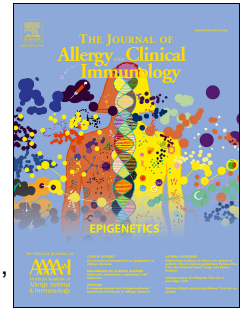


Accepted Manuscript

Novel locus for atopic dermatitis in African Americans and replication in European Americans

Berta Almoguera, Ph.D, Lyam Vazquez, M.D, Frank Mentch, Ph.D, Michael E. March, Ph.D, John J. Connolly, Ph.D, Peggy L. Peissig, Ph.D, James G. Linneman, BA, María del Carmen Plaza-Serón, Ph.D, Maria Pino-Yanes, Ph.D, Esteban G. Burchard, M.D, Murray Brilliant, Ph.D, Patrick Sleiman, Ph.D, Hakon Hakonarson, M.D, Ph.D



PII: S0091-6749(18)31580-X

DOI: <https://doi.org/10.1016/j.jaci.2018.10.038>

Reference: YMAI 13710

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

Received Date: 16 October 2017

Revised Date: 5 October 2018

Accepted Date: 19 October 2018

Please cite this article as: Almoguera B, Vazquez L, Mentch F, March ME, Connolly JJ, Peissig PL, Linneman JG, Plaza-Serón MdC, Pino-Yanes M, Burchard EG, Brilliant M, Sleiman P, Hakonarson H, Novel locus for atopic dermatitis in African Americans and replication in European Americans, *Journal of Allergy and Clinical Immunology* (2018), doi: <https://doi.org/10.1016/j.jaci.2018.10.038>.

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.

Novel locus for atopic dermatitis in African Americans and replication in European Americans

Berta Almoguera, Ph.D¹, Lyam Vazquez, M.D¹, Frank Mentch, Ph.D¹, Michael E. March, Ph.D¹, John J Connolly, Ph.D¹, Peggy L Peissig, Ph.D², James G Linneman, BA², María del Carmen Plaza-Serón, Ph.D³, Maria Pino-Yanes, Ph.D^{3,4,5}, Esteban G. Burchard, M.D^{6,7}, Murray Brilliant, Ph.D², Patrick Sleiman, Ph.D^{1,8}, Hakon Hakonarson, M.D, Ph.D^{1,8}

1. Center for Applied Genomics, The Children's Hospital of Philadelphia, Pennsylvania, USA

2. Marshfield Clinic Research Institute, Marshfield, Wisconsin, USA.

3. Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna (ULL), La Laguna, Tenerife, Spain.

4. Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna (ULL), Santa Cruz de Tenerife, Spain.

5. CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain.

6. Department of Medicine, University of California, San Francisco, CA, USA.

7. Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, USA.

8. Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

20

21 **Correspondence:**

22 Patrick Sleiman, PhD

23 Center for Applied Genomics, Children's Hospital of Philadelphia

24 Leonard Madlyn Abramson Research Center,

25 3615 Civic Center Boulevard, Suite 1216E, Philadelphia, PA 19104. USA

26 Phone: 267-426-7653; Fax: 267-426-0363

27 e-mail: sleimanp@email.chop.edu

28

29 **Keywords:** Atopic dermatitis; GWAS; African American; electronic health records

30

31 **Conflict of interest:** The authors declare no conflicts of interest

32

33 **Funding source:** The eMERGE Network was initiated and funded by NHGRI through the
34 following grants: U01HG006830 and U01HG8684 (Children's Hospital of Philadelphia);
35 U01HG006389 (Essentia Institute of Rural Health, Marshfield Clinic Research Foundation and
36 Pennsylvania State University). The study was additionally funded by an Institutional
37 Development Award to the Center for Applied Genomics from the Children's Hospital of
38 Philadelphia. MP-Y was funded by the Ramón y Cajal Program (RYC-2015-17205) by the

Spanish Ministry of Economy and Competitiveness. MP-Y was also supported by award number AC15/00015 by Instituto de Salud Carlos III through AES and EC within AAL framework and the SysPharmPedia grant awarded from the ERA-CoSysMed 1st Joint Transnational Call from the European Union under the Horizon 2020. E.G.B was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01HL128439, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, U54MD009523, 1R01MD010443 and the Tobacco-Related Disease Research Program under Award Number 24RT-0025. Marshfield Clinic Research Foundation efforts were supported by NIH NHGRI grant U01HG8701 and NCATS grant 1UL1TR002373.

Capsule summary: The present study provides the first genetic locus associated with AD in African Americans.

Abbreviations:

AA: African Americans

AD: Atopic dermatitis

CAG: Center for Applied Genomics

GWAS: Genome-wide association studies

- 60 EA: European Americans
- 61 eMERGE: electronic Medical Records and Genomics
- 62 eQTL: cis-expression quantitative trait locus
- 63 LD: linkage disequilibrium
- 64 PCA: Principal component analysis
- 65

To the editor:

Atopic dermatitis (AD) is a common chronic inflammatory skin disease that affects up to 30% of children and is highly heritable. To date, 31 genetic loci have been identified through GWAS, mapping to genes involved in immune regulation and skin barrier deficiencies (1). However, despite the higher prevalence of AD in African Americans (AA; 15.9% versus 9.7% in European ancestry) (2), this population has been underrepresented in GWAS and specific genetic risk factors remain elusive.

In this study, we report an AD GWAS in the largest AA sample studied to date. For case and control selection we developed, validated and implemented an AD phenotyping algorithm based on Electronic Health Record data mining, in collaboration with the eMERGE network. Cases were defined as subjects ≥ 2 months old with either: 1) ≥ 2 AD diagnoses (ICD9 code 691.8), on separate calendar days, and ≥ 1 prescription of AD-related medications (Table E1); or 2) ≥ 3 AD diagnoses, on separate calendar days. Subjects with scabies, Wiscott-Aldrich syndrome, allergic purpura, ichthyosis congenita and chromosomal anomalies were excluded. Controls were subjects with no ICD9 code for AD, no ICD9 codes for any skin condition, allergy, asthma, nutritional deficiencies or skin cancer (Table E2) and no relevant medications for AD (Table E1).

The algorithm was validated internally at the Center for Applied Genomics (CAG) and also externally at the Marshfield Clinic with positive predictive values over 92% for cases and controls at both sites (for details see Online Repository).

Implementation of the algorithm at CAG accrued 12,345 AD cases and controls with imputed genotyping data (Table E3), including 5,843 of AA ancestry (2,420 cases and 3,423 controls, Table E4), which were selected for the analysis. GWAS on 11,743,754 variants in the AA sample did not result in any variants surpassing genome wide significance, however, 150 variants at 49 loci showed nominal p -values $< 10^{-5}$ that were followed-up in an independent AA sample from the SAGE II study. Meta-analysis of the two AA samples resulted in a genome-wide significant association at rs3811419 (C>T) (Allele C: $\beta(\text{SE})=0.573(0.098)$; $p_{\text{META}}=5.64 \times 10^{-9}$; $\text{MAF}(\text{T})=3.7\%$; $p_{\text{CAG}}=1.48 \times 10^{-7}$; $p_{\text{SAGEII}}=0.011$) 1.5kb upstream of *RORC* (Figure 1, Table E5). This association also replicated in the 6,502 AD cases ($N=510$) and controls ($N=5,998$) of European ancestry that were identified using the same phenotyping algorithm ($\beta(\text{SE})=0.405(0.149)$; $p=6.61 \times 10^{-3}$; $\text{MAF}=11.6\%$).

We explored the potential eQTL effects of rs3811419 (C) and seven variants in LD ($r^2 > 0.8$) with it, in skin and blood, using HaploReg v4 (3), the NCBI Genotype-Tissue Expression (GTEx) version 6 data (4) and studies reporting eQTL effects in those tissues (see Online Repository). Seven of the eight SNPs investigated showed eQTL effects on the expression of *THEM4* in blood (top effect for rs3811418 (G); $p\text{-value}=6.22 \times 10^{-14}$) and five showed effects in skin (top effect for rs11582525 (G); $p\text{-value}=1.50 \times 10^{-5}$) (Table E6). Ferreira et al. recently reported association at this locus in a large meta-analysis of allergic disease in individuals of European ancestry (5). Their top variant, rs11204896, was independent of rs3811419 and did not replicate in our AA sample. Only nine of the 136 variants reported by Ferreira et al. showed nominal significance in the AA analysis ($p\text{-value} < 0.05$). Of those, seven had the same direction of effect (Table E7). It is noteworthy, however, that the authors report that rs11204896-C is also an eQTL

for *THEM4* (5), supporting a potential role for *THEM4* in atopy. Follow-up of associated variants might be needed in order to exclude the effect being due to LD rather than a causal role of *THEM4*.

eQTL effects were also found for *C2CD4D* (top effect for rs72692781 (T)=-0.21; p-value= 3.2×10^{-5}) in skin (Table E6). Interestingly, all variants queried also had weak eQTL effects over *FLG-AS1* (FLG antisense RNA 1) in skin (rs3811419 (T), effect=-0.23; p-value= 3.60×10^{-3}). *FLG-AS1* is a long noncoding RNA (lncRNA), located 500Kb from rs3811419, that overlaps some of the coding *FLG* and *FLG2* sequence on the antisense strand. Although the exact function of *FLG-AS1* has not been characterized to date, lncRNAs are increasingly recognized as key players in the regulation of gene expression in developmental and disease processes (reviewed in (6)). One of the most well-recognized pathophysiologic differences in AD of AA compared to Europeans and Asians is the absence of genetically determined FLG deficiency. While null *FLG* mutations are commonly found in European and Asian populations, they are very rare in individuals of AA ancestry and not considered a risk factor (7). The rs3811419-C risk allele would be associated with an increase of the expression of *FLG-AS1*, which could be translated in a decrease of FLG expression. This eQTL effect over *FLG-AS* could constitute an alternate mechanism leading to *FLG*-related skin barrier deficiency in AA that should be further explored.

To further characterize the functional consequences of the associated locus, we analyzed the methylation status of blood-derived DNA from 374 AA CAG subjects, generated on the Infinium

HumanMethylation450 BeadChip (details on the Online Repository). Methylation data were analyzed as the response variable in a linear regression, with the genotype of rs3811419 and the seven variants in LD with it ($r^2 > 0.8$) as the predictors, similar to an expression QTL analysis. Sex, age, and 10 PCs were included as covariates. The top association was for rs72692781 (T), a variant in LD with rs3811419 ($r^2 = 0.97$), and methylation of probe cg22228337 ($\beta(\text{SE}) = -0.187(0.026)$; $p = 6.70 \times 10^{-12}$) located 3.1 kb from the variant (chr1:151,802,807). Other associations passing the significance threshold of 10^{-6} were found for the top SNP rs3811419 (C) and methylation probe cg08477332 (chr1:153590243; $\beta(\text{SE}) = 1.133(0.225)$; $p = 8.05 \times 10^{-7}$); and for rs7540799 (T) and probes cg00975746 (chr1:161409970; $\beta(\text{SE}) = 0.619(0.129)$; $p = 2.30 \times 10^{-6}$) and cg00304520 (chr1: 161582581; $\beta(\text{SE}) = 0.984(0.208)$; $p = 3.28 \times 10^{-6}$). Based on data published by the BIOS consortium database, probe cg22228337 is an eQTM for *THEM5* (<https://genenetwork.nl/biosqtlbrowser/>), a gene almost exclusively expressed in skin (4). Information on the other probes was not available.

The strengths and limitations of the use of an EHR-based algorithm for case-control selection for genetic studies has been already described by our group (8). However, we could not validate the performance of the algorithm in terms of confirmation of known AD and allergy genetic loci reported in adults (Tables E7 and E8) (1): only 11q13.1 (*OVOL1*) replicated in the AA population. Interestingly, *OVOL1* was also one of the top significant genes showing trans ethnic effects after the MANTRA meta-analysis performed by Paternoster et al.(1). These results suggest that AD genetic factors described to date cannot be extrapolated to AA populations and specific genetic factors need to be uncovered.

154

155 In conclusion, we have identified the first genome wide significant association for AD in African
156 Americans and replicated the association in individuals of European ancestry. The top SNP at the
157 locus, rs3811419, is an eQTL for *THEM4*, a gene previously associated with allergy (5) and
158 *FLG-AS1*, a long non-coding RNA that overlaps the fillagrin gene. While we provide replication
159 of the associatiom in individuals of European ancestry, replication in an independent AA sample
160 and elucidation of the exact biological role of these variants in AD, is needed.

161

162 Berta Almoguera, Ph.D¹, Lyam Vazquez, M.D¹, Frank Mentch, Ph.D¹, Michael E. March, Ph.D¹,
163 John J Connolly, Ph.D¹, Peggy L Peissig, Ph.D², James G Linneman², BA, María del Carmen
164 Plaza-Serón, Ph.D³, Maria Pino-Yanes, Ph.D^{3,4,5}, Esteban G. Burchard, M.D^{6,7}, Murray Brilliant,
165 Ph.D², Patrick Sleiman, Ph.D^{1,8}, Hakon Hakonarson, M.D, Ph.D^{1,8}

166

167 1. Center for Applied Genomics, The Children's Hospital of Philadelphia, Pennsylvania, USA

168 2. Marshfield Clinic Research Institute, Marshfield, Wisconsin, USA.

169 3. Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and
170 Genetics, Universidad de La Laguna (ULL), La Laguna, Tenerife, Spain.

171 4. Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna (ULL),
172 Santa Cruz de Tenerife, Spain.

173 5. CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain.

174 6. Department of Medicine, University of California, San Francisco, CA, USA.

175 7. Department of Bioengineering and Therapeutic Sciences, University of California, San
176 Francisco, CA, USA.

177 8. Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania,
178 Philadelphia, PA, USA.

179

ACKNOWLEDGEMENTS

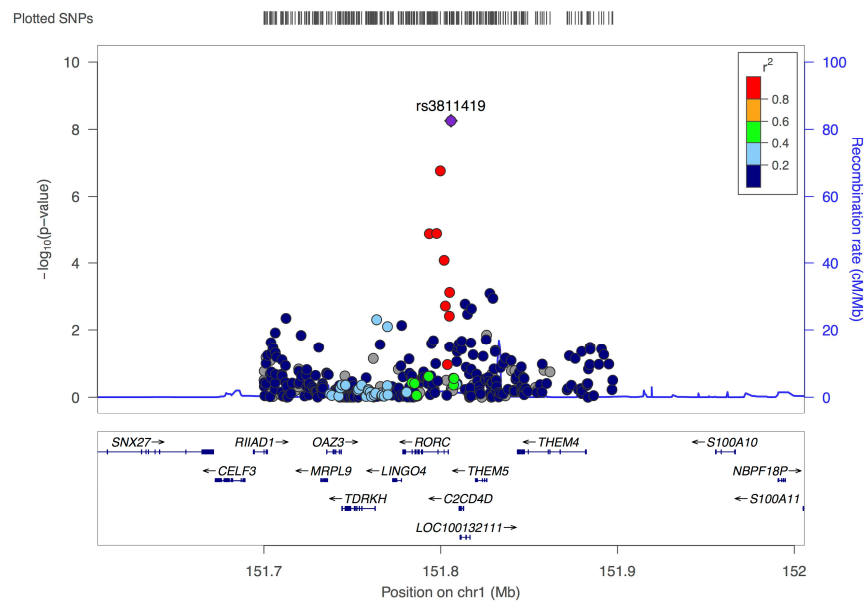
We would like to thank all patients and control subjects for their participation in the study and the eMERGE phenotyping group and researchers. The eMERGE Network was initiated and funded by NHGRI through the following grants: U01HG006830 and U01HG8684 (Children's Hospital of Philadelphia); U01HG006389 (Essentia Institute of Rural Health, Marshfield Clinic Research Foundation and Pennsylvania State University). The study was additionally funded by an Institutional Development Award to the Center for Applied Genomics from the Children's Hospital of Philadelphia. MP-Y was funded by the Ramón y Cajal Program (RYC-2015-17205) by the Spanish Ministry of Economy and Competitiveness. MP-Y was also supported by award number AC15/00015 by Instituto de Salud Carlos III through AES and EC within AAL framework and the SysPharmPedia grant awarded from the ERA-CoSysMed 1st Joint Transnational Call from the European Union under the Horizon 2020. E.G.B was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01HL128439, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, U54MD009523, 1R01MDS010443 and the Tobacco-Related Disease Research Program under Award Number 24RT-0025. Marshfield Clinic Research Foundation efforts were supported by NIH NHGRI grant U01HG8701 and NCATS grant 1UL1TR002373.

201 REFERENCES

- 202 1. Paternoster L SM, Waage J, Baurecht H, Hotze M, Strachan DP, Curtin JA, et al. Multi-
203 ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk
204 loci for atopic dermatitis. *Nature genetics*. 2015;47(12):1449-56.
- 205 2. Shaw TE, Currie GP, Koudelka CW, Simpson EL. Eczema prevalence in the United
206 States: data from the 2003 National Survey of Children's Health. *The Journal of investigative*
207 *dermatology*. 2011;131(1):67-73.
- 208 3. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation,
209 and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research*.
210 2012;40(Database issue):D930-4.
- 211 4. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue
212 Expression (GTEx) project. *Nature genetics*. 2013;45(6):580-5.
- 213 5. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared
214 genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nature*
215 *genetics*. 2017;49(12):1752-7.
- 216 6. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in Development and
217 Disease: Background, Mechanisms, and Therapeutic Approaches. *Physiol Rev*. 2016;96(4):1297-
218 325.
- 219 7. Margolis DJ, Gupta J, Apter AJ, Ganguly T, Hoffstad O, Papadopoulos M, et al.
220 Filaggrin-2 variation is associated with more persistent atopic dermatitis in African American
221 subjects. *The Journal of allergy and clinical immunology*. 2014;133(3):784-9.
- 222 8. Almoguera B, Vazquez L, Mentch F, Connolly J, Pacheco JA, Sundaresan AS, et al.
223 Identification of Four Novel Loci in Asthma in European American and African American
224 Populations. *American journal of respiratory and critical care medicine*. 2017;195(4):456-63.

Figure Legend

Figure 1. Regional association plot at the chromosome 1 locus in atopic dermatitis following meta-analysis in the African Americans. Statistical significance of each SNP plotted on the y-axis as a function of chromosome position on the x-axis. The top SNP at the locus, rs3811419, is shown in purple, the strength linkage disequilibrium (r^2) of surrounding SNPs with the top SNP is indicated by the colored dots, the scale is presented in the legend. Grey represents unknown r^2 . The fine scale recombination rate is shown on the right side of the plots in blue.



Online Supplement:**Methods****1. Samples and phenotype definition**

AD cases and controls for the study were selected from the biorepository at the Center for Applied Genomics (CAG), which has a collection of over 400,000 samples including over 90,000 internal pediatric samples genotyped using standard genome-wide arrays and linked to subjects' electronic health records (EHRs). All subjects have consented to analyses and EHR mining from the full longitudinal record, which has a mean duration >5.5 years/subject. Mean age of these subjects is 11 years and 47% are of EA ancestry, 43% AA and 10% from other ancestral groups.

The AD phenotype was defined using an algorithm that mined EHR data to identify cases and controls. Cases were defined as subjects two months old or older with either: 1) two or more diagnoses of AD (defined by the ICD9 code 691.8), on separate calendar days, and one or more prescriptions of AD related medications (Table E1); or 2) three or more diagnosis of AD (ICD9 code 691.8), on separate calendar days. Exclusion criteria for cases included scabies, Wiscott-Aldrich syndrome, allergic purpura, ichthyosis congenita and chromosomal anomalies. Controls were defined as subjects with no ICD9 code for AD, no ICD9 codes for any skin condition, allergy, asthma, nutritional deficiencies or skin cancer (Table E2) and no relevant medications for AD (Table E1). The algorithm was validated internally at CAG and also externally at a second eMERGE site, the Marshfield Clinic, by manual chart review of 100 randomly-selected subjects (50 cases and 50 controls). Positive predictive values (PPVs) for cases and controls were then calculated for the sample sets.

2. Genotype imputation and population stratification assessment

Imputed genotypes were extracted from 36,000 CAG subjects, genotyped on Illumina Human610-Quad version 1, Illumina 550K SNP array, Illumina Infinium Global Screening Array, Illumina OmniExpress arrays, and the Affymetrix Axiom Hu_SNP array (Table E3). Quality control of the data included removing variants with minor allele frequency (MAF) <1%, missing from >5% of samples, and samples missing >5% of SNPs. Imputation was performed on each chip type separately on the Michigan Imputation Server¹ against release 1.1 of the Haplotype Research Consortium reference panel². Variants with an 'R²' score <0.5 were removed.

For principal component analysis, we used Eigenstrat 3.0³ on a set of 130,000 linkage disequilibrium (LD)-pruned SNPs overlapping the Illumina Human610-Quad, Hap550, and OmniExpress arrays, that were extracted from the imputed files. Samples were separated into AA and EA based on the first two principal components using a k-means clustering analysis. Ten principal components were re-generated for each cohort and included as covariates to control for population stratification in each individual analysis. Cryptic relatedness and duplicated samples were assessed by pairwise Identity-By-Descent values with PLINK⁴, calculated on the same set of 130,000 SNPs and a random sample from each pair was removed.

3. Data analysis

The imputation was performed per chip type, as stated below and illustrated in Table E4. The implementation of the phenotyping algorithm in our biorepository identified an African American sample of 1,006 cases and 2,054 controls, genotyped on the Illumina Hap550 and Quad (sample "AA1"). The imputed genotypes generated in each of these

chips (Illumina Hap550 and Quad) was then merged together (using bcftools, <https://samtools.github.io/bcftools/bcftools.html>), subjected to quality control, and analyzed.

Since the medical record data linked to each DNA sample in our biorepository was updated during the course of this study, the algorithm was run a second time and identified 365 cases and 479 controls genotyped on the Illumina Hap550, Quad and OmniExpress chips. Imputed genotypes were merged together, subjected to QC and analyzed (sample “AA2”). The Illumina Hap550, Quad and OmniExpress chips have a large overlap in their content with over 300k SNPs, so we did not expect to have issues related with the merging. As shown in the table below, the case-control numbers per chip and the allele frequencies of the top SNP, rs3811419, in each sample were very similar across the three cohorts.

The third sample (“AA3”) consisted of 1,939 cases and controls identified in the two runs of the algorithm. These samples were specifically genotyped for this project with the Infinium Global Screening Array. This array includes 640,000 variants combining multi-ethnic genome-wide content, curated clinical research variants, and quality control (QC) markers. The genome-wide content was selected for high imputation accuracy at minor allele frequencies of >1% across all 26 1000 Genomes Project populations.

This third cohort was analyzed independently and results from the GWAS on each sample (AA1, AA2, and AA3) were meta-analyzed together. The genomic inflation factors calculated for each of the samples and for the meta-analysis showed no population stratification or inflation of the p-values

3.1. Association analysis

SNPTEST was used for the association analysis. Ten principal components, chip type, and gender were used as covariates in the analysis. Variants with MAF<1% and variants not meeting Hardy-Weinberg equilibrium in controls ($p\text{-value} < 5 \times 10^{-8}$) were removed. METAL was used for the meta-analyses of the AA, EA and the combined sample using an inverse variance fixed effects method, with control for genomic inflation. Additionally, genomic inflation factors were calculated to assess potential population stratification in the analysis (Figure E1).

Results from the meta-analysis published by Paternoster in 2015⁵ were used for replication of significant SNPs. These SNPs were additionally evaluated in the AA subset from the SAGE II study⁵. Briefly, the SAGE II cohort is a sample of 365 AD cases and 510 controls of self-reported AA ancestry, where subjects were genotyped with the Axiom LAT1 array (World Array 4, Affymetrix). Genotyped SNPs passing quality control were phased using SHAPEIT⁶ and imputed using IMPUTE2⁷ against the 1000 Genomes Project haplotype reference panel (Phase I v3). Association testing was performed with PLINK⁴ adjusting by sex and genetic ancestry.

3.2. eQTL effects

A cis-expression quantitative trait locus (eQTL) effect was investigated for rs3811419, and variants in LD ($r^2 > 0.8$), rs4845602, rs7540799, rs11582525, rs4845606, rs3811418, rs3811417, rs72692781 by mining HaploReg v4⁸, the NCBI Genotype-Tissue Expression (GTEx) version 6⁹, and in 37 studies reporting eQTL effects in target tissues on AD. The list of studies queried are listed below:

1. Andiappan et al., 2015; PMID: 26259071

2. Barreiro et al., 2012; PMID: 22233810
3. Battle et al., 2014; PMID: 24092820
4. Brumpton et al., 2016; PMID: 27155841
5. Caliskan et al., 2015; PMID: 25874939
6. Davenport et al., 2016; PMID: 26917434
7. Di Narzo et al., 2016; PMID: 27336838
8. Dimas et al., 2009; PMID: 19644074
9. Ding et al., 2010; PMID: 21129726
10. Dixon et al., 2007; PMID: 17873877
11. Fairfax et al., 2012; PMID: 22446964
12. Fairfax et al., 2014; PMID: 24604202
13. Fehrmann et al., 2011; PMID: 21829388
14. Ferraro et al., 2014; PMID: 24610777
15. Franco et al., 2013; PMID: 23878721
16. Grundberg et al., 2012; PMID: 22941192
17. Hao et al., 2012; PMID: 23209423
18. Huang et al., 2015; PMID: 25951796
19. Kasela et al., 2017; PMID: 28248954
20. Kim et al., 2014; PMID: 25327457
21. Kukurba et al., 2016; PMID: 27197214
22. Lappalainen et al., 2013; PMID: 24037378
23. Lee et al., 2014; PMID: 24604203
24. Lloyd Jones et al., 2017; PMID: 28065468
25. Luo et al., 2016; PMID: 26102239
26. Murphy et al., 2010; PMID: 20833654
27. Naranbhai et al., 2015; PMID: 26151758
28. Nedelec et al., 2016; PMID: 27768889
29. Peters et al., 2016; PMID: 27015630
30. Quach et al., 2016; PMID: 27768888
31. Raj et al., 2014; PMID: 24786080
32. Walsh et al., 2016; PMID: 27140173
33. Westra et al., 2013; PMID: 24013639
34. Yao et al., 2017; PMID: 28285768
35. Ye et al., 2014; PMID: 25214635
36. Zeller et al., 2010; PMID: 20502693
37. Zhernakova et al., 2017; PMID: 27918533

Cis-eQTL effects were explored for genes located within 1Mb from the locus (chr1:151305819-152305819), which included *PIP5K1A*, *PSMD4*, *ZNF687*, *PI4KB*, *RFX5*, *SELENBP1*, *PSMB4*, *POGZ*, *CGN*, *TUFT1*, *SNX27*, *CELF3*, *RIIAD1*, *MRPL9*, *OAZ3*, *TDRKH*, *LINGO4*, *RORC*, *C2CD4D*, *THEM5*, *THEM4*, *S100A10*, *S100A11*, *TCHHL1*, *TCHH*, *RPTN*, *HRNR*, *FLG*, *FLG2*, *FLG-AS*, *CRNN*, *LCE5A*. The threshold for significance was set as $p\text{-value}=7.4 \times 10^{-5}$ (7 SNPs, 3 tissues -Skin - Sun Exposed (Lower leg), Skin - Not Sun Exposed (Suprapubic), Whole blood- in 32 genes).

3.3. Analysis of methylation data

Methylation status on blood was ascertained for a subset of subjects recruited by the Center for Applied Genomics through the use of the Infinium HumanMethylation450 BeadChip Kit using manufacturer's protocols. Data were exported from the Illumina GenomeStudio as GenomeStudio output text files containing probe level summarized information. The GenomeStudio output files were loaded into the R statistical package (r-project.org) through the use of the lumi package¹⁰⁻¹³. The lumi package was used for preprocessing of data, which involved quantile color balance adjustment and background level correction and simple scaling normalization (ssn). M-value density and CpG-site Intensity were plotted after normalization, and aberrant chips were removed. Methylation probes known to overlap with common single nucleotide polymorphisms, 92667 probes identified in the IMA R package¹⁴, were excluded from the data set, resulting in a dataset containing 392910 probes. M-values (the log2 ratio between the methylated and unmethylated probe intensities) were extracted and stored as a matrix. Additive genotypes at rs3811419 and variants in LD ($r^2 \geq 0.8$) rs3811418, rs3811417, rs11582525, rs4845602, rs4845606, rs72692781, and rs7540799 for African American subjects were extracted from existing genotyping data using plink¹⁵. Methylation distribution for probe cg22228337 is represented in Figure E2.

Figure Legends

Figure E1. Quantile-quantile plots show $-\log_{10}(\text{p-value})$ of observed genome-wide association results against expected association results for atopic dermatitis. Genomic inflation factors are 1.03 for African Americans (A) and 1.01 for European Americans (B)

Figure E2. A plot of the distribution of the M-values at cg22228337, after removal of outliers, has been included. Outliers were defined as more than 3 standard deviations away from the mean.

References

1. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A et al. Next-generation genotype imputation service and methods. *Nature genetics* 2016; 48(10): 1284-1287.
2. Loh PR, Danecek P, Palamara PF, Fuchsberger C, Y AR, H KF et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nature genetics* 2016; 48(11): 1443-1448.
3. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* 2006; 38(8): 904-909.
4. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015; 4: 7.
5. Paternoster L SM, Waage J, Baurecht H, Hotze M, Strachan DP, Curtin JA, et al. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nature genetics* 2015; 47(12): 1449-1456.
6. 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. *Nature communications* 2014; 5: 3934.
7. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics* 2009; 5(6): e1000529.
8. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic acids research* 2012; 40(Database issue): D930-934.
9. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nature genetics* 2013; 45(6): 580-585.
10. Du P, Kibbe WA, Lin SM. nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. *Biol Direct* 2007; 2: 16.
11. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008; 24(13): 1547-1548.
12. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; 11: 587.
13. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic acids research* 2008; 36(2): e11.

14. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB et al. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. *Bioinformatics* 2012; 28(5): 729-730.
15. 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. *American journal of human genetics* 2007; 81(3): 559-575.
16. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nature genetics*. 2017;49(12):1752-7.
17. Ellinghaus D, Baurecht H, Esparza-Gordillo J, Rodriguez E, Matanovic A, Marenholz I et al. High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. *Nature genetics* 2013; 45(7): 808-812.
18. Esparza-Gordillo J, Weidinger S, Folster-Holst R, Bauerfeind A, Ruschendorf F, Patone G et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nature genetics* 2009; 41(5): 596-601.
19. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M et al. Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. *Nature genetics* 2012; 44(11): 1222-1226.
20. 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. *Nature genetics* 2011; 44(2): 187-192.
21. Esparza-Gordillo J, Schaarschmidt H, Liang L, Cookson W, Bauerfeind A, Lee-Kirsch MA et al. A functional IL-6 receptor (IL6R) variant is a risk factor for persistent atopic dermatitis. *The Journal of allergy and clinical immunology* 2013; 132(2): 371-377.
22. Sun LD, Xiao FL, Li Y, Zhou WM, Tang HY, Tang XF et al. Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. *Nature genetics* 2011; 43(7): 690-694.
23. Schaarschmidt H, Ellinghaus D, Rodriguez E, Kretschmer A, Baurecht H, Lipinski S et al. A genome-wide association study reveals 2 new susceptibility loci for atopic dermatitis. *The Journal of allergy and clinical immunology* 2015; 136(3): 802-806.

Tables

Table E1. Medications related to atopic dermatitis by subtype used for case inclusion and control exclusion by the algorithm. Compounds are presented by the generic name

1. Topical corticosteroids: Alclometasone, amcinonide, betamethasone, clobetasol, clocortolone, desonide, desoximetasone, diflorasone, fluocinonide, fluocinolone, flurandrenolide, fluticasone, halcinonide, halobetasol, hydrocortisone, mometasone, prednicarbate, triamcinolone
2. Oral corticosteroids: Dexamethasone, hydrocortisone, methylprednisolone, prednisone, prednisolone, triamcinolone
3. Topical calcineurin inhibitors: Cyclosporine, pimecrolimus, tacrolimus
4. Antihistamines: Cetirizine, clemastine, chlorpheniramine, cyproheptadine, desloratadine, diphenhydramine, fexofenadine, hydroxyzine, loratadine, mizolastine, promethazine
5. Other: Azathioprine, doxepin, emollients, IFN γ , methotrexate, mycophenolate mofetil, phototherapy (UVA, UVAB, BB-UVB, nb-UVB, PUVA)

Table E2. ICD9 codes applied by the atopic dermatitis algorithm for subject exclusion in controls

ICD9	Name
133.0	Scabies
172.0-173.99	Malignant melanoma of skin
232.x	Carcinoma in situ of skin
260.0-269.99	Nutritional deficiencies

279.x	Disorders involving the immune mechanism
287.0	Allergic purpura
288.1	Functional disorders of polymorphonuclear neutrophils
372.14	Other chronic allergic rhinoconjunctivitis
493.x	Asthma
477.x	Allergic rhinitis
684	Impetigo
690.0-698.99	Other inflammatory conditions of skin and subcutaneous tissue
708.x	Urticaria
757.1	Ichthyosis congenita
758.x	Chromosomal anomalies
782.1	Rash and other nonspecific skin eruption
995.2-995.3	Unspecified adverse effect of drug medicinal and biological substance not elsewhere classified; and unspecified allergy
V13.3	Personal Hx of diseases of skin and subcutaneous tissue
V14.x	Personal Hx of allergy to medicinal agents
V15.0x	Personal Hx of allergy, other than to medicinal agents
V10.82	Personal Hx of malignant melanoma of skin
V10.83	Personal Hx of malignant neoplasm of skin

Table E3. Arrays used for the genotyping of the African American and European American samples and total numbers

Array	African American	European American	Total

Hapmap550	1,993	2,500	4,493
Global Screening Array	1,939	269	2,208
HumanQuad610	1,764	2,823	4,587
OmniExpress-12v1	144	852	996
HumanOmniExpress-24v1	3	44	47
Axiom_Hu_SNP	0	14	14
Total	5,843	6,502	12,345

Table E4. Summary of cases and controls in each African American sample used in the GWAS per chip type, with case-control ratios, and minor allele frequency (MAF) of rs3811419 in the entire sample and in cases and controls.

	N Case	N Control	Case:Control	MAF	MAF cases	MAF controls
AA1	1,006	2,054		0.038	0.045	0.034
Hap550v1	143	289	2.02			
Hap550v3	386	836	2.16			
Quad	477	929	1.95			
AA2	365	479		0.035	0.048	0.025
Hap550v1	38	44	1.16			
Hap550v3	110	147	1.34			
Quad610	156	202	1.41			
OmniExpress	61	86	1.29			
AA3	1,049	890	0.85	0.036	0.041	0.030
GSA						

Table E5. Results from the meta-analysis in African Americans of the loci associated with AD with $p\text{-value} < 1 \times 10^{-5}$ in the CAG sample and followed-up in the SAGE II study. The effect is the coefficient of the logistic regression (beta). SE stands for standard error. Only the most significant SNP of the haplotype associated is presented in the table.

MarkerName	chr	bp	Risk allele	Freq	Effect (beta)	SE	P-value_meta-analysis
rs3811419	1	151805819	t	0.037	0.573	0.098	5.64E-09
rs114388740	18	28994249	a	0.043	-0.462	0.091	3.49E-07
rs74033913	15	98255864	a	0.860	0.264	0.052	4.56E-07
rs72832217	2	105546586	t	0.024	0.596	0.120	6.89E-07
rs6564272	16	73529760	t	0.322	-0.191	0.039	7.91E-07
rs8062972	16	9669993	t	0.781	-0.215	0.044	9.97E-07
rs75200181	2	161115896	t	0.963	-0.456	0.094	1.36E-06
rs2256769	1	61997226	a	0.024	-0.542	0.113	1.66E-06
rs140471137	2	112886213	t	0.024	0.555	0.117	1.93E-06
rs116345101	9	31518517	t	0.039	0.438	0.092	2.15E-06
rs2368497	2	241197917	a	0.710	-0.189	0.040	2.36E-06
rs61118733	8	61650755	t	0.953	-0.392	0.085	3.69E-06
rs7973708	12	5056924	a	0.529	0.170	0.037	4.12E-06

rs111977231	8	96089884	t	0.955	0.397	0.087	4.74E-06
rs12034401	1	9878444	t	0.035	0.473	0.103	4.81E-06
rs638037	11	97534229	a	0.463	0.162	0.036	7.69E-06
rs2520681	5	10024743	a	0.076	0.296	0.066	7.73E-06
rs2981011	3	128187427	c	0.211	-0.200	0.045	8.49E-06
rs7506383	18	55887866	a	0.347	-0.170	0.038	8.60E-06
rs3009192	1	120390354	a	0.109	-0.266	0.060	8.67E-06
rs2266023	1	120408549	t	0.149	-0.265	0.060	8.78E-06
rs844346	1	120402790	a	0.109	0.265	0.060	9.19E-06
rs79868636	3	62128518	a	0.048	0.380	0.086	9.28E-06
rs856658	1	120404295	a	0.109	-0.264	0.060	9.83E-06
rs860674	1	120402764	a	0.109	0.263	0.060	1.10E-05
rs6779664	3	179252544	a	0.294	-0.175	0.040	1.31E-05
rs7737797	5	38136423	a	0.737	-0.177	0.041	1.36E-05
rs79758205	7	97730565	t	0.026	0.496	0.114	1.39E-05
rs10820744	9	107671935	a	0.341	-0.168	0.039	1.48E-05
rs835578	1	120406583	t	0.109	-0.256	0.060	1.71E-05

rs959394	5	3706014	t	0.724	-0.172	0.040	1.88E-05
rs73362311	8	141571623	a	0.016	-0.647	0.151	1.89E-05
rs75546071	7	6583202	c	0.885	-0.243	0.057	2.00E-05
rs28573289	2	127924559	c	0.769	-0.181	0.043	2.03E-05
rs11233245	11	71059914	a	0.714	0.169	0.040	2.06E-05
rs2487571	1	120394328	a	0.108	-0.254	0.060	2.07E-05
rs61259525	8	95804462	a	0.843	0.215	0.051	2.16E-05
rs74653337	4	17239466	t	0.171	0.208	0.049	2.16E-05
rs835565	1	120400343	t	0.109	0.254	0.060	2.23E-05
rs3009190	1	120389422	t	0.108	0.248	0.059	2.27E-05
rs12664385	6	131693099	t	0.100	0.260	0.061	2.32E-05
rs77559038	14	97403754	c	0.032	-0.445	0.105	2.36E-05
rs2487572	1	120394844	a	0.109	-0.252	0.060	2.37E-05
rs12319868	12	114845013	a	0.024	-0.498	0.118	2.46E-05
rs6665826	1	172851231	a	0.550	0.154	0.037	2.80E-05
rs1997323	5	87220656	a	0.030	0.433	0.105	3.52E-05
rs74546512	11	17625409	t	0.053	0.334	0.081	3.86E-05

rs57606030	1	53609303	a	0.806	-0.187	0.046	3.90E-05
rs111983338	22	25713592	t	0.067	-0.299	0.073	4.43E-05
rs527481	11	62603912	a	0.071	0.302	0.074	4.77E-05
rs2718398	7	88367986	a	0.541	0.161	0.040	6.63E-05
rs526849	11	59767401	t	0.359	0.149	0.038	9.80E-05
rs74120641	10	21054053	a	0.824	-0.182	0.048	1.31E-04
rs80110481	7	124934300	a	0.095	-0.234	0.061	1.33E-04
rs185399466	3	145881031	t	0.026	0.434	0.116	1.78E-04
rs9271363	6	32586732	c	0.174	0.173	0.047	2.29E-04
rs7354817	1	10235712	a	0.149	-0.196	0.054	2.91E-04
rs79717344	3	134643076	a	0.960	0.325	0.094	5.43E-04

Table E6. eQTL effects surpassing the threshold for significance of $p\text{-value}=7.4\times 10^{-5}$ for rs3811419 y variants in LD. NA=Not available. *For eQTL effects queried in GTEx, the effect size magnitude has no biological interpretation

SNP (allele)	Gene Symbol	Effect*	P-Value	Tissue	Source
rs3811418 (G)	<i>THEM4</i>	NA	6.22x10-14	Whole blood	HaploReg (Westra, 2013)

rs7540799 (T)	<i>THEM4</i>	NA	7.01x10 ⁻¹²	Whole blood	HaploReg (Westra, 2013)
rs4845606 (A)	<i>THEM4</i>	NA	7.69x10 ⁻¹³	Whole blood	HaploReg (Westra, 2013)
rs3811417 (C)	<i>THEM4</i>	NA	5.50x10 ⁻¹³	Whole blood	HaploReg (Westra, 2013)
rs4845602 (A)	<i>THEM4</i>	NA	2.51x10 ⁻¹²	Whole blood	HaploReg (Westra, 2013)
rs11582525 (G)	<i>THEM4</i>	NA	2.68x10 ⁻¹⁰	Whole blood	HaploReg (Westra, 2013)
rs3811419 (T)	<i>THEM4</i>	NA	1.82x10 ⁻¹⁰	Whole blood	HaploReg (Westra, 2013)
rs11582525 (G)	<i>THEM4</i>	-0.27	1.50x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs72692781 (T)	<i>THEM4</i>	-0.26	2.00x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs3811417 (C)	<i>THEM4</i>	-0.26	2.50x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs72692781 (T)	<i>C2CD4D</i>	-0.21	3.20x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs11582525 (G)	<i>C2CD4D</i>	-0.24	5.50x10 ⁻⁰⁵	Skin - Not Sun Exposed (Suprapubic)	GTEx
rs3811418 (G)	<i>THEM4</i>	-0.26	6.40x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs7540799 (T)	<i>C2CD4D</i>	-0.20	7.30x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx
rs4845606 (A)	<i>THEM4</i>	-0.24	7.40x10 ⁻⁰⁵	Skin - Sun Exposed (Lower leg)	GTEx

Table E7. Replication of results reported by Ferreira et al in allergic disease in our CAG African American cohort. Only variants with p-value<0.05 in the African American sample are presented. Table adapted from Ferreira et al¹⁶.

Chr	Bp	SNP	Effect Allele (Ferreira et al, 2017)	OR	P-value	Gene context	Effect Allele in AA	Effect in AA (beta)	p-value AA	Direction of effect consistent between studies
10	8605553	rs11255753	T	1.041	2.00E-12	GATA3--[]SFTA1P	T	0.1624	4.62E-03	Yes
14	75968608	rs9323612	A	1.032	8.60E-09	JDP2-[]-BATF	A	0.1075	6.95E-03	Yes
5	131996500	rs848	A	1.068	1.50E-24	[IL13]	A	0.0964	0.01327	Yes
5	131989136	rs3091307	G	1.062	3.60E-21	RAD50-[]-IL13	A	-0.0988	0.01494	Yes
4	38798648	rs5743618	C	1.1	3.30E-58	[TLR1]	A	-0.1371	0.01579	Yes
10	8841669	rs2025758	T	1.041	4.70E-15	GATA3---[]---SFTA1P	T	-0.0964	0.02207	No
11	65551957	rs479844	G	1.038	1.60E-13	AP5B1-[]-OVOL1	A	-0.1071	0.02378	Yes
6	138195151	rs5029937	G	1.081	2.40E-08	[TNFAIP3]	T	0.0834	0.04561	No
5	110166083	rs6869502	T	1.079	6.40E-29	SLC25A46-[]--TSLP	A	-0.0777	0.04621	Yes
5	14610309	rs16903574	G	1.071	1.40E-12	[FAM105A]	C	0.3304	0.05289	No
2	242698640	rs34290285	G	1.082	4.00E-33	[D2HGDH]	A	0.1347	0.05571	No
17	38069076	rs921650	A	1.059	5.70E-30	[GSDMB]	A	0.0912	0.05865	Yes
5	35862841	rs7717955	C	1.073	9.10E-36	[IL7R]	T	-0.098	0.06095	Yes

5	176782218	rs13153019	C	1.035	1.30E-08	LMAN2-[]-RGS14	T	0.1131	0.07115	No
10	9064361	rs1444789	C	1.066	1.50E-22	GATA3---[]---SFTA1P	T	-0.0762	0.07281	Yes
12	56401085	rs10876864	G	1.047	1.40E-19	SUOX-[]-IKZF4	A	-0.0817	0.0851	Yes
1	173146921	rs4090390	A	1.048	1.30E-15	TNFSF18--[]-TNFSF4	A	0.0685	0.09003	Yes
7	20376018	rs10486391	A	1.03	6.80E-09	[ITGB8]	A	0.0696	0.09958	Yes
12	111932800	rs7137828	T	1.033	2.20E-10	[ATXN2]	T	0.105	0.1042	Yes
15	61068347	rs10519067	G	1.055	9.30E-13	[RORA]	A	-0.0662	0.1097	Yes
12	50345671	rs11169225	A	1.045	1.20E-11	[AQP2]	A	0.1471	0.1115	Yes
4	4775401	rs10033073	G	1.04	1.20E-10	STX18--[]-MSX1	A	-0.0681	0.1171	Yes
18	60009814	rs4574025	T	1.03	6.80E-09	[TNFRSF11A]	T	0.0606	0.1197	Yes
17	38149033	rs11652139	A	1.051	7.50E-22	[PSMD3]	A	-0.1186	0.1264	No
3	188402586	rs17607589	C	1.053	1.80E-14	[LPP]	T	-0.104	0.1332	Yes
1	154426970	rs2228145	C	1.038	4.30E-13	[IL6R]	A	-0.0869	0.1346	Yes
5	118684297	rs250308	T	1.031	4.00E-09	[TNFAIP8]	T	-0.062	0.1409	No
5	110401872	rs1837253	C	1.07	1.60E-31	SLC25A46--[]-TSLP	T	-0.0633	0.1418	Yes
9	6051399	rs343478	G	1.033	2.60E-10	RANBP6-[]-IL33	A	0.0574	0.1498	No
1	8482078	rs301806	T	1.048	1.80E-20	[RERE]	T	-0.0603	0.1521	No
6	32626403	rs34004019	A	1.101	3.80E-52	HLA-DQA1-[]-HLA-DQB1	A	-0.0674	0.159	No
16	11491007	rs12596613	C	1.032	6.10E-09	RMI2-[]-LITAF	C	-0.0602	0.1642	No

11	128158189	rs56129466	A	1.047	1.90E-13	KIRREL3-AS3---[]--ETS1	A	0.1001	0.1784	Yes
10	64382359	rs2893907	C	1.031	1.80E-09	[ZNF365]	A	-0.0723	0.1837	Yes
17	47398070	rs9889262	A	1.043	9.70E-16	[ZNF652]	A	0.0818	0.185	Yes
10	9049253	rs12413578	C	1.095	1.30E-27	GATA3---[]---SFTA1P	T	-0.1759	0.206	Yes
15	91045408	rs3540	G	1.036	3.30E-11	[IQGAP1]	A	-0.0503	0.2194	Yes
3	33069091	rs6776757	G	1.033	3.10E-10	[GLB1]	A	-0.0475	0.2265	Yes
7	50253897	rs17664743	A	1.042	6.20E-11	C7orf72-[]-IKZF1	A	0.0518	0.23	Yes
12	48196982	rs55726902	G	1.051	2.60E-16	[HDAC7]	A	0.1017	0.2333	No
10	104225832	rs10883723	C	1.03	1.60E-08	C10orf95-[]-ACTR1A	T	0.0592	0.2425	No
9	5064193	rs16922576	C	1.036	3.20E-10	[JAK2]	T	0.0613	0.2476	No
2	64836267	rs4671601	C	1.039	8.80E-09	[LOC339807]	T	-0.0624	0.2479	Yes
11	111470567	rs7130753	C	1.045	7.00E-15	LAYN-[]-SIK2	T	-0.0518	0.2493	Yes
10	8936162	rs11255968	C	1.09	7.70E-09	GATA3---[]---SFTA1P	T	0.0603	0.2586	No
5	110470137	rs6594499	C	1.075	4.60E-46	WDR36-[]-CAMK4	A	-0.0434	0.2773	Yes
12	57489709	rs1059513	T	1.084	1.00E-22	[STAT6]	T	0.0954	0.2791	Yes
12	121363724	rs6489785	T	1.043	1.60E-15	SPPL3-[]-HNF1A-AS1	T	0.0476	0.2832	Yes
7	20560996	rs6461503	T	1.039	1.70E-14	ITGB8--[]-ABCB5	T	-0.0455	0.2933	No
10	6094697	rs61839660	T	1.08	4.40E-19	[IL2RA]	T	-0.1432	0.3058	No
2	198950240	rs1064213	G	1.035	5.40E-12	[PLCL1]	A	0.0435	0.3117	No

11	76293758	rs7936323	A	1.088	2.20E-63	WNT11--[-]LRRC32	A	0.0408	0.3153	Yes
9	6208030	rs144829310	T	1.09	1.20E-35	RANBP6--[-]IL33	T	0.0818	0.3193	Yes
1	167431352	rs2988277	C	1.04	4.00E-14	[CD247]	T	0.0611	0.3209	No
6	157419508	rs9383820	C	1.037	1.20E-08	[ARID1B]	T	-0.0381	0.3361	Yes
6	31574525	rs28895016	C	1.097	9.40E-16	NCR3-[-]AIF1	T	0.1991	0.4032	No
4	123316076	rs4145717	T	1.059	9.20E-27	[ADAD1]	T	0.0502	0.4033	Yes
1	35681738	rs76167968	T	1.056	1.30E-08	SFPQ-[-]ZMYM4	T	0.0311	0.431	Yes
1	2510755	rs10910095	G	1.042	2.70E-08	TNFRSF14-[-]FAM213B	A	0.0609	0.4378	No
14	35761675	rs1048990	G	1.039	1.00E-08	[PSMA6]	C	-0.0656	0.4403	Yes
5	159909345	rs2910162	G	1.033	2.50E-09	MIR3142-[-]MIR146A	A	-0.0368	0.4508	Yes
20	62322699	rs6011033	G	1.047	3.50E-14	[RTEL1]	A	-0.0598	0.458	Yes
11	76343428	rs11236814	A	1.068	4.00E-14	WNT11--[-]LRRC32	A	0.0398	0.4731	Yes
11	76299431	rs55646091	A	1.179	2.30E-40	WNT11--[-]LRRC32	A	0.205	0.4914	Yes
4	103593898	rs227275	C	1.034	3.70E-11	[MANBA]	A	0.0262	0.5062	No
22	41816652	rs5758343	A	1.048	4.80E-14	TEF-[-]TOB2	A	0.0555	0.5204	Yes
2	102941311	rs10865050	G	1.13	7.00E-61	[IL18R1]	A	-0.0287	0.5215	Yes
3	187633268	rs519973	A	1.034	4.50E-10	BCL6--[-]LPP-AS2	A	0.0264	0.5326	Yes
2	112269127	rs13403656	A	1.045	2.20E-08	BCL2L11--[-]ANAPC1	A	-0.0286	0.537	No
1	151796742	rs11204896	C	1.063	2.40E-12	[RORC]	C	0.0566	0.5405	Yes

3	101242751	rs13088318	A	1.031	8.60E-09	FAM172BP-[]-TRMT10C	A	-0.0449	0.5446	No
14	103235012	rs9989163	A	1.029	1.90E-08	RCOR1-[]-TRAF3	A	0.0239	0.5448	Yes
8	128814091	rs6990534	A	1.042	6.40E-14	[MYC]	A	0.0256	0.5465	Yes
10	6074451	rs4747846	C	1.036	1.00E-11	[IL2RA]	C	-0.0267	0.5497	No
1	161185058	rs2070901	T	1.039	1.30E-11	NDUFS2-[]-FCER1G	T	0.0237	0.5499	Yes
2	143831599	rs74847330	A	1.048	1.80E-09	KYNU-[]-ARHGAP15	A	0.0642	0.5576	Yes
11	95425526	rs59593577	C	1.053	1.60E-11	SESN3--[]-FAM76B	T	0.0394	0.5591	No
15	67448363	rs56375023	A	1.073	8.20E-32	[SMAD3]	A	-0.0452	0.566	No
6	33647058	rs10947428	C	1.046	3.50E-13	[ITPR3]	T	-0.0218	0.5818	Yes
11	118743286	rs12365699	G	1.061	5.10E-18	DDX6-[]-CXCR5	A	0.0564	0.5943	No
15	41782684	rs12440045	C	1.033	4.90E-10	RTF1-[]-ITPKA	A	-0.0312	0.6129	Yes
3	188133336	rs60946162	T	1.041	8.60E-15	[LPP]	T	-0.0271	0.627	No
9	131613191	rs12551834	G	1.058	3.00E-09	C9orf114-[]-LRRRC8A	A	0.0459	0.649	No
6	31323012	rs2854001	A	1.059	1.20E-19	[HLA-B]	A	-0.0324	0.6497	No
14	68754417	rs2104047	T	1.042	1.60E-13	[RAD51B]	T	-0.0197	0.6627	No
19	33721455	rs10414065	C	1.098	6.10E-18	SLC7A10-[]-CEBPA	T	0.0472	0.6777	No
3	72394852	rs61192126	T	1.037	8.90E-11	LINC00870--[]-RYBP	T	-0.021	0.6804	No
9	123650534	rs10760123	T	1.032	5.20E-09	PHF19-[]-TRAF1	T	0.0185	0.6851	Yes
17	43430696	rs7214661	G	1.032	1.20E-08	MAP3K14-[]-	A	0.0131	0.7362	No

						ARHGAP27				
5	140925362	rs740474	C	1.034	5.60E-11	[DIAPH1]	T	-0.0172	0.7455	Yes
1	226914734	rs697852	A	1.041	1.60E-09	[ITPKB]	A	0.0144	0.7556	Yes
13	41173408	rs4943794	C	1.043	7.20E-12	[FOXO1]	C	0.0215	0.7657	Yes
1	25251923	rs760805	T	1.038	6.40E-13	[RUNX3]	A	0.0144	0.7917	No
16	11277358	rs11644510	C	1.072	6.10E-38	CLEC16A-[]-RMI2	T	-0.0102	0.7939	Yes
13	73627275	rs9573092	A	1.03	2.70E-08	PIBF1-[]-KLF5	A	0.012	0.8101	Yes
5	141494934	rs10068717	T	1.042	4.80E-15	[NDFIP1]	T	-0.0103	0.8198	No
21	44846426	rs76081789	T	1.065	1.30E-08	[SIK1]	T	-0.0091	0.8211	No
6	90987512	rs2134814	C	1.046	1.70E-17	[BACH2]	C	0.0104	0.8544	Yes
1	150265704	rs7512552	C	1.031	1.40E-09	C1orf54-[]-MRPS21	T	0.0091	0.8636	No
3	187793833	rs2030030	T	1.041	1.00E-08	BCL6--[]-LPP-AS2	T	0.0097	0.8767	Yes
5	131799626	rs3749833	C	1.039	3.30E-11	[C5orf56]	T	0.0095	0.88	No
6	106667535	rs9372120	G	1.042	4.20E-11	[ATG5]	T	0.0111	0.8854	No
18	52336175	rs4801001	T	1.031	5.90E-09	DYNAP-[]--RAB27B	T	-0.0053	0.9034	No
6	31351664	rs2507978	G	1.035	1.20E-10	HLA-B-[]-MICA	A	-0.0045	0.91	Yes
6	128294709	rs35469349	A	1.037	2.30E-10	[PTPRK]	A	-0.0067	0.9101	No
2	228707862	rs13384448	T	1.042	2.80E-12	CCL20-[]-DAW1	T	0.0043	0.9115	Yes
2	113590467	rs1143633	C	1.034	1.70E-10	[IL1B]	T	0.0051	0.9168	No

3	196372546	rs80064395	C	1.07	1.60E-12	FBXO45-[]-CEP19	T	-0.0093	0.9194	Yes
21	36467830	rs73205303	A	1.044	7.90E-10	[RUNX1]	A	-0.0095	0.9268	No
20	50157837	rs3787184	A	1.049	1.10E-12	[NFATC2]	A	-0.0038	0.9444	No
2	234115629	rs1057258	C	1.045	1.40E-10	[INPP5D]	T	-0.0027	0.9454	Yes
7	77018542	rs4296977	C	1.055	2.10E-13	[GSAP]	T	0.0067	0.9492	No
2	8442248	rs10174949	G	1.066	7.30E-31	[LINC00299]	A	-0.0037	0.9533	Yes
6	33046752	rs3097670	G	1.064	7.70E-12	[HLA-DPA1]	C	0.0051	0.9637	No
2	102926362	rs12470864	A	1.057	4.20E-26	IL1RL2-[]-IL18R1	A	-0.0021	0.9682	No
2	112388538	rs4848612	A	1.038	2.30E-10	BCL2L11--[]-ANAPC1	A	-0.0013	0.9761	No
5	40492655	rs7714574	T	1.032	5.90E-10	DAB2---[]-PTGER4	T	0.0006	0.9886	Yes
7	28156887	rs6977955	T	1.046	7.10E-13	[JAZF1]	T	0.0005	0.989	Yes
1	152285861	rs61816761	A	1.224	7.40E-21	[FLG]	NA	NA	NA	NA
1	152179152	rs12123821	T	1.111	6.80E-17	RPTN-[]-HRNR	NA	NA	NA	NA
17	4521473	rs71368508	C	1.124	2.00E-09	SMTNL2-[]-ALOX15	NA	NA	NA	NA
17	38764524	rs112401631	A	1.26	2.20E-26	CCR7-[]-SMARCE1	NA	NA	NA	NA
1	172700868	rs1102705	G	1.058	3.10E-10	FASLG-[]-TNFSF18	NA	NA	NA	NA
3	121652141	rs75557865	G	1.029	1.60E-08	[SLC15A2]	NA	NA	NA	NA
3	141321836	rs10663129	ACT	1.042	1.10E-13	[RASA2]	NA	NA	NA	NA
4	123454110	rs150254607	ATAT	1.078	4.60E-14	IL2-[]-IL21	NA	NA	NA	NA

5	110159879	rs1814576	C	1.121	1.40E-22	SLC25A46-[]--TSLP	NA	NA	NA	NA
6	29893575	rs9259819	G	1.036	2.40E-09	[HLA-J]	NA	NA	NA	NA
6	167390671	rs72033857	C	1.057	1.20E-09	RNASET2-[]-MIR3939	NA	NA	NA	NA
8	81292599	rs7824394	A	1.05	3.50E-20	MIR5708--[]--ZBTB10	NA	NA	NA	NA
10	9032555	rs72782676	C	1.303	3.20E-11	GATA3---[]---SFTA1P	NA	NA	NA	NA
12	123742692	rs63406760	T	1.047	3.00E-13	C12orf65-[]-CDK2AP1	NA	NA	NA	NA
14	38097001	rs111914382	TG	1.036	8.10E-09	FOXA1-[]--TTC6	NA	NA	NA	NA
17	38770641	rs11464691	TA	1.052	1.40E-21	CCR7-[]-SMARCE1	NA	NA	NA	NA
17	40414862	rs7207591	A	1.038	1.40E-09	[STAT5B]	NA	NA	NA	NA
20	52208356	rs2766678	G	1.057	5.00E-18	[ZNF217]	NA	NA	NA	NA

Table E8. p-values for reported SNPs in known AD loci in African Americans. In bold, p-value<0.05 is highlighted. AA stands for African Americans.

Chr	Bp	SNP	Effect Allele	OR	P-value	Gene context	Effect Allele in AA	Effect in AA (beta)	p-value AA	Direction of effect consistent between studies	Reference
10	6038853	rs6602364	A	0.98	0.137*	<i>CCDC80/CD200R1L</i>	A	-0.0074	0.854	Yes	5

10	64398466	rs10995251	G	1	0.744*	<i>OR10A3/NLRP10</i>	T	0.0264	0.665	Yes	5
10	64369999	rs2944542	C	0.94	1.2x10-6*	<i>ZNF365</i>	C	-0.0909	0.191	Yes	5
11	36432024	rs12295535	G	0.91	1.4X10-7*	<i>CD207/VAX2</i>	A	0.0591	0.229	Yes	5
11	36371757	rs2592555	A	0.93	1.4x10-9*	<i>MIR5708/ZBTB10</i>	A	-0.0533	0.188	Yes	5
11	7977161	rs4312054	T	1.11	1.5x10-13*	<i>IL18R1/IL18RAP</i>	--	--	--	--	5
11	7968359	rs878860	T	0.91	1.5x10-6*	<i>MICB</i>	--	--	--	--	5
11	65559266	rs10791824	G	1.08	1.5x10-9*	<i>IL15RA/IL2RA</i>	C	0.0277	0.486	Yes	5
11	65551957	rs479844	T	1.11	1.8x10-10*	<i>PPP2R3C</i>	T	0.0232	0.665	Yes	5
11	76301316	rs7927894	G	1.12	2.1x10-19*	<i>OVOL1</i>	A	-0.127	1.29x10-3	Yes	5
11	76281593	rs2212434	C	0.92	2.5x10-11*	<i>CLEC16A-DEX</i>	T	0.0744	0.059	Yes	5, 17
11	128187383	rs7127307	C	0.93	2.9x10-8	<i>IL7R/CAPSL</i>	T	0.04	0.525	Yes	5
14	35559126	rs2038255	A	0.93	3.4x10-8	<i>LINC00299/-</i>	A	-0.0821	0.145	Yes	5
16	11229589	rs2041733	A	1.09	3.5x10-8*	<i>PUS10</i>	A	0.0663	0.143	Yes	5
17	40528131	rs12951971	C	0.93	3.9x10-10*	<i>-/ETS1</i>	T	-0.0026	0.947	No	5
17	47440466	rs16948048	C	1.14	4.0x10-17*	<i>RAD50/IL13</i>	--	--	--	--	5
19	8789381	rs2164983	G	1.13	4.1x10-8	<i>STAT3</i>	--	--	--	--	5
1	150265704	rs7512552	R	1.08	4.2x10-9	<i>KIAA1109 (IL2)b</i>	--	--	--	--	5
1	152536650	rs61813875	T	1.09	4.6x10-13*	<i>C11orf30/LRRC32</i>	T	0.0255	0.549	Yes	5
1	154426970	rs2228145	T	1.07	4.9x10-8	<i>SFMBT1/RFT1</i>	T	0.0475	0.234	Yes	5

1	154419892	rs12730935	T	0.91	5.4x10-9*	<i>HLA-DRB1/HLA-DQA1</i>	T	0.0141	0.738	No	5
1	152178018	rs877776	G	1.61	5.6x10-29*	<i>FLG</i>	--	--	--	--	5, 18
20	52807221	rs16999165	A	0.39	6.1x10-11*	<i>IL6R</i>	--	--	--	--	5
20	62303115	rs4809219	C	0.9	7.0x10-13*	<i>RTEL1/TNFRSF6B</i>	A	0.0811	0.312	Yes	5
20	62309839	rs6010620	C	0.93	8.7x10-7*	<i>PRR5L</i>	--	--	--	--	5
2	71100105	rs112111458	T	0.93	9.1x10-10*	<i>C1orf51/MRPS21</i>	T	0.0091	0.864	No	5
2	61184651	rs4643526	C	1.28	5.85x10-20	<i>ZNF365</i>	T	-0.116	0.053	Yes	19
2	8495097	rs10199605	G	1.31	1.54x10-22	<i>OR10A3/NLRP10</i>	T	-0.01	0.836	Yes	19
2	103027103	rs6419573	T	1.19	1.65x10-8	<i>CYP24A1-PFDN4</i>	T	0.0192	0.807	Yes	19
2	102971865	rs13015714	G	1.27	8.36x10-18	<i>IL1RL1-IL18R1-IL-18RAP</i>	T	-0.0103	0.856	Yes	19
2	103094213	rs759382	G	1.25	2.77x10-16	<i>GLB1</i>	A	0.0425	0.387	No	19
2	167992286	rs6720763	G	1.29	1.56x10-19	<i>CCDC80</i>	T	-0.0101	0.815	Yes	19
3	53091164	rs7625909	T	1.4	8.38x10-20	<i>GP3M3</i>	T	-0.0746	0.082	No	19
3	33087200	rs6780220	G	1.18	7.83x10-9	<i>CARD11</i>	T	0.003	0.945	No	19
3	112376308	rs12634229	A	1.68	7.96x10-13	<i>PRR5L</i>	A	0.0795	0.228	Yes	17
3	112391174	rs1249910	G	1.17	2.92x10-9	<i>ZNF652</i>	A	-0.0671	0.091	Yes	17
4	123497697	rs17389644	C	2.06	6.01x10-11	<i>SLC9A4</i>	T	-0.0870	0.146	Yes	17
4	123243592	4:123243592	A	1.19	1.39x10-8	<i>IL2/IL21</i>	A	0.1032	0.289	Yes	17

5	35883734	rs10214237	A	0.89	1.1x10 ⁻¹³	<i>OVOL1</i>	A	-0.1071	0.024	Yes	20
5	109858821	rs7701890	A	1.16	7.1x10 ⁻⁹	<i>ACTL9/ADAMTS10</i>	A	0.0632	0.162	Yes	20
5	132049027	rs2897442	C	1.11	3.8x10 ⁻⁸	<i>IL-13-RAD50</i>	T	-0.0822	0.097	Yes	20
5	8789722	rs2918307	--	--	4.6x10 ⁻¹²	<i>KIF3A</i>	--	--	--	--	20
5	131991085	rs12188917	A	1.24	7.64x10 ⁻⁸	<i>C11orf30</i>	A	0.0163	0.693	Yes	18
6	32158319	rs176095	C	1.2	3.52x10 ^{-5*}	<i>LRRC32</i>	C	-0.0392	0.319	No	18
6	32575524	rs4713555	C	1.15	5.4x10 ⁻⁹	<i>IL6R</i>	A	-0.0868	0.135	Yes	21
6	31466217	rs145809981 (rs41293864)	G	1.17	3.00x10 ⁻⁸	<i>TNFRSF6B</i>	A	-0.0697	0.378	Yes	22
7	3128789	rs4722404	G	1.24	3.15x10 ⁻⁹	<i>TMEM232</i>	A	0.0502	0.26	No	22
8	81285892	rs6473227	C	1.29	4.37x10 ⁻⁸	<i>XIRP2(CYMA3)</i>	T	0.021	0.59	No	23
9	22373457	rs10738626	C	0.81	1.45x10 ⁻⁸	<i>DMRTA1</i>	T	0.0163	0.669	Yes	23

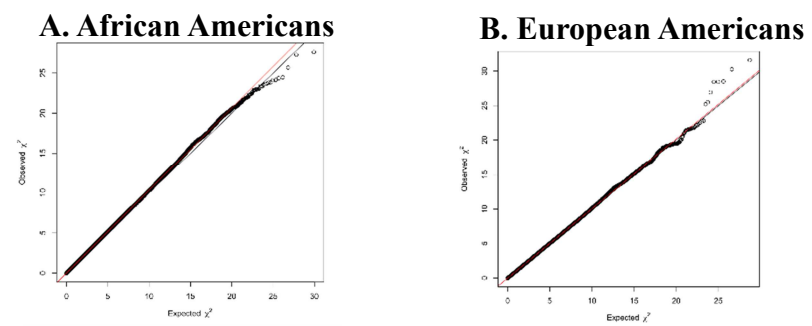
Figure E1

Figure E2