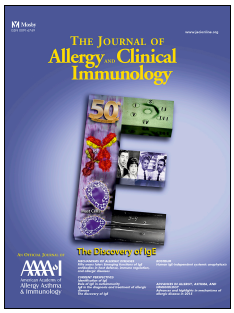
Accepted Manuscript



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

Yuka Asai, MD, MSc, Aida Eslami, PhD, C. Dorien van Ginkel, BSc, Loubna Akhabir, PhD, Ming Wan, BSc, George Ellis, Moshe Ben-Shoshan, MD, MSc, David Martino, PhD, Manuel A. Ferreira, PhD, Katrina Allen, MD PhD, Bruce Mazer, MD, Hans de Groot, MD, PhD, Nicolette W. de Jong, PhD, Roy N. Gerth van Wijk, MD, PhD, Anthony E.J. Dubois, MD PhD, Rick Chin, MSc, Steven Cheuk, MD, Joshua Hoffman, PhD, Eric Jorgensen, PhD, John S. Witte, PhD, Ronald B. Melles, MD, Xiumei Hong, MD, PhD, Xiaobin Wang, MD, MPH, ScD, Jennie Hui, PhD, Arthur W. (Bill) Musk, FRACP, Michael Hunter, PhD, Alan L. James, FRACP, Gerard H. Koppelman, MD, PhD, Andrew J. Sandford, PhD, Ann E. Clarke, MD, MSc, Denise Daley, PhD

PII: S0091-6749(17)31574-9

DOI: 10.1016/j.jaci.2017.09.015

Reference: YMAI 13040

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

Received Date: 18 September 2016

Revised Date: 5 September 2017

Accepted Date: 19 September 2017

Please cite this article as: Asai Y, Eslami A, van Ginkel CD, Akhabir L, Wan M, Ellis G, Ben-Shoshan M, Martino D, Ferreira MA, Allen K, Mazer B, de Groot H, de Jong NW, Gerth van Wijk RN, Dubois AEJ, Chin R, Cheuk S, Hoffman J, Jorgensen E, Witte JS, Melles RB, Hong X, Wang X, Hui J, Musk AW(B), Hunter M, James AL, Koppelman GH, Sandford AJ, Clarke AE, Daley D, 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, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2017.09.015.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please

note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Title: 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

Authors:

Yuka Asai MD, MSc, Division of Dermatology, Department of Medicine, Queen's University, Kingston, ON, Canada and Division of Experimental Medicine, Department of Medicine, McGill University, Montreal, QC, Canada.*
Aida Eslami PhD, Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada.*
C. Dorien van Ginkel BSc, University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology and University of Groningen, University Medical Center Groningen, GRIAC Research Institute, Groningen, the Netherlands.
Loubna Akhabir PhD, Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada.
Ming Wan BSc, Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada.
George Ellis, Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada.
Moshe Ben-Shoshan MD, MSc, Division of Allergy and Immunology, Department of Pediatrics, Montreal Children's Hospital, Research Institute of the McGill University Health Centre, Montreal, QC, Canada.
David Martino, PhD, Murdoch Children's Research Institute, Royal Children's Hospital, University of Melbourne, Australia.
Manuel A. Ferreira PhD, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
Katrina Allen, MD PhD, Murdoch Children's Research Institute, Royal Children's Hospital, University of Melbourne, Australia.
Bruce Mazer MD, Division of Allergy and Immunology, Department of Pediatrics, Montreal Children's Hospital, Research Institute of the McGill University Health Centre, Montreal, QC, Canada.
Hans de Groot MD, PhD, Department of Pediatric Allergology, Diaconessenhuis Voorburg, Reinier de Graaf Gasthuis, Delft, the Netherlands.
Nicolette W. de Jong PhD, Erasmus MC, Department of Allergology, Rotterdam, the Netherlands.
Roy N. Gerth van Wijk MD, PhD, Erasmus MC, Department of Allergology, Rotterdam, the Netherlands.
Anthony E. J. Dubois MD PhD, University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology and University of Groningen, University Medical Center Groningen, GRIAC Research Institute, Groningen, the Netherlands.
Rick Chin MSc, Division of Rheumatology, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.
Steven Cheuk MD, Calgary, AB, Canada.
Joshua Hoffman PhD, Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA.
Eric Jorgensen PhD, Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA.

ACCEPTED MANUSCRIPT

John S. Witte PhD, Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA.
Ronald B. Melles MD, Department of Ophthalmology, Kaiser Permanente Northern California, Redwood City Medical Center, Redwood City, CA, USA.
Xiumei Hong MD, PhD, Department of Population, Family and Reproductive Health, Center on the Early Life Origins of Disease, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA.
Xiaobin Wang MD, MPH, ScD, Department of Population, Family and Reproductive Health, Center on the Early Life Origins of Disease, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA.
Jennie Hui PhD, School of Population Health, University of Western Australia, Australia and Pathology and Laboratory Medicine, University of Western Australia, Perth, Australia.
Arthur W. (Bill) Musk FRACP, Department of Respiratory Medicine, Sir Charles Gairdner Hospital, and School of Population Health, University of Western Australia, Perth, Australia.
Michael Hunter PhD, School of Population Health, University of Western Australia, Perth, Australia.
Alan L. James FRACP, Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, and School of Medicine and Pharmacology, University of Western Australia, Perth, Australia.
Gerard H. Koppelman MD, PhD, University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology and University of Groningen, University Medical Center Groningen, GRIAC Research Institute, Groningen, the Netherlands.
Andrew J. Sandford PhD, Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada.
Ann E. Clarke MD[1], MSc, Division of Rheumatology, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.
Denise Daley PhD[1], Centre for Heart Lung Innovation, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada.

*Joint first authors, [1]Joint senior authors

Corresponding author: Denise Daley, Centre for Heart Lung Innovation, St. Paul's Hospital, Room 166 - 1081 Burrard Street, Vancouver, B.C. V6Z 1Y6. Phone: (604) 682-2344 x62860 Fax: (604) 806-9274, Email: denise.daley@hli.ubc.ca

Funding: This study was supported by the Allergy, Genes, and Environment Network of Centres of Excellence (AllerGen NCE), Canadian Institutes of Health Research, Canadian Research Chairs Program (salary award to D. Daley), Natural Sciences and Engineering Research Council and Michael Smith Foundation for Health Research/AllerGen NCE post-doctoral research fellowship awards (A. Eslami), Division of Experimental Medicine Entrance Fellowship (Y. Asai), Montreal Children's Hospital Foundation, McGill University Health Centre Foundation, Canadian Dermatology Foundation, Canadian Allergy, Asthma and Immunology Foundation, Canadian Society of Allergy and Clinical Immunology, the National Institutes of Health (grant #CA112355 awarded to J.S. Witte and RC2 AG036607 awarded to C.A. Schaefer and N.J. Risch), HealthWay, the National Health and Medical Research Council (Australia), the

Australian Research Council, and the US Department of Defense (grant # W81XWH-10-1-0487 to K. Allen). The authors C.D.van Ginkel, A.E.J. Dubois and G.H. Koppelman received an unrestricted grant from the Nutricia Research Foundation to complete the Dutch GENEVA cohort. The Dutch IDEAL cohort of N.W. de Jong, H. de Groot, R. Gerth van Wijk and A.E.J. Dubois was supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NOW) and partly funded by the Ministry of Economic Affairs (project number 11868) and Food Allergy Foundation, Siemens Healthcare Diagnostics, HAL Allergy, Intersnack the Netherlands B.V., ALK-Albello B.V., and the Netherlands Anaphylaxis Network. Funding sources had no role in study design, collection, analysis and interpretation of data or in the decision to submit or writing of the report.

Abstract:

Background: Peanut allergy (PA) is a complex disease with both environmental and genetic risk factors. Previously PA loci were identified in *FLG* and *HLA* in candidate gene studies, and loci in *HLA* in a genome-wide association study and meta-analysis.
Objective: To investigate genetic susceptibility to PA.
Methods: Eight hundred and fifty cases and 926 hyper-controls and >7.8 million genotyped and imputed single nucleotide polymorphisms (SNPs) were analyzed in a genome-wide association study to identify susceptibility variants for PA in the Canadian population. Meta-analysis of two phenotypes (PA and food allergy) was conducted using 7 studies from the Canadian, American (2), Australian, German and Dutch (2) populations.
Results: A SNP near *ITGA6* reached genome-wide significance with PA (p=1.80×10^{-8}), while SNPs associated with *SKAP1*, *MMP12/MMP13*, *CTNNA3*, *ARHGAP24*, *ANGPT4*, *c11orf30* (*EMSY*), and *EXOC4* reached a threshold suggestive of association (p≤1.49×10^{-6}). In the meta-analysis of PA, loci in or near *ITGA6*, *ANGPT4*, *MMP12/MMP13*, *c11orf30* and *EXOC4* were significant (p≤1.49×10^{-6}). When a phenotype of any food allergy was used for meta-analysis, the *c11orf30* locus reached genome-wide significance (p=7.50×10^{-11}), while SNPs associated with *ITGA6*, *ANGPT4*, *MMP12/MMP13*, *EXOC4* and additional *c11orf30* SNPs were suggestive (p≤1.49×10^{-6}). Functional annotation indicated *SKAP1* regulates expression of *CBX1*, which co-localizes with the EMSY protein coded by *c11orf30*.
Conclusion: This study identifies multiple novel loci as risk factors for PA and food allergy and establishes *c11orf30* as a risk locus for both peanut and food allergy. Multiple genes (*c11orf30*/*EMSY*, *SKAP1* and *CTNNA3*) identified by this study are involved in epigenetic regulation of gene expression.

Key messages:

- The *c11orf30* (*EMSY*) locus is a novel genetic risk factor for both peanut and food allergy, reaching genome-wide significance for food allergy in meta-analysis (p=7.50×10^{-11})
- New loci associated with *ANGPT4*, *MMP12/MMP13*, and *EXOC4* are suggestive of association with peanut allergy and food allergy
- Epigenetic mechanisms may be a new pathway in the pathogenesis of peanut allergy

Capsule summary:

C*11orf30* (*EMSY*) is a risk locus for food allergy, reaching genome-wide significance in meta-analysis (p=7.50×10^{-11}). Meta-analyses showed five loci suggestive of significance with peanut allergy. These 6 novel loci suggest epigenetic mechanisms.

Key words: peanut allergy; food allergy; genome-wide association study; meta-analysis; *EMSY*; *c11orf30*; epigenetics

Abbreviations: peanut allergy, PA; genome-wide association study, GWAS; skin prick test, SPT; single nucleotide polymorphism, SNP; human leukocyte antigen, HLA; confidence interval, CI; quantile-quantile, QQ; odds ratio, OR; copy number variant, CNV; LD, linkage disequilibrium; CanPAR: Canadian Peanut Allergy Registry; eQTL: expression quantitative trait loci

Introduction:

Peanut allergy (PA) is a main cause of anaphylaxis in North America.[1, 2] In Canada, the prevalence of PA is 1% overall, with a prevalence of 2.2% in children.[2] Self-reported prevalence of tree nut and peanut allergy in the United States was 2.1% in children,[1] while 3% of infants in an Australian study had a positive food challenge to peanut.[3] PA is highly heritable, with a concordance rate of 64% in monozygotic twins compared with 7% in dizygotic twins.[4] Family studies have found the risk of PA in individuals with a peanut-allergic sibling to be significantly higher than the general population, with odds ratios (ORs) ranging from 6.7 to 13.5.[5, 6]

The pathogenesis of PA involves both genetics and the environment. Involvement of environmental exposures is supported by: 1) findings that early oral exposure to peanut leads to development of tolerance[7, 8] 2) the differences in PA prevalence internationally [9-11] and 3) the rapid increase in disease prevalence reported in some studies that cannot be explained by genetic changes.[1]

Previous genetic work has found risk factors for PA in the innate and adaptive immune pathways, including human leukocyte antigen (*HLA*),[12-15] cluster of differentiation 14 (*CD14*),[16] interleukin 9 (*IL9*),[17] and filaggrin (*FLG*).[18, 19] Recently, genome-wide association studies (GWAS) of food allergy identified associations between PA and the *HLA* region.[20, 21] We have previously identified *HLA* and *FLG* associations with PA in a well-characterized group of Canadian peanut-allergic individuals from the Canadian Peanut Allergy Registry (CanPAR).[15, 18, 19] As a follow-up to this work, we conducted a GWAS of PA,[22] along with a meta-analysis of results from the previously published GWAS[20, 21] and other studies of food allergy. *HLA* variants were identified as significant risk factors for PA in the CanPAR GWAS (rs1049213, p=1.82×10^{-11}) and in meta-analysis (rs1063347, p=3.67×10^{-23}), as reported in a separate publication, where we have narrowed the locus to *HLA-DQB1* and show that its relationship to PA is independent of asthma.[22] Here we present novel non-*HLA* loci identified in a GWAS and meta-analysis in an additional 6 populations.

Methods:

Clinical characteristics: Inclusion criteria for CanPAR cases are found in the online repository (Table E1).[19]

GWAS: Salivary DNA was isolated from PA cases in CanPAR. Hyper-controls were self-reported Caucasians from the Busselton Health Study in Australia with no history of asthma, airway hyper-responsiveness, atopy, eczema, allergic rhinitis or food allergy who had blood-derived DNA, and assessment by methacholine challenge and skin prick testing.[23] Genotyping of 1974 individuals (987 cases, 987 controls) was conducted on the Illumina Omni 2.5M+Exome 8v1.1 chip (Génome Québec Innovation Centre, Montréal, Canada). Quality control (QC) including batch effects, single nucleotide polymorphism (SNP) and sample quality are described in the online repository (Figure E1a to c). A total of >7.8 million SNPs (1,388,588 genotyped and 6,441,607 imputed) and 1,776 individuals (850 cases, 926 controls) passed QC (Figure E2a and b). Details on imputation are in the online repository.

Two analyses were performed (related and unrelated), as examination of alleles identical by state and KING[24] kinship coefficients identified related cases (siblings) and controls (1st-3rd degree relatives). PC-AiR[25] and KING[24] (KING1.4; http://people.virginia.edu/~wc9c/KING/)[24] were used to estimate principal components (PC) and kinship coefficients for the related analysis. Association analyses were conducted using Stata®[26] with sandwich estimation to model the clustering of family genotypes with the addition of a family group identifier, 10 PCs to account for population stratification, and plate number to account for plate effects.

A secondary case-control study excluding related individuals was conducted with PLINKv1.07.[27] To make the sample unrelated, 160 individuals were excluded (14 cases, 146 controls); the youngest individual in each family was retained. The unrelated analysis was performed using 834 cases and 781 controls (N=1,615)(Figure E2b).

The analysis including related individuals is our primary analysis, as it has the largest sample size and greatest power. Rank order and odds ratio (OR) differences were evaluated between related and unrelated analyses. All subsequent analyses, including conditioning, were conducted using the related analysis. A *p*-value of 3.60×10^{-8} was considered as the threshold for genome-wide significance (Bonferroni correction), with 1.49×10^{-6} being suggestive evidence for association. We chose 1.49×10^{-6} as our threshold based upon the significance levels presented in two previously published PA GWAS studies.[20, 21]

Conditioning on *HLA*: Following the identification of multiple SNPs in the *HLA* region,[22] we conditioned on the top genotyped SNP (rs3134976) to investigate independence of signals from the rest of the genome and to determine the contribution of *HLA* associations to deviation from the expected line observed in the QQ plots (Figure 1a and b).

Meta-analysis: A meta-analysis was conducted using two phenotypes (PA, food allergy), including previously published PA GWAS results[20, 21] and unpublished data. CanPAR and six additional studies were included in the meta-analysis: two American studies: the Chicago Food Allergy Study (CFA)(N= 2,197; 316 PA cases)[20] and the Genetic Epidemiology Research on Aging (GERA) cohort (N=29,053; 5,108 self-reported food allergy),[28] the Australian HealthNuts study (N=221; 73 PA cases),[21] and the German Understanding Food Allergy (UFA) study[21]

(N=2,592; 205 PA cases) which contributed 21 previously published SNPs. Genotyping for SNPs was conducted in two Dutch studies: IDEAL and GENEVA (N=512; 138 PA cases). Both IDEAL and GENEVA include cases with general food allergy.[29] See Table E2 for full study and phenotype descriptions. The meta-analysis for PA included 1,582 PA cases and 5,446 controls, with over half the cases and controls coming from CanPAR. As GERA utilized self-reported food allergy phenotypes, with no additional diagnostic testing or history, the meta-analysis was completed both with and without GERA to evaluate the sensitivity of the meta-analysis results to stringent food allergy phenotyping. The meta-analysis for food allergy included 7,267 food allergy cases and 29,084 controls, with inclusion of GERA.

Fixed and random effects models evaluate heterogeneity but require point estimates and standard errors. As the CFA study provided *p*-value and sample sizes only, for the meta-analyses *p*-values were obtained using Stouffer's weighted z-score, which requires consistency in the direction of effects (for accurate *p*-value estimation) we were able to confirm that the direction of effect is the same for the CFA study as the investigators provided us with the case/control allele frequencies, which are consistent with the CanPar associations (Table 2).

Identification of expression quantitative trait loci (eQTL): The Genotype-Tissue Expression (gtexportal.org)[30] database was queried for the novel regions.

Results:

GWAS: SNPs in *HLA*[22] and an imputed SNP on chromosome (Chr) 2 close to integrin α 6 (*ITGA6*; rs115218289, $p=1.80\times10^{-8}$) (Table 1, E3, Figure 1a and b) reached genome-wide significance. Several SNPs with suggestive evidence for association were detected in novel loci (Tables 1, E3), including multiple SNPs located in Src kinase associated phosphoprotein 1 (*SKAP1*; Chr17), one located between matrix metallopeptidase 12 (*MMP12*) and *MMP13* (rs144897250, Chr11; $p=2.90\times10^{-7}$), multiple SNPs within catenin α3 (*CTNNA*3; Chr10), rs744597 near rho GTPase activating protein 24 (*ARHGAP24*; Chr4; $p=3.98\times10^{-7}$), rs523865 in angiopoietin 4 (*ANGPT4*; Chr20; $p=4.42\times10^{-7}$), multiple SNPs near chromosome 11 open reading frame (*c11orf30*; Chr11), also known as *EMSY*, and rs78048444, located in a region between coiled-coil-helix-coiled-coil-helix domain containing 3 (*CHCHD3*) and exocyst complex component 4 (*EXOC4*; Chr7; $p=5.44\times10^{-7}$).

No significant difference in OR for SNPs was noted between the unrelated and related analyses (Table 1). For two imputed SNPs (rs115218289 and rs144897250) with low (~2%) minor allele frequency (MAF), there were differences in the rank order between the related and unrelated analyses (Table E3), likely due to the low MAF.

Conditioning: After conditioning on the top genotyped *HLA* SNP (rs3134976, Figure 2a and b), the deviation observed in the QQ plot was largely resolved (Figure 1a and 2a); the residual deviation is primarily due to the number of SNPs supporting *SKAP1*, *CTNNA3*, and *c11orf30/EMSY* associations. Conditioning identified 16 additional SNPs near *SKAP1* and *CTNNA3* (rs139902172) (Table E4).

Meta-analysis for peanut allergy: We identified 85 SNPs in common between CanPAR and one or more of the previously reported PA GWAS.[20, 21] The top novel SNP identified in the meta-analysis for PA was rs115218289, located near *ITGA6*, which did not reach genome-wide significance but met the threshold suggestive for significance (p=9.16×10^{-8}) (Table 2, Table E5, full results in Table E6). Loci in *ANGPT4* (rs523865, p=1.54×10^{-7}), and intragenic SNPs (rs144897250, p=2.94×10^{-7}) near *MMP12/MMP13*, *c11orf30* (rs7936434, p=3.13×10^{-7}), and *EXOC4* (rs78048444, p=3.73×10^{-7}) were suggestive of significance (p≤1.49×10^{-6}) in meta-analysis for the phenotype of PA.

Meta-analysis of food allergy: Using the phenotype of "any food allergy" in all 6 populations, both with and without GERA, the top SNP identified in meta-analysis was rs7936434 near *c11orf30* (p=1.98×10^{-8} and p=7.50×10^{-11} with and without GERA, respectively) (Table 2a, Table E5). The SNPs associated with *ITGA6*, *ANGPT4*, and *MMP12/MMP13* were suggestive of significance (p≤1.49×10^{-6}) in meta-analysis for food allergy, but only if GERA was not included (Table 2a, Table E5). SNPs in *EXOC4*, *ARHGAP24*, *SKAP1* and *CTNNA3* were not suggestive of significance for food allergy.

Identification of eQTLs: Many SNPs identified near *SKAP1* by the CanPAR study were eQTLs, regulating expression of two genes – sorting nexin 1 (*SNX11*) and chromobox protein homolog 1 (*CBX1*) – in numerous tissues (sun-exposed skin, whole blood, transformed fibroblasts, testis, colon and thyroid). Results are presented for tissues relevant to PA and food allergy (sun-exposed skin, whole blood and transformed fibroblasts) with a p-value <1.0×10^{-6} (Table 3, Table E7). Little is known about *SNX11*; it belongs to a family of retrograde transport molecules[31] and its protein is involved in targeting cell surface molecules to the lysosome.[32] *CBX1* is a member of the highly conserved heterochromatin protein family that binds to histones via methylated lysine residues, mediating gene silencing and alternative splicing.[33, 34] It is believed that *CBX1* may play an important role in epigenetic regulation and gene expression.

Discussion:

This study identifies several novel loci for PA and food allergy. The effect sizes for the identified loci are large, with ORs ranging from 0.18 to 0.22 and 1.57 to 6.20; these effect sizes are particularly impressive for a complex disease and a small GWAS. The most significant novel PA locus from the CanPAR GWAS and PA meta-analysis was rs115218289, located in *ITGA6*. This was an imputed SNP with low MAF (0.021) with no directly genotyped SNPs supporting the association (Figure 1a). A second imputed SNP near *MMP12/MMP13* also had a low MAF (0.02). While significant or suggestive of significance in both the PA GWAS and the meta-analyses, the rank order of these two loci changed between the related and unrelated analyses, but no significant changes in the OR were observed. *SKAP1*, *CTNNA3*, and *ARHGAP24* were identified as suggestive for association with PA in the CanPAR GWAS, but were not suggestive for association with PA in the meta-analysis. This is likely due to the small sample size and limited number of contributing studies as evidenced by the minimal change in p-values for these loci in CanPAR GWAS data compared to the meta-analysis.

Loci in *ANGPT4*, *c11orf30* and *EXOC4* were suggestive of significance in meta-analysis for the phenotype of PA, but did not reach genome-wide significance, likely due to small PA sample

sizes, and heterogeneity in study designs (case-control versus family-based studies) and ascertainment criteria. Phenotype definition of both cases and controls differed in each population: cases were defined by food-challenge (IDEAL, GENEVA), food challenge and history with confirmatory testing (CanPAR, CFA, HealthNuts) or self-report (GERA) while controls ranged from hyper-controls (CanPAR), controls without food allergy (IDEAL, GENEVA), controls without positive prick test to foods (HealthNuts), population-based controls (CanPAR, GERA, UFA), and subjects with other food allergies (CFA). Ethnicity was diverse across studies with variable analytic methods used to control for it; age also differed due to the use of population-based controls. The Busselton cohort is a longitudinal dataset, the use of which is required for hyper-controls as they must be negative for allergic phenotypes, and inherently be older to ensure they will not develop eczema, asthma, or other allergic phenotypes. Our efforts to evaluate effect sizes across studies and populations were additionally hindered by differences in GWAS chip and imputation reference panels. *C11orf30* is a prime example: it was identified in the CanPAR GWAS with PA, but it only reaches genome-wide significance (p=7.50×10^{-11}) in the meta-analysis of food allergy. This is likely due to the small PA sample sizes in the other studies. This finding has important implications for other loci identified in the CanPAR GWAS: the lack of significance of the other loci in the meta-analysis should be interpreted with care, as there is insufficient power in the other studies to replicate CanPAR PA findings.

The identification of *c11orf30* as a susceptibility locus for food allergy fits within what is known about this area as it has been implicated in serum IgE levels[35] and asthma.[36] Recently, *c11orf30* was identified as a risk factor for eosinophilic esophagitis (OR, 2.22, p=5.38×10^{-10}),[37] a chronic allergic inflammatory disease of the esophagus that is mainly triggered by food proteins. Loci in this region have been significantly associated with atopic dermatitis[38] both with (rs2155219)[39, 40] and without (rs7927894) any other disease-related phenotypes such as asthma, allergic rhinoconjunctivitis, total serum IgE or family history of atopy.[41] These findings indicate that *c11orf30* is a risk factor for the atopic march, particularly those studies that investigated childhood eczema with later development of asthma.[38] *FLG* and *HLA-DQB1*, genes previously examined in CanPAR,[18, 19, 22] have similarly been found to be associated with childhood eczema and asthma.[38]

It could be argued that genome-wide significance of *c11orf30* in the food allergy meta-analysis with the addition of the GERA cohort indicates that the locus represents association with an allergic diathesis rather than with food allergy or PA specifically, due to the potential misclassification rate for self-reported food allergy in the GERA cohort.[42] The idea of *c11orf30* being associated with an allergic phenotype is supported by data that show that *c11orf30* is a risk factor for poly-sensitization to multiple allergens on prick testing.[43] *C11orf30* is associated with Crohn's disease and ulcerative colitis,[44, 45] autoimmune inflammatory bowel diseases (IBD) that are not classically part of the atopic march, although some epidemiologic data link eczema with IBD and other autoimmune diseases.[46, 47] Therefore, *c11orf30* could be a risk factor for immune dysfunction disorders in general.

Peanut allergy, food allergy, or atopy?

The underlying genetic model for PA is unknown: food allergy, asthma, eczema and allergic rhinitis may share common genetic susceptibility, while environmental factors determine which

specific atopic disease develops; alternatively, each specific allergy may have its own risk variants. The genetic model and study design impact the power to identify risk variants. The CanPAR GWAS subjects were recruited to investigate risk factors specific for PA, with the use of hyper-controls to increase power, in addition to a defined phenotype for allergy to peanut. The other food allergy case groups recruited on any food allergy or on egg allergy and PA;[20, 21, 29] this design may be more powerful if one assumes that all food allergies are influenced by the same genetic risk loci, rather than specific risk loci for specific foods.

Novel pathways in pathogenesis of food allergy:

The identification of *c11orf30* as a genetic risk locus for food allergy opens further research possibilities into the pathways to food allergy. Mechanistically, it is possible that the *c11orf30* region could be responsible for multiple phenotypes associated with the atopic march including food allergy, eczema, and asthma, and other autoimmune conditions. There are several potential ways *c11orf30* could exert its influence to result in a variety of clinical presentations, and are consistent with an epigenetic connection to allergic disease.[20, 48] The protein complex formed by KDM5A, a histone demethylase, and EMSY, the protein encoded by *c11orf30*, appears to increase gene transcription.[49] EMSY acts as a reader for trimethylation of lysine 4 on histone protein H3 (H3K4me3)[50] via its recruitment with other members of the EMSY complex to H3K4me3 marked promoters, where it appears to be positively correlated with transcriptional activity of target genes and cell proliferation.[49]

Along with members of the EMSY protein complex, EMSY co-localizes with CBX1.[33] Similar to its highly conserved *Drosophila* homolog protein dHP1, CBX1 protein has a binding affinity for trimethylation of lysine 9 on histone protein H3 (H3K9me3),[51] modifications that are linked to transcription repression.[52] In this study, *CBX1* was identified as being regulated by *SKAP1* as the same SNPs that are associated with PA also regulate the gene expression of *CBX1*, increasing the connections between the novel loci identified in the CanPAR GWAS. *SKAP1*, which has multiple SNPs identified in the CanPAR GWAS, has a known eQTL in *CBX1*, which encodes a protein that mediates gene silencing and alternative splicing,[33, 34] and whose product co-localizes with EMSY protein, encoded by *c11orf30*.

CTNNA3, also identified as a PA locus in this study, has an identified copy number variant (CNV; CNVR 6828297068284017)[53] in pediatric food allergy. It is also tied to histone modification as there is an enrichment of enhancer- and promoter- associated histone marker H3K4me1 (mono methylation of lysine 4 on histone protein 3) in C*TNNA3* CNVR.[53]

The identification of multiple genes (*c11orf30*, *SKAP1*, and *CTNNA3*) involved with histone-related proteins supports the hypothesis that epigenetic regulation is mechanistic in the development of allergy. Early exposure to peanut prevents development of PA,[7] and maintenance of peanut consumption promotes continued tolerance of peanut;[8] the identified theme of histone-related loci reveals a potential biological mechanism through which this epigenetic phenomenon may occur. This finding could pave the way for potential therapies for those already affected by PA.

Several SNPs in or near genes related to vascular and endothelial cell factors were identified in this study and could be involved in the pathophysiology of PA and food allergy by two putative mechanisms: 1) an endothelial barrier defect, promoting sensitization 2) endothelial cells acting as antigen presenting cells. Some evidence suggests lymphatic endothelial cells present self-antigens and help to regulate peripheral T-cell tolerance.[54] Several identified loci from this study have connections to both allergy and vascular regulation (*CTNNA3*),[55] permeability (*ANGPT4*)[56, 57] or endothelial cell function (*SKAP1*,[58] *EXOC4*[59]). Their true role in the pathogenesis of allergic disease requires further research.

Strengths and Limitations of the study:

Many of the strengths of this study are directly tied to its limitations. The CanPAR study is large with 850 cases after QC and is the largest of the included case groups by a factor of 3, but it is still small for a GWAS. There are no similar PA case groups of equivalent size; this complicates the replication of novel hits. The 6 studies included in the meta-analysis only contribute an additional 732 PA cases, which results in insufficient power to replicate CanPAR findings.

There are numerous sources of heterogeneity that increase the variance of effect size estimation and reduce power. These include study design, ascertainment criteria, case and control phenotype definition, ethnicity, age, and small sample sizes. Differences in ascertainment criteria (food allergy vs PA) may result in differences in the power of the studies to identify susceptibility genes for PA versus food allergy even in the presence of equivalent sample sizes. The impact of the ascertainment criteria on the power ultimately depends upon the underlying genetic model, and will differentially impact loci with peanut-specific susceptibility. The Dutch IMPACT and GENEVA samples are small and contribute point estimates for only 4 SNPs that were specifically genotyped for this study; the absence of genome-wide association data precludes the use of principal components to control for population stratification when generating the point estimates for IMPACT and GENEVA studies.

The importance of phenotyping food allergy is evident by the sizable changes in *p*-value when a large cohort of self-reported food allergy GERA subjects was added to the meta-analysis; it highlighted the potential difference in mechanism for *c11orf30*. The prevalence of food allergy in GERA subjects – which is self-reported with no corroborating clinical history or diagnostic tests – is 17.58%, which is much higher than other population-based studies in North America (6.4 to 7.5%),[2] indicating there is likely a high case misclassification rate. Misclassification is a known differential bias to the null and can result in false negatives, which can be problematic for relatively rare phenotypes such as PA and food allergy, as the misclassification rate can exceed the disease prevalence rate. This is particularly an issue in the use of universal controls, where genotyping of the controls has been performed in a separate experiment, and misclassification of cases and controls is common. Misclassification of an affected individual as a control is much less costly than the misclassification of a control as an affected individual.[60] A strength of the CanPAR study is that the cases and controls were genotyped in the same experiment. Additionally, using hyper-controls with comprehensive longitudinal lifetime history and in-depth phenotyping of asthma, atopy, skin prick tests, IgE, airway hyper-responsiveness and eczema phenotypes virtually eliminates misclassification of controls in the CanPAR study.

Careful consideration of potential misclassification needs to be made when using self-reported food allergy phenotypes in the absence of food challenges, skin prick tests and food-specific IgE levels. This has important implications in today's research environment where investigators are continually striving to maximize the utilization of existing GWAS data. We also see this conundrum with the common adoption of general population controls which may contribute to population stratification and can result in false positives, a differential bias away from the null; both CanPAR and UFA employed this approach. The use of hyper-controls in CanPAR increased the power to detect associations with PA, but makes the evaluation of confounders such as eczema or asthma difficult, as none of the controls express these phenotypes. The use of this study design with controls from a different population limits our ability to evaluate early life environmental exposures.

Direct genotyping was not conducted for rs7936434 (imputed), the most significant association observed in the *c11orf30/EMSY* region. While direct genotyping would ensure that the finding is not an imputation artifact it is unlikely that the observed association is due to artifact. The *c11orf30/EMSY* locus, reached genome-wide significance due to the contributions not only of CanPAR but two other studies (CFA and GERA). Additionally, the strong correlation of this locus with allergy and atopic conditions in multiple studies[35-41] serve to alleviate concerns that this association may be an imputation artifact.

Despite all limitations, this study robustly identified new loci for PA and food allergy with genome-wide significance (p=7.50×10^{-11}) across populations in a meta-analysis of unprecedented size (1,582 PA cases and 5,446 controls, and 7,627 food allergy cases and 29,084 food allergy controls). It is important to note the effect sizes for the loci identified are large, particularly for a complex disease, and that the finding of *c11orf30/EMSY* is robust as it demonstrates association across the studies despite the diversity of the studies. The eQTL data support these new loci and suggest new pathways in the pathogenesis of food allergy.

Conclusion and future directions:

Results of this study identify novel genetic risk factors for PA and food allergy. New pathways identified by this unbiased approach include *c11orf30/EMSY* in peanut and food allergy, and the importance of epigenetic mechanisms. It is evident that further work will require larger sample sizes, and international collaboration using well-phenotyped individuals for both food allergies and other atopic conditions. Functional work including studies of vascular and endothelial cell factors may be valuable. Future studies will need to examine gene-environment interactions, including duration, timing and mode of environmental exposure, and the role of CNVs, methylation, and histone modification.

Acknowledgements:

We wish to thank the subjects and parents who participated in this study and in the CanPAR registry, and their allergists for help in diagnostic documentation. We appreciate the help of our partners Food Allergy Canada (formerly Anaphylaxis Canada), Allergies Québec (formerly l'Association Québecois des Allergies Alimentaires), and the Allergy/Asthma Information Association with subject recruitment. The Genotype-Tissue Expression (GTEx) Project was

ACCEPTED MANUSCRIPT

supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 01/01/2017 to 02/28/2017. We acknowledge Basia Rogula and David Yin who assisted in quality control, JinCheol Choi who assisted with figure development and Jessica Que who assisted with SNP annotation. Whitney Steber, Chynace Van Lambalgen, Heather Waldhauser, Greg Shand, Elizabeth Turnbull, and Popi Panaritis assisted in data/DNA collection for CanPAR. We thank Peter Paré for his valuable comments and Reza Alizadehfar, Edmund Chan, and Peter Hull for their data collection and helpful discussions.

References:

1. Sicherer SH, Munoz-Furlong A, Godbold JH, Sampson HA. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J Allergy Clin Immunol 2010; 125:1322-6.
2. Soller L, Ben-Shoshan M, Harrington DW, Knoll M, Fragapane J, Joseph L, et al. Adjusting for nonresponse bias corrects overestimates of food allergy prevalence. J Allergy Clin Immunol Pract 2015; 3:291-3 e2.
3. Osborne NJ, Koplin JJ, Martin PE, Gurrin LC, Lowe AJ, Matheson MC, et al. Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. J Allergy Clin Immunol 2011; 127:668-76 e1-2.
4. Sicherer SH, Furlong TJ, Maes HH, Desnick RJ, Sampson HA, Gelb BD. Genetics of peanut allergy: a twin study. J Allergy Clin Immunol 2000; 106:53-6.
5. Liem JJ, Huq S, Kozyrskyj AL, Becker AB. Should Younger Siblings of Peanut-Allergic Children Be Assessed by an Allergist before Being Fed Peanut? Allergy Asthma Clin Immunol 2008; 4:144-9.
6. Hourihane JO, Dean TP, Warner JO. Peanut allergy in relation to heredity, maternal diet, and other atopic diseases: results of a questionnaire survey, skin prick testing, and food challenges. Bmj 1996; 313:518-21.
7. 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; 372:803-13.
8. 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; 374:1435-43.
9. Silva R, Gomes E, Cunha L, Falcao H. Anaphylaxis in children: a nine years retrospective study (2001-2009). Allergol Immunopathol (Madr) 2012; 40:31-6.
10. Hoffer V, Scheuerman O, Marcus N, Levy Y, Segal N, Lagovsky I, et al. Anaphylaxis in Israel: experience with 92 hospitalized children. Pediatr Allergy Immunol 2011; 22:172-7.
11. Asero R, Antonicelli L, Arena A, Bommarito L, Caruso B, Colombo G, et al. Causes of food-induced anaphylaxis in Italian adults: a multi-centre study. Int Arch Allergy Immunol 2009; 150:271-7.

12. Boehncke WH, Loeliger C, Kuehnl P, Kalbacher H, Bohm BO, Gall H. Identification of HLA-DR and -DQ alleles conferring susceptibility to pollen allergy and pollen associated food allergy. Clin Exp Allergy 1998; 28:434-41.
13. Howell WM, Turner SJ, Hourihane JO, Dean TP, Warner JO. HLA class II DRB1, DQB1 and DPB1 genotypic associations with peanut allergy: evidence from a family-based and case-control study. Clin Exp Allergy 1998; 28:156-62.
14. Dreskin SC. Do HLA genes play a role in the genetics of peanut allergy? Annals of Allergy, Asthma & Immunology 2006; 96:766-8.
15. Madore AM, Vaillancourt VT, Asai Y, Alizadehfar R, Ben-Shoshan M, Michel DL, et al. HLA-DQB1*02 and DQB1*06:03P are associated with peanut allergy. Eur J Hum Genet 2013; 21:1181-4.
16. Brough HA, Cousins DJ, Munteanu A, Wong YF, Sudra A, Makinson K, et al. IL-9 is a key component of memory TH cell peanut-specific responses from children with peanut allergy. J Allergy Clin Immunol 2014; 134:1329-38.e10.
17. Dreskin SC, Ayars A, Jin Y, Atkins D, Leo HL, Song B. Association of genetic variants of CD14 with peanut allergy and elevated IgE levels in peanut allergic individuals. Annals of Allergy, Asthma & Immunology 2011; 106:170-2.
18. Brown SJ, Asai Y, Cordell HJ, Campbell LE, Zhao Y, Liao H, et al. Loss-of-function variants in the filaggrin gene are a significant risk factor for peanut allergy. J Allergy Clin Immunol 2011; 127:661-7.
19. Asai Y, Greenwood C, Hull PR, Alizadehfar R, Ben-Shoshan M, Brown SJ, et al. Filaggrin gene mutation associations with peanut allergy persist despite variations in peanut allergy diagnostic criteria or asthma status. J Allergy Clin Immunol 2013; 132:239-42.
20. Hong X, Hao K, Ladd-Acosta C, Hansen KD, Tsai HJ, Liu X, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. Nat Commun 2015; 6:6304.
21. Martino DJ, Ashley S, Koplin J, Ellis J, Saffery R, Dharmage SC, et al. Genome-wide association study of peanut allergy reproduces association with amino acid polymorphisms in HLA-DRB1. Clin Exp Allergy 2016.
22. Asai Y, Eslami A, van Ginkel CD, Akhabir L, Wan M, Yin D, et al. Canadian genome-wide association study and meta-analysis confirm HLA as a risk factor for peanut allergy. J Allergy Clin Immunol 2017; submitted.
23. James AL, Knuiman MW, Divitini ML, Hui J, Hunter M, Palmer LJ, et al. Changes in the prevalence of asthma in adults since 1966: the Busselton health study. Eur Respir J 2010; 35:273-8.
24. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010; 26:2867-73.
25. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet Epidemiol 2015; 39:276-93.
26. StataCorp. Stata Statistical Software: Release 11. College Station, Texas: StataCorp LP, 2009.

ACCEPTED MANUSCRIPT

27. 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. Am J Hum Genet 2007; 81:559-75.
28. Hoffmann TJ, Kvale MN, Hesselson SE, Zhan Y, Aquino C, Cao Y, et al. Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. Genomics 2011; 98:79-89.
29. van Ginkel CD, Flokstra-de Blok BM, Kollen BJ, Kukler J, Koppelman GH, Dubois AE. Loss-of-function variants of the filaggrin gene are associated with clinical reactivity to foods. Allergy 2015; 70:461-4.
30. Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013; 45:580-5.
31. Xu J, Xu T, Wu B, Ye Y, You X, Shu X, et al. Structure of sorting nexin 11 (SNX11) reveals a novel extended phox homology (PX) domain critical for inhibition of SNX10-induced vacuolation. J Biol Chem 2013; 288:16598-605.
32. Li C, Ma W, Yin S, Liang X, Shu X, Pei D, et al. Sorting Nexin 11 Regulates Lysosomal Degradation of Plasma Membrane TRPV3. Traffic 2016; 17:500-14.
33. Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF, et al. EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. Cell 2003; 115:523-35.
34. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. Trends Genet 2015; 31:274-80.
35. Li X, Ampleford EJ, Howard TD, Moore WC, Li H, Busse WW, et al. The C11orf30-LRRC32 region is associated with total serum IgE levels in asthmatic patients. J Allergy Clin Immunol 2012; 129:575-8, 8.e1-9.
36. Li X, Hastie AT, Hawkins GA, Moore WC, Ampleford EJ, Milosevic J, et al. eQTL of bronchial epithelial cells and bronchial alveolar lavage deciphers GWAS-identified asthma genes. Allergy 2015; 70:1309-18.
37. Sleiman PM, Wang ML, Cianferoni A, Aceves S, Gonsalves N, Nadeau K, et al. GWAS identifies four novel eosinophilic esophagitis loci. Nat Commun 2014; 5:5593.
38. Manz J, Rodriguez E, ElSharawy A, Oesau EM, Petersen BS, Baurecht H, et al. Targeted Resequencing and Functional Testing Identifies Low-Frequency Missense Variants in the Gene Encoding GARP as Significant Contributors to Atopic Dermatitis Risk. J Invest Dermatol 2016; 136:2380-6.
39. Marenholz I, Esparza-Gordillo J, Ruschendorf F, Bauerfeind A, Strachan DP, Spycher BD, et al. Meta-analysis identifies seven susceptibility loci involved in the atopic march. Nat Commun 2015; 6:8804.
40. Weidinger S, Willis-Owen SA, Kamatani Y, Baurecht H, Morar N, Liang L, et al. A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. Hum Mol Genet 2013; 22:4841-56.
41. Greisenegger EK, Zimprich F, Zimprich A, Gleiss A, Kopp T. Association of the chromosome 11q13.5 variant with atopic dermatitis in Austrian patients. Eur J Dermatol 2013; 23:142-5.
42. Rona RJ, Keil T, Summers C, Gislason D, Zuidmeer L, Sodergren E, et al. The prevalence of food allergy: a meta-analysis. J Allergy Clin Immunol 2007; 120:638-46.
43. Amaral AF, Minelli C, Guerra S, Wjst M, Probst-Hensch N, Pin I, et al. The locus C11orf30 increases susceptibility to poly-sensitization. Allergy 2015; 70:328-33.

44. Dubinsky MC, Kugathasan S, Kwon S, Haritunians T, Wrobel I, Wahbeh G, et al. Multidimensional prognostic risk assessment identifies association between IL12B variation and surgery in Crohn's disease. Inflamm Bowel Dis 2013; 19:1662-70.
45. Waterman M, Xu W, Stempak JM, Milgrom R, Bernstein CN, Griffiths AM, et al. Distinct and overlapping genetic loci in Crohn's disease and ulcerative colitis: correlations with pathogenesis. Inflamm Bowel Dis 2011; 17:1936-42.
46. Schmitt J, Schwarz K, Baurecht H, Hotze M, Folster-Holst R, Rodriguez E, et al. Atopic dermatitis is associated with an increased risk for rheumatoid arthritis and inflammatory bowel disease, and a decreased risk for type 1 diabetes. J Allergy Clin Immunol 2016; 137:130-6.
47. Kim M, Choi KH, Hwang SW, Lee YB, Park HJ, Bae JM. Inflammatory bowel disease is associated with an increased risk of inflammatory skin diseases: A population-based cross-sectional study. J Am Acad Dermatol 2017; 76:40-8.
48. Song Y, Liu C, Hui Y, Srivastava K, Zhou Z, Chen J, et al. Maternal allergy increases susceptibility to offspring allergy in association with TH2-biased epigenetic alterations in a mouse model of peanut allergy. J Allergy Clin Immunol 2014; 134:1339-45.e7.
49. Varier RA, Carrillo de Santa Pau E, van der Groep P, Lindeboom RG, Matarese F, Mensinga A, et al. Recruitment of the Mammalian Histone-modifying EMSY Complex to Target Genes Is Regulated by ZNF131. J Biol Chem 2016; 291:7313-24.
50. Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell 2010; 142:967-80.
51. Kaustov L, Ouyang H, Amaya M, Lemak A, Nady N, Duan S, et al. Recognition and specificity determinants of the human cbx chromodomains. J Biol Chem 2011; 286:521-9.
52. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007; 129:823-37.
53. Li J, Fung I, Glessner JT, Pandey R, Wei Z, Bakay M, et al. Copy Number Variations in CTNNA3 and RBFOX1 Associate with Pediatric Food Allergy. J Immunol 2015; 195:1599-607.
54. Hirosue S, Vokali E, Raghavan VR, Rincon-Restrepo M, Lund AW, Corthesy-Henrioud P, et al. Steady-state antigen scavenging, cross-presentation, and CD8+ T cell priming: a new role for lymphatic endothelial cells. J Immunol 2014; 192:5002-11.
55. Folmsbee SS, Budinger GR, Bryce PJ, Gottardi CJ. The cardiomyocyte protein alphaT-catenin contributes to asthma through regulating pulmonary vein inflammation. J Allergy Clin Immunol 2016.
56. Kesler CT, Pereira ER, Cui CH, Nelson GM, Masuck DJ, Baish JW, et al. Angiopoietin-4 increases permeability of blood vessels and promotes lymphatic dilation. Faseb j 2015; 29:3668-77.
57. Yao JH, Cui M, Li MT, Liu YN, He QH, Xiao JJ, et al. Angiopoietin1 inhibits mast cell activation and protects against anaphylaxis. PLoS One 2014; 9:e89148.
58. Jiang L, Hu J, Feng J, Han D, Yang C. Substrate stiffness of endothelial cells directs LFA-1/ICAM-1 interaction: A physical trigger of immune-related diseases? Clin Hemorheol Microcirc 2016; 61:633-43.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

59. Barkefors I, Fuchs PF, Heldin J, Bergstrom T, Forsberg-Nilsson K, Kreuger J. Exocyst complex component 3-like 2 (EXOC3L2) associates with the exocyst complex and mediates directional migration of endothelial cells. J Biol Chem 2011; 286:24189-99.
60. Edwards BJ, Haynes C, Levenstien MA, Finch SJ, Gordon D. Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. BMC Genet 2005; 6:18.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Tables and Figure Legends for: 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

Table 1: Most significant SNPs from 10 genomic regions identified in the Canadian Peanut Allergy Registry genome-wide association study, listed by order of significance*

						Related analysis (850 cases, 926 controls)				Unrelated analysis (834 cases, 781 controls)					
SNP	Chr	Position	Allele	MAF	Source of SNPs	OR	LCI	UCI	P	OR	LCI	UCI	P	P#	Gene/Nearest Gene
rs115218289	2	173265750	A/C	0.02	Imputed	0.18	0.10	0.32	1.80×10^{-8}	0.20	0.09	0.46	1.39×10^{-4}	8.04×10^{-1}	(298kb)*DLX2* | (26kb)*ITGA6*
rs72827854	17	46460525	T/C	0.09	Imputed	2.16	1.61	2.90	2.60×10^{-7}	2.08	1.50	2.87	9.00×10^{-6}	8.58×10^{-1}	*SKAP1*
rs144897250	11	102750264	A/C	0.02	Imputed	6.20	3.09	12.45	2.90×10^{-7}	6.72	2.72	16.64	3.79×10^{-5}	8.90×10^{-1}	(5kb)*MMP12* | (63kb)*MMP13*
rs7475217	10	68444013	T/C	0.38	Genotyped	1.64	1.35	1.98	3.58×10^{-7}	1.56	1.28	1.90	9.19×10^{-6}	7.37×10^{-1}	*CTNNA3*
rs744597	4	86337028	A/G	0.40	Genotyped	0.61	0.50	0.74	3.98×10^{-7}	0.63	0.52	0.77	3.91×10^{-6}	8.01×10^{-1}	*ARHGAP24*
rs523865	20	894881	C/T	0.23	Genotyped	0.57	0.46	0.71	4.42×10^{-7}	0.57	0.45	0.71	1.19×10^{-6}	9.49×10^{-1}	*ANGPT4*
rs7936434	11	76293805	C/G	0.49	Imputed	1.58	1.32	1.90	5.17×10^{-7}	1.58	1.31	1.91	2.73×10^{-6}	9.85×10^{-1}	(30kb)*C11orf30* | (43kb)*LOC101928813*
rs78048444	7	132832218	C/T	0.02	Genotyped	0.22	0.12	0.39	5.44×10^{-7}	0.23	0.11	0.46	4.57×10^{-5}	9.35×10^{-1}	(65kb)*CHCHD3* | (106kb)*EXOC4*
rs56151068	17	46381431	T/C	0.10	Imputed	2.06	1.54	2.76	9.58×10^{-7}	1.97	1.44	2.70	2.34×10^{-5}	8.39×10^{-1}	*SKAP1;LOC101927148*
rs139462954	17	46523678	A/AC	0.09	Imputed	2.06	1.54	2.76	1.23×10^{-6}	1.97	1.43	2.71	2.92×10^{-5}	8.37×10^{-1}	*LOC101927166*

*Nearest gene was used to determine genomic location. Chr: chromosome; Allele: minor allele/major allele; MAF: minor allele frequency; OR: odds ratio; LCI: lower 95% confidence interval; UCI: upper 95% confidence interval; P: *p*-value; #comparing related and unrelated analyses odds ratio

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2: Meta-analysis of Canadian, American, Australian, German, and Dutch populations for association with peanut and food allergy phenotypes

			Peanut allergy							Food allergy						
SNP†	**Chr**	**Allele**	**P-CanPAR*[a]**	**P – CFA[b]**	**P - HealthNuts*[c]**	**P-UFA*[d]**	**P - IDEAL/GENEVA Case-control[e]**	**P - Dutch GENEVA Family - based[f]**	**Pmeta_PA[g]**	**P –CFA[h]**	**P - IDEAL/GENEVA Case-control[i]**	**P - GENEVA Family study[j]**	**P - GERA[k]**	**Pmeta[l]**	**Pmeta (without GERA)[m]**	**Gene/Nearest Gene**
rs115218289^	2	A/C	1.80×10^{-8}	NA	6.77×10^{-1}	NA	7.18×10^{-1}	3.54×10^{-1}	9.16×10^{-8}	NA	1.84×10^{-1}	1.90×10^{-1}	5.25×10^{-1}	2.91×10^{-2}	2.38×10^{-8}	(298kb)*DLX2*| (26kb)*ITGA6*
rs523865#	20	C/T	4.42×10^{-7}	NA	NA	NA	8.33×10^{-1}	1.63×10^{-2}	1.54×10^{-7}	NA	2.03×10^{-1}	2.60×10^{-2}	2.66×10^{-1}	9.29×10^{-3}	4.09×10^{-8}	*ANGPT4*
rs144897250^	11	A/C	2.90×10^{-7}	NA	3.84×10^{-1}	NA	NA	NA	2.94×10^{-7}	NA	NA	NA	5.89×10^{-1}	6.83×10^{-2}	2.94×10^{-7}	(5kb)*MMP12*| (63kb)*MMP13*
rs7936434^	11	C/G	5.17×10^{-7}	3.66×10^{-2}	1.43×10^{-1}	NA	NA	NA	3.13×10^{-7}	5.89×10^{-5}	NA	NA	4.13×10^{-4}	1.98×10^{-8}	7.50×10^{-11}	(30kb)*C11orf30*| (43kb) *LOC101928813*
rs78048444#	7	C/T	5.44×10^{-7}	NA	2.13×10^{-1}	NA	6.46×10^{-1}	3.54×10^{-1}	3.73×10^{-7}	NA	8.29×10^{-1}	7.87×10^{-1}	6.97×10^{-1}	8.88×10^{-2}	2.53×10^{-6}	(65kb)*CHCHD3*| (106kb)*EXOC4*
rs744597#	4	A/G	3.98×10^{-7}	1.15×10^{-1}	5.57×10^{-1}	NA	8.20×10^{-2}	9.62×10^{-1}	1.63×10^{-6}	5.19×10^{-1}	1.13×10^{-1}	2.69×10^{-1}	3.63×10^{-1}	1.29×10^{-2}	1.42×10^{-5}	*ARHGAP24*
rs72827854^	17	T/C	2.60×10^{-7}	1.36×10^{-1}	3.55×10^{-1}	NA	NA	NA	3.43×10^{-6}	5.83×10^{-1}	NA	NA	7.69×10^{-1}	9.27×10^{-2}	7.47×10^{-5}	*SKAP1*
rs55765969^	17	T/C	1.23×10^{-6}	1.97×10^{-1}	2.81×10^{-1}	NA	NA	NA	1.45×10^{-5}	6.59×10^{-1}	NA	NA	3.77×10^{-1}	3.18×10^{-2}	1.97×10^{-4}	*LOC101927166*
rs56151068^	17	T/C	9.58×10^{-7}	2.51×10^{-1}	3.57×10^{-1}	NA	NA	NA	2.33×10^{-5}	7.33×10^{-1}	NA	NA	4.45×10^{-1}	4.45×10^{-2}	2.66×10^{-4}	*SKAP1; LOC101927148*
rs71193762^	10	A/G	3.77×10^{-7}	8.63×10^{-1}	3.62×10^{-1}	NA	NA	NA	2.73×10^{-4}	6.87×10^{-1}	NA	NA	7.06×10^{-1}	8.83×10^{-2}	1.41×10^{-4}	*CTNNA3*

† SNPs in this table were selected to represent each of the 10 genomic regions identified in the CanPAR GWAS (see Table E4 for full results).

Shaded rows indicate suggestive significance ($p \leq 1.49 \times 10^{-6}$) in peanut allergy.

* Used for both - Peanut allergy (PA) and food allergy (FA)

[a]P-value from CanPAR (N=1,776), [b]P-value from Chicago FoodAllergy Study (N=2,197), [c]P-value from Australian HealthNutstudy (N=221), [d]P-value from German Understanding of Food Allergy study (N=2,592)

[e]Number of individuals 226, 229, 227, and 217 for rs115218289, rs523865, rs78048444, and rs744597, respectively, corrected for atopic dermatitis (AD), Asthma (As), and rhinoconjunctivitis (RC)

[f]P-value from Dutch GENEVA family study, number of informative families 20, 112, 21, and 112 for rs115218289, rs523865, rs78048444, and rs744597, respectively

[g]P-value from Stouffer's weighted z-score meta-analysis method for peanut allergy

[h]P-value from Chicago Food Allergy Study (N=2,197)

[i]P-value from Dutch IDEAL and GENEVA case-control study, number of individuals for SNPs 479, 487, 482, and 466 for rs115218289, rs523865, rs78048444, and rs744597, respectively, corrected for atopic dermatitis (AD), Asthma (As), and rhinoconjunctivitis (RC)

ACCEPTED MANUSCRIPT

[j]P-value from Dutch GENEVA, number of informative families for SNPs 26, 196, 37, and 214 for rs115218289, rs523865, rs78048444, and rs744597, respectively
[k]P-value from GERA food allergy study (N=29,053)
[l]P-value from Stouffer's weighted z-score meta-analysis method for food allergy
[m]P-value from Stouffer's weighted z-score meta-analysis method for food allergy without GERA study
Chr: chromosome; Allele: minor allele/major allele; P: *p*-value
[^] imputed SNP from CanPAR
[#] genotyped SNP from CanPAR

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3: Top Canadian Peanut Allergy Registry SNPs associated with PA and Expression Quantitative Trait Loci (eQTL) for each locus and tissue type

SNP	Chr	Position	Nearest gene	Tissue	*p*-value	Gene symbol
rs4491576	17	46408636	*SKAP1*	whole blood	1.05×10^{-10}	*SNX11*
rs16956501	17	46497274	*SKAP1*	skin, sun exposed lower leg	4.48×10^{-10}	*SNX11*
rs139462954	17	46523678	LOC101927166	skin, sun exposed lower leg	8.94×10^{-10}	*SNX11*
rs139462954	17	46523678	LOC101927166	cells, transformed fibroblasts	7.36×10^{-7}	*CBX1*

Chr: chromosome; P: *p*-value

ACCEPTED MANUSCRIPT

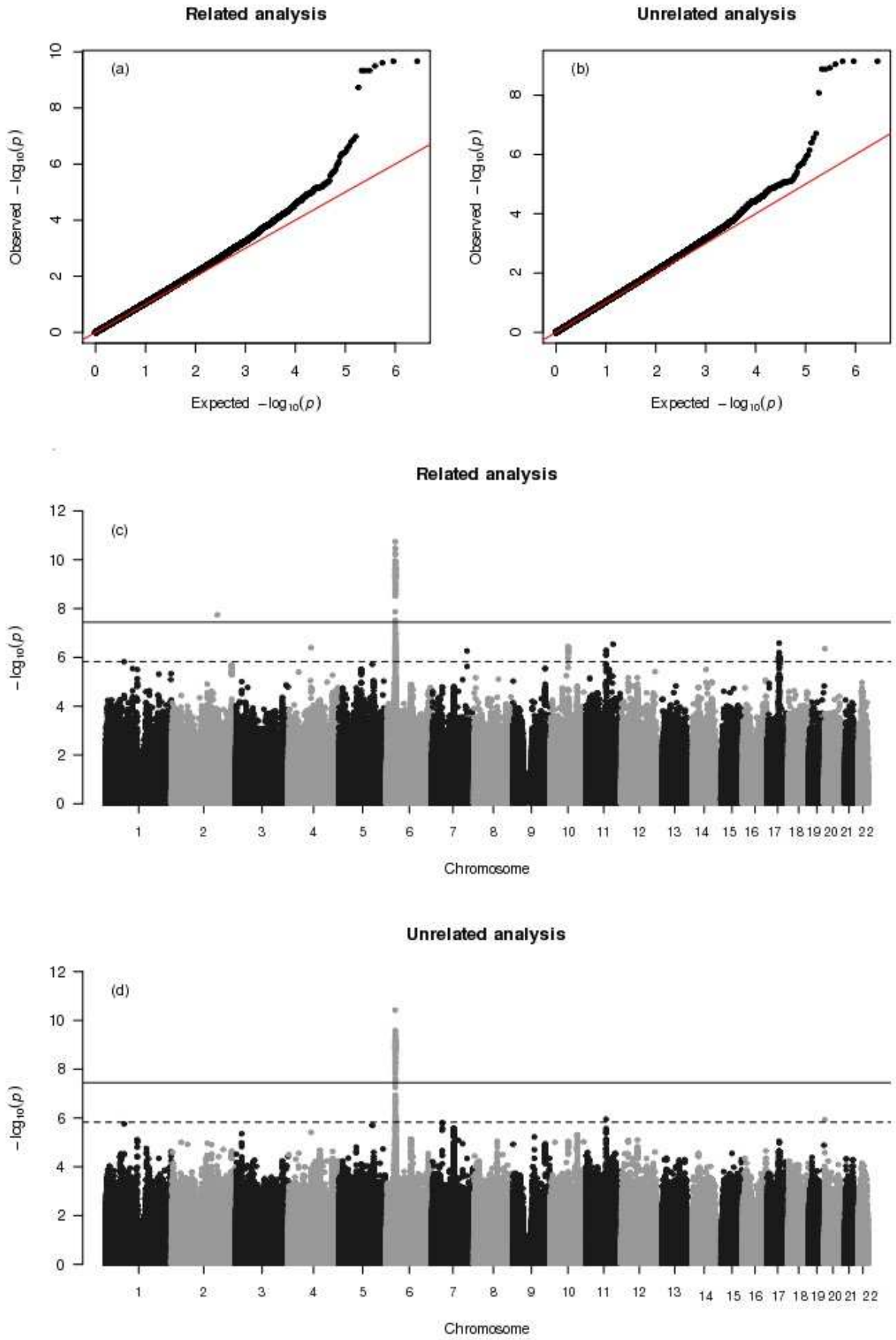
ACCEPTED MANUSCRIPT

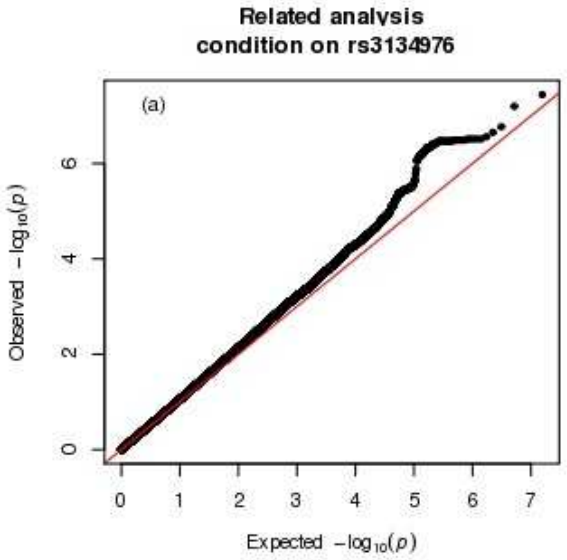
FIGURE LEGENDS:

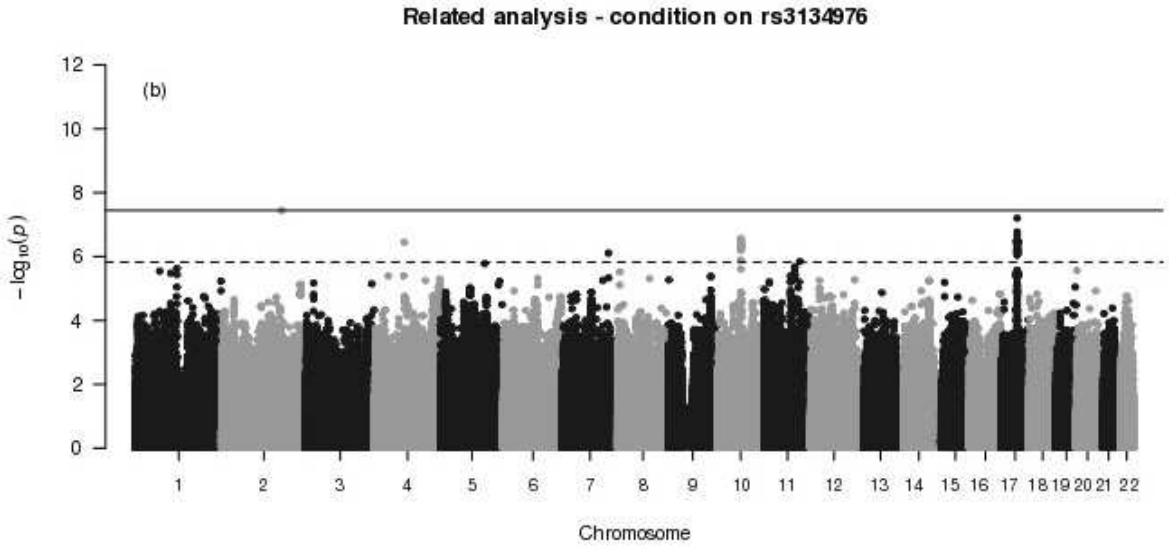
Figure 1: QQ and Manhattan plots of related and unrelated analyses. a) QQ plot of the expected distribution of test statistics *(x*-axis) versus observed *p*-values *(y*-axis) for related (left) and unrelated analysis (right). **b)** Manhattan plot: SNPs in 850 PA cases and 926 hyper-controls for the related (upper) and unrelated (lower) analyses. The *x*-axis denotes the genomic location and the *y*-axis the association level. The solid line indicates the threshold for genome-wide significance ($p \leq 3.60 \times 10^{-8}$) and the dashed line indicates the suggestive association significance threshold ($p \leq 1.49 \text{x} 10^{-6}$).

Figure 2: QQ and Manhattan plots of related and unrelated analyses, conditioned on the top *HLA* SNP (rs3134976). a) QQ plot of the expected distribution of test statistics (*x*-axis) versus observed *p*-values (*y*-axis) for related analysis, conditioned on rs3134976. SNPs in complete linkage with rs3134976 were excluded. **b)** Manhattan plot: SNPs in 850 PA cases and 926 hyper-controls for the related analysis, conditioned on rs3134976. The *x*-axis denotes genomic location and the *y*-axis association level. The solid line indicates the threshold for genome-wide significance ($p \leq 3.60 \times 10^{-8}$) and the dashed line indicates the suggestive association significance threshold ($p \leq 1.49 \text{x} 10^{-6}$).

ACCEPTED MANUSCRIPT


Related analysis
Unrelated analysis
(a)
(b)
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
Related analysis
(c)
Unrelated analysis
(d)
$-\log_{10}(p)$
Chromosome


Related analysis
condition on rs3134976
(a)
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
0
2
4
6
0
1
2
3
4
5
6
7


Related analysis - condition on rs3134976
(b)
$-\log_{10}(p)$
Chromosome
0
2
4
6
8
10
12
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

ACCEP

ACCEPTED MANUSCRIPT

Online Repository Material for: 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

METHODS:

Genome-wide association study (GWAS): Inclusion criteria for the PA cases[E1]are shown in Table E1. Hyper-controls: individuals with no history of asthma, airway hyperresponsiveness, atopy, eczema, allergic rhinitis or food allergy selected from the Busselton Health Study from Busselton, Australia.[E2] Hyper-controls were selected to maximize the potential of capturing any potential risk variants given the small number of cases for a GWAS; this was weighed against the possibility of population stratification given the recruitment of cases and controls from different countries, and differential SNP call rates from saliva versus blood–derived samples. PA case DNA was extracted from saliva and controls from blood.[E2] Genotyping of 987 cases and 987 controls (N=1974) was conducted on the Illumina Omni 2.5M+Exome 8v1.1 chip at the Génome Québec Innovation Centre (Montréal, Canada).

Quality control (QC) included evaluation of batch effects, single nucleotide polymorphism (SNP) and sample quality. Testing for batch effects identified statistically significant differences in call rates between plates and cases and controls but no interaction between the two (Figure E1a to c). Plate number was included as a covariate to account for these differences, which increased the *p*-values but did not change the ranking of the top SNPs. To further investigate confounding by DNA source, genotypes were clustered in the cases and controls jointly (which is ideal), then cases and controls were clustered separately and genotype calls between joint and separated clustering were compared. No significant differences were observed and the joint genotype calls were used for the analysis. (Figure E1a and b and c)

SNPs with a call rate of less than 98% (N= 56,988), ≥2 Mendelian errors (N=2), and monomorphic SNPs (N=483,553) were excluded. SNPs with 1 Mendelian error (N=312) and departure from Hardy-Weinberg equilibrium (HWE), with $p<10^{-8}$ (N=424), were flagged but retained in the analysis; SNPs with ≥2 of these analysis flags were excluded (N=19). It should be noted that none of the SNPs associated in this study with either PA or food allergy were flagged during the analysis. Due to the design of the chip, 94,443 SNPs were included two or more times: 90,296 duplicates (45,148x2), 4,131 triplicates (1,377x3), and 16 quadruplicates (4x4). Of the duplicated SNPs, those with more than one discordant genotype (N=1,900) and those with the lowest call rate were excluded (N=41,521). Analysis was restricted to SNPs with a minor allele frequency (MAF) >2%; a total of 1,388,588 SNPs were included 8 genotyping failures (4 cases, 4 controls); 24 gender errors (19 cases, 5 controls) and 49 samples with call rates <98% (26 cases, 23 controls) were removed (Figure E2a). One case subject withdrew consent, and 5 cases failed to meet clinical criteria set out in the protocol. Identity by state and multidimensional scaling as implemented in PLINKv1.07 (http://pngu.mgh.harvard.edu/purcell/plink/)[E3] was used to identify 13 duplicated case samples and 71 population outliers (56 cases, 15 controls) which were removed from the analysis. Heterozygosity outliers (>3 standard deviations) were identified and removed (13 cases, 14 controls). A primary related analysis and a secondary unrelated analysis were completed (Figure E2a and b). After completion of QC 1,776 individuals (850 cases, 926 controls) were available for the primary related analysis and 1,615 (834 cases, 781 controls) in the secondary unrelated case-control analysis (Figure E2a and b).

Imputation: Genotypic imputation was performed using reference data from 1000 Genomes Project (phase 3). Haplotypes were phased using SHAPEIT2

ACCEPTED MANUSCRIPT

(https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html).[E4] Genotype data from 1,388,588 directly genotyped variants were used to impute unmeasured variants, using IMPUTEv2.3.2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html).[E5] QC for imputed data was performed by filtering for information score < 80%, certainty < 98%, and MAF < 0.02 (among unrelated individuals) resulting in 6,441,607 SNPs that passed all QC.

Meta-analysis: Case and control groups used for the meta-analysis are described in Table E2. Analysis of published studies has been previously described.[E6-10] Due to differences in genotyping platforms and imputation reference panels, it was not possible to evaluate all SNPs in all studies.

Dutch case groups (IDEAL and GENEVA): Two Dutch case groups were included, GENEVA and IDEAL. IDEAL[E11] was designed to study the diagnosis of cashew allergy in children sensitized to cashew. GENEVA was initiated to study the genetics of food allergy[E12] and includes children who had a double-blind, placebo-controlled food challenge (DBPCFC) and have clinical data for a variety of foods. DNA was collected by blood or saliva depending on the clinical routine. Genotyping was performed by competitive allele-specific PCR using KASPar genotyping chemistry (LGC Genomic, Teddington, UK). Associations between the SNPs and DBPCFC outcomes were analyzed in trios of the GENEVA cohort using a family-based design. The case-control analysis of both GENEVA and IDEAL was performed by logistic regression using SPSS 20.0 (IBM, Chicago, IL, USA). Atopic co-morbidities (asthma, eczema, and rhinitis) were used in the model as confounders. FBAT software version 2.0.4 was used for family-based analyses under the additive model and for testing for Mendelian errors. All analyses were performed for food allergy in general and for peanut separately. For detailed information regarding cases and controls, see Table E2. Complete trios (father, mother and child) with a non-western ethnicity were only included in FBAT analysis. For all tests, a two-tailed significance level of $p < 0.05$ was used unless otherwise specified. QC was performed using Haploview (https://www.broadinstitute.org/haploview/haploview), removing SNPs with a call rate below 90% or deviating from HWE. Samples were excluded from analyses if the call rate was ≤50% or if trios had ≥ 2 Mendelian errors.[E12]

References:

E1. Asai Y, Greenwood C, Hull PR, Alizadehfar R, Ben-Shoshan M, Brown SJ, et al. Filaggrin gene mutation associations with peanut allergy persist despite variations in peanut allergy diagnostic criteria or asthma status. J Allergy Clin Immunol 2013; 132:239-42.

E2. James AL, Knuiman MW, Divitini ML, Hui J, Hunter M, Palmer LJ, et al. Changes in the prevalence of asthma in adults since 1966: the Busselton health study. Eur Respir J 2010; 35:273-8.

E3. 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. Am J Hum Genet 2007; 81:559-75.

E4. Delaneau O, Marchini J, Genomes Project C, Genomes Project C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun 2014; 5:3934.

E5. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009; 5:e1000529.
E6. Hong X, Hao K, Ladd-Acosta C, Hansen KD, Tsai HJ, Liu X, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. Nat Commun 2015; 6:6304.
E7. Martino DJ, Ashley S, Koplin J, Ellis J, Saffery R, Dharmage SC, et al. Genome-wide association study of peanut allergy reproduces association with amino acid polymorphisms in HLA-DRB1. Clin Exp Allergy 2016.
E8. Banda Y, Kvale MN, Hoffmann TJ, Hesselson SE, Ranatunga D, Tang H, et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics 2015; 200:1285-95.
E9. Hoffmann TJ, Kvale MN, Hesselson SE, Zhan Y, Aquino C, Cao Y, et al. Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. Genomics 2011; 98:79-89.
E10. Kvale MN, Hesselson S, Hoffmann TJ, Cao Y, Chan D, Connell S, et al. Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. Genetics 2015; 200:1051-60.
E11. van der Valk JP, Gerth van Wijk R, Dubois AE, de Groot H, Reitsma M, Vlieg-Boerstra B, et al. Multicentre Double-Blind Placebo-Controlled Food Challenge Study in Children Sensitised to Cashew Nut. PLoS One 2016; 11:e0151055.
E12. van Ginkel CD, Flokstra-de Blok BM, Kollen BJ, Kukler J, Koppelman GH, Dubois AE. Loss-of-function variants of the filaggrin gene are associated with clinical reactivity to foods. Allergy 2015; 70:461-4.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Online Repository Tables and Figure Legends for: 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

Table E1: Case definition of peanut allergy in the Canadian Peanut Allergy Registry

	History		**Skin prick test to peanut**		**Peanut-specific IgE**	**# of subjects (N=850)**
1	Oral food challenge					27
2	A convincing history of an allergic reaction*	**AND**	≥ 3 mm	**AND**	≥ 0.35 kU/L	314
3	A convincing history of an allergic reaction*	**AND**	≥ 3 mm			354
4	A convincing history of an allergic reaction*	**AND**			≥ 0.35 kU/L	45
5	Uncertain history of an allergic reaction&	**AND**	≥ 3 mm	**AND**	≥ 15 kU/L	46
6	No history of a reaction^	**AND**	≥ 3 mm	**AND**	≥ 15 kU/L	64

*Convincing history includes: 2 mild symptoms or signs, **OR** 1 moderate or 1 severe symptom or sign, **AND** occurring within 120 minutes after known peanut contact or ingestion; & Uncertain history includes: 1 mild symptom or sign occurring within 120 minutes after known peanut contact or ingestion, **OR** 1 moderate or 1 severe symptom or sign but lacking information on time or mode of peanut contact; ^No history of a reaction; individuals were advised to avoid peanut due to testing **AND/OR** affected sibling, **OR** have no history of peanut exposure. A mild reaction was defined by pruritus, itchy throat, urticaria, flushing, or rhinoconjunctivitis. A moderate reaction was defined by angioedema, coughing, stridor, tight throat, voice change, nausea, abdominal pain, vomiting, or difficulty breathing. A severe reaction was defined by wheezing, cyanosis, or circulatory collapse.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table E2: Summary of case and control groups

Location	Cases	Controls	Study design	Age	Ethnicity	Genotyping chip
Canadian Peanut Allergy Registry (CanPAR) (PA: N=1,776, FA: N=1,776)	PA (N=987, N=850 after quality control), see Table E1 for inclusion criteria. Self-reported FA and PA (N=636 after quality control)	Controls for PA and FA (N=987; 926 after quality control): non-atopic, no history of asthma, airway hyper-responsiveness, atopy, eczema, allergic rhinitis or FA (Australia, Busselton Health Study)	Case-control and related analysis, recruited on PA	Cases: Mean age in years (SD): 12(6) Range: 1-63years Hyper-controls: Mean age in years (SD): 49(25) Range : 6-93years	Self-identified Caucasian	Illumina Omni 2.5M+Exome 8v1.1
American (Chicago Food Study)[E6] (PA: N=2,197, FA: N=2,197)	PA (N=316), convincing history of clinical allergic reaction on ingestion to peanut and evidence of sensitization to peanut (peanut-specific IgE≥0.10kU/L and/or a positive SPT ≥3 mm); food allergy (N=671)	Controls for PA (N=1,881): 144 non-allergic non-sensitized controls and 1737 controls of uncertain phenotypes Controls for FA (N=1,526): 144 non-allergic non-sensitized controls and 1,382 controls of uncertain phenotypes	Case-control, and family-based, recruited on food allergy	Mean age in years (SD): European: 5.8 (3.8) Non-European: 6.0 (3.8) Age range of children in discovery: 0-21 years	Multi-ethnic	Illumina HumanOmni1-Quad BeadChip
American GERA cohort (Genetic Epidemiology Research on Aging)[E8-10] (FA: N= 29,053)	Total cohort members: N=103,067 Data used for the analysis is a nested case-control Cases: Self-reported food allergy in medical chart: (N=5,108)	Controls for FA (N= 23,945): Controls without self-reported FA	Cohort study	Age range: 18 to >100 years, average age 63 years	Multi-ethnic	Affymetrix Axiom with 674,517 SNPs
German (Understanding Food allergy)[E7] ** (PA and FA: N=2,592)	PA (N=205), +DBPCFC or a history of a severe allergic reaction to peanuts plus specific sensitization to peanut protein (>0.35 kUl^{-1}); food allergy (N=205)	Controls for PA and FA (N= 2,387): unrelated individuals from the German population based Heinz Nixdorf Recall Study [E12]	Case-control recruited on food allergy	Cases: No information available Controls: adults aged 45-75 years	Self-reported German ancestry	HumanOmniExpressExome-8 v1.2 (Cases) HumanOmniExpress-12 v1.1 plus HumanExome-12 v1 or HumanOmni1M-4 v1 plus HumanExome-12 v1. (Controls)
Australian (Health Nuts	PA (N=73), + oral challenge to peanut or parent report of	Controls for PA and FA (N=148): SPT negative to	Case-control, recruited on PA	Mean age in months (SD):	Multi-ethnic	Illumina Omni 2.5+8

ACCEPTED MANUSCRIPT

Study) (PA and FA: N=221)[E7]	clear history of immediate reaction in the past 2 months (N=3); food allergy (N=73)	egg white, peanut, sesame, shrimp, cow's milk, cashew, almond, hazelnut, soy and wheat, with negative oral challenge to peanut	and egg allergy	PA cases: 12.8 (0.8) Non-atopic controls 12.6 (0.67)		
Dutch IDEAL study [$] (PA: N=27, FA: N=138)	PA (N=13): +DBPCFC to peanut FA in general (N=115): +DBPCFC to any food; peanut (N=13), cashew (N=111), cow's milk (N=4), hazelnut (N=5), hens egg (N=7), walnut (N=2) and sesame seed (N=1). Some subjects have a +DBPCFC for multiple foods.	Controls for PA (N=14): -DBPCFC to peanut Controls for FA (N=23): only -DBPCFC to any food including: peanut (N=14), cashew (N=23), hazelnut (N=1) and walnut (N=1). Some subjects have a -DBPCFC for multiple foods	Case-control, recruited on cashew allergy	Age range: peanut-tested children: 1-18years Mean (SD) age in years: peanut-tested children: 9(7) years	Self-reported Western ancestry	Competitive allele-specific PCR using KASPar genotyping chemistry, under contract by LGC Genomic (LGC, Teddington, UK).
Dutch GENEVA study [$] (PA: N=215, FA: N=375)	PA (N=125): +DBPCFC to peanut FA in general (N=246): +DBPCFC to any food; peanut (N=125), cashew (N=22), cow's milk (N=68), hazelnut (N=39), hens egg (N=43), walnut (N=19), soy (N=10), almond (N=1), buckwheat (N=1), sesame seed (N=2), pine nut (N=2) and brazil nut (N=1). Some subjects have a +DBPCFC for multiple foods. There were no + DBPCFCs for wheat, lupin seed, pistachio and macadamia.	Controls for PA (N=90): -DBPCFC to peanut Controls for FA (N=129): -DBPCFC to any food including: peanut (N=90), cashew (N=5), cow's milk (N=45), hazelnut (N=12), hens egg (N=27), walnut (N=6), soy (N=3), almond (N=2), wheat (N=3), lupin seed (N=2), pistachio (N=2) and macadamia (N=1). Some subjects have a -DBPCFC for multiple foods	Family-based, recruited based on any food allergy	Age range: cashew-tested children: 2-18years Mean (SD) age in years: cashew-tested children: 9(9)	Multi-ethnic	Competitive allele-specific PCR using KASPar genotyping chemistry, under contract by LGC Genomic (LGC, Teddington, UK).

****** Only previously published data were available from this study for the following SNPs: rs57144668, rs4240433, rs6928827, rs73220497, rs16870788, rs6763069, rs10018666, rs2439871, rs864481, rs10812871, rs73971133, rs6686894, rs12142904, rs7131777, rs10474468, rs8077351, rs11700330, rs6584390, rs7300806, rs9362681, rs17555239

PA: peanut allergy; FA: food allergy, DBPCFC: double-blind, placebo-controlled food challenge, SPT: skin prick test

[$] for the case-control analyses, members from both the IDEAL and GENEVA were used

Table E3: Canadian Peanut Allergy Registry (CanPAR) related analysis, genome-wide associations (n=61) with p-value ≤1.49x10^{-6*}

							Related analysis (850 cases, 926 controls)				Unrelated analysis (834 cases, 781 controls)				
SNP	Chr	Position	r2$^{\#}$	Allele	MAF'	Source of SNPs	OR	LCI	UCI	P	OR	LCI	UCI	P	Gene/Nearest Gene
rs115218289	2	173265750	-	A/C	0.02	Imputed	0.18	0.1	0.32	1.80x10^{-8}	0.2	0.09	0.46	1.39x10^{-4}	(298kb)*DLX2* | (26kb)*ITGA6*
rs744597	4	86337028	-	A/G	0.4	Genotyped	0.61	0.5	0.74	3.98x10^{-7}	0.63	0.52	0.77	3.91x10^{-6}	*ARHGAP24*
rs78048444	7	132832218	-	C/T	0.02	Genotyped	0.22	0.12	0.39	5.44x10^{-7}	0.23	0.11	0.46	4.57x10^{-5}	(65kb)*CHCHD3* | (106kb)*EXOC4*
rs7475217	10	68444013	-	T/C	0.38	Genotyped	1.64	1.35	1.98	3.58x10^{-7}	1.56	1.28	1.9	9.19x10^{-6}	*CTNNA3*
rs71193762	10	68437687	0.9961	A/G	0.38	Imputed	1.63	1.35	1.98	3.77x10^{-7}	1.56	1.28	1.9	9.42x10^{-6}	*CTNNA3*
rs2394283	10	68432211	0.9967	G/T	0.38	Imputed	1.64	1.36	1.99	4.42x10^{-7}	1.56	1.28	1.91	1.08x10^{-5}	*CTNNA3*
rs12779828	10	68437537	0.9973	C/T	0.38	Imputed	1.63	1.35	1.97	4.66x10^{-7}	1.55	1.28	1.89	1.08x10^{-5}	*CTNNA3*
rs2394298	10	68462751	0.9970	G/A	0.38	Imputed	1.62	1.34	1.96	5.73x10^{-7}	1.55	1.27	1.88	1.42x10^{-5}	*CTNNA3*
rs34018067	10	68447638	0.9998	A/G	0.38	Imputed	1.62	1.34	1.96	5.73x10^{-7}	1.54	1.27	1.88	1.42x10^{-5}	*CTNNA3*
rs7911791	10	68459353	0.9997	A/T	0.38	Imputed	1.62	1.34	1.96	6.04x10^{-7}	1.54	1.27	1.88	1.42x10^{-5}	*CTNNA3*
rs11815638	10	68450128	0.9997	G/A	0.38	Imputed	1.62	1.34	1.96	6.04x10^{-7}	1.54	1.27	1.88	1.49x10^{-5}	*CTNNA3*
rs11816216	10	68451379	0.9997	T/A	0.38	Imputed	1.62	1.34	1.96	6.04x10^{-7}	1.54	1.27	1.88	1.49x10^{-5}	*CTNNA3*
rs61866029	10	68455505	0.9997	C/T	0.38	Imputed	1.62	1.34	1.96	6.04x10^{-7}	1.54	1.27	1.88	1.49x10^{-5}	*CTNNA3*
rs7910177	10	68450290	0.9997	T/G	0.38	Imputed	1.62	1.34	1.96	6.04x10^{-7}	1.54	1.27	1.88	1.49x10^{-5}	*CTNNA3*
rs7894410	10	68466106	0.9959	G/A	0.38	Imputed	1.62	1.34	1.95	7.42x10^{-7}	1.54	1.26	1.88	1.71x10^{-5}	*CTNNA3*
rs748704	10	68481094	0.9285	C/T	0.37	Imputed	1.6	1.32	1.93	1.06x10^{-6}	1.52	1.25	1.86	3.05x10^{-5}	*CTNNA3*
rs144897250	11	102750264	-	A/C	0.02	Imputed	6.2	3.09	12.45	2.90x10^{-7}	6.72	2.72	16.64	3.79x10^{-5}	(5kb)*MMP12* | (63kb)*MMP13*
rs7936434	11	76293805	-	C/G	0.49	Imputed	1.58	1.32	1.9	5.17x10^{-7}	1.58	1.31	1.91	2.73x10^{-6}	(30kb)*C11orf30* | (43kb)*LOC101928813*
rs7931483	11	76302067	0.9831	A/C	0.49	Imputed	1.58	1.32	1.89	5.44x10^{-7}	1.58	1.3	1.91	3.16x10^{-6}	(38kb)*C11orf30* | (35kb)*LOC101928813*
rs7936070	11	76293527	0.9933	T/G	0.49	Imputed	1.57	1.31	1.88	7.81x10^{-7}	1.57	1.3	1.9	4.03x10^{-6}	(30kb)*C11orf30* | (43kb)*LOC101928813*
rs7936312	11	76293726	0.9933	T/G	0.49	Imputed	1.57	1.31	1.88	7.81x10^{-7}	1.57	1.3	1.9	4.03x10^{-6}	(30kb)*C11orf30* | (43kb)*LOC101928813*
rs7936323	11	76293758	0.9933	A/G	0.49	Imputed	1.57	1.31	1.88	7.81x10^{-7}	1.57	1.3	1.9	4.03x10^{-6}	(30kb)*C11orf30* | (43kb)*LOC101928813*
rs72827854	17	46460525	-	T/C	0.09	Imputed	2.16	1.61	2.9	2.60x10^{-7}	2.08	1.5	2.87	9.00x10^{-6}	*SKAP1*
rs200314279	17	46446542	0.9404	A/T	0.1	Imputed	2.06	1.55	2.74	6.36x10^{-7}	1.94	1.42	2.66	3.32x10^{-5}	*SKAP1*

rs4491576	17	46408636	0.9513	T/A	0.1	Imputed	2.09	1.56	2.79	$6.70x10^{-7}$	2.03	1.48	2.78	$1.03x10^{-5}$	*SKAP1*
rs17623518	17	46438150	0.9674	A/G	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs17695179	17	46370309	0.9674	T/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs111396120	17	46445042	0.9674	A/G	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs143328356	17	46427578	0.9674	TA/T	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs17623125	17	46410661	0.9674	G/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs17623416	17	46430565	0.9674	C/A	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs2175156	17	46442786	0.9674	G/A	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs56151068	17	46381431	0.9674	T/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1;LOC101927148*
rs72827825	17	46392352	0.9674	T/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs72827848	17	46445747	0.9674	A/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs16956001	17	46451808	0.9674	G/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs72827851	17	46452495	0.9674	A/C	0.1	Imputed	2.06	1.54	2.76	$9.58x10^{-7}$	1.97	1.44	2.7	$2.34x10^{-5}$	*SKAP1*
rs16955960	17	46445993	0.9674	G/T	0.1	Genotyped	2.06	1.54	2.75	$1.01x10^{-6}$	1.97	1.44	2.7	$2.40x10^{-5}$	*SKAP1*
rs56093336	17	46468579	0.9675	T/C	0.09	Imputed	2.06	1.54	2.75	$1.01x10^{-6}$	1.97	1.44	2.7	$2.44x10^{-5}$	*SKAP1*
rs143892933	17	46478338	0.9674	A/AT	0.09	Imputed	2.06	1.54	2.75	$1.06x10^{-6}$	1.97	1.44	2.7	$2.55x10^{-5}$	*SKAP1*
rs16956501	17	46497274	0.9670	C/G	0.09	Imputed	2.06	1.54	2.75	$1.12x10^{-6}$	1.97	1.43	2.69	$2.67x10^{-5}$	*SKAP1*
rs55912545	17	46487324	0.9672	C/A	0.09	Imputed	2.06	1.54	2.75	$1.12x10^{-6}$	1.97	1.44	2.7	$2.55x10^{-5}$	*SKAP1*
rs55641965	17	46487300	0.9672	G/T	0.09	Imputed	2.06	1.54	2.75	$1.12x10^{-6}$	1.97	1.44	2.7	$2.55x10^{-5}$	*SKAP1*
rs72827805	17	46349432	0.9271	C/T	0.1	Imputed	2.04	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs112147946	17	46350146	0.9273	C/T	0.1	Imputed	2.05	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs72827806	17	46351066	0.9273	A/C	0.1	Imputed	2.05	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs17694670	17	46356471	0.9274	T/C	0.1	Imputed	2.05	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs72827810	17	46355550	0.9274	G/A	0.1	Imputed	2.05	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs17621689	17	46358952	0.9274	G/T	0.1	Imputed	2.05	1.53	2.73	$1.23x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs16957085	17	46532956	0.9410	A/G	0.09	Imputed	2.06	1.54	2.76	$1.23x10^{-6}$	1.97	1.44	2.71	$2.79x10^{-5}$	*LOC101927166*
rs56311919	17	46528350	0.9411	T/C	0.09	Imputed	2.06	1.54	2.77	$1.23x10^{-6}$	1.97	1.44	2.71	$2.92x10^{-5}$	*LOC101927166*
rs55765969	17	46528015	0.9411	T/C	0.09	Imputed	2.06	1.54	2.76	$1.23x10^{-6}$	1.97	1.44	2.71	$2.92x10^{-5}$	*LOC101927166*

ACCEPTED MANUSCRIPT

rs139462954	17	46523678	0.9433	A/AC	0.09	Imputed	2.06	1.54	2.76	$1.23x10^{-6}$	1.97	1.43	2.71	$2.92x10^{-5}$	*LOC101927166*
rs138636532	17	46386930	0.9642	G/GAT	0.09	Imputed	2.06	1.54	2.75	$1.23x10^{-6}$	1.97	1.43	2.7	$2.79x10^{-5}$	*SKAP1*
rs17694404	17	46348354	0.9268	A/G	0.1	Imputed	2.04	1.53	2.73	$1.30x10^{-6}$	1.95	1.43	2.66	$2.67x10^{-5}$	*SKAP1*
rs147813436	17	46344017	0.9268	T/C	0.1	Imputed	2.04	1.53	2.73	$1.30x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs17694168	17	46340934	0.9269	A/G	0.1	Imputed	2.04	1.53	2.73	$1.30x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs17694092	17	46334294	0.9269	G/A	0.1	Imputed	2.04	1.53	2.73	$1.30x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs72825596	17	46334340	0.9269	A/G	0.1	Imputed	2.04	1.53	2.73	$1.30x10^{-6}$	1.95	1.43	2.66	$2.55x10^{-5}$	*SKAP1*
rs10514944	17	46398352	0.9635	A/G	0.09	Imputed	2.05	1.53	2.74	$1.37x10^{-6}$	1.96	1.43	2.69	$2.92x10^{-5}$	*SKAP1*
rs523865	20	894881	-	C/T	0.23	Genotyped	0.57	0.46	0.71	$4.42x10^{-7}$	0.57	0.45	0.71	$1.19x10^{-6}$	*ANGPT4*

**p*-value obtained in related analysis. [∀]Table does not include SNPs in *HLA*. ɜ taken from related analyses. # r2 indicates LD with the most significant SNP for each locus, which is always the first entry. Chr: chromosome; Allele: minor allele/major allele; MAF: minor allele frequency; OR: odds ratio; LCI: lower 95% confidence interval; UCI: upper 95% confidence interval; P: *p*-value

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table E4: SNPs with *p*-values ≤1.49x10^{-6} after conditioning on SNP rs3134976 in the related analysis, ordered by the conditioning *p*-value.

				Related analysis (850 cases, 926 controls)				Related analysis conditioned on rs3134976 (850 cases, 926 controls)				
SNP	CHR	Position	Allele	OR	LCI	UCI	P	OR	LCI	UCI	P	Gene/Nearest Gene
rs115218289	2	173265750	A/C	0.18	0.10	0.32	1.80x10^{-8}	0.16	0.08	0.31	3.59x10^{-8}	(298kb)*DLX2*|(26kb)*ITGA6*
rs72827854	17	46460525	T/C	2.16	1.61	2.90	2.60x10^{-7}	2.29	1.69	3.08	6.30x10^{-8}	*SKAP1*
rs200314279	17	46446542	A/T	2.06	1.55	2.74	6.36x10^{-7}	2.18	1.63	2.92	1.70x10^{-7}	*SKAP1*
rs4491576	17	46408636	T/A	2.09	1.56	2.79	6.70x10^{-7}	2.19	1.63	2.95	2.22x10^{-7}	*SKAP1*
rs7475217	10	68444013	T/C	1.64	1.35	1.98	3.58x10^{-7}	1.65	1.36	1.99	2.75x10^{-7}	*CTNNA3*
rs112147946	17	46350146	C/T	2.05	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.91	3.06x10^{-7}	*SKAP1*
rs17621689	17	46358952	G/T	2.05	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.91	3.06x10^{-7}	*SKAP1*
rs17694670	17	46356471	T/C	2.05	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.91	3.06x10^{-7}	*SKAP1*
rs72827806	17	46351066	A/C	2.05	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.91	3.06x10^{-7}	*SKAP1*
rs72827810	17	46355550	G/A	2.05	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.91	3.06x10^{-7}	*SKAP1*
rs147813436	17	46344017	T/C	2.04	1.53	2.73	1.30x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs17694092	17	46334294	G/A	2.04	1.53	2.73	1.30x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs17694168	17	46340934	A/G	2.04	1.53	2.73	1.30x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs17694404	17	46348354	A/G	2.04	1.53	2.73	1.30x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs71193762	10	68437687	A/G	1.63	1.35	1.98	3.77x10^{-7}	1.64	1.36	1.98	3.22x10^{-7}	*CTNNA3*
rs72825596	17	46334340	A/G	2.04	1.53	2.73	1.30x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs72827805	17	46349432	C/T	2.04	1.53	2.73	1.23x10^{-6}	2.16	1.61	2.90	3.22x10^{-7}	*SKAP1*
rs111396120	17	46445042	A/G	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs143328356	17	46427578	TA/T	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs16956001	17	46451808	G/C	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs17623125	17	46410661	G/C	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs17623416	17	46430565	C/A	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs17623518	17	46438150	A/G	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs17695179	17	46370309	T/C	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*
rs2175156	17	46442786	G/A	2.06	1.54	2.76	9.58x10^{-7}	2.16	1.61	2.91	3.40x10^{-7}	*SKAP1*

ACCEPTED MANUSCRIPT

rs56151068	17	46381431	T/C	2.06	1.54	2.76	$9.58x10^{-7}$	2.16	1.61	2.91	$3.40x10^{-7}$	*SKAP1;LOC101927148*
rs72827825	17	46392352	T/C	2.06	1.54	2.76	$9.58x10^{-7}$	2.16	1.61	2.91	$3.40x10^{-7}$	*SKAP1*
rs72827848	17	46445747	A/C	2.06	1.54	2.76	$9.58x10^{-7}$	2.16	1.61	2.91	$3.40x10^{-7}$	*SKAP1*
rs72827851	17	46452495	A/C	2.06	1.54	2.76	$9.58x10^{-7}$	2.16	1.61	2.91	$3.40x10^{-7}$	*SKAP1*
rs16955960	17	46445993	G/T	2.06	1.54	2.75	$1.01x10^{-6}$	2.16	1.61	2.91	$3.40x10^{-7}$	SKAP1
rs56093336	17	46468579	T/C	2.06	1.54	2.75	$1.01x10^{-6}$	2.16	1.61	2.91	$3.58x10^{-7}$	*SKAP1*
rs744597	4	86337028	A/G	0.61	0.50	0.74	$3.98x10^{-7}$	0.60	0.49	0.73	$3.58x10^{-7}$	*ARHGAP24*
rs143892933	17	46478338	A/AT	2.06	1.54	2.75	$1.06x10^{-6}$	2.16	1.60	2.90	$3.77x10^{-7}$	*SKAP1*
rs2394283	10	68432211	G/T	1.64	1.36	1.99	$4.42x10^{-7}$	1.65	1.36	2.00	$3.77x10^{-7}$	*CTNNA3*
rs12779828	10	68437537	C/T	1.63	1.35	1.97	$4.66x10^{-7}$	1.63	1.35	1.98	$3.98x10^{-7}$	*CTNNA3*
rs16956501	17	46497274	C/G	2.06	1.54	2.75	$1.12x10^{-6}$	2.15	1.60	2.90	$3.98x10^{-7}$	*SKAP1*
rs35261737*	17	46333537	ATG/A	2.03	1.52	2.71	$1.67x10^{-6}$	2.15	1.60	2.89	$3.98x10^{-7}$	*SKAP1*
rs55641965	17	46487300	G/T	2.06	1.54	2.75	$1.12x10^{-6}$	2.16	1.60	2.90	$3.98x10^{-7}$	*SKAP1*
rs55912545	17	46487324	C/A	2.06	1.54	2.75	$1.12x10^{-6}$	2.16	1.60	2.90	$3.98x10^{-7}$	*SKAP1*
rs138636532	17	46386930	G/GAT	2.06	1.54	2.75	$1.23x10^{-6}$	2.15	1.60	2.90	$4.42x10^{-7}$	*SKAP1*
rs2394298	10	68462751	G/A	1.62	1.34	1.96	$5.73x10^{-7}$	1.63	1.35	1.97	$4.42x10^{-7}$	*CTNNA3*
rs11815638	10	68450128	G/A	1.62	1.34	1.96	$6.04x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs11816216	10	68451379	T/A	1.62	1.34	1.96	$6.04x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs34018067	10	68447638	A/G	1.62	1.34	1.96	$5.73x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs61866029	10	68455505	C/T	1.62	1.34	1.96	$6.04x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs7910177	10	68450290	T/G	1.62	1.34	1.96	$6.04x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs7911791	10	68459353	A/T	1.62	1.34	1.96	$6.04x10^{-7}$	1.63	1.35	1.97	$4.66x10^{-7}$	*CTNNA3*
rs72825595*	17	46330118	C/T	2.03	1.52	2.70	$1.51x10^{-6}$	2.13	1.59	2.86	$4.66x10^{-7}$	*SKAP1*
rs10514944	17	46398352	A/G	2.05	1.53	2.74	$1.37x10^{-6}$	2.14	1.59	2.89	$5.17x10^{-7}$	*SKAP1*
rs202084258	17	46369506	A/AAAG	2.00	1.50	2.67	$2.36x10^{-6}$	2.12	1.58	2.85	$5.17x10^{-7}$	*SKAP1*
rs7894410	10	68466106	G/A	1.62	1.34	1.95	$7.42x10^{-7}$	1.63	1.34	1.97	$5.44x10^{-7}$	*CTNNA3*
rs139462954	17	46523678	A/AC	2.06	1.54	2.76	$1.23x10^{-6}$	2.15	1.59	2.90	$5.73x10^{-7}$	*LOC101927166*
rs16957085	17	46532956	A/G	2.06	1.54	2.76	$1.23x10^{-6}$	2.15	1.59	2.90	$5.73x10^{-7}$	*LOC101927166*
rs55765969	17	46528015	T/C	2.06	1.54	2.76	$1.23x10^{-6}$	2.15	1.59	2.90	$5.73x10^{-7}$	*LOC101927166*
rs56311919	17	46528350	T/C	2.06	1.54	2.77	$1.23x10^{-6}$	2.15	1.59	2.90	$5.73x10^{-7}$	*LOC101927166*

ACCEPTED MANUSCRIPT

rs17620866*	17	46328398	G/C	2.03	1.51	2.72	$2.60x10^{-6}$	2.14	1.59	2.90	$6.36x10^{-7}$	*SKAP1*
rs748704	10	68481094	C/T	1.60	1.32	1.93	$1.06x10^{-6}$	1.61	1.34	1.94	$6.36x10^{-7}$	*CTNNA3*
rs2036520*	17	46328970	A/G	2.02	1.51	2.72	$2.60x10^{-6}$	2.14	1.59	2.89	$6.70x10^{-7}$	*SKAP1*
rs2036521*	17	46328976	G/T	2.02	1.51	2.72	$2.60x10^{-6}$	2.14	1.59	2.89	$6.70x10^{-7}$	*SKAP1*
rs2271264*	17	46507417	A/G	2.04	1.52	2.72	$1.51x10^{-6}$	2.12	1.58	2.86	$6.70x10^{-7}$	*SKAP1*
rs55751709*	17	46531175	C/G	2.04	1.52	2.73	$1.59x10^{-6}$	2.13	1.58	2.87	$6.70x10^{-7}$	*LOC101927166*
rs56371910*	17	46530731	T/C	2.04	1.52	2.73	$1.59x10^{-6}$	2.13	1.58	2.87	$6.70x10^{-7}$	*LOC101927166*
rs9907565*	17	46535345	G/A	2.03	1.52	2.72	$1.75x10^{-6}$	2.12	1.58	2.86	$7.05x10^{-7}$	*LOC101927166*
rs17693878*	17	46328110	G/A	2.02	1.51	2.71	$2.87x10^{-6}$	2.14	1.58	2.88	$7.42x10^{-7}$	*SKAP1*
rs56103412*	17	46542302	A/T	2.03	1.52	2.72	$1.75x10^{-6}$	2.12	1.58	2.86	$7.42x10^{-7}$	*LOC101927166*
rs9899997*	17	46542719	G/T	2.03	1.52	2.72	$1.75x10^{-6}$	2.12	1.58	2.86	$7.42x10^{-7}$	*LOC101927166*
rs78048444	7	132832218	C/T	0.22	0.12	0.39	$5.44x10^{-7}$	0.19	0.10	0.37	$7.81x10^{-7}$	(65kb)*CHCHD3*|(106kb)*EXOC4*
rs112560794*	17	46509723	A/G	2.02	1.51	2.70	$1.94x10^{-6}$	2.11	1.57	2.83	$8.65x10^{-7}$	(2kb)*SKAP1*|(12kb)*LOC101927166*
rs3809744*	17	46507918	C/G	2.02	1.51	2.70	$1.94x10^{-6}$	2.11	1.57	2.83	$8.65x10^{-7}$	(<1kb)*SKAP1*|(14kb)*LOC101927166*
rs72827869*	17	46506699	A/G	2.02	1.51	2.70	$1.94x10^{-6}$	2.11	1.57	2.83	$8.65x10^{-7}$	*SKAP1*
rs72827873*	17	46515784	C/T	2.02	1.51	2.70	$1.94x10^{-6}$	2.11	1.57	2.83	$8.65x10^{-7}$	(8kb)*SKAP1*|(6kb)*LOC101927166*
rs139902172*	10	68453196	T/C	1.60	1.32	1.94	$1.59x10^{-6}$	1.61	1.33	1.96	$1.30x10^{-6}$	*CTNNA3*
rs144897250	11	102750264	A/C	6.20	3.09	12.45	$2.90x10^{-7}$	5.86	2.86	12.03	$1.44x10^{-6}$	(5kb)*MMP12*|(63kb)*MMP13*

* SNPs are significant only after conditioning on rs3134976; Allele: minor allele/major allele; OR: odds ratio; LCI: lower 95% confidence interval; UCI: upper 95% confidence interval; P: *p*-value

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table E5: Meta-analysis of Canadian, American, Australian, German, and Dutch populations for association with peanut and food allergy phenotypes for previously published association SNPs

					Peanut allergy							Food allergy						
SNP†	Chr	Position	Allele	Source of SNPs	P-CanPAR*[a]	P – CFA[b]	P - HealthNuts*[c]	P-UFA*[d]	P - IDEAL/ GENEVA Case-control[e]	P - Dutch GENEVA Family -based[f]	Pmeta_PA[g]	P – CFA[h]	P - IDEAL/ GENEVA Case-control[i]	P – GENEVA Family study[j]	P - GERA[k]	Pmeta[l]	Pmeta (without GERA)[m]	Gene/Nearest Gene
rs862942	14	26492233	C/T	Hong *et al.*	$1.62x10^{-1}$	$3.00x10^{-6}$	$1.07x10^{-1}$	NA	NA	NA	$3.14x10^{-6}$	$2.70x10^{-4}$	NA	NA	$5.21x10^{-1}$	$4.64x10^{-2}$	$8.96x10^{-5}$	(973kb)*STXBP6*|(423kb)*NOVA1*
rs10878354	12	66384885	A/G	Hong *et al.*	$1.50x10^{-1}$	$5.10x10^{-6}$	$4.62x10^{-1}$	NA	NA	NA	$1.05x10^{-5}$	$3.10x10^{-2}$	NA	NA	$4.41x10^{-1}$	$9.55x10^{-2}$	$7.65x10^{-3}$	(25kb)*HMGA2*|(75kb)*RNA5SP362*
rs4584173	8	135336557	G/A	Hong *et al.*	$8.34x10^{-1}$	$3.60x10^{-6}$	$3.77x10^{-1}$	NA	NA	NA	$2.22x10^{-4}$	$3.10x10^{-2}$	NA	NA	$6.64x10^{-2}$	$1.68x10^{-2}$	$5.74x10^{-2}$	(422kb)*LOC101927822*|(153kb)*ZFAT*
rs57144668	4	186704292	T/C	Martino *et al.*	$2.26x10^{-1}$	$9.00x10^{-3}$	$3.37x10^{-5}$	$2.30x10^{-1}$	NA	NA	$3.24x10^{-4}$	NA	NA	NA	$1.39x10^{-1}$	$2.01x10^{-2}$	$1.03x10^{-2}$	*SORBS2*
rs4240433	9	132008809	G/A	Martino *et al.*	$4.88x10^{-2}$	$6.36x10^{-1}$	$6.62x10^{-6}$	$2.60x10^{-1}$	NA	NA	$5.33x10^{-3}$	$5.14x10^{-1}$	NA	NA	$6.09x10^{-1}$	$8.60x10^{-2}$	$3.87x10^{-3}$	(68kb)*IER5L*|(36kb)*LOC101929331*
rs6928827	6	144315219	A/G	Martino *et al.*	$1.68x10^{-1}$	$4.30x10^{-1}$	$9.54x10^{-8}$	$3.90x10^{-1}$	NA	NA	$8.07x10^{-3}$	$9.72x10^{-2}$	NA	NA	$1.38x10^{-1}$	$6.88x10^{-3}$	$1.67x10^{-3}$	*PLAGL1*
rs73220497	7	115842729	T/G	Martino *et al.*	$8.57x10^{-1}$	$5.84x10^{-1}$	$2.73x10^{-5}$	$3.00x10^{-2}$	NA	NA	$1.23x10^{-2}$	$5.31x10^{-1}$	NA	NA	$8.43x10^{-2}$	$7.75x10^{-3}$	$1.09x10^{-2}$	(43kb)*TFEC*|(8kb)*TES*
rs16870788	8	104883146	G/A	Martino *et al.*	$4.47x10^{-1}$	$7.04x10^{-1}$	$3.02x10^{-5}$	$9.00x10^{-2}$	NA	NA	$1.61x10^{-2}$	$5.26x10^{-1}$	NA	NA	$8.82x10^{-1}$	$2.14x10^{-1}$	$1.08x10^{-2}$	*RIMS2*
rs6763069	3	180686365	T/A	Martino *et al.*	$6.17x10^{-1}$	$4.20x10^{-1}$	$1.58x10^{-5}$	$1.50x10^{-1}$	NA	NA	$1.71x10^{-2}$	$7.91x10^{-1}$	NA	NA	$8.67x10^{-1}$	$2.92x10^{-1}$	$3.79x10^{-2}$	*FXR1*
rs10018666	4	10004805	C/T	Martino *et al.*	$5.22x10^{-1}$	$6.46x10^{-1}$	$3.68x10^{-8}$	$3.00x10^{-1}$	NA	NA	$2.62x10^{-2}$	$6.43x10^{-1}$	NA	NA	$8.51x10^{-1}$	$2.55x10^{-1}$	$2.61x10^{-2}$	*SLC2A9*
rs2439871	11	20123190	C/G	Martino *et al.*	$6.53x10^{-1}$	$2.10x10^{-1}$	$2.17x10^{-5}$	$4.90x10^{-1}$	NA	NA	$3.26x10^{-2}$	NA	NA	NA	$3.53x10^{-1}$	$1.33x10^{-1}$	$8.35x10^{-2}$	*NAV2*
rs864481	5	179411289	T/C	Martino *et al.*	$5.96x10^{-1}$	$6.92x10^{-2}$	$4.63x10^{-5}$	$8.80x10^{-1}$	NA	NA	$3.29x10^{-2}$	$6.62x10^{-2}$	NA	NA	$5.55x10^{-1}$	$1.43x10^{-1}$	$3.20x10^{-2}$	*RNF130*
rs10812871	9	28757900	A/G	Martino *et al.*	$2.67x10^{-1}$	$1.78x10^{-1}$	$3.84x10^{-5}$	$9.40x10^{-1}$	NA	NA	$3.37x10^{-2}$	$8.85x10^{-1}$	NA	NA	$7.45x10^{-1}$	$3.58x10^{-1}$	$1.50x10^{-1}$	*LINGO2*
rs73971133	18	76652861	T/C	Martino *et al.*	$5.96x10^{-1}$	$4.20x10^{-1}$	$2.60x10^{-5}$	$3.60x10^{-1}$	NA	NA	$3.99x10^{-2}$	$6.43x10^{-1}$	NA	NA	$1.64x10^{-1}$	$3.90x10^{-2}$	$6.30x10^{-2}$	(235kb)*LOC101928018*|(84kb)*LOC645321*
rs6686894	1	165082110	G/A	Martino *et al.*	$6.60x10^{-1}$	$2.53x10^{-1}$	$3.56x10^{-7}$	$7.50x10^{-1}$	NA	NA	$4.65x10^{-2}$	$4.76x10^{-1}$	NA	NA	$2.48x10^{-1}$	$7.17x10^{-2}$	$8.08x10^{-2}$	(261kb)*PBX1*|(89kb)*LMX1A*
rs12142904	1	192351266	G/A	Martino *et al.*	$2.46x10^{-1}$	$8.40x10^{-1}$	$5.00x10^{-6}$	$4.70x10^{-1}$	NA	NA	$4.79x10^{-2}$	NA	NA	NA	$6.60x10^{-1}$	$2.13x10^{-1}$	$2.34x10^{-2}$	(15kb)*RGS21*|(194kb)*RGS1*
rs7131777	12	21594028	T/C	Martino *et al.*	$1.21x10^{-1}$	$9.20x10^{-1}$	$4.16x10^{-5}$	$8.10x10^{-1}$	NA	NA	$8.21x10^{-2}$	$9.37x10^{-2}$	NA	NA	$6.64x10^{-2}$	$5.12x10^{-3}$	$8.41x10^{-3}$	*PYROXD1*
rs10474468	5	75659270	C/T	Martino *et al.*	$3.73x10^{-1}$	$8.60x10^{-1}$	$4.63x10^{-5}$	$4.80x10^{-1}$	NA	NA	$8.41x10^{-2}$	NA	NA	NA	$6.81x10^{-2}$	$1.53x10^{-2}$	$4.79x10^{-2}$	(10kb)*SV2C*|(40kb)*IQGAP2*
rs8077351	17	2545473	C/T	Martino *et al.*	$4.12x10^{-1}$	$9.46x10^{-1}$	$3.08x10^{-5}$	$4.20x10^{-1}$	NA	NA	$8.75x10^{-2}$	$4.09x10^{-1}$	NA	NA	$2.28x10^{-1}$	$4.38x10^{-2}$	$3.24x10^{-2}$	*PAFAH1B1*
rs11700330	20	52489424	T/C	Martino *et al.*	$9.52x10^{-1}$	$9.00x10^{-1}$	$2.75x10^{-6}$	$2.80x10^{-1}$	NA	NA	$1.06x10^{-1}$	NA	NA	NA	$8.41x10^{-2}$	$2.15x10^{-2}$	$6.04x10^{-2}$	(279kb)*ZNF217*|(71kb)*BCAS1*
rs6584390	10	102476167	C/T	Martino *et al.*	$5.55x10^{-1}$	$6.59x10^{-1}$	$3.76x10^{-5}$	$7.50x10^{-1}$	NA	NA	$1.35x10^{-1}$	$6.37x10^{-1}$	NA	NA	$5.35x10^{-1}$	$2.24x10^{-1}$	$1.31x10^{-1}$	(162kb)*HIF1AN*|(19kb)*PAX2*
rs7300806	12	68603179	A/G	Martino *et al.*	$8.26x10^{-1}$	$5.62x10^{-1}$	$1.24x10^{-5}$	$7.50x10^{-1}$	NA	NA	$1.53x10^{-1}$	$5.17x10^{-1}$	NA	NA	$7.66x10^{-1}$	$3.65x10^{-1}$	$1.42x10^{-1}$	*IL26*
rs9362681	6	90476452	C/T	Martino *et al.*	$8.81x10^{-1}$	$5.66x10^{-1}$	$1.36x10^{-5}$	$7.00x10^{-1}$	NA	NA	$1.54x10^{-1}$	$2.90x10^{-1}$	NA	NA	$7.46x10^{-2}$	$1.90x10^{-2}$	$8.88x10^{-2}$	*MDN1*

ACCEPTED MANUSCRIPT

rs17555239	15	25840403	T/C	Martino *et al.*	$7.57x10^{-1}$	$5.57x10^{-1}$	$2.78x10^{-5}$	$8.80x10^{-1}$	NA	NA	$1.79x10^{-1}$	$2.23x10^{-1}$	NA	NA	$3.74x10^{-1}$	$1.23x10^{-1}$	$8.88x10^{-2}$	(156kb)*UBE3A*|(83kb)*ATP10A*

† Replication SNPs from two previously published PA GWAS studies:3 SNPs from Table 1 from Hong *et al.*[19] and 21 SNPs from Supplementary Table 4 from Martino *et al.*[20]
Shaded rows indicate suggestive significance ($p \leq 1.49x10^{-6}$) in peanut allergy.
* Used for both - Peanut allergy (PA) and food allergy (FA)
[a]P-value from CanPAR(N=1,776), [b]P-value from Chicago FoodAllergy Study (N=2,197), [c]P-value from Australian HealthNutstudy (N=221), [d]P-value from German Understanding of Food Allergy study (N=2,592)
[e]Dutch IDEAL/GENEVA Case-control study
[f] Dutch GENEVA family study
[g]P-value from Stouffer's weighted z-score meta-analysis method for peanut allergy
[h]P-value from Chicago Food Allergy Study (N=2,197)
[i]Dutch IDEAL and GENEVA case-control study
[j]Dutch GENEVA family study
[k]P-value from GERA food allergy study (N=29,053)
[l]P-value from Stouffer's weighted z-score meta-analysis method for food allergy
[m]P-value from Stouffer's weighted z-score meta-analysis method for food allergy without GERA study
Chr: chromosome; Allele: minor allele/major allele; P: *p*-value

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table E6: Meta-analysis of Canadian, American, Australian, German, and Dutch populations for association with peanut and food allergy phenotypes, by peanut allergy meta-analysis *p*-value.

					Peanut allergy							Food allergy						
SNP†	Chr	Position	Allele	Source of SNPs	P-CanPAR*[a]	P – CFA[b]	P - HealthNuts*[c]	P-UFA*[d]	P - IDEAL/ GENEVA Case-control[e]	P - GENEVA Family-based[f]	Pmeta_PA[g]	P – CFA[h]	P - IDEAL/ GENEVA Case-control[i]	P - GENEVA Family - based[j]	P - GERA[k]	Pmeta[l]	Pmeta (without GERA)[m]	Gene/Nearest Gene
rs115218289	2	173265750	A/C	CanPAR - Imputed	$1.80x10^{-8}$	NA	$6.77x10^{-1}$	NA	$7.18x10^{-1}$	$3.54x10^{-1}$	$9.16x10^{-8}$	NA	$1.84x10^{-1}$	$1.90x10^{-1}$	$5.25x10^{-1}$	$2.91x10^{-2}$	$2.38x10^{-8}$	(298kb)*DLX2*|(26kb)*ITGA6*
rs523865	20	894881	C/T	CanPAR - Genotyped	$4.42x10^{-7}$	NA	NA	NA	$8.33x10^{-1}$	$1.63x10^{-2}$	$1.54x10^{-7}$	NA	$2.03x10^{-1}$	$2.60x10^{-2}$	$2.66x10^{-1}$	$9.29x10^{-3}$	$4.09x10^{-8}$	*ANGPT4*
rs144897250	11	102750264	A/C	CanPAR - Imputed	$2.90x10^{-7}$	NA	$3.84x10^{-1}$	NA	NA	NA	$2.94x10^{-7}$	NA	NA	NA	$5.89x10^{-1}$	$6.83x10^{-2}$	$2.94x10^{-7}$	(5kb)*MMP12*|(63kb)*MMP13*
rs7936434	11	76293805	C/G	CanPAR - Imputed	$5.17x10^{-7}$	$3.66x10^{-2}$	$1.43x10^{-1}$	NA	NA	NA	$3.13x10^{-7}$	$5.89x10^{-5}$	NA	NA	$4.13x10^{-4}$	$1.98x10^{-8}$	$7.50x10^{-11}$	(30kb)*C11orf30*|(43kb)*LOC101928813*
rs7936312	11	76293726	T/G	CanPAR - Imputed	$7.81x10^{-7}$	$3.66x10^{-2}$	$1.15x10^{-1}$	NA	NA	NA	$3.60x10^{-7}$	$5.89x10^{-5}$	NA	NA	$2.28x10^{-4}$	$9.00x10^{-9}$	$8.94x10^{-11}$	(30kb)*C11orf30*|(43kb)*LOC101928813*
rs7936323	11	76293758	A/G	CanPAR - Imputed	$7.81x10^{-7}$	$3.66x10^{-2}$	$1.20x10^{-1}$	NA	NA	NA	$3.69x10^{-7}$	$5.89x10^{-5}$	NA	NA	$2.38x10^{-4}$	$9.69x10^{-9}$	$9.22x10^{-11}$	(30kb)*C11orf30*|(43kb)*LOC101928813*
rs78048444	7	132832218	C/T	CanPAR - Genotyped	$5.44x10^{-7}$	NA	$2.13x10^{-1}$	NA	$6.46x10^{-1}$	$3.54x10^{-1}$	$3.73x10^{-7}$	NA	$8.29x10^{-1}$	$7.87x10^{-1}$	$6.97x10^{-1}$	$8.88x10^{-2}$	$2.53x10^{-6}$	(65kb)*CHCHD3*|(106kb)*EXOC4*
rs7936070	11	76293527	T/G	CanPAR - Imputed	$7.81x10^{-7}$	$3.57x10^{-2}$	$1.91x10^{-1}$	NA	NA	NA	$4.77x10^{-7}$	$6.12x10^{-5}$	NA	NA	$4.97x10^{-6}$	$5.86x10^{-11}$	$1.40x10^{-10}$	(30kb)*C11orf30*|(43kb)*LOC101928813*
rs744597	4	86337028	A/G	CanPAR - Genotyped	$3.98x10^{-7}$	$1.15x10^{-1}$	$5.57x10^{-1}$	NA	$8.20x10^{-2}$	$9.62x10^{-1}$	$1.63x10^{-6}$	$5.19x10^{-1}$	$1.13x10^{-1}$	$2.69x10^{-1}$	$3.63x10^{-1}$	$1.29x10^{-2}$	$1.42x10^{-5}$	*ARHGAP24*
rs862942	14	26492233	C/T	Hong *et al.*	$1.62x10^{-1}$	$3.00x10^{-6}$	$1.07x10^{-1}$	NA	NA	NA	$3.14x10^{-6}$	$2.70x10^{-4}$	NA	NA	$5.21x10^{-1}$	$4.64x10^{-2}$	$8.96x10^{-5}$	(973kb)*STXBP6*|(423kb)*NOVA1*
rs72827854	17	46460525	T/C	CanPAR - Imputed	$2.60x10^{-7}$	$1.36x10^{-1}$	$3.55x10^{-1}$	NA	NA	NA	$3.43x10^{-6}$	$5.83x10^{-1}$	NA	NA	$7.69x10^{-1}$	$9.27x10^{-2}$	$7.47x10^{-5}$	*SKAP1*
rs10878354	12	66384885	A/G	Hong *et al.*	$1.50x10^{-1}$	$5.10x10^{-6}$	$4.62x10^{-1}$	NA	NA	NA	$1.05x10^{-5}$	$3.10x10^{-2}$	NA	NA	$4.41x10^{-1}$	$9.55x10^{-2}$	$7.65x10^{-3}$	(25kb)*HMGA2*|(75kb)*RNA5SP362*
rs17695179	17	46370309	T/C	CanPAR - Imputed	$9.58x10^{-7}$	$1.66x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$1.07x10^{-5}$	$7.46x10^{-1}$	NA	NA	$4.98x10^{-1}$	$5.43x10^{-2}$	$2.78x10^{-4}$	*SKAP1*
rs55765969	17	46528015	T/C	CanPAR - Imputed	$1.23x10^{-6}$	$1.97x10^{-1}$	$2.81x10^{-1}$	NA	NA	NA	$1.45x10^{-5}$	$6.59x10^{-1}$	NA	NA	$3.77x10^{-1}$	$3.18x10^{-2}$	$1.97x10^{-4}$	*LOC101927166*
rs56311919	17	46528350	T/C	CanPAR - Imputed	$1.23x10^{-6}$	$1.97x10^{-1}$	$2.81x10^{-1}$	NA	NA	NA	$1.45x10^{-5}$	$6.59x10^{-1}$	NA	NA	$3.78x10^{-1}$	$3.18x10^{-2}$	$1.97x10^{-4}$	*LOC101927166*
rs16957085	17	46532956	A/G	CanPAR - Imputed	$1.23x10^{-6}$	$1.97x10^{-1}$	$2.81x10^{-1}$	NA	NA	NA	$1.45x10^{-5}$	$6.59x10^{-1}$	NA	NA	$3.82x10^{-1}$	$3.24x10^{-2}$	$1.97x10^{-4}$	*LOC101927166*
rs16956501	17	46497274	C/G	CanPAR - Imputed	$1.12x10^{-6}$	$2.13x10^{-1}$	$2.78x10^{-1}$	NA	NA	NA	$1.56x10^{-5}$	$7.10x10^{-1}$	NA	NA	$4.02x10^{-1}$	$3.63x10^{-2}$	$2.27x10^{-4}$	*SKAP1*
rs56093336	17	46468579	T/C	CanPAR - Imputed	$1.01x10^{-6}$	$2.10x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$1.71x10^{-5}$	$7.01x10^{-1}$	NA	NA	$6.61x10^{-1}$	$8.64x10^{-2}$	$2.41x10^{-4}$	*SKAP1*
rs55641965	17	46487300	G/T	CanPAR - Imputed	$1.12x10^{-6}$	$2.10x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$1.81x10^{-5}$	$7.01x10^{-1}$	NA	NA	$4.06x10^{-1}$	$3.79x10^{-2}$	$2.53x10^{-4}$	*SKAP1*
rs55912545	17	46487324	C/A	CanPAR - Imputed	$1.12x10^{-6}$	$2.10x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$1.81x10^{-5}$	$7.01x10^{-1}$	NA	NA	$4.06x10^{-1}$	$3.79x10^{-2}$	$2.53x10^{-4}$	*SKAP1*
rs111396120	17	46445042	A/G	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.96x10^{-1}$	$5.34x10^{-2}$	$2.65x10^{-4}$	*SKAP1*
rs17623416	17	46430565	C/A	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.45x10^{-1}$	$4.45x10^{-2}$	$2.65x10^{-4}$	*SKAP1*
rs17623518	17	46438150	A/G	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.96x10^{-1}$	$5.33x10^{-2}$	$2.65x10^{-4}$	*SKAP1*
rs2175156	17	46442786	G/A	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.96x10^{-1}$	$5.34x10^{-2}$	$2.65x10^{-4}$	*SKAP1*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

rs72827848	17	46445747	A/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.95x10^{-1}$	$5.32x10^{-2}$	$2.65x10^{-4}$	*SKAP1*
rs17623125	17	46410661	G/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.32x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.37x10^{-1}$	$4.33x10^{-2}$	$2.65x10^{-4}$	*SKAP1*
rs72827825	17	46392352	T/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$2.33x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.34x10^{-1}$	$4.26x10^{-2}$	$2.66x10^{-4}$	*SKAP1*
rs16956001	17	46451808	G/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$2.33x10^{-5}$	$7.33x10^{-1}$	NA	NA	$6.41x10^{-1}$	$8.34x10^{-2}$	$2.66x10^{-4}$	*SKAP1*
rs72827851	17	46452495	A/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$2.33x10^{-5}$	$7.33x10^{-1}$	NA	NA	$6.41x10^{-1}$	$8.34x10^{-2}$	$2.66x10^{-4}$	*SKAP1*
rs56151068	17	46381431	T/C	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$2.33x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.45x10^{-1}$	$4.45x10^{-2}$	$2.66x10^{-4}$	*SKAP1;LOC101927148*
rs16955960	17	46445993	G/T	CanPAR - Genotyped	$1.01x10^{-6}$	$2.51x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$2.39x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.96x10^{-1}$	$5.37x10^{-2}$	$2.72x10^{-4}$	*SKAP1*
rs139462954	17	46523678	A/AC	CanPAR - Imputed	$1.23x10^{-6}$	$1.97x10^{-1}$	NA	NA	NA	NA	$2.66x10^{-5}$	$6.59x10^{-1}$	NA	NA	$3.77x10^{-1}$	$3.87x10^{-2}$	$3.56x10^{-4}$	*LOC101927166*
rs10514944	17	46398352	A/G	CanPAR - Imputed	$1.37x10^{-6}$	$2.51x10^{-1}$	$3.59x10^{-1}$	NA	NA	NA	$2.86x10^{-5}$	$7.33x10^{-1}$	NA	NA	$4.25x10^{-1}$	$4.29x10^{-2}$	$3.18x10^{-4}$	*SKAP1*
rs143892933	17	46478338	A/AT	CanPAR - Imputed	$1.06x10^{-6}$	$2.19x10^{-1}$	NA	NA	NA	NA	$2.96x10^{-5}$	$6.77x10^{-1}$	NA	NA	$5.01x10^{-1}$	$6.14x10^{-2}$	$3.54x10^{-4}$	*SKAP1*
rs143328356	17	46427578	TA/T	CanPAR - Imputed	$9.58x10^{-7}$	$2.51x10^{-1}$	NA	NA	NA	NA	$3.63x10^{-5}$	$7.33x10^{-1}$	NA	NA	$5.87x10^{-1}$	$8.30x10^{-2}$	$4.16x10^{-4}$	*SKAP1*
rs4491576	17	46408636	T/A	CanPAR - Imputed	$6.70x10^{-7}$	$3.21x10^{-1}$	$5.34x10^{-1}$	NA	NA	NA	$4.21x10^{-5}$	$8.54x10^{-1}$	NA	NA	$5.40x10^{-1}$	$6.88x10^{-2}$	$4.49x10^{-4}$	*SKAP1*
rs72827806	17	46351066	A/C	CanPAR - Imputed	$1.23x10^{-6}$	$3.24x10^{-1}$	$3.52x10^{-1}$	NA	NA	NA	$4.44x10^{-5}$	$7.62x10^{-1}$	NA	NA	$5.01x10^{-1}$	$5.69x10^{-2}$	$3.33x10^{-4}$	*SKAP1*
rs112147946	17	46350146	C/T	CanPAR - Imputed	$1.23x10^{-6}$	$3.24x10^{-1}$	$3.56x10^{-1}$	NA	NA	NA	$4.47x10^{-5}$	$7.62x10^{-1}$	NA	NA	$4.97x10^{-1}$	$5.62x10^{-2}$	$3.35x10^{-4}$	*SKAP1*
rs72827810	17	46355550	G/A	CanPAR - Imputed	$1.23x10^{-6}$	$3.24x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$4.47x10^{-5}$	$7.62x10^{-1}$	NA	NA	$4.92x10^{-1}$	$5.53x10^{-2}$	$3.35x10^{-4}$	*SKAP1*
rs17694670	17	46356471	T/C	CanPAR - Imputed	$1.23x10^{-6}$	$3.24x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$4.47x10^{-5}$	$7.62x10^{-1}$	NA	NA	$4.92x10^{-1}$	$5.53x10^{-2}$	$3.35x10^{-4}$	*SKAP1*
rs17621689	17	46358952	G/T	CanPAR - Imputed	$1.23x10^{-6}$	$3.24x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$4.47x10^{-5}$	$7.62x10^{-1}$	NA	NA	$4.92x10^{-1}$	$5.54x10^{-2}$	$3.35x10^{-4}$	*SKAP1*
rs17694092	17	46334294	G/A	CanPAR - Imputed	$1.30x10^{-6}$	$3.24x10^{-1}$	$3.48x10^{-1}$	NA	NA	NA	$4.53x10^{-5}$	$7.62x10^{-1}$	NA	NA	$5.01x10^{-1}$	$5.72x10^{-2}$	$3.39x10^{-4}$	*SKAP1*
rs72825596	17	46334340	A/G	CanPAR - Imputed	$1.30x10^{-6}$	$3.24x10^{-1}$	$3.48x10^{-1}$	NA	NA	NA	$4.53x10^{-5}$	$7.62x10^{-1}$	NA	NA	$5.10x10^{-1}$	$5.89x10^{-2}$	$3.39x10^{-4}$	*SKAP1*
rs17694168	17	46340934	A/G	CanPAR - Imputed	$1.30x10^{-6}$	$3.24x10^{-1}$	$3.50x10^{-1}$	NA	NA	NA	$4.54x10^{-5}$	$7.62x10^{-1}$	NA	NA	$5.17x10^{-1}$	$6.03x10^{-2}$	$3.39x10^{-4}$	*SKAP1*
rs147813436	17	46344017	T/C	CanPAR - Imputed	$1.30x10^{-6}$	$3.24x10^{-1}$	$3.57x10^{-1}$	NA	NA	NA	$4.60x10^{-5}$	$7.62x10^{-1}$	NA	NA	$5.57x10^{-1}$	$6.87x10^{-2}$	$3.43x10^{-4}$	*SKAP1*
rs17694404	17	46348354	A/G	CanPAR - Imputed	$1.30x10^{-6}$	$3.98x10^{-1}$	$3.26x10^{-1}$	NA	NA	NA	$6.70x10^{-5}$	$8.83x10^{-1}$	NA	NA	$4.93x10^{-1}$	$6.04x10^{-2}$	$4.99x10^{-4}$	*SKAP1*
rs72827805	17	46349432	C/T	CanPAR - Imputed	$1.23x10^{-6}$	$4.04x10^{-1}$	$3.55x10^{-1}$	NA	NA	NA	$7.10x10^{-5}$	$8.59x10^{-1}$	NA	NA	$5.59x10^{-1}$	$7.37x10^{-2}$	$4.70x10^{-4}$	*SKAP1*
rs4584173	8	135336557	G/A	Hong *et al.*	$8.34x10^{-1}$	$3.60x10^{-6}$	$3.77x10^{-1}$	NA	NA	NA	$2.22x10^{-4}$	$3.10x10^{-2}$	NA	NA	$6.64x10^{-2}$	$1.68x10^{-2}$	$5.74x10^{-2}$	(422kb)*LOC101927822*|(153kb)*ZFAT*
rs71193762	10	68437687	A/G	CanPAR - Imputed	$3.77x10^{-7}$	$8.63x10^{-1}$	$3.62x10^{-1}$	NA	NA	NA	$2.73x10^{-4}$	$6.87x10^{-1}$	NA	NA	$7.06x10^{-1}$	$8.83x10^{-2}$	$1.41x10^{-4}$	*CTNNA3*
rs2394283	10	68432211	G/T	CanPAR - Imputed	$4.42x10^{-7}$	$8.63x10^{-1}$	$3.31x10^{-1}$	NA	NA	NA	$2.79x10^{-4}$	$6.87x10^{-1}$	NA	NA	$7.51x10^{-1}$	$9.96x10^{-2}$	$1.44x10^{-4}$	*CTNNA3*
rs12779828	10	68437537	C/T	CanPAR - Imputed	$4.66x10^{-7}$	$8.63x10^{-1}$	$3.29x10^{-1}$	NA	NA	NA	$2.85x10^{-4}$	$6.87x10^{-1}$	NA	NA	$7.56x10^{-1}$	$1.01x10^{-1}$	$1.47x10^{-4}$	*CTNNA3*
rs57144668	4	186704292	T/C	Martino *et al.*	$2.26x10^{-1}$	$9.00x10^{-3}$	$3.37x10^{-5}$	$2.30x10^{-1}$	NA	NA	$3.24x10^{-4}$	NA	NA	NA	$1.39x10^{-1}$	$2.01x10^{-2}$	$1.03x10^{-2}$	*SORBS2*
rs748704	10	68481094	C/T	CanPAR - Imputed	$1.06x10^{-6}$	$8.23x10^{-1}$	$3.05x10^{-1}$	NA	NA	NA	$3.53x10^{-4}$	$8.10x10^{-1}$	NA	NA	$7.86x10^{-1}$	$1.27x10^{-1}$	3.37x10-4	*CTNNA3*
rs11815638	10	68450128	G/A	CanPAR - Imputed	$6.04x10^{-7}$	$8.58x10^{-1}$	$4.23x10^{-1}$	NA	NA	NA	$3.70x10^{-4}$	$8.41x10^{-1}$	NA	NA	$7.08x10^{-1}$	$1.05x10^{-1}$	3.48x10-4	*CTNNA3*

ACCEPTED MANUSCRIPT

rs7475217	10	68444013	T/C	CanPAR - Genotyped	3.58x10-7	$9.75x10^{-1}$	$3.28x10^{-1}$	NA	NA	NA	$3.72x10^{-4}$	$6.62x10^{-1}$	NA	NA	$7.73x10^{-1}$	$1.01x10^{-1}$	1.16x10-4	*CTNNA3*
rs34018067	10	68447638	A/G	CanPAR - Imputed	5.73x10-7	$9.63x10^{-1}$	$3.64x10^{-1}$	NA	NA	NA	$4.73x10^{-4}$	$6.59x10^{-1}$	NA	NA	$7.27x10^{-1}$	$9.50x10^{-2}$	$1.56x10^{-4}$	*CTNNA3*
rs11816216	10	68451379	T/A	CanPAR - Imputed	6.04x10-7	$9.32x10^{-1}$	$4.34x10^{-1}$	NA	NA	NA	$4.86x10^{-4}$	$5.82x10^{-1}$	NA	NA	$7.28x10^{-1}$	$9.22x10^{-2}$	$1.30x10^{-4}$	*CTNNA3*
rs7910177	10	68450290	T/G	CanPAR - Imputed	6.04x10-7	$9.78x10^{-1}$	$3.39x10^{-1}$	NA	NA	NA	$4.89x10^{-4}$	$6.07x10^{-1}$	NA	NA	$6.65x10^{-1}$	$7.70x10^{-2}$	$1.24x10^{-4}$	*CTNNA3*
rs2394298	10	68462751	G/A	CanPAR - Imputed	5.73x10-7	$9.78x10^{-1}$	$3.69x10^{-1}$	NA	NA	NA	$5.01x10^{-4}$	$6.07x10^{-1}$	NA	NA	$7.68x10^{-1}$	$1.02x10^{-1}$	$1.27x10^{-4}$	*CTNNA3*
rs7911791	10	68459353	A/T	CanPAR - Imputed	6.04x10-7	$9.78x10^{-1}$	$3.66x10^{-1}$	NA	NA	NA	$5.11x10^{-4}$	$6.07x10^{-1}$	NA	NA	$7.81x10^{-1}$	$1.05x10^{-1}$	$1.30x10^{-4}$	*CTNNA3*
rs61866029	10	68455505	C/T	CanPAR - Imputed	6.04x10-7	$9.78x10^{-1}$	$3.68x10^{-1}$	NA	NA	NA	$5.13x10^{-4}$	$6.07x10^{-1}$	NA	NA	$7.74x10^{-1}$	$1.04x10^{-1}$	$1.30x10^{-4}$	*CTNNA3*
rs7894410	10	68466106	G/A	CanPAR - Imputed	7.42x10-7	$9.78x10^{-1}$	$3.84x10^{-1}$	NA	NA	NA	$5.79x10^{-4}$	$6.07x10^{-1}$	NA	NA	$7.93x10^{-1}$	$1.11x10^{-1}$	$1.49x10^{-4}$	*CTNNA3*
rs4240433	9	132008809	G/A	Martino *et al.*	$4.88x10^{-2}$	$6.36x10^{-1}$	$6.62x10^{-6}$	$2.60x10^{-1}$	NA	NA	$5.33x10^{-3}$	$5.14x10^{-1}$	NA	NA	$6.09x10^{-1}$	$8.60x10^{-2}$	$3.87x10^{-3}$	(68kb)*IER5L*|(36kb)*LOC101929331*
rs6928827	6	144315219	A/G	Martino *et al.*	$1.68x10^{-1}$	$4.30x10^{-1}$	$9.54x10^{-8}$	$3.90x10^{-1}$	NA	NA	$8.07x10^{-3}$	$9.72x10^{-2}$	NA	NA	$1.38x10^{-1}$	$6.88x10^{-3}$	$1.67x10^{-3}$	*PLAGL1*
rs73220497	7	115842729	T/G	Martino *et al.*	$8.57x10^{-1}$	$5.84x10^{-1}$	$2.73x10^{-5}$	$3.00x10^{-2}$	NA	NA	$1.23x10^{-2}$	$5.31x10^{-1}$	NA	NA	$8.43x10^{-2}$	$7.75x10^{-3}$	$1.09x10^{-2}$	(43kb)*TFEC*|(8kb)*TES*
rs16870788	8	104883146	G/A	Martino *et al.*	$4.47x10^{-1}$	$7.04x10^{-1}$	$3.02x10^{-5}$	$9.00x10^{-2}$	NA	NA	$1.61x10^{-2}$	$5.26x10^{-1}$	NA	NA	$8.82x10^{-1}$	$2.14x10^{-1}$	$1.08x10^{-2}$	*RIMS2*
rs6763069	3	180686365	T/A	Martino *et al.*	$6.17x10^{-1}$	$4.20x10^{-1}$	$1.58x10^{-5}$	$1.50x10^{-1}$	NA	NA	$1.71x10^{-2}$	$7.91x10^{-1}$	NA	NA	$8.67x10^{-1}$	$2.92x10^{-1}$	$3.79x10^{-2}$	*FXR1*
rs10018666	4	10004805	C/T	Martino *et al.*	$5.22x10^{-1}$	$6.46x10^{-1}$	$3.68x10^{-8}$	$3.00x10^{-1}$	NA	NA	$2.62x10^{-2}$	$6.43x10^{-1}$	NA	NA	$8.51x10^{-1}$	$2.55x10^{-1}$	$2.61x10^{-2}$	*SLC2A9*
rs2439871	11	20123190	C/G	Martino *et al.*	$6.53x10^{-1}$	$2.10x10^{-1}$	$2.17x10^{-5}$	$4.90x10^{-1}$	NA	NA	$3.26x10^{-2}$	NA	NA	NA	$3.53x10^{-1}$	$1.33x10^{-1}$	$8.35x10^{-2}$	*NAV2*
rs864481	5	179411289	T/C	Martino *et al.*	$5.96x10^{-1}$	$6.92x10^{-2}$	$4.63x10^{-5}$	$8.80x10^{-1}$	NA	NA	$3.29x10^{-2}$	$6.62x10^{-2}$	NA	NA	$5.55x10^{-1}$	$1.43x10^{-1}$	$3.20x10^{-2}$	*RNF130*
rs10812871	9	28757900	A/G	Martino *et al.*	$2.67x10^{-1}$	$1.78x10^{-1}$	$3.84x10^{-5}$	$9.40x10^{-1}$	NA	NA	$3.37x10^{-2}$	$8.85x10^{-1}$	NA	NA	$7.45x10^{-1}$	$3.58x10^{-1}$	$1.50x10^{-1}$	*LINGO2*
rs73971133	18	76652861	T/C	Martino *et al.*	$5.96x10^{-1}$	$4.20x10^{-1}$	$2.60x10^{-5}$	$3.60x10^{-1}$	NA	NA	$3.99x10^{-2}$	$6.43x10^{-1}$	NA	NA	$1.64x10^{-1}$	$3.90x10^{-2}$	$6.30x10^{-2}$	(235kb)*LOC101928018*|(84kb)*LOC645321*
rs6686894	1	165082110	G/A	Martino *et al.*	$6.60x10^{-1}$	$2.53x10^{-1}$	$3.56x10^{-7}$	$7.50x10^{-1}$	NA	NA	$4.65x10^{-2}$	$4.76x10^{-1}$	NA	NA	$2.48x10^{-1}$	$7.17x10^{-2}$	$8.08x10^{-2}$	(261kb)*PBX1*|(89kb)*LMX1A*
rs12142904	1	192351266	G/A	Martino *et al.*	$2.46x10^{-1}$	$8.40x10^{-1}$	$5.00x10^{-6}$	$4.70x10^{-1}$	NA	NA	$4.79x10^{-2}$	NA	NA	NA	$6.60x10^{-1}$	$2.13x10^{-1}$	$2.34x10^{-2}$	(15kb)*RGS21*|(194kb)*RGS1*
rs7131777	12	21594028	T/C	Martino *et al.*	$1.21x10^{-1}$	$9.20x10^{-1}$	$4.16x10^{-5}$	$8.10x10^{-1}$	NA	NA	$8.21x10^{-2}$	$9.37x10^{-2}$	NA	NA	$6.64x10^{-2}$	$5.12x10^{-3}$	$8.41x10^{-3}$	*PYROXD1*
rs10474468	5	75659270	C/T	Martino *et al.*	$3.73x10^{-1}$	$8.60x10^{-1}$	$4.63x10^{-5}$	$4.80x10^{-1}$	NA	NA	$8.41x10^{-2}$	NA	NA	NA	$6.81x10^{-2}$	$1.53x10^{-2}$	$4.79x10^{-2}$	(10kb)*SV2C*|(40kb)*IQGAP2*
rs8077351	17	2545473	C/T	Martino *et al.*	$4.12x10^{-1}$	$9.46x10^{-1}$	$3.08x10^{-5}$	$4.20x10^{-1}$	NA	NA	$8.75x10^{-2}$	$4.09x10^{-1}$	NA	NA	$2.28x10^{-1}$	$4.38x10^{-2}$	$3.24x10^{-2}$	*PAFAH1B1*
rs11700330	20	52489424	T/C	Martino *et al.*	$9.52x10^{-1}$	$9.00x10^{-1}$	$2.75x10^{-6}$	$2.80x10^{-1}$	NA	NA	$1.06x10^{-1}$	NA	NA	NA	$8.41x10^{-2}$	$2.15x10^{-2}$	$6.04x10^{-2}$	(279kb)*ZNF217*|(71kb)*BCAS1*
rs6584390	10	102476167	C/T	Martino *et al.*	$5.55x10^{-1}$	$6.59x10^{-1}$	$3.76x10^{-5}$	$7.50x10^{-1}$	NA	NA	$1.35x10^{-1}$	$6.37x10^{-1}$	NA	NA	$5.35x10^{-1}$	$2.24x10^{-1}$	$1.31x10^{-1}$	(162kb)*HIF1AN*|(19kb)*PAX2*
rs7300806	12	68603179	A/G	Martino *et al.*	$8.26x10^{-1}$	$5.62x10^{-1}$	$1.24x10^{-5}$	$7.50x10^{-1}$	NA	NA	$1.53x10^{-1}$	$5.17x10^{-1}$	NA	NA	$7.66x10^{-1}$	$3.65x10^{-1}$	$1.42x10^{-1}$	*IL26*
rs9362681	6	90476452	C/T	Martino *et al.*	$8.81x10^{-1}$	$5.66x10^{-1}$	$1.36x10^{-5}$	$7.00x10^{-1}$	NA	NA	$1.54x10^{-1}$	$2.90x10^{-1}$	NA	NA	$7.46x10^{-2}$	$1.90x10^{-2}$	$8.88x10^{-2}$	*MDN1*
rs17555239	15	25840403	T/C	Martino *et al.*	$7.57x10^{-1}$	$5.57x10^{-1}$	$2.78x10^{-5}$	$8.80x10^{-1}$	NA	NA	$1.79x10^{-1}$	$2.23x10^{-1}$	NA	NA	$3.74x10^{-1}$	$1.23x10^{-1}$	$8.88x10^{-2}$	(156kb)*UBE3A*|(83kb)*ATP10A*
rs138636532	17	46386930	G/GAT	CanPAR - Imputed	$1.23x10^{-6}$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	$1.23x10^{-6}$	*SKAP1*
rs200314279	17	46446542	A/T	CanPAR - Imputed	$6.36x10^{-7}$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	$6.36x10^{-7}$	*SKAP1*

ACCEPTED MANUSCRIPT

rs7931483	11	76302067	A/C	CanPAR - Imputed	$5.44x10^{-7}$	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	$5.44x10^{-7}$	(38kb)*C11orf30*|(35kb)*LOC101928813*

†This table includes imputed and genotyped SNPs in CanPAR GWAS with p-value $\leq 1.49x10^{-6}$ in peanut allergy analysis. Replication SNPs from two previously published food allergy GWAS studies are also included, 3 SNPs from Table 1 from Hong *et al*.[E6] and 21 SNPs from Supplementary Table 4 from Martino *et al*.[E7]

Shaded rows indicate suggestive significance ($p \leq 1.49x10^{-6}$) in peanut allergy.

* Used for both - Peanut allergy (PA) and food allergy (FA)

[a]P-value from CanPAR (N=1,776), [b]P-value from Chicago Food Allergy Study (N=2,197), [c]P-value from HealthNuts study (N=221), [d]P-value from Understanding Food Allergy peanut allergy study (N=2,593)

[e]P-value from Dutch peanut allergy case-control study (IDEAL/GENEVA), number of individuals for SNPs 226, 229, 227, and 217 for rs115218289, rs523865, rs78048444, and rs744597, respectively, corrected for atopic dermatitis (AD), Asthma (As), and rhinoconjunctivitis (RC)

[f]P-value from Dutch peanut allergy family study (GENEVA), number of informative families for SNPs 20, 112, 21, and 112 for rs115218289, rs523865, rs78048444, and rs744597, respectively

[g]P-value from Stouffer's weighted z-score meta-analysis method for peanut allergy

[h]P-value from Chicago Food Allergy Study (N=2,197)

[i]P-value from Dutch food allergy case-control study, number of individuals for SNPs 479, 487, 482, and 466 for rs115218289, rs523865, rs78048444, and rs744597, respectively, corrected for atopic dermatitis (AD), Asthma (As), and rhinoconjunctivitis (RC)

[j]P-value from Dutch food allergy family study (GENEVA), number of informative families for SNPs 26, 196, 37, and 214 for rs115218289, rs523865, rs78048444, and rs744597, respectively

[k]P-value from GERA food allergy study (N=29,053)

[l]P-value from Stouffer's weighted z-score meta-analysis method for food allergy

[m]P-value from Stouffer's weighted z-score meta-analysis method for food allergy without GERA study

Chr: chromosome; Allele: minor allele/major allele; P: *p*-value

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table E7: Canadian Peanut Allergy Registry SNPs associated with PA and Expression Quantitative Trait Loci (eQTL)

SNP	Chr	Position	Nearest gene	Tissue	*p*-value	Gene symbol
rs72827854	17	46460525	*SKAP1*	skin, sun exposed lower leg	4.66×10^{-10}	*SNX11*
rs4491576	17	46408636	*SKAP1*	skin, sun exposed lower leg	7.23×10^{-10}	*SNX11*
rs4491576	17	46408636	*SKAP1*	whole blood	1.05×10^{-10}	*SNX11*
rs111396120	17	46445042	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs143328356	17	46427578	*SKAP1*	skin, sun exposed lower leg	6.13×10^{-10}	*SNX11*
rs16956001	17	46451808	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs17623125	17	46410661	*SKAP1*	skin, sun exposed lower leg	4.79×10^{-10}	*SNX11*
rs17623416	17	46430565	*SKAP1*	skin, sun exposed lower leg	4.80×10^{-10}	*SNX11*
rs17623518	17	46438150	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs17696179	17	46370309	*SKAP1*	skin, sun exposed lower leg	6.14×10^{-10}	*SNX11*
rs2175156	17	46442786	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs56151068	17	46381431	*SKAP1*	skin, sun exposed lower leg	5.78×10^{-10}	*SNX11*
rs72827825	17	46392352	*SKAP1*	skin, sun exposed lower leg	5.21×10^{-10}	*SNX11*
rs72827848	17	46445747	*SKAP1*	skin, sun exposed lower leg	5.38×10^{-10}	*SNX11*
rs72827851	17	46452495	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs16955960	17	46445993	*SKAP1*	skin, sun exposed lower leg	4.76×10^{-10}	*SNX11*
rs56093336	17	46468579	*SKAP1*	skin, sun exposed lower leg	4.77×10^{-10}	*SNX11*
rs143892933	17	46478338	*SKAP1*	skin, sun exposed lower leg	1.07×10^{-9}	*SNX11*
rs16956501	17	46497274	*SKAP1*	skin, sun exposed lower leg	4.48×10^{-10}	*SNX11*
rs55641965	17	46487300	*SKAP1*	skin, sun exposed lower leg	4.81×10^{-10}	*SNX11*
rs55912545	17	46487324	*SKAP1*	skin, sun exposed lower leg	4.81×10^{-10}	*SNX11*
rs112147946	17	46350146	*SKAP1*	skin, sun exposed lower leg	2.16×10^{-9}	*SNX11*
rs139462954	17	46523678	LOC101927166	cells, transformed fibroblasts	7.36×10^{-7}	*CBX1*
rs139462954	17	46523678	LOC101927166	skin, sun exposed lower leg	8.94×10^{-10}	*SNX11*
rs16957085	17	46532956	LOC101927166	cells, transformed fibroblasts	7.36×10^{-7}	*CBX1*
rs16957085	17	46532956	LOC101927166	skin, sun exposed lower leg	8.94×10^{-10}	*SNX11*
rs17621689	17	46358952	*SKAP1*	skin, sun exposed lower leg	1.81×10^{-9}	*SNX11*
rs17694670	17	46356471	*SKAP1*	skin, sun exposed lower leg	1.83×10^{-9}	*SNX11*
rs55765969	17	46528015	LOC101927166	cells, transformed fibroblasts	7.36×10^{-7}	*CBX1*
rs55765969	17	46528015	LOC101927166	skin, sun exposed lower leg	8.94×10^{-10}	*SNX11*
rs56311919	17	46528350	LOC101927166	cells, transformed fibroblasts	7.36×10^{-7}	*CBX1*
rs56311919	17	46528350	LOC101927166	skin, sun exposed lower leg	8.94×10^{-10}	*SNX11*
rs72827805	17	46349432	*SKAP1*	skin, sun exposed lower leg	2.93×10^{-8}	*SNX11*
rs72827806	17	46351066	*SKAP1*	skin, sun exposed lower leg	2.09×10^{-9}	*SNX11*
rs72827810	17	46355550	*SKAP1*	skin, sun exposed lower leg	1.84×10^{-9}	*SNX11*
rs147813436	17	46344017	*SKAP1*	skin, sun exposed lower leg	9.94×10^{-9}	*SNX11*
rs17694092	17	46334294	*SKAP1*	skin, sun exposed lower leg	4.83×10^{-9}	*SNX11*

ACCEPTED MANUSCRIPT

rs17694168	17	46340934	*SKAP1*	skin, sun exposed lower leg	4.86×10^{-9}	*SNX11*
rs17694404	17	46348354	*SKAP1*	skin, sun exposed lower leg	7.26×10^{-8}	*SNX11*
rs72825596	17	46334340	*SKAP1*	skin, sun exposed lower leg	4.83×10^{-9}	*SNX11*
rs10514944	17	46398352	*SKAP1*	skin, sun exposed lower leg	5.18×10^{-10}	*SNX11*

Chr: chromosome; P: *p*-value

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

FIGURE LEGENDS:

Figure E1: Plate effects: a) distribution of sample call rates, mean DNA concentrations (ng/µl), as well as plating, transport and genotyping dates, **b)** Pair-wise testing of call rates reveals significant *p* values for all plates (Bonferroni-corrected), *p*-values in red are significant at α=0.05. Plates 9, 12, 13, 14, 17 and 18 show the most significant differences when compared to the other plates and seem to be responsible for the significant batch effects. We investigated dates of plating and genotyping, DNA concentration and proportion of cases from a previous candidate gene study, but none of these potential confounders explained the observed plate effects. **c)** Sample call rate for cases and controls per plate. Results of two-way ANOVA show significant case/control effect ($<2\times10^{-16}$). These results show that there is no interaction effect between plate and group study (case/control) (p=0.515).

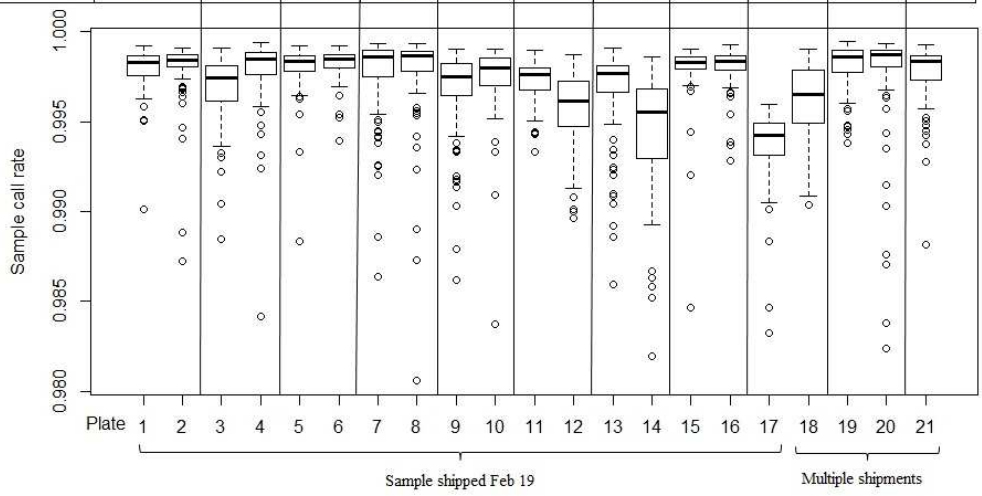
Figure E2: Flowcharts depicting exclusion criteria: a) Work flow for SNP exclusion in unrelated (left) and related (right) studies, **b)** Flowchart of subject exclusion criteria for unrelated (left) and related (right) analyses

ACCEPTED MANUSCRIPT

Figure E1: Plate effects.

a)

Plating dates	Jun 2nd						Jun 11th														
Cases concentration	67	57	76	68	71	66	68	63	62	50	57	65	68	54	52	45	52	60	53	59	69
Controls concentration	49	47	24	25	24	25	29	27	26	29	29	28	28	27	27	30	29	34	25	26	27
Genotyping dates	8-Jun		10-Jun		16-Jun		17-Jun		18-Jun		13-Jul		29-Jun		6-Jul		7-Jul		8-Jul		9-Jul



b)

Plates	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	0.21303	0.11803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	1	1	0.11121	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	1	1	0.00628	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	1	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	1	1	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
9	0.01082	0.00536	1	0.33862	0.00499	0.00018	0.39527	0.10993	-	-	-	-	-	-	-	-	-	-	-	-
10	1	1	1	1	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
11	1	1	1	1	1	0.16071	1	1	1	1	-	-	-	-	-	-	-	-	-	-
12	1.90E-12	5.30E-13	0.00137	1.30E-09	4.70E-13	1.90E-15	2.00E-09	3.40E-10	0.0436	1.70E-06	1.90E-05	-	-	-	-	-	-	-	-	-
13	0.02243	0.01143	1	0.60298	0.01068	0.00043	0.69635	0.20202	1	1	1	0.02372	-	-	-	-	-	-	-	-
14	< 2e-16	< 2e-16	1.00E-12	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	4.70E-10	< 2e-16	1.00E-15	0.17311	1.70E-10	-	-	-	-	-	-	-
15	1	1	0.19368	1	1	1	1	1	0.00982	1	1	1.80E-12	0.02036	< 2e-16	-	-	-	-	-	-
16	1	1	0.11226	1	1	1	1	1	0.00505	1	1	4.80E-13	0.0108	< 2e-16	1	-	-	-	-	-
17	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	4.30E-07	< 2e-16	1	< 2e-16	< 2e-16	-	-	-	-
18	2.70E-06	1.00E-06	1	0.00033	9.50E-07	1.30E-08	0.00043	8.90E-05	1	0.03328	0.23498	1	1	1.40E-05	2.50E-06	9.70E-07	2.90E-13	-	-	-
19	1	1	0.02887	1	1	1	1	1	0.00102	1	0.57621	2.90E-14	0.00232	< 2e-16	1	1	< 2e-16	1.10E-07		-
20	1	1	1	1	1	1	1	1	0.38921	1	1	2.10E-09	0.68512	< 2e-16	1	1	< 2e-16	0.00043	1	-

c)

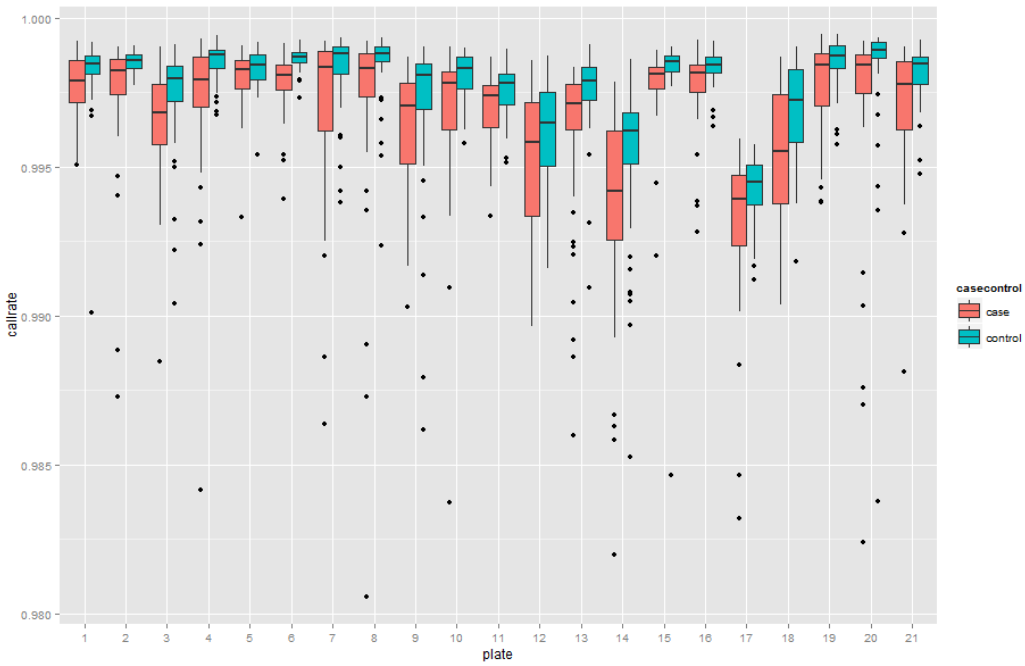
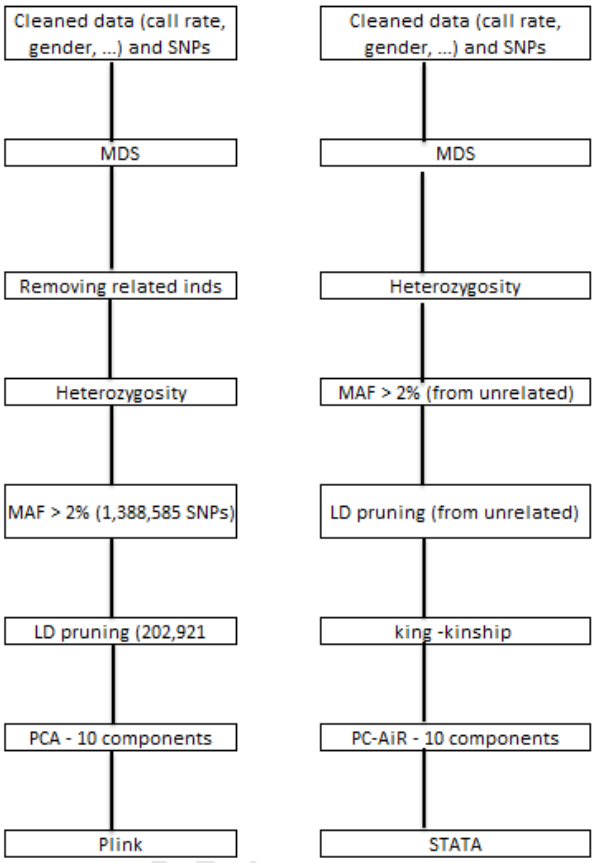
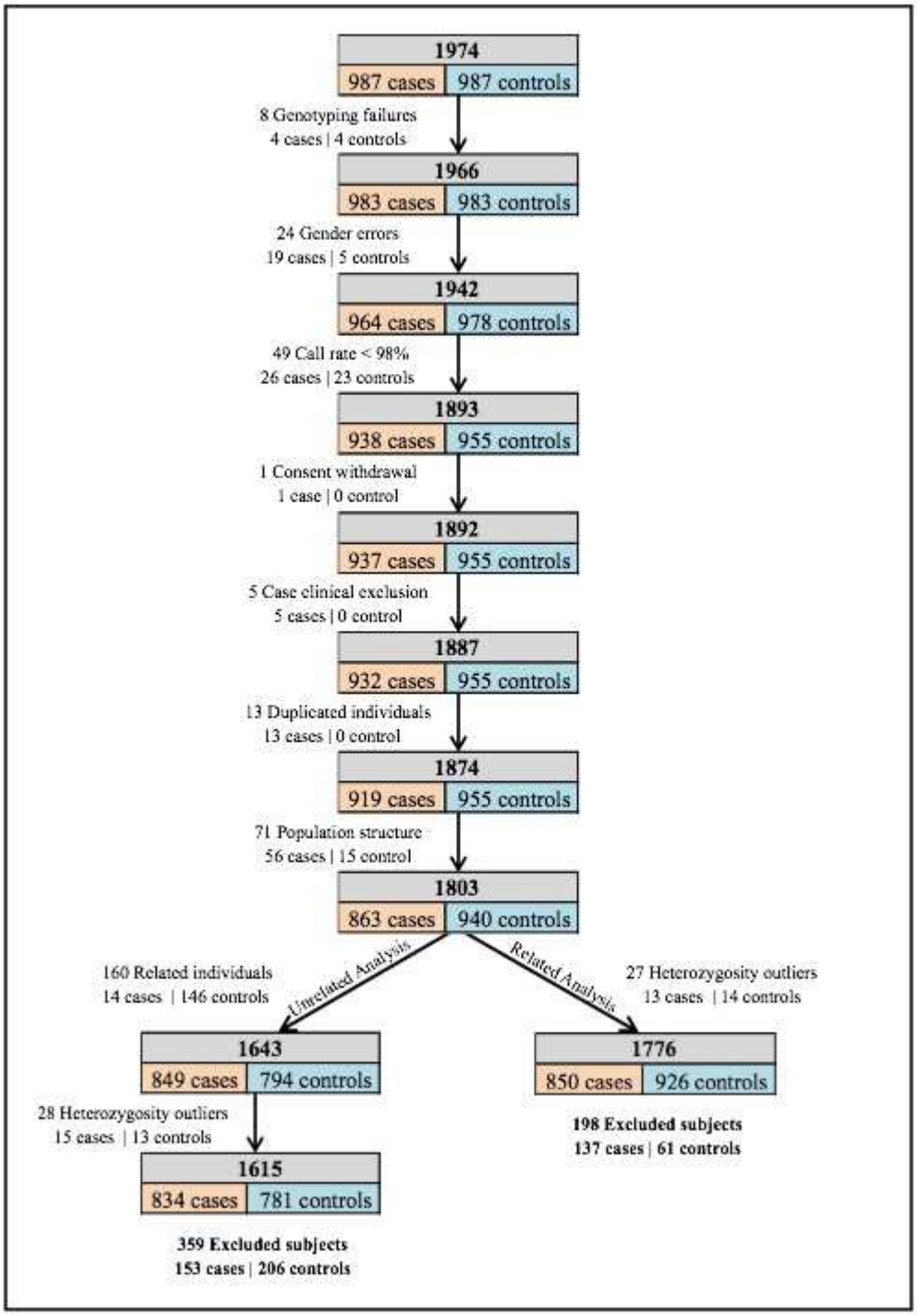


Figure E2: Flowcharts depicting exclusion criteria
a)



ACCEPTED MANUSCRIPT


1974
987 cases
987 controls
8 Genotyping failures
4 cases | 4 controls
1966
983 cases
983 controls
24 Gender errors
19 cases | 5 controls
1942
964 cases
978 controls
49 Call rate < 98%
26 cases | 23 controls
1893
938 cases
955 controls
1 Consent withdrawal
1 case | 0 control
1892
937 cases
955 controls
5 Case clinical exclusion
5 cases | 0 control
1887
932 cases
955 controls
13 Duplicated individuals
13 cases | 0 control
1874
919 cases
955 controls
71 Population structure
56 cases | 15 control
1803
863 cases
940 controls
Unrelated Analysis
Related Analysis
160 Related individuals
14 cases | 146 controls
27 Heterozygosity outliers
13 cases | 14 controls
1643
849 cases
794 controls
1776
850 cases
926 controls
28 Heterozygosity outliers
15 cases | 13 controls
198 Excluded subjects
137 cases | 61 controls
1615
834 cases
781 controls
359 Excluded subjects
153 cases | 206 controls