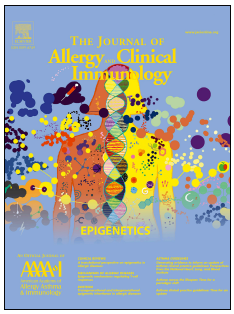
Accepted Manuscript



The *MALT1* locus and Peanut Avoidance in the Risk for Peanut Allergy

Alexandra Winters, PhD, Henry T. Bahnson, MPH, Ingo Ruczinski, PhD, Meher P. Boorgula, MS, Claire Malley, MS, Ali R. Keramati, MD, Sameer Chavan, MS, David Larson, PhD, Karen Cerosaletti, PhD, Peter H. Sayre, MD, PhD, Marshall Plaut, MD, George Du Toit, MB, BCh, FRCPCH, Gideon Lack, MB, BCh, FRCPCH, Kathleen C. Barnes, PhD, Gerald T. Nepom, MD, PhD, Rasika A. Mathias, ScD, for the Immune Tolerance Network LEAP Study Team

PII: S0091-6749(19)30279-9

DOI: https://doi.org/10.1016/j.jaci.2019.02.016

Reference: YMAI 13908

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 4 December 2018

Revised Date: 25 January 2019

Accepted Date: 19 February 2019

Please cite this article as: Winters A, Bahnson HT, Ruczinski I, Boorgula MP, Malley C, Keramati AR, Chavan S, Larson D, Cerosaletti K, Sayre PH, Plaut M, Du Toit G, Lack G, Barnes KC, Nepom GT, Mathias RA, for the Immune Tolerance Network LEAP Study Team, The *MALT1* locus and Peanut Avoidance in the Risk for Peanut Allergy, *Journal of Allergy and Clinical Immunology* (2019), doi: https://doi.org/10.1016/j.jaci.2019.02.016.

The *MALT1* locus and Peanut Avoidance in the Risk for Peanut Allergy

Alexandra Winters, PhD,[1] Henry T Bahnson, MPH,[2,3] Ingo Ruczinski, PhD,[4] Meher P Boorgula, MS,[5] Claire Malley, MS,[1] Ali R Keramati, MD,[1] Sameer Chavan, MS,[5] David Larson, PhD,[3] Karen Cerosaletti, PhD,[2] Peter H Sayre, MD, PhD,[3,6] Marshall Plaut, MD,[7] George Du Toit, MB, BCh, FRCPCH,[8] Gideon Lack, MB, BCh, FRCPCH,[8] Kathleen C Barnes, PhD,[5] Gerald T Nepom, MD, PhD,[3] Rasika A Mathias, ScD,[1,*] for the Immune Tolerance Network LEAP Study Team.

* Address correspondence to: Rasika A Mathias, ScD. rmathias@jhmi.edu. 5501 Hopkins Bayview Circle, Baltimore, MD 21224. 410-550-2487.

1. Department of Medicine, Johns Hopkins University, Baltimore, MD, USA.
2. Benaroya Research Institute, Seattle, WA, USA.
3. Immune Tolerance Network, San Francisco, CA, USA.
4. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.
5. Department of Medicine, University of Colorado, Anschutz, CO, USA.
6. Department of Medicine, University of California, San Francisco, CA, USA.
7. Allergy, Asthma and Airway Biology Branch, DAIT, NIAID, Bethesda, MD, USA.
8. Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, and Guy's and St. Thomas' NHS Foundation Trust, London, United Kingdom.

This research was performed as a project of the Immune Tolerance Network, an international clinical research consortium headquartered at the Benaroya Research Institute, and supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health under Award Number UM1AI109565.

Conflict of Interest

Dr. Lack reports holding stock and stock options in DBV Technologies. No other potential conflict of interest relevant to this article was reported.

Capsule Summary: We identified a strong association between peanut allergy and the *MALT1* locus in LEAP participants in the peanut avoidance group with 58.6% of carriers developing peanut allergy at 60 months as compared to 12.7% of non-carriers.

Key words: Peanut allergy, *MALT1*, GWAS, food allergy, immunogenetics, early allergen exposure, IgE

To the Editor:

The Learning Early about Peanut Allergy (LEAP) trial (1, 2) motivated a change in pediatric guidelines for the early introduction of dietary peanut as an effective strategy for the prevention of peanut allergy. LEAP participants were presumed to be at increased risk for peanut allergy(3), and dietary introduction of peanut protein beginning in the first four to eleven months of life significantly decreased the frequency of peanut allergy later in childhood and modulated the immune response to peanuts in this at-risk group (1). To identify the genetic determinants of peanut allergy in the LEAP participants, whole genome sequencing (WGS) was performed (***Supplementary Information Sections 1-2, Table S1***). Following published standards for WGS data (4) (***Supplementary Information Sections 3-4***), there were 542 per protocol LEAP participants available for genome-wide genetic association tests, including 49 with peanut allergy, which was defined as a positive result on a double-blind placebo-controlled oral food challenge at 60 months of age (***Supplementary Information Section 1)***. The de-convolution of genetic ancestry aligns well with self-reported race/ethnicity (***Supplementary Information Section 5, Fig S1***). Given the high success of early introduction of dietary peanut in the LEAP trial, 48 peanut allergy participants were from the peanut avoidance arm and only one from the consumption arm (***Table S1***). Therefore, genome-wide association was assessed for peanut allergy in the 275 participants from the avoidance arm (N=48 peanut allergic/N=227 non peanut allergic) on a total of 4,444,069 single nucleotide variants (SNVs, ***Fig 1***, ***Supplementary Information Section 6, Fig S2***) in a discovery analysis. Subsequent follow up of the peak genetic signal(s) was extended to include the participants from the consumption arm (N=267) with immunological quantitative traits of importance (***Fig 2***) to facilitate the examination of the identified genetic loci in the context of the intervention.

The peak association for peanut allergy in the avoidance group was observed on chromosome 18 (***Fig 1A*)** mapping to the mucosa-associated lymphoid tissue lymphoma translocation (*MALT1*) gene (***Fig 1C, Table S2***). The region (Chr18:56337602..56456191) includes strong regulatory signatures for MALT1 expression as well as the expression of the intergenic non-coding RNA (lincRNA), RP11-108P20.1 in the Genotype-Tissue Expression (GTEx) data (***Figs S3, S4***). However, the specific set of SNVs with p-value $<10^{-5}$ for peanut allergy only have eQTL signatures for MALT1 (***Table S3***). The peak associated SNV was

rs57265082 with an estimated Odds Ratio (OR) of 10.99, minor allele frequency (MAF) of 5.6%, and $p = 6.49 \times 10^{-8}$. Gene-based analysis was performed across rare exonic SNVs (MAF ≤5%) using SKAT (***Tables S4*** and ***S5***). There was nominal association with either all rare exonic SNVs ($p = 0.0830$) or all rare damaging exonic SNVs ($p = 0.0828$); however, with the inclusion of the peak WGS variant, rs57265082, the gene-based evidence was very strong ($p = 1.89 \times 10^{-10}$). Conditioning on the peak SNV, rs57265082, shows that the observed common variant signal is a single genetic locus within the region (***Fig S5***). There are overall strong differences in the clinical profiles of the *MALT1* risk allele carriers compared to non-carriers within the peanut avoidance participants (***Table S6***). *MALT1* is not associated with baseline selection criteria of egg allergy or eczema ($p = 0.3241$, and $p=0.1626$ respectively in the avoidance group), and the association between peanut allergy and *MALT1* is independent of these baseline selections (***Table S7***). We observe no association between the key filaggrin variant, R501X, documented to play a role in eczema and peanut allergy ($p=0.4014$ and MAF of 3.6% in the avoidance group), but recognize that our sample size of N=275 may be underpowered for this.

We observe a weaker association with rs57265082 to sensitization (at 60 months, sensitization is defined as those with peanut-specific IgE ≥ 0.1 kU/liter) in the peanut avoidance group (OR = 4.55, $p = 0.0011$). Additionally, the *MALT1* locus remains significantly associated with peanut allergy ($p = 0.0003$), even within the subset of sensitized participants in the peanut avoidance group (***Fig 2A*),** supporting its role as a genetic risk factor for allergy and not only sensitization. With the inclusion of the LEAP participants from the consumption arm (N=267), *MALT1* was found to be significantly associated with an IgE response to multiple specific peanut allergenic protein components Ara h1, Ara h2, and Ara h3 at 60 months ($p = 1.11 \times 10^{-5}$, ***Fig 2B, Supplementary Section 7***) in the full set of LEAP participants adjusting for intervention.

When examining specific IgE to peanut, as well as the three major allergenic components of peanut, we observe a progressive divergence in the upper end of the IgE distributions in *MALT1* carriers (***Fig 2C***) with two key observations to note. First, the intervention with peanut exposure effectively reduced peanut-specific IgE *irrespective* of carrier status (truncated distributions in ***Fig 2C, bottom panel***). Second, within the avoidance group, the levels of peanut-specific IgE between the carriers and non-carriers is markedly different; rs57265082 carriers

within the peanut avoidance group had the highest peanut-specific IgE levels as compared to non-carriers (***Fig 2C, upper panel***). The mean titers of peanut-specific IgE were significantly different between carriers vs non-carriers and by treatment group (interaction $p = 1.86x10^{-5}$), even after adjusting for the baseline differences in peanut-specific IgE (***Fig S6A***). Importantly, this effect of *MALT1* on peanut-specific IgE in the peanut avoidance group is independent of total IgE (**Fig S6B**, $p=2.03x10^{-5}$ for peanut-specific IgE and p=0.366 for total IgE). Finally, the additional value of knowing rs57265082 carrier status in predicting an individual's likelihood of allergy was evaluated, and rs57265082 was found to be an independent predictor of allergy in the avoidance group (***Fig S7***).

In this first report of the genetics of peanut allergy within the LEAP study, a key biological candidate, the *MALT1* gene, is implicated as an independent risk factor for peanut allergy in the context of peanut avoidance. These associations are irrespective of sensitization status (in **Fig S7**, sensitization at baseline is defined by skin prick positivity, and in **Fig 2A**, sensitization at 60 months sensitization is defined as peanut-specific IgE $\geq$ 0.1 kU/liter), supporting a relationship with progression to symptomatic allergy after peanut sensitization, a disease pattern that is inhibited by early and continuous consumption of peanuts. *MALT1* encodes a paracaspase that functions as a critical part of the *CARMA1-BCL10-MALT1* (CBM) complex, causing NF-KappaB activation in B and T cells in response to an antigen binding to the B or T cell receptor(5). In T cells, this forms part of the signaling cascade leading to T cell activation(6) and involves the two MALT1 isoforms, MALT1A and MALT1B(7). Given that our top SNVs affect MALT1 expression, it is possible that these variants may predispose an individual to greater allergic disease by altering MALT1 expression or affecting the ratio of MALT1A to MALT1B, thus increasing Th2 differentiation after antigen presentation. Additional genes encoding other members of the CBM complex do not show evidence for association within our discovery data (***Fig S8***).

MALT1 has not been implicated in prior genetic studies, and we are also unable to replicate prior published associations (***Table S8***) (8, 9). It is important to note that the prior studies compare non-allergic controls to peanut-allergic subjects (8, 9), and the genetic associations identified in these likely represent risk of allergic sensitization and not specifically

peanut allergy. In contrast, the LEAP study included only participants who were at high risk for peanut allergy, many of whom were sensitized at baseline, and this unique ascertainment of the LEAP study facilitates our ability to test specifically for the risk of peanut allergy. Yet another singular advantage of the LEAP study is that we are able to interrogate the avoidance group (high incidence of peanut allergy) and contrast this to the consumption group (low incidence of peanut allergy) using quantitative immunological markers to identify the genetic determinants of peanut allergy that are relevant in the absence of peanut exposure. This homogeneity of exposure (i.e. avoidance) and ascertainment (i.e. baseline risk factors) within LEAP account for the ability to detect a strong association with *MALT1* despite the limited sample size of N=275 in the discovery analysis; in fact the p-value of 6.49×10^{-8} for the single variant tests is near the Bonferroni threshold for GWAS significance (5×10^{-8}), and our gene-based analysis results in a p = 1.89×10^{-10}. Targeted genotyping of rs57265082 on additional LEAP participants, including the non per protocol participants, does not change the results from the discovery sample (***Supplementary Section 8, Table S9***). Furthermore, of the seven participants within the consumption arm that had peanut allergy at baseline, three were *MALT1* carriers (unadjusted OR for peanut allergy at baseline in the LEAP consumption group = 5.3, p = 0.0188 using a Pearson chi-square test). However, the lack of a suitable population to use as a replication group is a major limitation of this study, and additional replication will be important to follow up on these associations observed within LEAP.

One striking observation is the differing effect of *MALT1* carrier status on peanut-specific IgE patterns between the two intervention arms in LEAP. The introduction of dietary peanut as a strategy for the prevention of peanut allergy is equally effective within carriers and non-carriers. However, our results indicate that within the LEAP participants, *MALT1* carriers from the peanut avoidance group have the highest risk for peanut allergy (58.6% of carriers of the *MALT1* variant in the avoidance group go on to get peanut allergy in contrast to only 12.7% of the non-carriers, ***Table S6***). Coupled with the observations that 1) the acquisition of additional peanut antigen target specificities in the IgE response is markedly increased in the *MALT1* carriers and 2) this peanut-specific IgE response is independent of total IgE, our findings support a genotype-phenotype relationship that implicates the MALT1 pathway in the allergic immune pathogenesis of peanut allergy.

Alexandra Winters, PhD[1]
Henry T Bahnson, MPH[2,3]
Ingo Ruczinski, PhD[4]
Meher P Boorgula, MS[5]
Claire Malley, MS[1]
Ali R Keramati, MD[1]
Sameer Chavan, MS[5]
David Larson, PhD[3]
Karen Cerosaletti, PhD[2]
Peter H Sayre, MD, PhD[3,6]
Marshall Plaut, MD[7]
George Du Toit, MB, BCh, FRCPCH[8]
Gideon Lack, MB, BCh, FRCPCH[8]
Kathleen C Barnes, PhD[5]
Gerald T Nepom, MD, PhD[3]
Rasika A Mathias, ScD[1]
for the Immune Tolerance Network LEAP Study Team

1. Department of Medicine, Johns Hopkins University, Baltimore, MD, USA.
2. Benaroya Research Institute, Seattle, WA, USA.
3. Immune Tolerance Network, San Francisco, CA, USA.
4. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.
5. Department of Medicine, University of Colorado, Anschutz, CO, USA.
6. Department of Medicine, University of California, San Francisco, CA, USA.
7. Allergy, Asthma and Airway Biology Branch, DAIT, NIAID, Bethesda, MD, USA.
8. Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, and Guy's and St. Thomas' NHS Foundation Trust, London, United Kingdom.

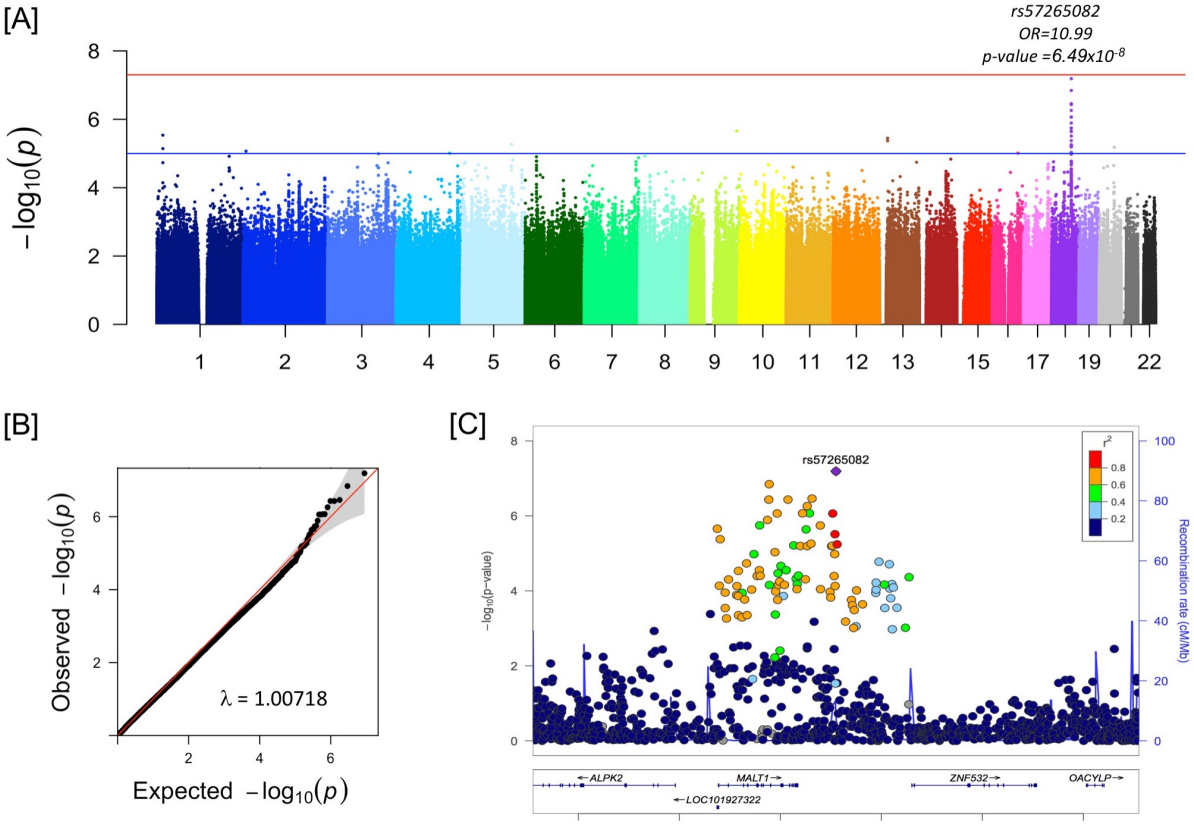
1. Du Toit G, Roberts G, Sayre PH, Bahnson HT, Radulovic S, Santos AF, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. N Engl J Med. 2015;372(9):803-13.
2. Togias A, Cooper SF, Acebal ML, Assa'ad A, Baker JR, Jr., Beck LA, et al. Addendum guidelines for the prevention of peanut allergy in the United States: Report of the National Institute of Allergy and Infectious Diseases-sponsored expert panel. Ann Allergy Asthma Immunol. 2017;118(2):166-73 e7.
3. Du Toit G, Roberts G, Sayre PH, Plaut M, Bahnson HT, Mitchell H, et al. Identifying infants at high risk of peanut allergy: the Learning Early About Peanut Allergy (LEAP) screening study. J Allergy Clin Immunol. 2013;131(1):135-43 e1-12.
4. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, et al. A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. Nat Commun. 2016;7:12522.
5. Thome M. CARMA1, BCL-10 and MALT1 in lymphocyte development and activation. Nat Rev Immunol. 2004;4(5):348-59.
6. Vickery BP, Chin S, Burks AW. Pathophysiology of food allergy. Pediatr Clin North Am. 2011;58(2):363-76, ix-x.
7. Meininger I, Griesbach RA, Hu D, Gehring T, Seeholzer T, Bertossi A, et al. Alternative splicing of MALT1 controls signalling and activation of CD4(+) T cells. Nat Commun. 2016;7:11292.
8. Asai Y, Eslami A, van Ginkel CD, Akhabir L, Wan M, Yin D, et al. A Canadian genome-wide association study and meta-analysis confirm HLA as a risk factor for peanut allergy independent of asthma. J Allergy Clin Immunol. 2018;141(4):1513-6.
9. Asai Y, Eslami A, van Ginkel CD, Akhabir L, Wan M, Ellis G, et al. Genome-wide association study and meta-analysis in multiple populations identifies new loci for peanut allergy and establishes C11orf30/EMSY as a genetic risk factor for food allergy. J Allergy Clin Immunol. 2018;141(3):991-1001.

Figure 1: Genome-wide association with peanut allergy at 60 months in N=275 LEAP participants in the peanut avoidance group. **Panel A** is the genome-wide Manhattan plot for N=4,444,069 SNVs with MAF ≥ 2%, missingness < 5%, and Hardy-Weinberg Equilibrium p ≥ 10^{-6}. **Panel B** is the quantile-quantile plot of the same data as Panel A. **Panel C** shows the peak association region on chromosome 18.

Figure 2: **Panel A** shows the proportion of allergic and non-allergic LEAP participants at 60 months of age by treatment group and *MALT1* carrier status in all LEAP participants (left) and the sensitized group (right, defined as those with peanut-specific IgE ≥ 0.1 kU/liter at 60 months). **Panel B** shows the IgE response to Ara h1, Ara h2, and Ara h3 over the course of the LEAP study in all participants (N=542). Ara h status was imputed to 0 for all participants with peanut-specific IgE <0.1. **Panel C** is the proportion density plots showing the relative distribution of peanut-specific IgE and IgE to Ara h1, Ara h2, and Ara h3 between the *MALT1* Carrier and Non-carrier groups at 60 months of age. The horizontal reference line at 12% indicates the proportion of the population with at least one *MALT1* risk allele, which illustrates a null distribution with equal proportions of individuals at all titer levels between the carriers and non-carriers. Ara h status was imputed to -2 (log10) for all participants with peanut-specific IgE <0.1 kU/liter. For all panels, imputed genotypes were used for 7 individuals missing allele calls at rs57265082, and Non-carriers were defined as having at least one copy of the T allele (due to the low MAF).

Acknowledgements

This research was performed as a project of the Immune Tolerance Network, an international clinical research consortium headquartered at the Benaroya Research Institute, and supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health under Award Number UM1AI109565. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Drs. Daniel Rotrosen and Alkis Togias from the NIAID for their review of the manuscript, Ms. Monica Campbell from the University of Colorado for her help preparing samples for genotyping, and all the LEAP participants who took part in the study. Dr. Lack reports holding stock and stock options in DBV Technologies. No other potential conflict of interest relevant to this article was reported.



Supplementary Tables

The *MALT1* locus and Peanut Avoidance in the Risk for Peanut Allergy

Alexandra Winters, PhD,[1] Henry T Bahnson, MPH,[2,3] Ingo Ruczinski, PhD,[4] Meher P Boorgula, MS,[5] Claire Malley, MS,[1] Ali R Keramati, MD,[1] Sameer Chavan, MS,[5] David Larson, PhD,[3] Karen Cerosaletti, PhD,[2] Peter H Sayre, MD, PhD,[3,6] Marshall Plaut, MD,[7] George Du Toit, MB, BCh, FRCPCH,[8] Gideon Lack, MB, BCh, FRCPCH,[8] Kathleen C Barnes, PhD,[5] Gerald T Nepom, MD, PhD,[3] Rasika A Mathias, ScD,[1,*] for the Immune Tolerance Network LEAP Study Team.

* Address correspondence to: Rasika A Mathias, ScD. rmathias@jhmi.edu. 5501 Hopkins Bayview Circle, Baltimore, MD 21224. 410-550-2487.

1. Department of Medicine, Johns Hopkins University, Baltimore, MD, USA.
2. Benaroya Research Institute, Seattle, WA, USA.
3. Immune Tolerance Network, San Francisco, CA, USA.
4. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.
5. Department of Medicine, University of Colorado, Anschutz, CO, USA.
6. Department of Medicine, University of California, San Francisco, CA, USA.
7. Allergy, Asthma and Airway Biology Branch, DAIT, NIAID, Bethesda, MD, USA.
8. Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, and Guy's and St. Thomas' NHS Foundation Trust, London, United Kingdom.

Table S1: Clinical characteristics of the LEAP participants with WGS data.

	All Subjects	Avoidance Group	Consumption Group
N	542	275	267
Age in months at baseline, mean (SD)	7.80 (1.74)	7.89 (1.71)	7.72 (1.78)
N with female sex (%)	226 (42)	100 (36)	126 (47)
N with egg allergy at baseline (%)	345 (65)	179 (65)	166 (62)
N with eczema baseline (%)	480 (89)	241 (88)	239 (90)
SCORAD at baseline, mean (SD)	33.89 (18.79)	34.34 (19.31)	33.42 (18.26)
N with positive skin prick test to peanut at baseline (%)	83 (15)	50 (18)	33 (12)
Log peanut-specific IgG4 at 30 months, mean (SD)	2.57 (0.75)	2.12 (0.45)	3.04 (0.71)
Log peanut-specific IgG4 at 60 months, mean (SD)	2.60 (0.69)	2.28 (0.56)	2.923 (0.66)
N with peanut-specific IgG4 at 60 months >70 (%)	371 (68)	145 (53)	226 (85)
Log peanut-specific IgE at 60 months, mean (SD)	-0.91 (1.02)	-0.88 (1.13)	-0.94 (0.90)
Log peanut-specific IgG4:IgE ratio at 60 months, mean (SD)	3.13 (0.98)	2.79 (1.02)	3.48 (0.80)
Log Eosinophilia at 60 months, mean (SD)	-0.44 (0.37)	-0.41 (0.36)	-0.47 (0.38)
SCORAD at 60 months, mean (SD)	7.02 (10.78)	7.82 (11.63)	6.19 (9.79)
N with positive skin prick test to peanut at 60 months (%)	117 (22)	76 (28)	41 (15)
N Peanut Allergic at 60 months (%)	49 (9)	48 (17)	1 (0.4)

Table S2: Overview of 35 SNVs with p<10E-05 for association with peanut allergy in the N=275 peanut avoidance LEAP participants. * Alternate allele frequency (AAF) is provided as the additive model is coded for 0/1/2 copies of the alternate allele. **Permutation p-value is based on 10 million permutations using joint permutation of allergy status and model covariates.

		SNV Annotation								
SNV	**Position**	**Function**	**Gene (left,right)**	**GTEx eQTL**	**Reference/Alternate allele***	**OR**	**SE**	**P-value**	**Permutation P-value****	**AAF***
rs80151145	chr1:17717066	intronic	PADI6	EIF1AXP1	T/C	6.1630	0.3890	2.91E-06	2.90E-06	8%
rs145726195	chr1:17771192	intergenic	RCC2,ARHGEF10L		C/T	9.6269	0.5047	7.22E-06	5.70E-06	4%
rs72777284	chr2:8126454	intergenic	LINC00298,LINC00299		C/T	10.8121	0.5349	8.55E-06	8.60E-06	3%
rs12643843	chr4:154178697	UTR5	TRIM2	TRIM2	A/G	3.8705	0.3061	9.82E-06	3.00E-06	34%
rs74326323	chr5:141470422	intergenic	GNPDA1,NDFIP1		C/T	6.0359	0.3952	5.40E-06	2.40E-06	7%
rs11243375	chr9:134238611	intergenic	PLPP7,PRRC2B		T/C	12.0176	0.5253	2.21E-06	7.00E-07	4%
rs73156540	chr13:23254040	intergenic	LINC00540,BASP1P1		C/A	5.0656	0.3500	3.56E-06	2.40E-06	10%
rs73156541	chr13:23254052	intergenic	LINC00540,BASP1P1		A/G	4.9743	0.3489	4.26E-06	3.20E-06	10%
rs2526073	chr16:73994048	intergenic	LINC01568,LOC101928035		G/T	3.7257	0.2975	9.79E-06	3.70E-06	31%
rs4940418	chr18:56337602	UTR3	LOC101927322	MALT1	T/C	8.8243	0.4600	2.21E-06	1.60E-06	5%
rs4940419	chr18:56340438	intronic	MALT1	MALT1	G/C	9.0487	0.4788	4.21E-06	2.80E-06	4%
rs79002421	chr18:56379546	intronic	MALT1	MALT1	G/A	5.0116	0.3375	1.80E-06	9.00E-07	12%
rs77714205	chr18:56387301	intronic	MALT1	MALT1	T/G	8.7742	0.4486	1.29E-06	1.30E-06	5%
rs140235750	chr18:56388568	intronic	MALT1	MALT1	T/C	9.4843	0.4425	3.70E-07	5.00E-07	5%
rs77767290	chr18:56388996	intronic	MALT1	MALT1	G/A	12.1475	0.4747	1.44E-07	1.00E-07	4%
rs143282859	chr18:56394683	intronic	MALT1	MALT1	A/G	5.6584	0.391	9.31E-06	7.90E-06	9%
rs75022589	chr18:56396837	intronic	MALT1	MALT1	T/C	10.0751	0.4695	8.66E-07	7.00E-07	4%
rs77267911	chr18:56407555	intronic	MALT1	MALT1	A/G	9.4843	0.4425	3.70E-07	5.00E-07	5%
rs4940749	chr18:56413217	intronic	MALT1	MALT1	G/T	5.1256	0.3615	6.16E-06	4.20E-06	11%
rs4940750	chr18:56419779	intergenic	MALT1,ZNF532	MALT1	G/A	6.5355	0.4159	6.37E-06	7.50E-06	7%
rs76653504	chr18:56421922	intergenic	MALT1,ZNF532	MALT1	T/G	10.0982	0.4698	8.58E-07	7.00E-07	4%
rs8093073	chr18:56425466	intergenic	MALT1,ZNF532		C/T	7.5888	0.4289	2.30E-06	2.20E-06	6%
rs141101728	chr18:56426592	intergenic	MALT1,ZNF532	MALT1	G/T	9.5720	0.4510	5.50E-07	6.00E-07	6%
rs144838979	chr18:56426622	intergenic	MALT1,ZNF532	MALT1	G/A	7.8716	0.4573	6.42E-06	8.40E-06	5%

SNV	Position	SNV Annotation: Function	SNV Annotation: Gene (left,right)	SNV Annotation: GTEx eQTL	Reference/Alternate allele*	OR	SE	P-value	Permutation P-value**	AAF*
rs8095265	chr18:56428866	intergenic	MALT1,ZNF532	MALT1	T/C	5.1686	0.3337	8.56E-07	4.00E-07	13%
rs76767530	chr18:56430356	intergenic	MALT1,ZNF532	MALT1	G/A	9.4119	0.4935	5.54E-06	4.80E-06	5%
rs116985291	chr18:56431371	intergenic	MALT1,ZNF532	MALT1	C/T	10.6175	0.4636	3.48E-07	4.00E-07	4%
rs955405	chr18:56439552	intergenic	MALT1,ZNF532	MALT1	C/T	9.2614	0.4663	1.81E-06	1.70E-06	4%
rs59853704	chr18:56450773	intergenic	MALT1,ZNF532	MALT1	A/C	6.5751	0.4173	6.40E-06	8.20E-06	6%
rs4940751	chr18:56451388	intergenic	MALT1,ZNF532	MALT1	C/T	7.4946	0.4460	6.29E-06	5.40E-06	5%
rs73959507	chr18:56451862	intergenic	MALT1,ZNF532	MALT1	G/A	10.4314	0.4767	8.69E-07	9.00E-07	4%
rs4940752	chr18:56454092	intergenic	MALT1,ZNF532	MALT1	G/A	10.8576	0.5113	3.09E-06	2.30E-06	3%
rs57265082	chr18:56455080	intergenic	MALT1,ZNF532		G/T	10.9881	0.4435	6.49E-08	1.00E-07	6%
rs8096360	chr18:56456191	intergenic	MALT1,ZNF532		G/C	6.3083	0.4061	5.76E-06	6.90E-06	7%
rs6031719	chr20:43303010	intergenic	LINC01260,KCNK15-AS1	RP11-445H22.3; ADA; WISP2	T/C	3.8179	0.2973	6.60E-06	1.90E-06	43%

Table S3: GTEx v7 was mined to identify all cis eQTLs mapping to Chr18:56200000..56700000. All eQTLs mapping to SNPs with p-value $<10^{-5}$ for peanut allergy are listed in this table, and these specific SNPs only show an eQTL signature for *MALT1*.

SNV Annotation				LEAP Peanut Allergy in Avoidance Group				GTEx v7 Results		
SNV	Function	Gene (left,right)	Reference/ Alternate allele*	OR	SE	P-value	AAF*	P-Value	Tissue	Gene Symbol
rs4940418	UTR3	LOC101927322	T/C	8.8243	0.46	2.21E-06	5%	2.70E-07	Cells - Transformed fibroblasts	MALT1
								2.40E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs4940419	intronic	MALT1	G/C	9.0487	0.4788	4.21E-06	4%	1.80E-07	Cells - Transformed fibroblasts	MALT1
								1.60E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs79002421	intronic	MALT1	G/A	5.0116	0.3375	1.80E-06	12%	8.60E-09	Muscle - Skeletal	MALT1
rs77714205	intronic	MALT1	T/G	8.7742	0.4486	1.29E-06	5%	1.60E-07	Cells - Transformed fibroblasts	MALT1
								1.60E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs140235750	intronic	MALT1	T/C	9.4843	0.4425	3.70E-07	5%	1.60E-07	Cells - Transformed fibroblasts	MALT1
								1.60E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs77767290	intronic	MALT1	G/A	12.1475	0.4747	1.44E-07	4%	1.80E-07	Cells - Transformed fibroblasts	MALT1
								1.60E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs143282859	intronic	MALT1	A/G	5.6584	0.391	9.31E-06	9%	1.90E-06	Skin - Sun Exposed (Lower leg)	MALT1
rs75022589	intronic	MALT1	T/C	10.0751	0.4695	8.66E-07	4%	1.80E-07	Cells - Transformed fibroblasts	MALT1
								1.70E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs77267911	intronic	MALT1	A/G	9.4843	0.4425	3.70E-07	5%	1.60E-07	Cells - Transformed fibroblasts	MALT1
								3.00E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs4940749	intronic	MALT1	G/T	5.1256	0.3615	6.16E-06	11%	7.50E-08	Muscle - Skeletal	MALT1
rs4940750	intergenic	MALT1,ZNF532	G/A	6.5355	0.4159	6.37E-06	7%	1.40E-06	Cells - Transformed fibroblasts	MALT1
								1.50E-05	Muscle - Skeletal	MALT1
								1.10E-06	Skin - Sun Exposed (Lower leg)	MALT1
rs76653504	intergenic	MALT1,ZNF532	T/G	10.0982	0.4698	8.58E-07	4%	1.80E-07	Cells - Transformed fibroblasts	MALT1
								1.60E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs8093073	intergenic	MALT1,ZNF532	C/T	7.5888	0.4289	2.30E-06	6%	-	-	-
rs141101728	intergenic	MALT1,ZNF532	G/T	9.572	0.451	5.50E-07	6%	1.30E-06	Cells - Transformed fibroblasts	MALT1
								1.10E-05	Muscle - Skeletal	MALT1
								1.20E-06	Skin - Sun Exposed (Lower leg)	MALT1

SNV	SNV Annotation			LEAP Peanut Allergy in Avoidance Group				GTEx v7 Results		
	Function	Gene (left,right)	Reference/ Alternate allele*	OR	SE	P-value	AAF*	P-Value	Tissue	Gene Symbol
rs144838979	intergenic	MALT1,ZNF532	G/A	7.8716	0.4573	6.42E-06	5%	1.10E-06	Cells - Transformed fibroblasts	MALT1
								1.10E-05	Muscle - Skeletal	MALT1
								1.30E-06	Skin - Sun Exposed (Lower leg)	MALT1
rs8095265	intergenic	MALT1,ZNF532	T/C	5.1686	0.3337	8.56E-07	13%	3.30E-08	Muscle - Skeletal	MALT1
rs76767530	intergenic	MALT1,ZNF532	G/A	9.4119	0.4935	5.54E-06	5%	6.00E-06	Cells - Transformed fibroblasts	MALT1
								5.30E-06	Muscle - Skeletal	MALT1
								2.50E-05	Skin - Sun Exposed (Lower leg)	MALT1
rs116985291	intergenic	MALT1,ZNF532	C/T	10.6175	0.4636	3.48E-07	4%	2.50E-07	Cells - Transformed fibroblasts	MALT1
rs955405	intergenic	MALT1,ZNF532	C/T	9.2614	0.4663	1.81E-06	4%	8.40E-07	Cells - Transformed fibroblasts	MALT1
rs59853704	intergenic	MALT1,ZNF532	A/C	6.5751	0.4173	6.40E-06	6%	3.30E-06	Cells - Transformed fibroblasts	MALT1
rs4940751	intergenic	MALT1,ZNF532	C/T	7.4946	0.446	6.29E-06	5%	7.30E-06	Cells - Transformed fibroblasts	MALT1
rs73959507	intergenic	MALT1,ZNF532	G/A	10.4314	0.4767	8.69E-07	4%	3.60E-06	Cells - Transformed fibroblasts	MALT1
rs4940752	intergenic	MALT1,ZNF532	G/A	10.8576	0.5113	3.09E-06	3%	1.30E-05	Cells - Transformed fibroblasts	MALT1
rs57265082	intergenic	MALT1,ZNF532	G/T	10.9881	0.4435	6.49E-08	6%	-	-	-
rs8096360	intergenic	MALT1,ZNF532	G/C	6.3083	0.4061	5.76E-06	7%	-	-	-

Table S4: Landscape of variation identified through WGS in N=542 LEAP participants (chr18:56338653-56418303).

	N	MAF < 1%*	MAF 1-5%	MAF >5%
Total variants	842	591	76	175
Known	678	428	76	174
Novel	164	163	0	1
3' UTR	27	17	2	6
3' UTR - known	20	12	2	6
3' UTR - novel	5	5	0	0
5' UTR & Regulatory region	9	7	1	1
5' UTR & Regulatory region - known	7	5	1	1
5' UTR & Regulatory region - novel	2	2	0	0
Downstream gene variant	12	8	0	4
Downstream gene variant - known	10	6	0	4
Downstream gene variant - novel	2	2	0	0
Intronic	704	494	63	147
Intronic - known	566	357	63	146
Intronic - novel	138	137	0	1
Intronic & Regulatory region	82	56	9	17
Intronic & Regulatory region - known	67	41	9	17
Intronic & Regulatory region - novel	15	15	0	0
Missense	5	5	0	0
Missense - known	4	4	0	0
Missense - novel	1	1	0	0
Missense & Regulatory region	1	1	0	0
Missense & Regulatory region - known	0	0	0	0
Missense & Regulatory region - novel	1	1	0	0
Missense & Splice	1	0	1	0
Missense & Splice - known	1	0	1	0
Missense & Splice - novel	0	0	0	0
Intron & Splice	1	1	0	0
Intron & Splice - known	1	1	0	0
Intron & Splice - novel	0	0	0	0
Synonymous	2	2	0	0
Synonymous - known	2	2	0	0
Synonymous - novel	0	0	0	0

*Removed 84 variants monophoric in N=542, 32 known and 52 novel

Table S5: SKAT results for association between rare coding variants in MALT1 with peanut allergy in the peanut avoidance group.

	N Variants in test	Average MAF	Minimum MAF	Maximum MAF	P-value Liu
Exonic SNVs in MALT1 including peak variant rs57265082	8	0.0125	0.0018	0.0531	1.89E-10
Exonic SNVs in MALT1 without peak	7	0.0068	0.0018	0.0364	0.0829
Damaging exonic variants in MALT1 including peak	6	0.0161	0.0018	0.0531	1.89E-10
Damaging exonic variants in MALT1 without peak	5	0.0087	0.0018	0.0364	0.0828

Table S6: Clinical characteristics by MALT1 carrier status in the peanut avoidance group (Note: N=273 with genotype).

	MALT1 Carrier	MALT1 Non-carrier	P-value
N	29	244	-
N with Peanut Allergy (%)	17 (59)	31 (13)	4.034E-09
N with SPT wheal size at 60 months >0 (%)	20 (69)	56 (23)	5.51E-07
N with Peanut-Specific IgE at 60 months >0.1 (%)	21 (72)	96 (39)	0.0023
Log 10 Peanut-Specific IgE at 60 months quantitative, mean (SD)	0.13 (0.27)	-1.01 (0.06)	0.0003
Log 10 Peanut-Specific IgG4 at 30 months quantitative, mean (SD)	2.38 (-0.11)	2.08 (-0.03)	0.0129
N with Peanut-Specific IgG4 at 60 months >70 (%)	24 (83)	120 (49)	0.0024
Log 10 Peanut-Specific IgG4 at 60 months quantitative, mean (SD)	2.69 (0.13)	2.23 (0.03)	0.0019
Log 10 Peanut-Specific IgG4:IgE ratio at 60 months quantitative, mean (SD)	2.18 (0.20)	2.86 (0.06)	0.0032
Log 10 Eosinophilia at 60 months, mean (SD)	-0.39 (0.07)	-0.41 (0.02)	0.7490
SCORAD at 60 months, mean (SD)	9.20 (2.46)	7.72 (0.73)	0.5660

Table S7: Tests for association for all SNVs with suggestive evidence for peanut allergy in the peanut avoidance group to evaluate if [A] these associations with peanut allergy are independent of baseline risk factors, and [B] these SNVs are directly associated with baseline risk factors themselves. Tests for association for all SNVs with suggestive evidence for peanut allergy in the peanut avoidance group to evaluate if [A] these associations with peanut allergy are independent of baseline risk factors, and [B] these SNVs are directly associated with baseline risk factors themselves.

	[A] Peanut Allergy Independent of baseline risk factors			[B] Not a marker of baseline Egg Allergy or Eczema			
	Peanut Allergy in Avoidance Group	Peanut Allergy in Avoidance Group Adjusted for Eczema at Baseline	Peanut Allergy in Avoidance Group adjusted for Egg Allergy at Baseline	Egg Allergy at baseline in Avoidance Group	Egg Allergy at baseline in Consumption Group	Eczema at baseline in Avoidance Group	Eczema at baseline in Consumption Group
rs57265082	6.49E-08	1.76E-07	3.27E-08	0.32407	0.28032	0.16255	0.01020
rs77767290	1.44E-07	3.73E-07	9.42E-08	0.69109	0.94061	0.23274	0.03065
rs116985291	3.48E-07	9.32E-07	2.50E-07	0.74142	0.86799	0.21408	0.07263
rs140235750	3.70E-07	1.07E-06	3.36E-07	0.96469	0.96034	0.17942	0.04722
rs77267911	3.70E-07	1.07E-06	3.36E-07	0.96469	0.84039	0.17942	0.05623

	[A] Peanut Allergy Independent of baseline risk factors			[B] Not a marker of baseline Egg Allergy or Eczema			
	Peanut Allergy in Avoidance Group	Peanut Allergy in Avoidance Group Adjusted for Eczema at Baseline	Peanut Allergy in Avoidance Group adjusted for Egg Allergy at Baseline	Egg Allergy at baseline in Avoidance Group	Egg Allergy at baseline in Consumption Group	Eczema at baseline in Avoidance Group	Eczema at baseline in Consumption Group
rs141101728	5.50E-07	9.61E-07	3.07E-07	0.35267	0.37058	0.36815	0.21584
rs8095265	8.56E-07	1.88E-06	4.59E-07	0.64571	0.78020	0.34276	0.31390
rs76653504	8.58E-07	2.17E-06	5.55E-07	0.67175	0.97197	0.22000	0.04653
rs75022589	8.66E-07	2.17E-06	7.50E-07	0.98817	0.97197	0.22297	0.04653
rs73959507	8.69E-07	2.19E-06	4.52E-07	0.47032	0.52608	0.23063	0.32688
rs2499652	1.21E-06	1.64E-06	1.54E-06	0.73496	0.53902	0.42142	0.20933
rs77714205	1.29E-06	3.43E-06	1.04E-06	0.90426	0.96034	0.18807	0.04722
rs79002421	1.80E-06	3.56E-06	1.04E-06	0.83289	0.69658	0.31008	0.02736
rs955405	1.81E-06	4.75E-06	1.09E-06	0.57923	0.74492	0.20457	0.69949
rs4940418	2.21E-06	3.88E-06	1.54E-06	0.78264	0.84064	0.49458	0.05715
rs11243375	2.21E-06	5.11E-06	2.47E-06	0.69159	0.65427	0.99660	0.80626
rs8093073	2.30E-06	3.56E-06	2.49E-06	0.96589	0.49776	0.45990	0.14067
rs80151145	2.91E-06	2.79E-06	2.85E-06	0.91413	0.99155	0.66963	0.21917
rs4940752	3.09E-06	6.35E-06	1.52E-06	0.40129	0.11880	0.32990	0.63587
rs73156540	3.56E-06	5.26E-06	1.93E-06	0.36000	0.09860	0.41417	0.29150
rs4940419	4.21E-06	8.87E-06	3.09E-06	0.90029	0.97197	0.25824	0.04653
rs73156541	4.26E-06	5.82E-06	3.28E-06	0.68992	0.05191	0.53889	0.25213
rs74326323	5.40E-06	1.17E-05	2.52E-06	0.54248	0.72785	0.15914	0.99692
rs76767530	5.54E-06	6.31E-06	7.54E-06	0.60513	0.80676	0.72272	0.51885
rs8096360	5.76E-06	1.07E-05	1.59E-06	0.14738	0.21190	0.22959	0.18942
rs6686904	6.02E-06	6.56E-06	6.75E-06	0.92912	0.61523	0.95389	0.16580
rs4940749	6.16E-06	1.03E-05	2.29E-06	0.35523	0.51141	0.43560	0.10013
rs4940751	6.29E-06	1.48E-05	3.93E-06	0.74184	0.35705	0.17322	0.65691
rs4940750	6.37E-06	1.15E-05	5.41E-06	0.90044	0.79101	0.27762	0.04061
rs59853704	6.40E-06	8.34E-06	4.94E-06	0.91194	0.96825	0.50245	0.72532
rs144838979	6.42E-06	1.06E-05	4.15E-06	0.45705	0.14416	0.32917	0.07405
rs6031719	6.60E-06	8.69E-06	5.77E-06	0.72423	0.61296	0.54460	0.33014
rs145726195	7.22E-06	1.25E-05	2.59E-05	0.10477	0.88954	0.29312	0.18718
rs4240808	8.38E-06	1.24E-05	5.36E-06	0.97484	0.02081	0.23468	0.48312
rs72777284	8.55E-06	7.88E-06	1.52E-05	0.50950	0.44998	0.72426	0.99758
rs143282859	9.31E-06	1.56E-05	4.40E-06	0.56400	0.51060	0.37024	0.10603
rs9391248	9.53E-06	1.17E-05	1.29E-05	0.31194	0.50803	0.92843	0.40557
rs2526073	9.79E-06	1.80E-05	7.37E-06	0.89649	0.57255	0.08201	0.01493
rs12643843	9.82E-06	1.18E-05	5.59E-06	0.60620	0.79375	0.80134	0.38702

57 **Table S8**. Association of top variants from Asai et. al in LEAP peanut avoidance group (N=275) and LEAP all participants (N=542).

							LEAP Avoidance Group[a]			LEAP All		
	SNP	**Chr**	**MAF**	**OR**	**P value[b]**	**Gene/nearest gene**	**MAF**	**OR**	**P - value**	**MAF**	**OR**	**P - value**
Asai et al. March 2018 JACI genome-wide	rs115218289	2	0.02	0.18	$1.80x10^{-8}$	(298 kb)DLX2 | (26 kb)ITGA6	0.02	0.530	5.75E-01	0.02	0.53	5.73E-01
	rs72827854	17	0.09	2.16	$2.60x10^{-7}$	SKAP1	0.09	0.490	1.60E-01	0.09	0.47	1.36E-01
	rs144897250	11	0.02	6.2	$2.90x10^{-7}$	(5 kb)MMP12 | (63 kb)MMP13						
	rs7475217	10	0.38	1.64	$3.58x10^{-7}$	CTNNA3	0.26	0.710	3.31E-01	0.27	0.77	4.48E-01
	rs744597	4	0.40	0.61	$3.98x10^{-7}$	ARHGAP24	0.35	0.810	4.14E-01	0.35	0.83	4.54E-01
	rs523865	20	0.23	0.57	$4.42x10^{-7}$	ANGPT4	0.25	1.290	3.88E-01	0.24	1.29	3.88E-01
	rs7936434	11	0.49	1.58	$5.17x10^{-7}$	(30 kb)C11orf30 |(43 kb)LOC101928813	0.33	2.110	6.14E-02	0.34	1.85	1.09E-01
	rs78048444	7	0.02	0.22	$5.44x10^{-7}$	(65 kb)CHCHD3 |(106 kb)EXOC4	0.02	2.890	1.68E-01	0.03	2.61	1.84E-01
	rs56151068	17	0.10	2.06	$9.58x10^{-7}$	SKAP1; LOC101927148	0.09	0.480	1.58E-01	0.09	0.47	1.34E-01
	rs139462954	17	0.09	2.06	$1.23x10^{-6}$	LOC101927166						
Asai et al. April 2018 JACI on HLA	rs1063347	6			$3.67x10^{-23}$	*HLA-DQB1*						
	rs3134975	6			$7.01x10^{-23}$	(18kb)*HLA-DQB1*|(57kb)*HLA-DQA2*						
	rs1049213	6			$2.30x10^{-21}$	*HLA-DQB1*						
	rs9275596	6			$1.15x10^{-21}$	(47kb)*HLA-DQB1*|(28kb)*HLA-DQA2*	0.34	1.550	8.37E-02	0.32	1.6	3.95E-02
	rs3134976	6			$6.86x10^{-21}$	(18kb)*HLA-DQB1*|(57kb)*HLA-DQA2*	0.21	1.570	1.23E-01	0.19	1.64	5.79E-02
	rs2858305	6			$2.25x10^{-18}$	(36kb)*HLA-DQB1*|(39kb)*HLA-DQA2*	0.38	1.550	7.75E-02	0.35	1.65	2.80E-02
	rs2858309	6			$2.26x10^{-18}$	(34kb)*HLA-DQB1*|(40kb)*HLA-DQA2*	0.38	1.550	7.75E-02	0.35	1.64	2.92E-02
	rs2856717	6			$2.43x10^{-18}$	(36kb)*HLA-DQB1*|(39kb)*HLA-DQA2*	0.38	1.550	7.75E-02	0.35	1.65	2.80E-02
	rs2858320	6			$2.48x10^{-18}$	(27kb)*HLA-DQB1*|(48kb)*HLA-DQA2*						
	rs7192	6			$1.94x10^{-18}$	*HLA-DRA*	0.36	1.550	5.23E-02	0.35	1.52	5.20E-02
	rs9275227	6			$1.43x10^{-15}$	(26kb)*HLA-DQB1*|(49kb)*HLA-DQA2*						
	rs2858332	6			$2.63x10^{-13}$	(47kb)*HLA-DQB1*|(28kb)*HLA-DQA2*	0.49	1.660	4.40E-02	0.46	1.78	1.19E-02
	rs3129890	6			$9.73x10^{-3}$	(1kb)*HLA-DRA*|(71kb)*HLA-DRB5*	0.25	1.270	3.26E-01	0.26	1.19	4.60E-01
	rs154975	6			$8.21x10^{-2}$	(29kb)*LOC100294145*|(2kb)*HLA-DMB*	0.36	1.290	2.81E-01	0.37	1.17	4.68E-01

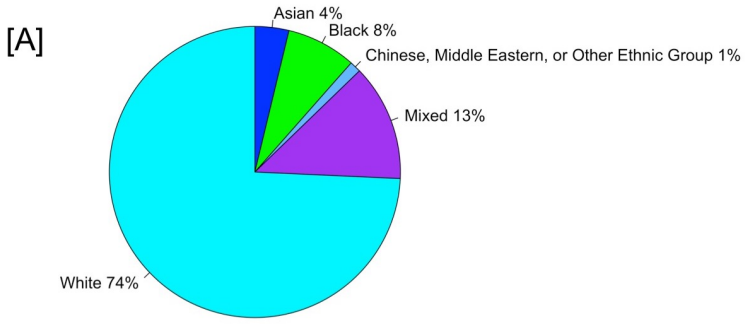
[a] WGS data for variants not in the HLA region and Omni 2.5 chip data for variants in the HLA region, as the HLA region was not well called by traditional sequence variant calling algorithms [b] Meta-analysis p-value reported for HLA variants [c] treatment included as covariate in model to adjustment for treatment.

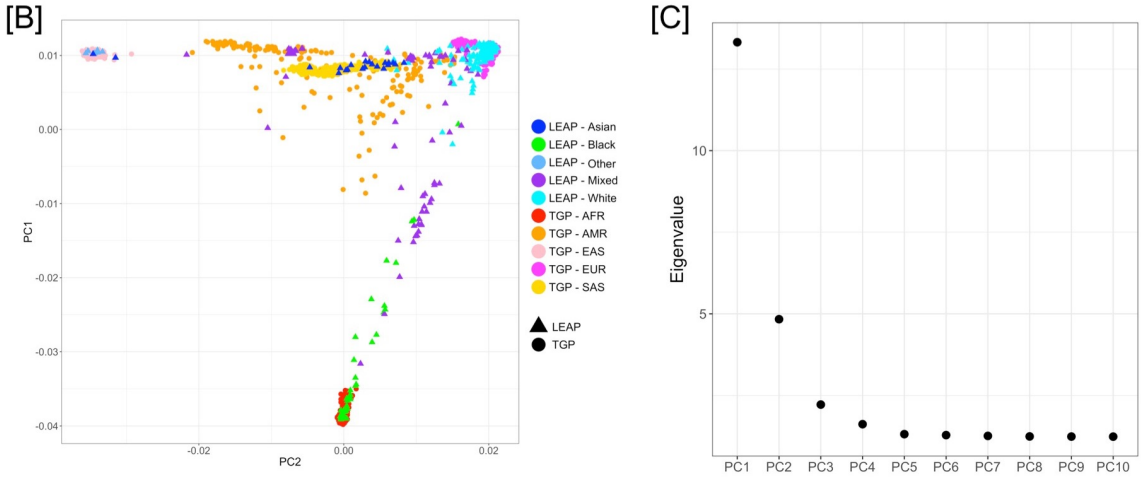
Table S9. Comparison of per protocol and intention to treat (ITT) individuals for association of MALT1 top variant rs57265082 with peanut allergy in the peanut avoidance group. * WGS genotype data, ** WGS and MGB Pleiades assay genotype data.

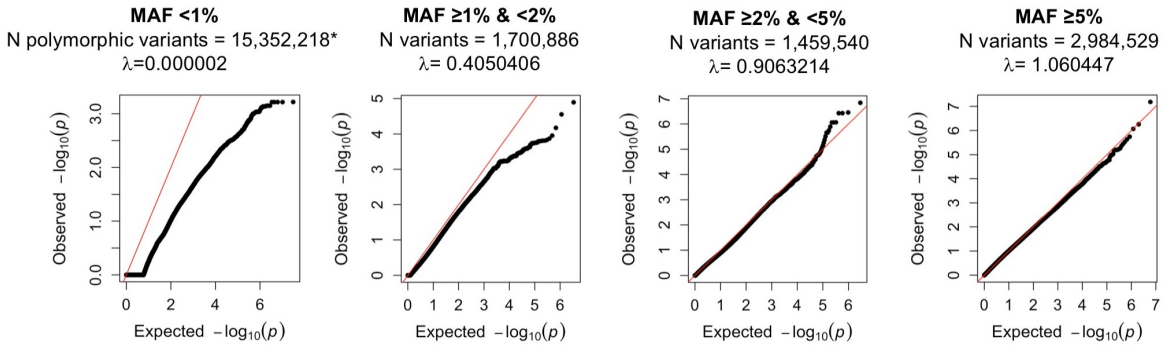
	Peanut Avoidance			
	N	**N with peanut allergy**	**OR**	**P - value**
Original per protocol group adjusted for PCs 1-5, age, and sex	273*	48	10.99	6.49E-08
Original per protocol group adjusted for age and sex only	273*	48	9.65	9.40E-08
All per protocol individuals adjusted for age and sex only	280**	51	9.64	9.04E-08
All ITT individuals adjusted for age and sex only	299**	54	8.15	1.28E-07

Table S10. rs57265082 genotypes by peanut allergy status and intervention group

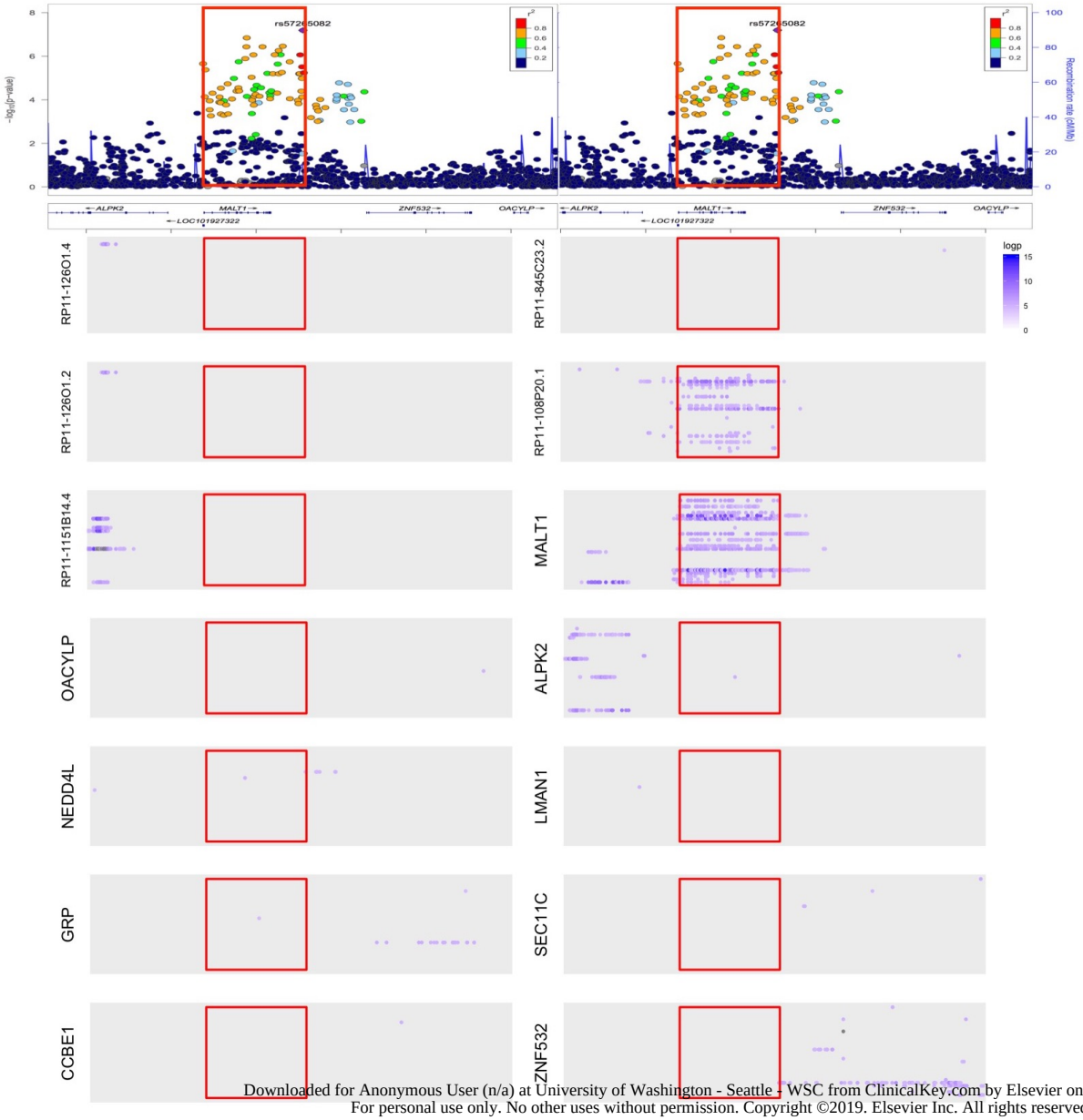
		Peanut Avoidance N=275		**Peanut Consumption N=267**	
		Allergic N=48	**Non-allergic N=227**	**Allergic N=1**	**Non-allergic N=266**
Imputed Data	GG	31	213	1	233
	GT	17	14	0	30
	TT	0	0	0	3
	missing	0	0	0	0
Genotyped Data	GG	31	213	1	233
	GT	17	12	0	25
	TT	0	0	0	3
	missing	0	2	0	5

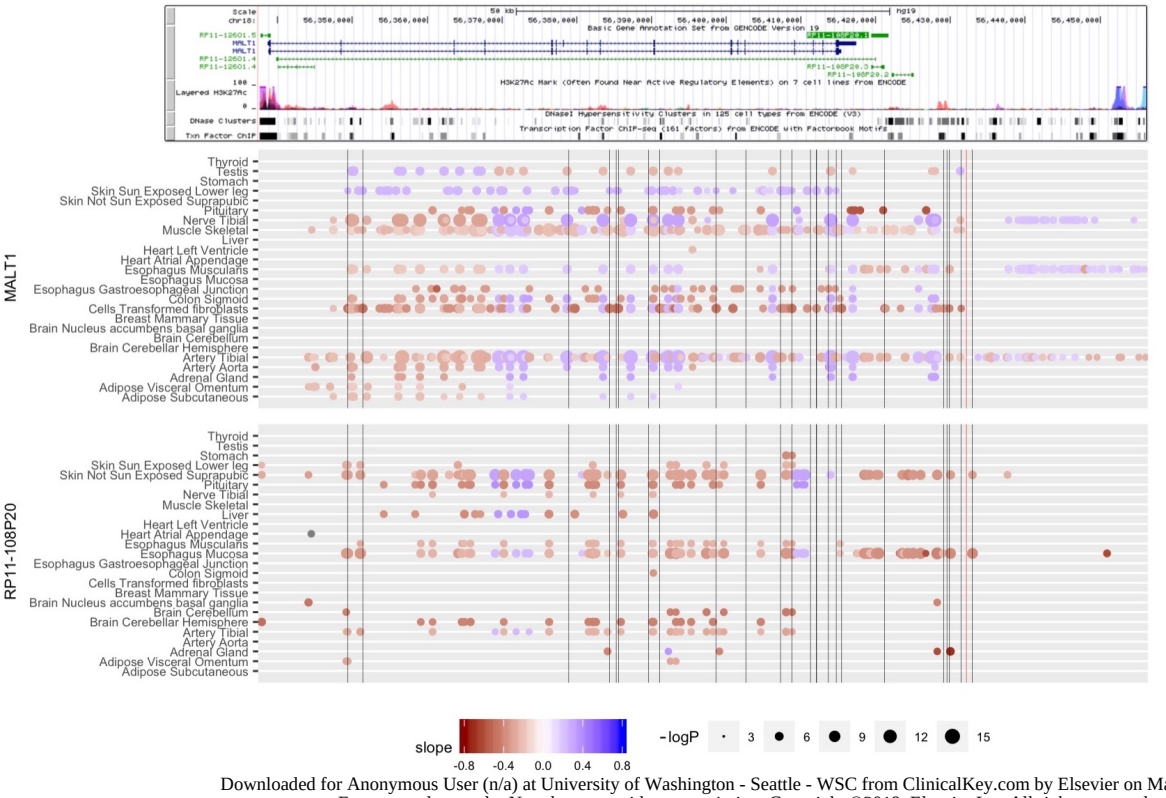


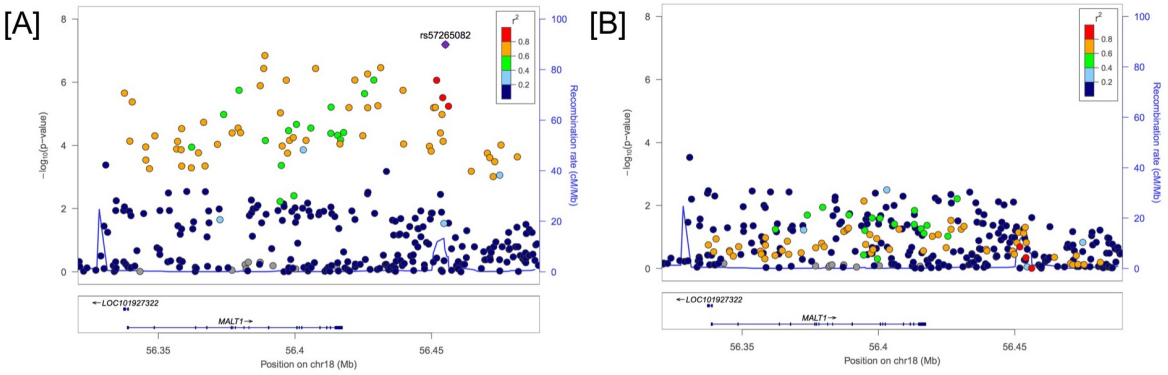


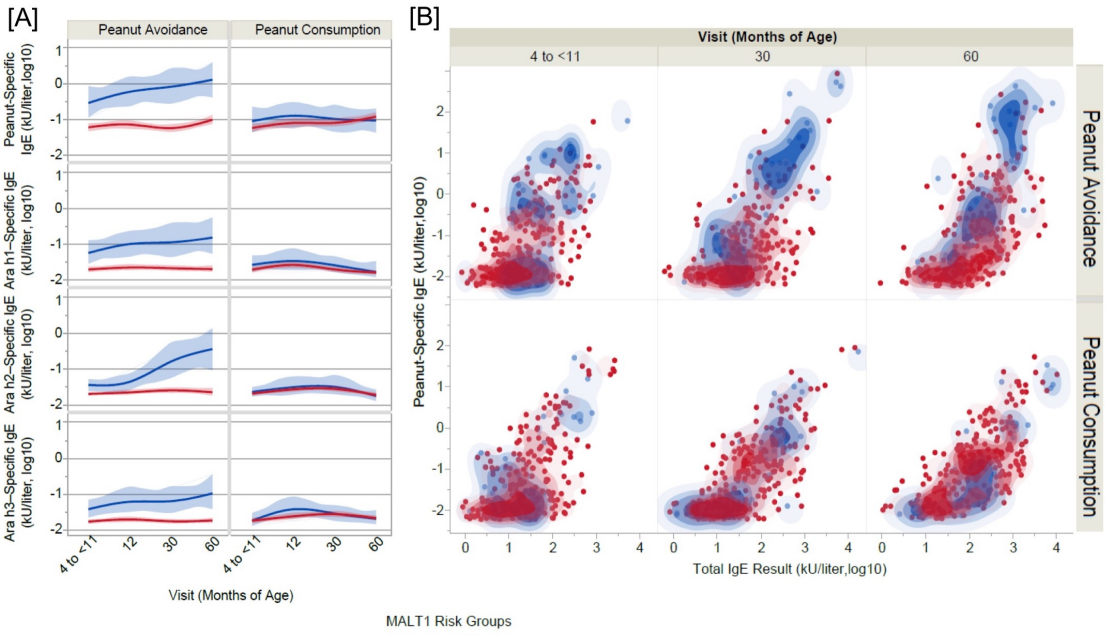


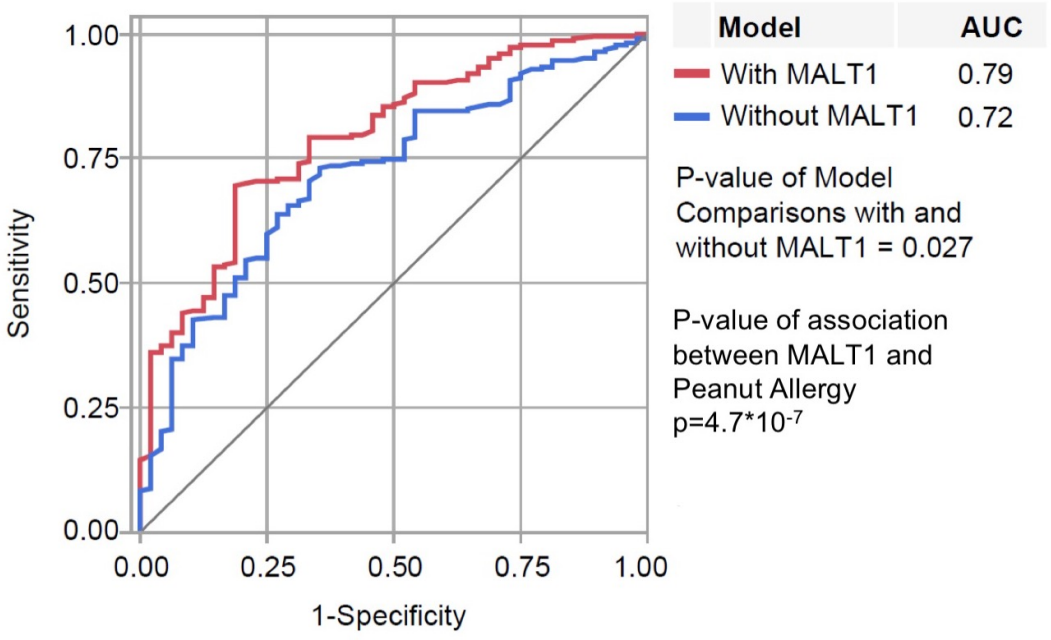
*Excluding 8,907,642 monomorphic variants

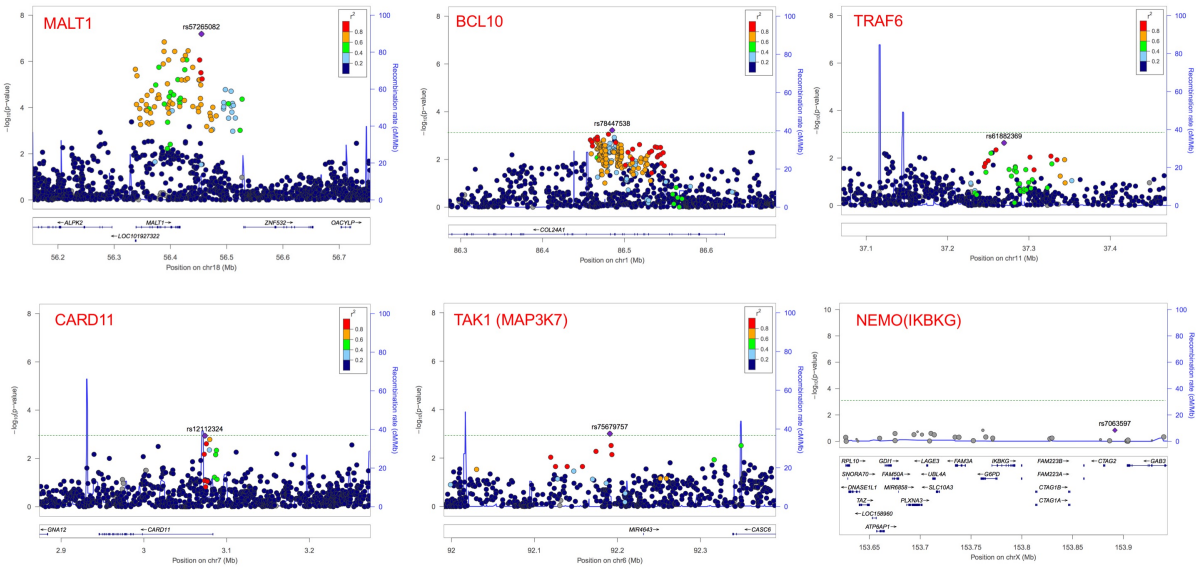












Supplementary Notes

The *MALT1* locus and Peanut Avoidance in the Risk for Peanut Allergy

Alexandra Winters, PhD,[1] Henry T Bahnson, MPH,[2,3] Ingo Ruczinski, PhD,[4] Meher P Boorgula, MS,[5] Claire Malley, MS,[1] Ali R Keramati, MD,[1] Sameer Chavan, MS,[5] David Larson, PhD,[3] Karen Cerosaletti, PhD,[2] Peter H Sayre, MD, PhD,[3,6] Marshall Plaut, MD,[7] George Du Toit, MB, BCh, FRCPCH,[8] Gideon Lack, MB, BCh, FRCPCH,[8] Kathleen C Barnes, PhD,[5] Gerald T Nepom, MD, PhD,[3] Rasika A Mathias, ScD,[1,*] for the Immune Tolerance Network LEAP Study Team.

* Address correspondence to: Rasika A Mathias, ScD. rmathias@jhmi.edu. 5501 Hopkins Bayview Circle, Baltimore, MD 21224. 410-550-2487.

1. Department of Medicine, Johns Hopkins University, Baltimore, MD, USA.
2. Benaroya Research Institute, Seattle, WA, USA.
3. Immune Tolerance Network, San Francisco, CA, USA.
4. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.
5. Department of Medicine, University of Colorado, Anschutz, CO, USA.
6. Department of Medicine, University of California, San Francisco, CA, USA.
7. Allergy, Asthma and Airway Biology Branch, DAIT, NIAID, Bethesda, MD, USA.
8. Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, and Guy's and St. Thomas' NHS Foundation Trust, London, United Kingdom.

Material and Methods

1. Genomic Research on the Learning Early about Peanut Allergy (LEAP) study

The LEAP study, described in detail previously (1), was a randomized trial examining the effect of early peanut exposure on infants at high risk for peanut allergy. The study (1) found that early exposure to peanuts greatly decreased the likelihood of peanut allergy among both skin prick test negative and skin prick test positive participants. Among the skin prick test negative participants in the intention-to-treat population, 13.7% of the peanut avoidance group was allergic to peanuts at 60 months of age compared to 1.9% of the peanut consumption group. This finding was mirrored in the skin prick test positive intention-to-treat population, where 35.3% of the peanut avoidance group was allergic to peanuts at 60 months of age compared to 10.6% of the peanut consumption group.

All participants that completed the LEAP study, followed the protocol to avoid or consume peanuts according to their randomization assignment, and consented to genetic studies were included in our sample for this work (N=556). The final sample set, following quality control as defined in **Supplementary Section 3-4** below was N=542, as described in **Supplementary Table S1**. Cases were defined as participants with a positive oral food challenge (OFC) to peanuts at 60 months of age. The OFC was a double-blind placebo-controlled challenge consisting of a single 5g dose of peanut protein for participants who were not suspected to be allergic on the basis of skin prick test values and other clinical characteristics and a total of 9.4 g of peanut protein in increasing increments for the participants in LEAP that were suspected to be allergic. In participants where the OFC was unavailable or the results were inconclusive, allergy was determined using clinical factors including history of peanut specific exposure, peanut-specific IgE levels, and skin prick test values(1).

2. Informed consent

The LEAP (1, 2) study was approved by the institutional review board at the National Research Ethics Service Committee London–Fulham. In addition, informed written consent was obtained from the parent or guardian of all participants.

3. Whole genome sequencing

Whole genome sequencing to a depth of 30X was performed using the HiSeq X System, using HiSeq X HD SBS Kit reagents. Illumina HiSeq Control Software (HCS), and Real-Time Analysis (RTA) was used with the HiSeq X sequencers for real-time image analysis and base calling. Assembly of each individual genome was performed using the Isaac aligner (3). The Starling small variant caller (formerly called the Isaac Variant Caller (3)) was used to call both SNVs and small indels to yield a genome variant file (gVCF) that includes variants along with quality metrics. As described previously (4, 5), multi-sample VCF files for each chromosome were generated from single-sample VCF files (removing indels) provided by Illumina. As shown in previous work(5), single-sample and multi-sample calling have a similar accuracy for these Illumina whole genome sequence calls. As is standard practice(4), we filtered variants with GQX < 30 and DP < 7 and regions of segmental duplication. Single-nucleotide variant (SNV) annotation was performed using the ANNOVAR package (6).

Seven of our 542 participants did not have a genotype available for rs57265082. In order to include these individuals in our follow-up analyses, we used the Michigan Imputation server (7) with the 1000 Genomes Project Phase 3 database as the reference panel and Eagle v.2.3 for phasing to impute missing genotypes for these individuals. Quality of the imputation was high for this region, with an overall imputation r^2=0.98 for rs57265082 (**Supplementary Table S10**).

4. Sample based quality control

All samples with sequencing were also genotyped on the Illumina Omni 2.5 Array as described previously [5,6]. On the basis of concordance between the sequence and array genotype calls, we removed 11 of the 556 samples from the dataset for poor quality DNA. We removed 1 additional sample for discrepancy between self-reported and genetic sex. Identity by descent (IBD) was run on a set of linkage disequilibrium (LD) pruned SNVs using PLINK 1.9 (8); 2 samples were sequenced as identical (Z2>0.97) but not confirmed to be monozygotic twins and were both removed from the analysis subset. The final sample size passing quality control used for analysis was N=542.

5. Ancestry de-convolution of the LEAP study participants

We implemented protocols similar to those established for the 1000 Genomes Project (TGP) reference populations in previous work (4) including the same set of 2179 TGP individuals and a set of 218,340 LD pruned SNVs. We used the smartpca program, a part of the EIGENSOFT package (9), to perform Principal Components Analysis (PCA) (**Supplementary Figure S1B**). The final association models included the first five principal components for ancestry based on the scree plot shown in **Supplementary Figure S1C**.

6. Tests for association

For the primary outcome, we tested the hypothesis that peanut allergy among the peanut avoidance group participants was associated with SNV genotype under an additive model including 5 PCs, age in months at entrance into the LEAP study, and sex. Similar analyses were done to test for the association between SNV genotype and allergy adjusted for baseline risk factors, with baseline risk factors themselves, sensitization at 60 months, and for association with the filaggrin variant R501X (**Supplementary Table S7**). Analyses were run using PLINK 1.9 and PSEQ, a part of the PLINK/SEQ package ((8), https://atgu.mgh.harvard.edu/plinkseq/) .

Only biallelic SNVs were evaluated for association. Several QC filters were investigated including overall minor allele frequency (MAF), overall variant missingness, differential variant missingness between cases and controls for the final outcome of peanut allergy at 60 months of age, Hardy-Weinberg Equilibrium (HWE) among all participants, HWE among cases (that is those participants who were allergic to peanuts at 60 months of age), and HWE among controls (those not allergic to peanut at 60 months). To address issues of small sample size we performed a detailed review of results by MAF. QQ plots for all SNVs analyzed are presented in **Supplementary Figure S2** by categories of MAF (less than 1, 1-2, 2-5, and greater than or equal to 5). We observed based on the QQ plots that our tests for association do not violate expected distributions for all SNVs with an MAF range greater than or equal to 2%. Therefore, the final variant list included only variants with overall MAF greater than or equal to 2%, less than 5% overall missingness, and overall HWE greater than or equal to 10^{-6}, which yields a total of 4,444,069 variants. Additionally, for all variants with p<10E-05, we derived permutation p-values, based on 10 million permutations, and added these permutation p-values to **Supplementary Table S2**. We jointly permuted the outcome and the covariates, thereby also preserving the SNV linkage disequilibrium. These empirical permuted p-values generally were in agreement with the p-values derived from the z-statistic in logistic regression model. Both types of p-values reflected the fact that obtaining a test statistic at least as extreme as in the observed data by random chance is very unlikely.

6b. Exploring association results

After observing the association results, we explored the question of whether any of the variants with suggestive evidence for association with peanut allergy in the peanut avoidance group (**Supplementary Table S2**) were associated with either baseline risk factors or egg allergy

and eczema at baseline. We found that adjustment for either egg allergy at baseline or eczema at baseline did not meaningfully change the association between our top variants and peanut allergy in the peanut avoidance group (**Supplementary Table S7**). Additionally, we did not see a significant association between our top variants and egg allergy or eczema in either the peanut avoidance or consumption groups, suggesting that the association we see between *MALT1* and peanut allergy is not a marker for these baseline traits.

7. Analysis of quantitative correlates of peanut allergy with *MALT1* carrier status

To test the association of *MALT1* carrier status at rs57265082 and component spreading of the IgE response to the specific peanut protein components Ara h1, Ara h2, and Ara h3 at 60 months, we used the Cochran-Mantel-Haenszel chi-square test. This test looks at the association between component spreading and *MALT1* across (or controlling for) the treatment assignment using both groups. Using a multivariate ordinal logistic regression model, there is a significant interaction effect between *MALT1* carrier status and the randomized intervention when regressed on component spreading at 60 months to Ara h1, Ara h2, and Ara h3 (p-value = 0.0053). Ara h status was imputed to 0 for all participants with peanut-specific IgE <0.1.

To evaluate the predictive value of *MALT1* carrier status for peanut allergy in the LEAP avoidance Group (N=275), we used a multivariate logistic regression model including *MALT1* rs57265082 genotype, SCORAD, peanut SPT, and egg allergy at baseline compared to the model omitting *MALT1* as a predictor, and the AUC of each model was compared using a likelihood ratio chi-square test.

In all cases *MALT1* carrier status was modeled under the dominant model, comparing those with the GG genotype (Non-carriers) to those carrying at least one T allele (Carriers with genotype TG or TT). Because of the small number of TT genotypes, other models were not

appropriate (**Supplementary Table S10**). Additionally, for all analyses including the quantitative correlates of peanut allergy, the imputed version of rs57265082 was used to include 7 participants with genotype missing.

8. Inclusion of intention-to-treat (ITT) participants

We repeated the analysis of the association of peanut allergy with our top SNV, rs57265082, with the addition of the non per protocol ITT participants from the LEAP study. These non per protocol participants were not included in the original per protocol group, because they deviated from the protocol of their assigned group (peanut avoidance or peanut consumption). This was done to show that our results were not sensitive to their exclusion from the discovery data.

All ITT participants without WGS were genotyped for *MALT1* variant rs57265082 using the MGB Pleiades assay (ELITech Group). The assay was validated using 4 CEPH DNAs, 2 with GG genotype and 2 with GT genotype. As no CEPH DNAs with TT genotype were available, an additional validation was run with 88 healthy control individuals with genotypes as follows: 73 GG, 14 GT, and 1 TT. In addition, each genotyping plate contained 2 wells each of GT and TT DNA used as a reference for genotype calling. Finally, genotypes were confirmed by Sanger sequencing using the SimpleSeq dye termination kit (Eurofins), and sequences traces were manually read to confirm GT and TT genotypes.

A similar association was seen for peanut allergy with the *MALT1* variant rs57265082 in the total peanut avoidance group consisting of both the per protocol and ITT participants (N=299; $p=1.28x10^{-7}$) (**Supplementary Table S9**).

REFERENCES

1. Du Toit G, Roberts G, Sayre PH, Bahnson HT, Radulovic S, Santos AF, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. N Engl J Med. 2015;372(9):803-13.
2. Du Toit G, Sayre PH, Roberts G, Sever ML, Lawson K, Bahnson HT, et al. Effect of Avoidance on Peanut Allergy after Early Peanut Consumption. N Engl J Med. 2016;374(15):1435-43.
3. Raczy C, Petrovski R, Saunders CT, Chorny I, Kruglyak S, Margulies EH, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics. 2013;29(16):2041-3.
4. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, et al. A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. Nat Commun. 2016;7:12522.
5. Shringarpure SS, Mathias RA, Hernandez RD, O'Connor TD, Szpiech ZA, Torres R, et al. Using genotype array data to compare multi- and single-sample variant calls and improve variant call sets from deep coverage whole-genome sequencing data. Bioinformatics. 2017;33(8):1147-53.
6. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
7. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284-7.
8. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
9. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006;2(12):e190.
10. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348(6235):648-60.

Supplementary Figure S1: Panel A shows the proportion of LEAP participants (N=556) with WGS by self-reported race/ethnicity. **Panel B** shows the principal components analysis of the N=556 LEAP sequenced participants comparing the self-reported race/ethnicity to 1000 Genomes participants, and **panel C** is the scree plot of eigenvalues for the first ten principal components of ancestry for all N=542 LEAP participants used in analysis.

Supplementary Figure S2: Quantile-quantile plots of the genome-wide tests for association on all variants stratified by minor allele frequency (MAF). Tests for association were performed under the additive model including age, sex and the first 5 principal components to adjust for population stratification.

Supplementary Figure S3: GTEx v7 (10) was mined to identify all cis eQTLs mapping to Chr18:56200000..56700000. A total of 9 protein coding gene transcripts and 5 long non-coding RNAs were identified with at least one cis-eQTL within the region. The red box indicates the region of peak association to peanut allergy (the region with SNVs having $p<10^{-5}$). The region includes strong eQTL signatures for *MALT1* and the long intergenic non-coding RNA (lincRNA) RP11-108P20.1. A detailed overview of these eQTLs for *MALT1* and RP11-108P20.1 are in **Fig S4**.

Supplementary Figure S4: GTEx v7 (10) was mined to identify all cis eQTLs mapping to Chr18:56200000..56700000. There are two transcripts with strong eQTLs within the region of peak association Chr18:56337602..56456191: the protein coding *MALT1* and lincRNA RP11-108P20.1. The region shows eQTLs to both transcripts across a total of 25 tissue types within

GTEx. The specific set of SNVs with $p<10^{-5}$ for peanut allergy (marked with vertical black lines, red line = rs57265082) only show eQTL signatures to *MALT1* (see **Table S3**) in transformed fibroblasts, skeletal muscle, and sun exposed skin (lower leg), and not to RP11-108P20.1.

Supplementary Figure S5: Association in the *MALT1* region with peanut allergy at 60 months in N=275 LEAP participants in the peanut avoidance group. **Panel A** shows the locus zoom plot of the peak association on chromosome 18 in the *MALT1* gene region from Fig 1C, zoomed in to the region of LD with the top variant rs57265082. **Panel B** is the same region but with p-values for association conditioned on rs57265082, showing a loss of association in that region after conditioning on the top variant.

Supplementary Figure S6: Panel A is the distribution of peanut-specific IgE and the three major allergenic components of peanut (Ara h1, Ara h2, and Ara h3) by the randomized assignment and presence or absence of the *MALT1* risk allele; 95% confidence intervals are calculated using bootstrap sampling, independently at each assessment. **Panel B** shows the association between peanut-specific IgE, total IgE, and *MALT1* carrier status in a contour scatter plot. In a multivariate logistic regression model associating *MALT1* carrier status with peanut-specific and total IgE, only peanut-specific IgE is significantly associated with *MALT1* carrier status ($p = 2.03x10^{-5}$ for peanut-specific IgE and $p = 0.3661$ for total IgE).

Supplementary Figure S7: Predictive value of *MALT1* carrier status for peanut allergy in LEAP avoidance group (N=275). Using a multivariate logistic regression model, *MALT1* is a significant predictor of peanut allergy independent of SCORAD, peanut SPT, and egg allergy at

baseline ($p=4.7*10^{-7}$). This multivariate model was repeated after omitting *MALT1* as a predictor, and the AUC of each model was compared using a likelihood ratio chi-square test. The logistic regression model controlling for the baseline covariates in addition to *MALT1* was significantly more predictive of peanut allergy than the multivariate model without *MALT1* ($p=0.027$).

Supplementary Figure S8: Locus zoom plots of *MALT1* pathway genes (**green** dotted line is a Bonferroni threshold of $p \sim 2x10^{-3}$).