



Published in final edited form as:

J Allergy Clin Immunol. 2023 April ; 151(4): 1137–1142.e4. doi:10.1016/j.jaci.2022.11.008.

Epidermal Differentiation Complex Genetic Variation in Atopic Dermatitis and Peanut Allergy

Michelle F. Huffaker, MD^{1,*}, Kanika Kanchan, PhD^{2,*}, Henry T. Bahnson^{3,4}, Ingo Ruczinski⁵, Gautam Shankar², Donald Y. M. Leung, MD PhD⁶, Carolyn Baloh, MD^{1,7}, George Du Toit, MB BCh⁸, Gideon Lack, MB BCh⁸, Gerald T. Nepom, MD, PhD^{3,4}, Rasika A. Mathias, ScD²

¹The Immune Tolerance Network, San Francisco, CA

²Division of Allergy and Clinical Immunology, Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA.

³The Immune Tolerance Network, Seattle, WA

⁴Benaroya Research Institute at Virginia Mason, Seattle, WA, USA.

⁵Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.

⁶Division of Pediatric Allergy and Immunology, National Jewish Health, Denver, CO.

⁷Division of Allergy and Clinical Immunology, Brigham and Women's Hospital, Boston, MA

⁸The Department of Pediatric Allergy, Division of Asthma, Allergy and Lung Biology, King's College London, and Guy's and St Thomas' NHS Foundation Trust, London, UK.

Corresponding Author: Michelle F. Huffaker, Immune Tolerance Network, University of California, San Francisco, Diabetes Center, 513 Parnassus Avenue, HSW 11, Box 0534, San Francisco, CA 94143-0534, mhuffaker@immunetolerance.org.
*Co-first authors

Disclosure Statement

Michelle F. Huffaker, MD: no relevant disclosures or conflicting interests.

Kanika Kanchan, PhD: no relevant disclosures or conflicting interests.

Henry T. Bahnson: reports contract work paid to institution, Benaroya Research Institute, from DBV Technologies, MYOR <https://www.myorcare.com/>, King's College London, and Stanford University; and additional salary support paid by King's College London and Stanford University.

Ingo Ruczinski: no relevant disclosures or conflicting interests.

Gautam Shankar: no relevant disclosures or conflicting interests.

Donald Y. M. Leung, MD PhD: consulted for Genentech, Sanofi-Genzyme, Leo Pharma, and Incyte Corporation.

Carolyn Baloh, MD: no relevant disclosures or conflicting interests.

George Du Toit, MB BCh: reports grants from National Institute of Allergy and Infectious Diseases (NIAID, NIH), Food Allergy & Research Education (FARE), MRC & Asthma UK Centre, UK Dept of Health through NIHR, Action Medical Research and National Peanut Board. Scientific Advisory Board member Aimmune. Investigator on pharma-sponsored allergy studies (Aimmune, and DBV Technologies). Scientific advisor to Aimmune, DBV and Novartis.

Gideon Lack, MB BCh: reports grants from National Institute of Allergy and Infectious Diseases (NIAID, NIH), other from Food Allergy & Research Education (FARE), other from MRC & Asthma UK Centre, other from UK Dept of Health through NIHR, other from National Peanut Board (NPB), other from The Davis Foundation, during the conduct of the study; shareholder in DBV Technologies, and Mighty Mission Me, personal fees from Novartis, personal fees from Sanofi-Genzyme, personal fees from Regeneron, personal fees from ALK-Abello, personal fees from Lurie Children's Hospital, outside the submitted work.

Gerald T. Nepom, MD, PhD: no relevant disclosures or conflicting interests.

Rasika A. Mathias, ScD: no relevant disclosures or conflicting interests.

Clinical Trial registration

LEAP: [NCT00329784](https://clinicaltrials.gov/ct2/show/study/NCT00329784)

Abstract

Background: Deleterious variation in the epidermal differentiation complex (EDC) on chromosome 1 is a well-known genetic determinant of atopic dermatitis (AD) and has been associated with risk of peanut allergy (PA) in population-based studies.

Objective: To determine the effect of genetic variation in the EDC on AD trajectory and risk of PA in early life.

Methods: Genome sequencing was used to measure genetic variation in the EDC in the Learning Early About Peanut (LEAP) study participants. Association tests were done to identify gene- and variant-level predicted deleterious variation associated with AD severity using SCORAD (N=559) at baseline and each follow-up visit, as well as PA and food allergy in peanut avoiders (N=275). Predicted deleterious variants included missense variants that were frameshift insertion, frameshift deletion, stop-gain, or stop-loss. Associations between variant load, SCORAD, and PA were tested using linear and generalized linear regression models.

Results: *FLG*, *FLG2*, *HRNR* and *TCHH1* harbored the most predicted deleterious variation (30, 6, 3 and 1 variants, respectively). *FLG* variants were associated with SCORAD at all timepoints; four variants-- R1798X, R501X, S126X, and S761fs-- drove the association with SCORAD at each time point, and higher variant load was associated with greater AD severity over time. There was an association between these variants and PA, which remained significant independent of baseline AD severity (OR=2.63, 95% CI 1.11–6.01, p=0.02).

Conclusions: Variation in *FLG* predicted to be deleterious is associated with AD severity at baseline and longitudinally, and has an association with PA independent of baseline severity.

Capsule Summary

We identified variation predicted to be deleterious in the epidermal differentiation complex associated with atopic dermatitis severity in the LEAP participants and found an independent association with deleterious variation in *FLG* and development of peanut allergy.

Keywords

Atopic Dermatitis; Eczema; Peanut Allergy; Genetics; Filaggrin; Epidermal Differentiation Complex

INTRODUCTION

The filaggrin gene (*FLG*) is a well-described genetic determinant of atopic dermatitis (AD), (1) and deleterious variants in *FLG* have been associated with early onset severe disease.(2–4) *FLG* is located within the epidermal differentiation complex (EDC) on chromosome 1, a cluster of genes critical for skin barrier function, and variants in this region outside of those in *FLG* have also been implicated in atopic dermatitis.(5, 6) To our knowledge, no studies have examined the association of genetic variants across the EDC with the trajectory of AD over the first five years of life.

Deleterious variants in *FLG* have also been associated with peanut allergy (PA) (7–10), however, there are gaps in our understanding of this relationship, particularly as to whether

genetic risk is conferred in the absence of clinically apparent skin barrier dysfunction. A recent review examined prior work in detail,(11) and the association of *FLG* with food allergy in consideration of AD is unclear. Some studies demonstrate an association with food allergy independent of AD and/or AD severity (7, 11) while others do not find an independent association between the *FLG* variants and food allergy.(11–13) Therefore, a better understanding of the role of *FLG* in the first few years of life during the time of allergy development, and longitudinal evaluation of AD severity in this timeframe is needed to clarify this relationship between genetics, atopic dermatitis, and food allergy.

The Learning Early About Peanut (LEAP) study enrolled infants 4 to 11 months of age with egg allergy and/or severe atopic dermatitis.(14) Participants were randomized to either peanut consumption or avoidance, and consumption proved to be highly protective against the development of peanut allergy. Whole genome sequencing (WGS) was previously performed on the LEAP study participants.(15) In this population enriched for atopic dermatitis, followed longitudinally over the first several years of life, and with challenge proven peanut allergic status, we have the unique opportunity to interrogate the relationship between genetic determinants of atopic dermatitis, AD severity, and peanut allergy. Here, we report the genetic predictors of atopic dermatitis in the LEAP participants and their relationship with AD severity and the development of peanut allergy.

RESULTS AND DISCUSSION

The landscape of exonic single nucleotide variants (SNVs), and small frameshift insertion and deletions (INDELs) in the EDC in LEAP participants is shown in Figure 1a. The four genes with the greatest number of exonic variants are *FLG*, *FLG2*, *HRNR*, and *TCHH1*, and these are also the genes with the highest number of predicted deleterious variants (30, 6, 3 and 1 variants, respectively). The carrier rate and minor allele frequency (MAF) of these predicted deleterious variants are presented in Supplementary Table E1. The frequency of the predicted deleterious variants are in general higher in LEAP than that reported in the Exome Aggregation Consortium (ExAC) database (Figure 1b and 1c); given LEAP study inclusion criteria of severe atopic dermatitis and the known importance of the EDC, and specifically *FLG*, this pattern is anticipated. The LEAP participants were majority (74%) European ancestry by self-report and genetic ancestry.(15)

Gene-based association tests were performed to identify predicted deleterious variants associated with AD severity using SCORAD at each study time point in both the peanut consumers and peanut avoiders (N=559) for the four genes *FLG*, *FLG2*, *HRNR*, and *TCHH1* individually, and a combined test across all four (Figure 2). We found that predicted deleterious variants in the EDC were significantly associated with higher SCORAD across all time points ($p < 0.001$), and this association was driven by variation in *FLG* gene. A gene-based test including 30 *FLG* variants was associated with SCORAD at all time points, and with SCORAD intercept (i.e., the starting point of someone's AD severity at baseline) (Figure 2). No association was seen with the slope (i.e., the general trajectory over the course (p -value 0.29)). To identify SCORAD driving variants in *FLG* at each study time point, we performed leave-one-out (LOO) analysis and identified multiple *FLG* variants to drive the gene-based signal (Figure 3a). LOO analysis for SCORAD intercept also revealed

an association between the *FLG* variants and the SCORAD intercept (p-value 2.58e-04) (Figure 3b), implicating an association that begins within the first year of life. Further, to estimate the effect of *FLG* predicted deleterious variants on SCORAD, we performed single variant association test between *FLG* variants and SCORAD at 30 months (Figure 3a) and SCORAD intercept (Figure 3b). For both SCORAD at 30 months and SCORAD intercept the variants that had largest contribution to the gene-based signal (SCORAD at 30 months: R1798X, R501X, S126X, and S761fs) and (SCORAD intercept: R2477X, R1798X, R501X, and S126X) also had significant association and effect sizes at the variant level.

Four *FLG* variants-- R1798X, R501X, S126X, and S761fs-- were found to drive gene-based significance at all time points and for the intercept (Figure 3c). These four variants are located in exon 3, the region of the gene comprised of 10–12 tandem repeats of what becomes the filaggrin peptide (Figure 4). R501X, S761fs, R1798X, have previously been implicated in early onset of severe disease.(4) To our knowledge, this the first report of S126X. An association between variant load across these 4 variants and SCORAD was tested using a linear regression model. We found 461 individuals had no deleterious alleles, 90 had 1 deleterious allele from any of the four variants, and 8 had 2; no individuals carried greater than 2 deleterious alleles. Those with 2 variants had the highest SCORAD at baseline, and while all study participants were noted to have an improvement in SCORAD over the timeline of the trial, those with 2 variant alleles maintained the highest severity across all time points (Figure 5).

Similar gene-based association tests using all predicted deleterious variants with PA in the peanut avoiders (N=275), did not yield statistically significant associations (Figure 2). However, when we performed the gene-based test limiting to only the four SCORAD driving variants in *FLG* (R1798X, R501X, S126X, and S761fs), we found a significant association with PA (p=0.02). To calculate an odds ratio for PA risk from these four variants, we performed a logistic regression model for PA with carrier status defined as the presence of any of the four variants, and found an association with PA (OR=3.13, 95% CI 1.36–7.03, p=0.006). This association decreased but remained significant when adjusting for SCORAD at baseline (OR=2.63, 95% CI 1.11–6.01, p=0.02). Similar to the findings by other groups, there does appear to be a residual effect of *FLG* on PA risk after adjusting for AD and/or AD severity at baseline entry into the study.(7–9, 11)

There are several limitations to consider. First, the analysis was performed on variation in the EDC and did not consider variation outside this region. Second, there are genetic loci associated with atopic dermatitis outside the EDC, including variation in genes for *IL4*, *IL-4* receptor, and *IL-13*, and others that could be considered in future analyses.(16) A third limitation is that the present work does not address environmental factors that may have contributed to the development and severity of AD, and is limited to genetic analyses. Here, we are using genome sequencing at the standard depth of 30X, and while we have shown the value of using WGS at this depth to study the EDC previously (17), we recognize that we many have missed some EDC variants given complexity of sequencing this region.(18) Another important limitation to consider is that it is difficult to tease apart the independent effects of the driving variants versus AD without these driving variants, as the population

was selected for severe AD. An additional limitation is that while the variants are predicted to be deleterious, functional studies are needed to understand the impact on the skin barrier.

To our knowledge, this is the first report utilizing WGS to understand genetic risk on the trajectory of atopic dermatitis over the first several years of life. We found that participants carrying deleterious variation in the EDC, specifically *FLG*, had higher SCORAD across the first 5 years of life. The slope of improvement in AD severity over time, however, was similar in those with and without deleterious variants. This is also the first report utilizing WGS and well characterized atopic dermatitis over the first 5 years of life to understand the relationship between *FLG* and PA. The data suggest that there may be an association between *FLG* and PA independent of baseline AD severity. Identifying individuals with deleterious variation in *FLG* early may be important for mitigating peanut allergy risk through early oral introduction. Furthermore, future research is needed to understand environmental factors that may contribute to AD development, severity, and the interplay of these environmental factors with underlying genetic risk. It remains unknown if the risk of food allergy in carriers of *FLG* variants may be reduced through early intervention and improvement of AD control.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Drs. Daniel Rotrosen and Alkis Togias from the National Institute of Allergy and Infectious Diseases for their thoughtful review of the manuscript. We are grateful to all the LEAP study participants who took part in the study.

Funding Statement

This research was performed as a project of the Immune Tolerance Network, an international clinical research consortium headquartered at the Benaroya Research Institute, and supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award no. UM1AI109565. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations

AD	Atopic Dermatitis
EDC	Epidermal Differentiation Complex
ExAC	Exome Aggregation Consortium ExAC
FLG	Filaggrin gene
INDELs	Insertions and Deletions
LEAP	Learning Early About Peanut Allergy Study
LOO	Leave One Out
MAF	Minor Allele Frequency

PA	Peanut Allergy
SCORAD	Scoring of Atopic Dermatitis
SKAT-O	SNP-set (Sequence) Kernel Association Test, Optimal
SNV	Single Nucleotide Variant
OR	Odds Ratio
PC	Principal Components
PCA	Principal Components Analysis
WGS	Whole Genome Sequencing

REFERENCES

1. Irvine AD, McLean WH, Leung DY. Filaggrin mutations associated with skin and allergic diseases. *N Engl J Med*. 2011;365(14):1315–27. [PubMed: 21991953]
2. Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, et al. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J Invest Dermatol*. 2007;127(3):564–7. [PubMed: 16990802]
3. Hoyer A, Rehbinder EM, Fardig M, Asad S, Lodrup Carlsen KC, Endre KMA, et al. Filaggrin mutations in relation to skin barrier and atopic dermatitis in early infancy. *Br J Dermatol*. 2022;186(3):544–52. [PubMed: 34698386]
4. Smieszek SP, Welsh S, Xiao C, Wang J, Polymeropoulos C, Birznieks G, et al. Correlation of age-of-onset of Atopic Dermatitis with Filaggrin loss-of-function variant status. *Sci Rep*. 2020;10(1):2721. [PubMed: 32066784]
5. Mischke D, Korge BP, Marenholz I, Volz A, Ziegler A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex (“epidermal differentiation complex”) on human chromosome 1q21. *J Invest Dermatol*. 1996;106(5):989–92. [PubMed: 8618063]
6. Paternoster L, Standl M, Chen CM, Ramasamy A, Bonnelykke K, Duijts L, et al. Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. *Nat Genet*. 2011;44(2):187–92. [PubMed: 22197932]
7. 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(3):661–7. [PubMed: 21377035]
8. 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(1):239–42. [PubMed: 23684069]
9. Brough HA, Simpson A, Makinson K, Hankinson J, Brown S, Douiri A, et al. Peanut allergy: effect of environmental peanut exposure in children with filaggrin loss-of-function mutations. *J Allergy Clin Immunol*. 2014;134(4):867–75 e1. [PubMed: 25282568]
10. Kalb B, Marenholz I, Jeanrenaud AC, Meixner L, Arnau-Soler A, Rosillo-Salazar OD, et al. Filaggrin loss-of-function mutations are associated with persistence of egg and milk allergy. *J Allergy Clin Immunol*. 2022.
11. Kanchan K, Clay S, Irizar H, Bunyavanich S, Mathias RA. Current insights into the genetics of food allergy. *J Allergy Clin Immunol*. 2021;147(1):15–28. [PubMed: 33436162]
12. Tan HT, Ellis JA, Koplin JJ, Matheson MC, Gurrin LC, Lowe AJ, et al. Filaggrin loss-of-function mutations do not predict food allergy over and above the risk of food sensitization among infants. *J Allergy Clin Immunol*. 2012;130(5):1211–3 e3. [PubMed: 22964107]

13. Venkataraman D, Soto-Ramirez N, Kurukulaaratchy RJ, Holloway JW, Karmaus W, Ewart SL, et al. Filaggrin loss-of-function mutations are associated with food allergy in childhood and adolescence. *J Allergy Clin Immunol*. 2014;134(4):876–82 e4. [PubMed: 25174864]
14. Du Toit G, Roberts G, Sayre PH, Bahnson HT, Radulovic S, Santos AF, et al. Randomized trial of peanut consumption in infants at risk for peanut allergy. *N Engl J Med*. 2015;372(9):803–13. [PubMed: 25705822]
15. Winters A, Bahnson HT, Ruczinski I, Boorgula MP, Malley C, Keramati AR, et al. The MALT1 locus and peanut avoidance in the risk for peanut allergy. *J Allergy Clin Immunol*. 2019;143(6):2326–9. [PubMed: 30825465]
16. Loset M, Brown SJ, Saunes M, Hveem K. Genetics of Atopic Dermatitis: From DNA Sequence to Clinical Relevance. *Dermatology*. 2019;235(5):355–64. [PubMed: 31203284]
17. Bin L, Malley C, Taylor P, Preethi Boorgula M, Chavan S, Daya M, et al. Whole genome sequencing identifies novel genetic mutations in patients with eczema herpeticum. *Allergy*. 2021;76(8):2510–23. [PubMed: 33548076]
18. Margolis DJ, Mitra N, Wubbenhorst B, Nathanson KL. Filaggrin sequencing and bioinformatics tools. *Arch Dermatol Res*. 2020;312(2):155–8. [PubMed: 31372728]
19. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401–4. [PubMed: 22588877]
20. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):pl1. [PubMed: 23550210]

Clinical Implications

Variation predicted to be deleterious in epidermal differentiation complex is associated with atopic dermatitis severity during the first five years of life. Deleterious variation in *FLG* is independently associated with peanut allergy.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

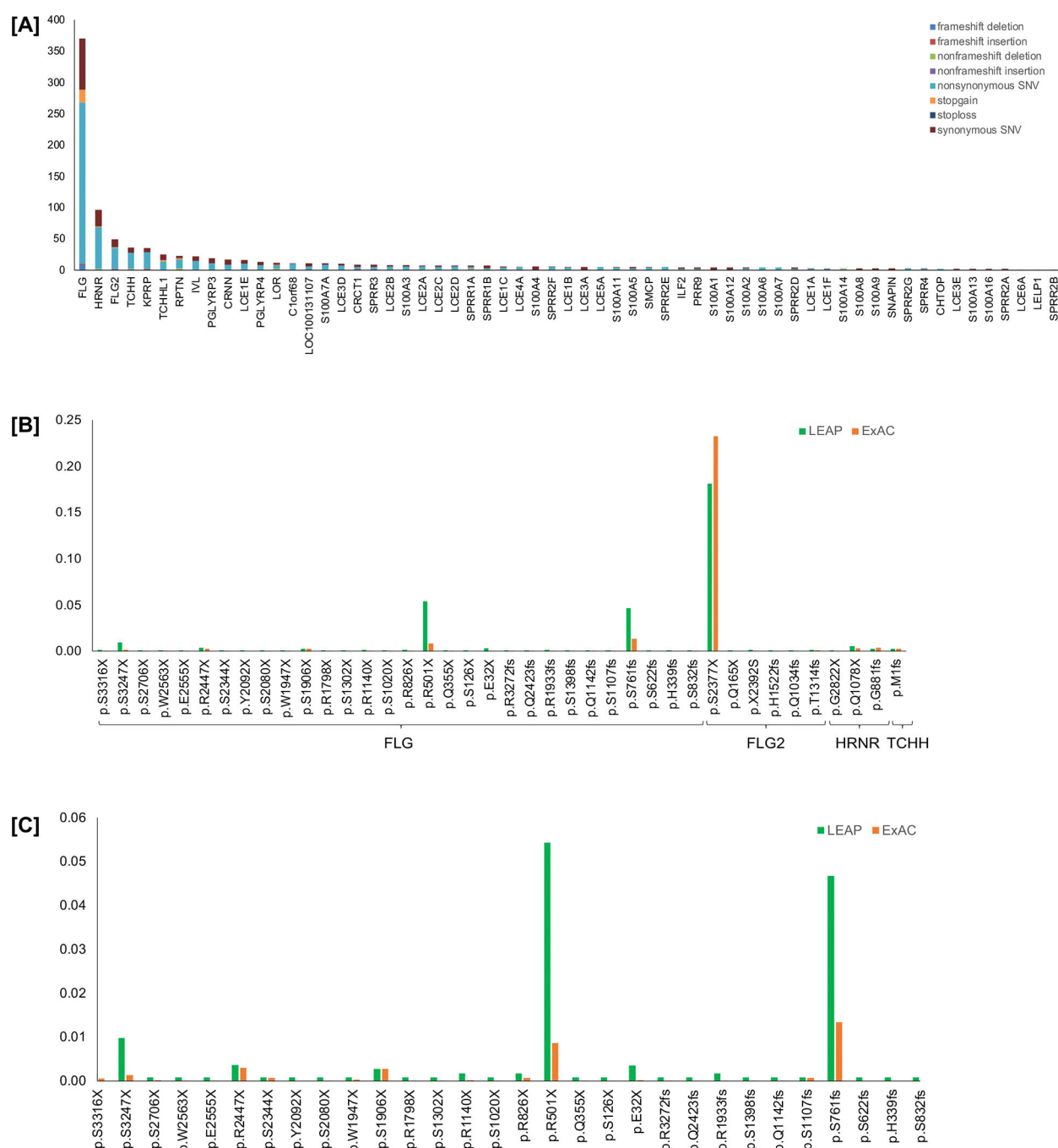
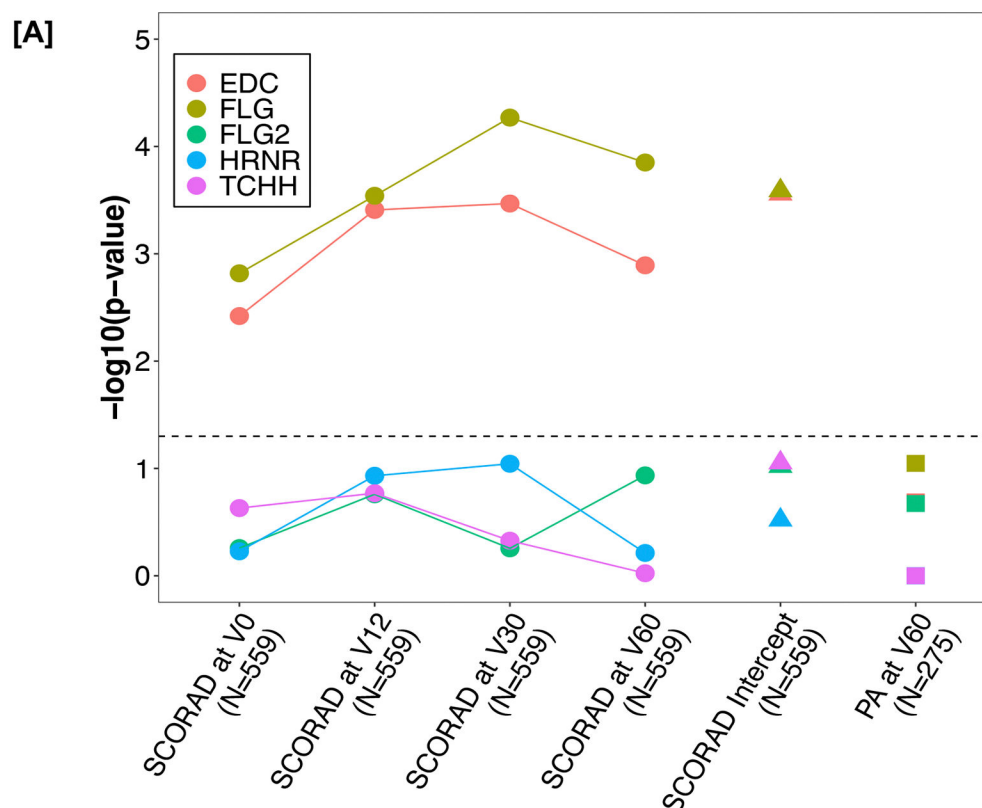


Figure 1: EDC region exonic variations and their frequencies in the LEAP participants.

[A] Exonic variations in the EDC genes in the LEAP Participants. Genes are ordered on the total number of exonic variants observed, and bars show the representation of each class of variation. **[B]** Frequencies of EDC deleterious variants in LEAP participants (green) and ExAC database (orange). **[C]** FLG deleterious variant frequencies in LEAP participants (green) and ExAC database (orange).



[B]

Gene set (# included variants)	SCORAD in N = 559 LEAP subjects from both intervention arms					PA in N = 275 LEAP subjects from avoidance arm at V60
	V0	V12	V30	V60	SCORAD Intercept	
EDC (40)	3.80E-03	3.90E-04	3.40E-04	1.28E-03	2.77E-04	0.21
FLG (30)	1.52E-03	2.88E-04	5.38E-05	1.41E-04	2.58E-04	0.09
FLG2 (6)	0.55	0.17	0.56	0.12	0.10	0.21
HRNR (3)	0.59	0.12	0.09	0.61	0.30	1.00
TCHH (1)	0.23	0.17	0.47	0.95	0.09	1.00

Figure 2: Association of deleterious variants in the EDC region with SCORAD and Peanut Allergy.

For SCORAD, association was tested at all visits (baseline, 12, 30 and 60 months of age) during the trial, and with SCORAD intercept (starting point of AD severity) in the full group of participants (n=559). For PA, the association was tested in the participants of the peanut avoidance group (n=275) at 60 months. Panel A is a graphical illustration of the p-values presented in Panel B. The p-values for SCORAD at all visits are illustrated as points, for SCORAD intercept as triangles, and for PA at 60 months as squares. V0 = baseline visit, V12 = 12 months visit, V30 = 30 months visit, and V60 = 60 months visit.

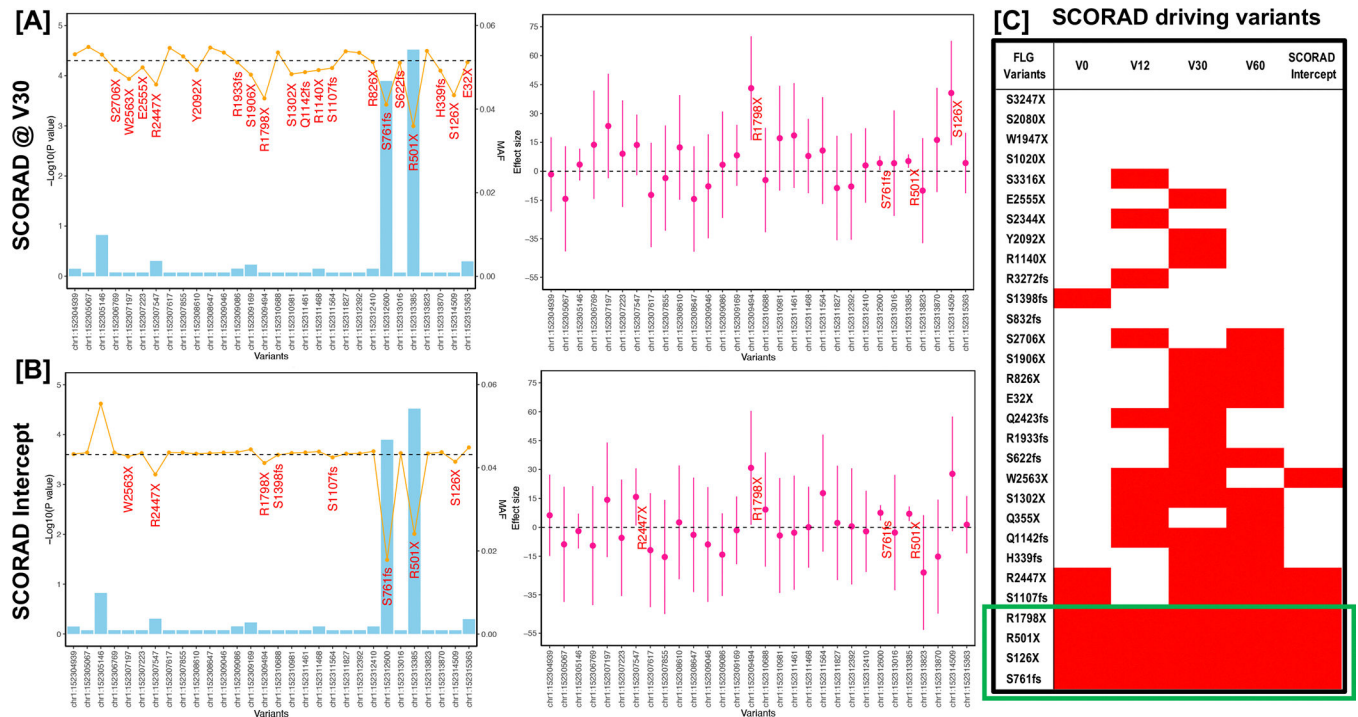


Figure 3: Deleterious variants in FLG driving association with SCORAD.

[A] SCORAD at 30 months, and **[B]** SCORAD Intercept for each subject derived from a model of SCORAD over time. Left panel shows the gene-based association results where the horizontal dashed line shows the p-value when all variants are included. Each orange dot represents the new gene-based p-value when a single variant (denoted by the x-axis label) is left out. The blue bars represent the minor allele frequency (noted on the y-axis on the right). Every instance where the orange dot falls under the horizontal line implicates a variant as a driving variant. Right panel shows the single variant association results where the horizontal dashed line shows there is no linear relationship ($\beta=0$) between the response and the predictor variables. Each pink dot represents the effect size (β) for the variants and the pink lines show the 95% CIs. The variants showing significant association ($p<0.05$) are highlighted. **[C]** The red box flag variants identified to be a driving variant for the FLG gene-based association test for SCORAD at each visit and SCORAD intercept. Four variants (green box) were consistently score driving for each timepoint and SCORAD intercept.

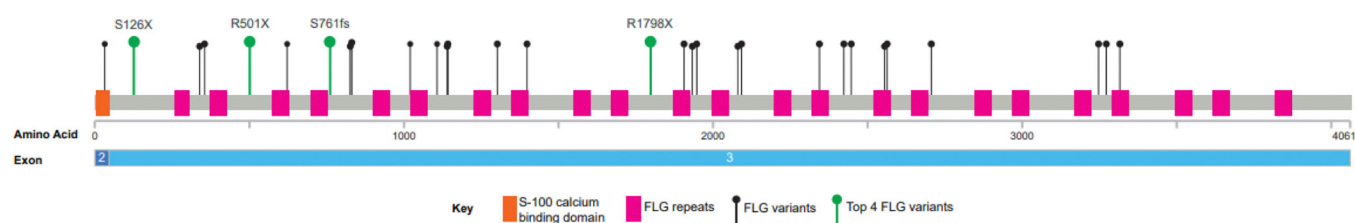


Figure 4: Lollipop Plot of four deleterious variants driving association with SCORAD at every study time point and for the model intercept.
Figure developed using the cBio Cancer Genomics Portal.(19, 20)

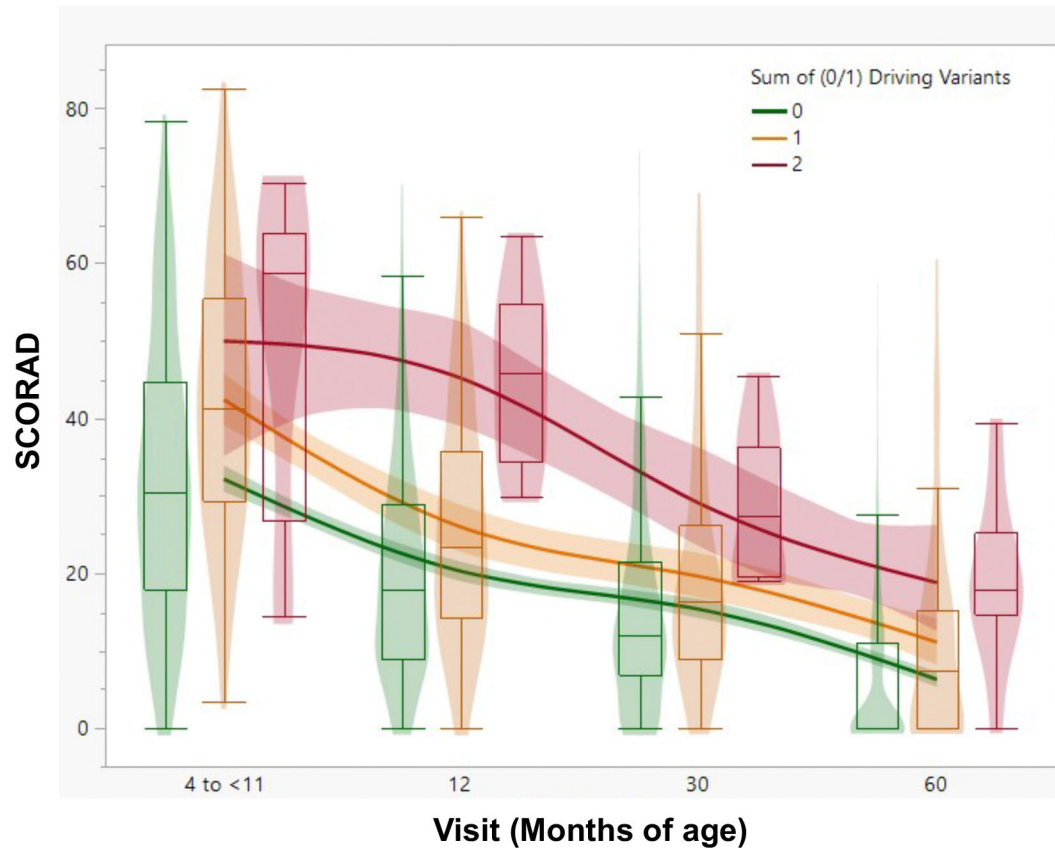


Figure 5: SCORAD over time for LEAP participants with 0, 1, or 2 driving variants.

The unadjusted mean trajectory of SCORAD is shown using a smoothed cubic spline with bootstrapped 95% CIs stratified by the number of driving Variants. No study subjects were found to have more than 2 driving variants across the final four selected.