20541 with allergic rhinitis risk: a meta-analysis. Gene. 2013;521(2):222-6.

11. Sadeghnejad A, Karmaus W, Hasan Arshad S, Ewart S. IL13 gene polymorphism association with cord serum immunoglobulin E. Pediatric allergy and immunology. 2007;18(4):288-92.

12. Yang K, Chang JC, Chuang H, Liang HM, Kuo HC, Lee YS, et al. Gene–gene and gene–environment interactions on IgE production in prenatal stage. Allergy. 2010;65(6):731-9.

13. Bottema RW, Reijmerink NE, Kerkhof M, Koppelman GH, Stelma FF, Gerritsen J, et al. Interleukin 13, CD14, pet and tobacco smoke influence atopy in three Dutch cohorts: the allergenic study. European Respiratory Journal. 2008;32(3):593-602.

14. Hong X, Tsai H-J, Liu X, Arguelles L, Kumar R, Wang G, et al. Does genetic regulation of IgE begin in utero? Evidence from T H 1/T H 2 gene polymorphisms and cord blood total IgE. Journal of Allergy and Clinical Immunology. 2010;126(5):1059-67. e1.

15. Rosenbaum D, Hawkins G, Zhu L, Chen Y, Bleecker E, Meyers D, et al. The rare IL-13R105Q variant, strongly enhances IL-13 protein activity (HYP6P. 268). The Journal of Immunology. 2014;192(1 Supplement):118.13-.13.

16. Vladich FD, Brazille SM, Stern D, Peck ML, Ghittoni R, Vercelli D. IL-13 R130Q, a common variant associated with allergy and asthma, enhances effector mechanisms essential for human allergic inflammation. The Journal of clinical investigation. 2005;115(3):747-54.

17. Howard TD, Whittaker PA, Zaiman AL, Koppelman GH, Xu J, Hanley MT, et al. Identification and association of polymorphisms in the interleukin-13 gene with asthma and

atopy in a Dutch population. American Journal of Respiratory Cell and Molecular Biology. 2001;25(3):377-84.

18. 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;113(3):489-95.

19. 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. Journal of Allergy and Clinical Immunology. 2011;127(3):668-76. e2.

20. Koplin JJ, Wake M, Dharmage SC, Matheson M, Tang ML, Gurrin LC, et al. Cohort Profile: The HealthNuts Study: Population prevalence and environmental/genetic predictors of food allergy. International journal of epidemiology. 2015;44(4):1161-71.

21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263-5.

22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics. 2007;81(3):559-75.

23. Vuillermin P, Saffery R, Allen KJ, Carlin JB, Tang ML, Ranganathan S, et al. Cohort Profile: The Barwon Infant Study. Int J Epidemiol. 2015;44(4):1148-60. Epub 2015/04/02.

24. Lowe AJ, Lodge CJ, Allen KJ, Abramson MJ, Matheson MC, Thomas PS, et al. Cohort Profile: Melbourne Atopy Cohort study (MACS). International journal of epidemiology. 2016:dyw011.

25. Zurzolo GA, Allen KJ, Taylor SL, Shreffler WG, Baumert JL, Tang ML, et al. Peanut Allergen Threshold Study (PATS): validation of eliciting doses using a novel single-dose challenge protocol. Allergy, Asthma & Clinical Immunology. 2013;9(1):35.

26. Tang ML, Ponsonby A-L, Orsini F, Tey D, Robinson M, Su EL, et al. Administration of a probiotic with peanut oral immunotherapy: a randomized trial. Journal of Allergy and Clinical Immunology. 2015;135(3):737-44. e8.

27. Peters RL, Allen KJ, Dharmage SC, Tang ML, Koplin JJ, Ponsonby A-L, et al. Skin prick test responses and allergen-specific IgE levels as predictors of peanut, egg, and sesame allergy in infants. Journal of Allergy and Clinical Immunology. 2013;132(4):874-80.

28. Bousman C, Yung A, Pantelis C, Ellis J, Chavez R, Nelson B, et al. Effects of NRG1 and DAOA genetic variation on transition to psychosis in individuals at ultra-high risk for psychosis. Translational psychiatry. 2013;3(4):e251.

29. Sampson HA, van Wijk RG, Bindslev-Jensen C, Sicherer S, Teuber SS, Burks AW, et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology-European Academy of Allergy and Clinical Immunology PRACTALL consensus report. Journal of Allergy and Clinical Immunology. 2012;130(6):1260.

Table 1: Cochran-Mantel-Haenszel 2x2xK associations between *IL13* SNPs and food allergy phenotypes in the discovery sample, with strata defined by reported parental country of birth A) 367 food allergic cases and 156 non-allergic controls only B) 199 food sensitised, tolerant and 156 non-allergic controls C) 367 food allergic cases and 199 food sensitised-tolerant cases

		A. Food allergic cases vs non-food allergic controls				*B. Food sensitised-tolerant vs non-food allergic controls*				*C. Food allergic cases vs sensitised-tolerant cases*			
SNP	A1	P	OR	L95	U95	P	OR	L95	U95	P	OR	L95	U95
rs1295686	A	**0.003**	1.75	1.20	2.53	0.06	1.48	0.98	2.23	0.23	1.18	0.90	1.55
rs2243297	A	0.13	2.10	0.80	5.51	0.23	1.91	0.67	5.42	0.72	1.12	0.61	2.04
rs1295687	G	0.19	1.63	0.78	3.41	0.05	2.19	1.01	4.76	0.25	0.77	0.49	1.21
rs2243211	A	0.22	1.45	0.80	2.64	0.87	0.94	0.48	1.87	0.08	1.52	0.95	2.44
rs1295683	T	0.25	1.32	0.82	2.13	0.32	1.30	0.78	2.19	0.88	1.03	0.72	1.46
rs2243248	G	0.51	1.21	0.69	2.11	0.37	0.74	0.38	1.42	0.03	1.68	1.05	2.69
rs2243300	T	0.51	1.19	0.70	2.02	0.43	0.78	0.42	1.45	0.11	1.43	0.92	2.23
rs1800925	T	0.60	1.11	0.75	1.63	0.84	1.05	0.68	1.60	0.74	1.05	0.78	1.43
rs3091307	G	0.62	1.10	0.75	1.60	0.67	1.10	0.72	1.66	0.90	1.02	0.76	1.37

*P is the p-value determined by CMH testing with clusters defined by parental country of birth, A1 is the effect allele, OR is the odds ratio and L95 and U95 the lower and upper 95% confidence intervals.

Table 2 A. Logistic regression analysis of SNP-food allergy associations in the replication sample, adjusted for ancestry by principal components. 203 food allergic cases vs 330 non-food allergic controls. **B.** Meta-analysis of discovery and replication samples modelled with random and fixed effects. 570 food allergic cases vs 486 non-food allergic controls

	A. Replication analysis: food allergic cases vs non-food allergic controls					*B. Meta-analysis – Discovery and Replication*					
SNP	A1	OR	L95	U95	P	P	P(R)	OR	OR(R)	Q	I
rs1295686	A	1.37	1.03	1.82	0.03	0.0005	0.0006	1.50	1.50	0.31	3.32
rs1295687	C	1.10	0.68	1.78	0.70	0.30	0.30	1.24	1.24	0.38	0.00

*__A)__ A1 is the reference allele; OR is the odds ratio; L95 and U95 are the lower and upper confidence intervals; P is the p-value obtained from logistic regression **B)** P is the P-value calculated by fixed-effects meta-analysis model; P(R) is the P-value determined from random-effects meta-analysis model. OR is the odds ratio determined from fixed-effects; OR(R) is the odds ratio determined from random-effects. The Q value is Cochrane's P-value, a measure of effect size heterogeneity between studies. I is the I-index (0-100), another measure of effect size heterogeneity between studies.

Supplementary material for: *Genetic Variation at the Th2 Immune Gene IL13 is Associated with IgE-mediated Paediatric Food Allergy*

Sarah Ashley[1,2], Hern-Tze Tina Tan[1,3], Rachel Peters[1,4], Katrina Allen[1,4,9,10], Peter Vuillermin[1,6,7], Shyamali C Dharmage[1,8], Mimi LK Tang[1,4,9], Jennifer Koplin[1,4], Adrian Lowe[1,8], Anne-Louise Ponsonby[1,4], John Molloy[1,7], Melanie C Matheson[1,4], Richard Saffery[1,2,4], Justine A Ellis[1,4,11]*, David Martino[1,4,5]* and the HealthNuts team.

Supplemental methods

Use of Genome-wide SNP data to assess accuracy of ancestry strata defined by parental country of birth in the discovery sample

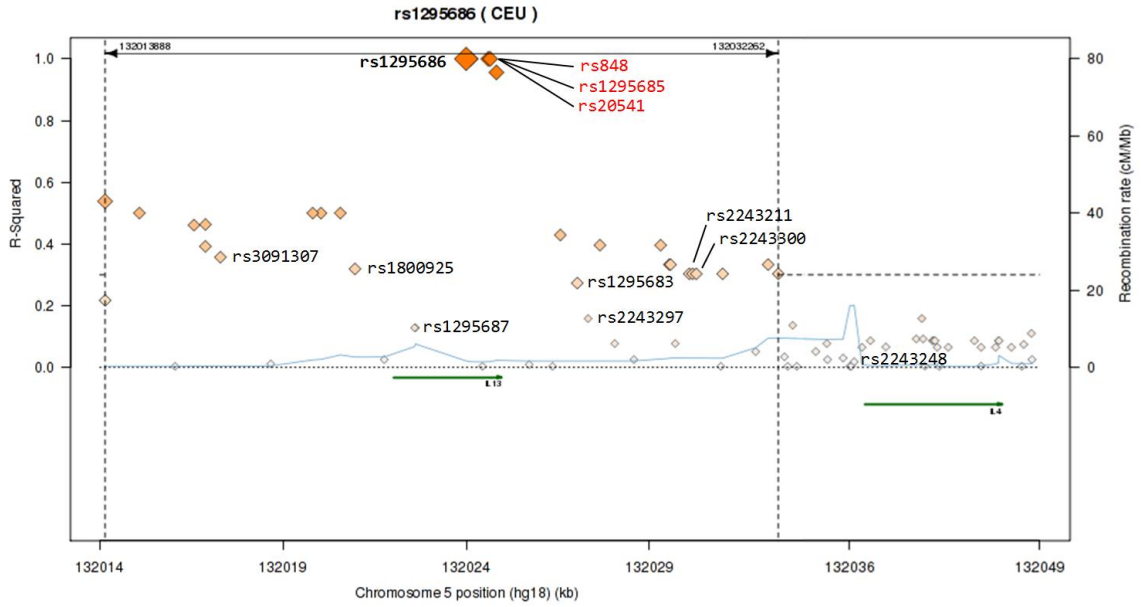
Genome-wide SNP data was available on 344 of the individuals of the discovery sample. This data was used to assess the accuracy of ancestral strata determined by reported parental country of birth. The Qiagen FlexiGene kit was used to extract DNA from peripheral blood according to manufacturers instructions. DNA was genotyped with the Illumina HumanOmni 2.5-8 SNP array at the Australian Genome Research Facility. SNPs were excluded if they had a minor allele frequency (MAF) of <1%, or MAF that deviated significantly from those reported in Europeans of the 1000 Genomes Project ($P<0.001$) (1000 Genomes Project Consortium, 2015), a call rate of <95%, and/or deviation from the Hardy Weinberg Equilibrium (HWE) with a P-value less than 10^{-6}. This resulted in a total of 389,427 SNPs post-quality control. Unrelated individuals were included in the analysis with call rates >95%. A multidimensional scaling (MDS) analysis of the identity-by-state (IBS) distance between individuals from the HealthNuts study and samples (of known ancestry) from the 1000 Genomes project was used to infer ancestry of the HealthNuts sample. Stringent cut-offs were visually determined based on clusters of different ancestral populations from the 1000 Genomes project. These genetically inferred ancestry strata were then compared to

ancestry strata defined by parental country of birth on the same HealthNuts participants to assess accuracy of these parent-reported categories.

1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. Nature, 526(7571), 68-74.

Supplementary table 1: Clinical characteristics of 722 (Post-QC) infants included in the discovery cohort.

	Non-allergic controls (NA) N=156	Food sensitised, tolerant cases (FS) N=199	Food Allergy cases (FA) N=367	P-value
Infant demographics				
Age in months at recruitment (mean, SD)	12.6 (0.65)	12.6 (0.68)	12.7(0.75)	0.18
Gender (% male)	47.4%	54.3%	58.9%	0.03
Reported ethnicity				
Asian	3.3%	17%	14%	8.8×10^{-6}
European	86.8%	66%	61.5%	6.0×10^{-7}
Mixed European/Asian	9.9%	24.1%	24.4%	0.002
Infant clinical characteristics				
History of eczema	25.0%	37.7%	56.4%	2.2×10^{-7}
Food sensitisation				
Egg sensitisation	0%	29.6%	93.7%	-
Sesame sensitisation	0%	5.5%	9.0%	-
Peanut sensitisation	0%	30.2%	49.9%	-
Food allergy				
Egg allergy	0%	0%	89.3%	-
Sesame allergy	0%	0%	6.7%	-
Peanut allergy	0%	0%	32.2%	-
Family characteristics				
Any siblings	54.5%	48.7%	44.3%	0.009
Asthma				
Maternal asthma	19.9%	21.1%	22.9%	0.43
Paternal asthma	18.6%	15.6%	19.1%	0.72
Sibling asthma	10.9%	9.6%	10.1%	0.83
Hay fever				
Maternal hay fever	22.4%	15.1%	24.0%	0.93
Paternal hay fever	35.3%	28.1%	37.9%	0.30
Sibling hay fever	5.8%	3.5%	6.5%	0.50
Eczema				
Maternal eczema	22.4%	14.1%	22.1%	0.35
Paternal eczema	14.7%	10.1%	12.3%	0.59
Sibling eczema	22.4%	16.6%	17.2%	0.22
Food allergy				
Maternal food allergy	12.2%	6.0%	6.3%	0.04
Paternal food allergy	11.5%	4.5%	4.4%	0.004
Sibling food allergy	10.9%	5.5%	6.8%	0.18



Supplementary Figure 1: LD plot, in reference to SNP rs1295686, of the genotyped region covered by the tag-SNPs (labelled) of the discovery study. Figure generated from 1000 Genomes CEU data using the Broad Institute's SNAP facility. Three additional known allergy associated variants, rs848, rs1295685 and rs20541, have also been labelled in red to demonstrate their shared LD with our food allergy associated variant rs1295685.

Supplementary Table 2: Demographics of the 533 infants of the replication sample group

	Non-allergic controls (NA) N=330	Food Allergy cases (FA) N=203	P-value
Infant demographics			
Age in years at recruitment (mean, SD)	1.1(0.1)	4.1(4.0)	<0.001
Gender (% male)	51.9%	55.3%	0.47
Reported ethnicity			
Asian	0%	7.5%	-
European	78.5%	66.8%	0.001
Mixed European/Asian	21.5%	25.6%	0.23
Infant clinical characteristics			
History of eczema	33.0%	41.9%	0.003
Food allergy			
Egg allergy	0%	72.5%	-
Sesame allergy	0%	2.8%	-
Peanut allergy	0%	62.1%	-

Supplementary Table 3: CMH analysis for food allergy in individuals without eczema

SNP	*A1*	*P*	*OR*	*L95*	*U95*
rs1295686	A	0.008	1.82	1.17	2.83
rs2243297	A	0.08	2.45	0.89	6.77
rs1295687	G	0.16	1.99	0.77	5.14
rs1295683	T	0.16	1.51	0.84	2.69
rs2243300	T	0.29	1.40	0.75	2.62
rs2243211	A	0.33	1.40	0.71	2.73
rs2243248	G	0.57	1.21	0.63	2.35
rs3091307	G	0.73	1.08	0.70	1.66
rs1800925	T	0.77	1.07	0.69	1.66

*A1 is the reference allele, P is the p-value for the CMH test, OR is the corresponding odds ratio and, L95 and U95 are the lower and upper limits of the 95% confidence interval.

Supplementary Table 4: CMH analysis for food allergy in the discovery sample, adjusted for the allergic parameters in the infant and family found to be significantly different across the three groups (see Supplementary Table 1).

SNP	*A1*	*P*	*OR*	*L95*	*U95*
rs1295686	A	0.004	1.73	1.19	2.51
rs2243297	A	0.16	2.00	0.76	5.26
rs1295687	G	0.18	1.66	0.79	3.48
rs1295683	T	0.24	1.33	0.83	2.14
rs2243211	A	0.28	1.39	0.76	2.55
rs2243248	G	0.61	1.16	0.66	2.03
rs2243300	T	0.62	1.14	0.67	1.95
rs1800925	T	0.79	1.06	0.72	1.55
rs3091307	G	0.80	1.05	0.72	1.53

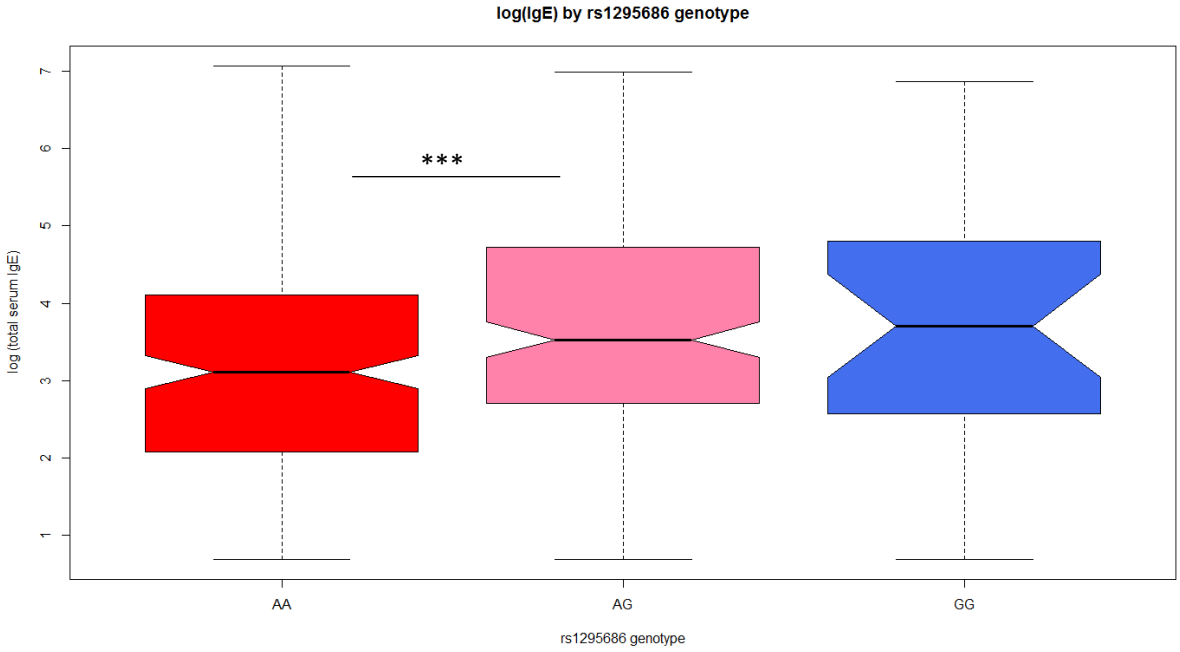
*CMH analysis adjusted for sex, eczema, presence of siblings and maternal and/or paternal food allergy. 356 food allergic cases and 152 non-allergic controls.

Supplementary Table 5: PLINK association test for *IL13* variants and log transformed total serum IgE, adjusted for food allergy in the discovery cohort

Log(total serum IgE)

SNP	A1	BETA	SE	L95	U95	P
rs1295686	A	0.38	0.10	0.17	0.58	0.0003
rs3091307	G	0.29	0.11	0.08	0.50	0.007
rs1800925	T	0.26	0.11	0.05	0.48	0.02
rs1295683	T	0.29	0.13	0.04	0.55	0.03
rs2243300	T	0.11	0.16	-0.20	0.43	0.48
rs2243297	A	0.13	0.23	-0.32	0.57	0.57
rs1295687	G	-0.07	0.18	-0.43	0.28	0.68
rs2243211	A	-0.03	0.17	-0.35	0.29	0.86
rs2243248	G	0.02	0.17	-0.31	0.35	0.89

*A1 is the reference allele, BETA is the beta value for the qualitative association test, SE is the standard error, L95 and U95 are the lower and upper limits of the 95% confidence interval and P is the P-value



Supplementary figure 2: Box plot of log transformed IgE levels by rs1295686 genotype in the discovery cohort. The y-axis is the log transformed total serum IgE values, with each separate genotype of rs1295686 presented on the x-axis. Genotype AA (red) is the major allele homozygous group and GG (blue) are the group with carriage of both minor alleles. Tukey multiple comparisons of means p=0.0003 between the major allele homozygous group and heterozygous group.