Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis

Genetic association studies have identified 21 loci associated with atopic dermatitis risk predominantly in populations of European ancestry. To identify further susceptibility loci for this common, complex skin disease, we performed a meta-analysis of >15 million genetic variants in 21,399 cases and 95,464 controls from populations of European, African, Japanese and Latino ancestry, followed by replication in 32,059 cases and 228,628 controls from 18 studies. We identified ten new risk loci, bringing the total number of known atopic dermatitis risk loci to 31 (with new secondary signals at four of these loci). Notably, the new loci include candidate genes with roles in the regulation of innate host defenses and T cell function, underscoring the important contribution of (auto)immune mechanisms to atopic dermatitis pathogenesis.

Atopic dermatitis (eczema) is a common inflammatory skin disease affecting 15–30% of children and 5–10% of adults. Its pathogenesis involves skin barrier abnormalities and T cell–driven cutaneous inflammation. Atopic dermatitis has considerable genetic contributions, with heritability estimates of up to 90% in Europeans. The strongest known risk factors are null mutations of the FLG gene (encoding filaggrin), resulting in epidermal barrier deficiency. Genome-wide association studies (GWAS) have identified 20 additional loci (ten in European, eight in Japanese and two in Chinese populations), mostly implicated in immune dysregulation. Genetic modeling suggests that further loci may be identified with well-powered GWAS. We therefore carried out a multi-ancestry meta-analysis of 26 studies comprising 21,399 cases and 95,464 controls imputed to the 1000 Genomes Project Phase 1 reference panel (Supplementary Table 1 and Supplementary Note). We analyzed 15,539,996 variants with minor allele frequency (MAF) ≥1%.

A fixed-effects meta-analysis of the 22 European studies identified 21 genome-wide significant ($P < 5 \times 10^{-8}$) loci (Fig. 1, Table 1 and Supplementary Figs. 1–4), and a multi-ancestry meta-analysis identified an additional six loci with log$_{10}$ (Bayes factor) > 6.1, four of which (10q21.2, 6p21.33, 11p13 and 2p13.3) also showed nominal association in the European analysis (Table 1). These 27 loci included all 11 loci previously associated with atopic dermatitis in Europeans and five loci originally reported in Japanese. Three loci identified in Japanese (6p21.33, 10q21.2 and 2q12.1) were also strongly associated in the European analysis, whereas two (3q13.2 and 11p15.4) might represent Japanese-specific signals (Supplementary Figs. 1 and 2), with the European confidence intervals ruling out all but very small effects (odds ratio (OR) < 1.03; Table 1). Furthermore, a locus originally reported in a Chinese GWAS (20q13.33) showed association in Europeans. We identified 11 new loci for atopic dermatitis. Four (11q24.3, 10p15.1, 8q21.13 and 2p25.1) were previously associated with self-reported allergy, and another (8q21.13) was associated with asthma. Two newly identified variants (5p13.2 and 2p25.1) showed statistically significant evidence of heterogeneity between European and non-European studies (Cochran’s Q test $P > 0.01$; Supplementary Table 2). Both variants showed little evidence for association in non-Europeans (particularly Japanese; Supplementary Fig. 2). Nevertheless, studies with phenotype definition based on a dermatological exam tended to report larger effect sizes than studies using self-reporting (Supplementary Fig. 4), which is to be expected, assuming a moderate degree of phenotypic misclassification in the latter. The inclusion of studies using self-reporting is therefore likely to bias estimates of effect size toward the null, and this should be borne in mind when interpreting the odds ratios from our study. Given that the primary aim of GWAS is the discovery of new loci, the increase in sample size achieved by including these studies is so large that any potential detrimental effect on statistical power is more than outweighed, and the expected direction of bias means that there is unlikely to be an issue of spurious findings (corroborated by Supplementary Fig. 4).

Seven of the 21 established loci for asthma, seven of the ten loci for allergic sensitization and six of 14 loci for self-reported allergy showed association with atopic dermatitis ($P < 0.05$), all with consistent directions of effect, supporting the notion of common atopic mechanisms in atopic dermatitis and allergy (Supplementary Table 2). However, several studies used here have contributed to multiple GWAS, which may have biased the observed overlap upward. Nevertheless, a substantial proportion of the loci associated with other atopic conditions appeared not to be strongly associated with atopic dermatitis.

A full list of authors and affiliations appears at the end of the paper.

Received 20 March; accepted 25 September; published online 19 October 2015; doi:10.1038/ng.3424
Twenty-one of the 27 atopic dermatitis–associated loci have previously been implicated in other immune-mediated traits (Supplementary Table 4), most notably inflammatory bowel disease (IBD) and psoriasis. We therefore compared significant results from GWAS of IBD, psoriasis, ankylosing spondylitis, multiple sclerosis, rheumatoid arthritis and type 1 diabetes with the results from our present study of atopic dermatitis. Of 163 established IBD risk variants, 39 reached \( P < 0.05 \) for atopic dermatitis (8.1 expected, \( P = 2.4 \times 10^{-16} \); Supplementary Table 5), with 35 showing the same direction of effect (sign-test \( P = 0.0001 \), consistent with the observational association between the two diseases). Of the 36 known psoriasis-associated variants, 15 reached \( P < 0.05 \) for atopic dermatitis (1.8 expected, \( P = 6 \times 10^{-11} \); Supplementary Table 6), with ten showing the same direction of effect (sign-test \( P = 0.30 \)). However, these two conditions rarely co-occur clinically and the most strongly associated genetic variants show opposite directions of effect. Therefore, our results, suggesting a more complex genetic relationship, might warrant further investigation. SNPs robustly associated with other autoimmune diseases were also more likely to be nominally associated with atopic dermatitis than expected by chance, but there was little evidence of any consistency in direction of effect (Supplementary Tables 7–10). These findings did not appear to be affected by contamination from the use of common controls across studies. Analyses performed excluding common cohorts yielded similar results (data not shown).

Conditional analysis showed evidence for secondary independent signals at four known atopic dermatitis loci (2q12.1, 4q27, 11p13 and 5q31.1; Supplementary Table 11), one of which (5q31.1) has been previously reported. In the epidermal differentiation complex (1q21.2-3; where FLG is located), the signals near MRPS21 (rs7512552) and IL6R (rs12730935 or the known functional mutation rs2228145) were independent from the European analysis; not associated in the European-only analysis. (multi-ancestry MANTRA meta-analysis of all studies. Arrows mark variants derived from the same direction of effect (sign-test \( P < 0.0001 \), consistent with the observational association between the two diseases28–30. Of the 36

For replication, we selected the lead SNPs from the 11 new loci, nine candidate SNPs from the MAGENTA analysis (with \( P < 1 \times 10^{-5} \) mapping to gene sets with FDR \( \leq 0.01 \)) related to innate immune signaling and T cell polarization was observed (Supplementary Fig. 5).

For replication, we selected the lead SNPs from the 11 new loci, nine candidate SNPs from the MAGENTA analysis (with \( P < 1 \times 10^{-5} \) mapping to gene sets with FDR \( < 0.05 \)) and three SNPs representing potentially new secondary signals. These SNPs were investigated in 18 studies (32,059 cases and 228,628 controls; Supplementary Table 1). In the European studies, 11 of the 20 new loci reached a Bonferroni-corrected significance threshold (\( \alpha = 0.0025 \) with one-sided tests in a fixed-effects analysis (Table 2). However, one of these loci showed evidence of heterogeneity (10p15.1, \( P = 0.041 \)) and was not significantly associated in a random-effects analysis (\( P = 0.019 \); Supplementary Table 15). Two of the SNPs selected by gene set enrichment analysis reached genome-wide significance in the combined analysis (representing the 2q37.1 and 12q15 loci). A random-effects analysis of all the replication cohorts (European and other ancestries) gave broadly consistent results (although only six SNPs reached genome-wide significance), with no clear population-specific effects (Supplementary Fig. 6 and Supplementary Table 16).

All three secondary signals showed significant association in the replication phase conditional analysis (Supplementary Table 11).

As a preliminary step toward understanding the functional underpinnings of the atopic dermatitis genetic associations, we established a ‘credible set’ of SNPs (all with strong association) for each locus (as described in the Online Methods). We reviewed the functional annotations of these SNPs in Encyclopedia of DNA Elements (ENCODE) Consortium and Roadmap Epigenomics Consortium data, evaluated expression quantitative trait locus (eQTL) effects in MuTHER, reviewed evidence of differential expression and surveyed relevant mouse mutants (Supplementary Tables 17–21 and Supplementary Note). Regions of DNase I hypersensitivity from the ENCODE and Roadmap Epigenomics data were strongly enriched for atopic dermatitis association in comparison to the rest of the genome (Supplementary Fig. 7 and Supplementary Table 22), particularly in immune cells (naïve T cells and helper T cells type 1 (Th1) and type 17 (Th17); \( P < 0.0001 \)); this enrichment was observed well below the genome-wide significance threshold, indicating the presence of additional undetected risk variants. We observed multiple cis-eQTLs (Bonferroni-corrected \( P < 7 \times 10^{-8} \)) in lymphoblastoid cell lines (LCLs) or skin (Supplementary Tables 17 and 19).
The most significant were two variants from the credible set at 2p13.3, which were strong eQTLs for SRT117 and SRT117, respectively; the previous association with FLG is within 250 kb. In LD with the known functional variant rs2228145 (\(r^2 = 0.86\), a nearby SNP (rs6872776; at 123,184,411 bp) in LD (\(r^2 = 0.97\) in the 1000 Genomes Project) showed similar association (log10(Bayes factor) = 7.21, European fixed-effects \(P = 3 \times 10^{-9}\)). The most significant were two variants from the credible set at 2p13.3, which were strong eQTLs for CD207 (langerin) in skin infections, which may be driven by impaired skin barrier function and differences in langerin function might contribute to this dysregulated cutaneous immunity.

The table shows the index variant for loci associated with \(P < 5 \times 10^{-8}\) in the European analysis or log10 (Bayes factor) > 6.1 in the multi-ancestry MANTRA analysis. Previous top trait associations with these loci are also listed. \(P\) values and –log10 (Bayes factor) values in bold indicate genome-wide significant results. AD, atopic dermatitis; A, asthma; AS, allergic sensitization; SRA, self-reported allergy; AR, allergic rhinitis; A+HF, asthma and hay fever combined; EA/OA, effect allele/other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; \(n\), sample size, BF, Bayes factor. Traits in parentheses did not reach genome-wide significance in the previous studies but were reported as suggestively associated.

The most significant were two variants from the credible set at 2p13.3, which were strong eQTLs for CD207 (langerin) in skin infections, which may be driven by impaired skin barrier function and differences in langerin function might contribute to this dysregulated cutaneous immunity.

The most significant were two variants from the credible set at 2p13.3, which were strong eQTLs for CD207 (langerin) in skin infections, which may be driven by impaired skin barrier function and differences in langerin function might contribute to this dysregulated cutaneous immunity.
Table 2: Replication results for the ten genomewide significant loci and loci identified in the MAMTA gene set enrichment analysis

| SNP          | Chr | Gene* | OS   | Effect size | OR (95% CI) | p value |
|--------------|-----|-------|------|-------------|-------------|---------|
| rs10199605   | 2q  | PUS10 | G/A  | 0.32        | 1.07 (1.04–1.10) | 10^{-8}   |
| rs2227483    | 12q | IFNG  | T/C  | 0.02        | 1.35 (1.19–1.54) | 10^{-6}   |
| rs6473227    | 8q  | ZBTB10| A/G  | 0.24        | 0.93 (0.91–0.96) | 10^{-6}   |
| rs7127637    | 11q | IL15RA| C/T  | 0.05        | 1.09 (1.06–1.13) | 10^{-6}   |
| rs14061755   | 17q | SFMBT1| C/T  | 0.04        | 0.99 (0.98–1.01) | 10^{-4}   |
| rs7127637    | 11q | IL15RA| C/T  | 0.05        | 1.09 (1.06–1.13) | 10^{-6}   |
| rs14061755   | 17q | SFMBT1| C/T  | 0.04        | 0.99 (0.98–1.01) | 10^{-4}   |
| rs10199605   | 2q  | PUS10 | G/A  | 0.32        | 1.07 (1.04–1.10) | 10^{-8}   |
| rs2227483    | 12q | IFNG  | T/C  | 0.02        | 1.35 (1.19–1.54) | 10^{-6}   |
| rs6473227    | 8q  | ZBTB10| A/G  | 0.24        | 0.93 (0.91–0.96) | 10^{-6}   |
| rs7127637    | 11q | IL15RA| C/T  | 0.05        | 1.09 (1.06–1.13) | 10^{-6}   |
| rs14061755   | 17q | SFMBT1| C/T  | 0.04        | 0.99 (0.98–1.01) | 10^{-4}   |
| rs10199605   | 2q  | PUS10 | G/A  | 0.32        | 1.07 (1.04–1.10) | 10^{-8}   |
| rs2227483    | 12q | IFNG  | T/C  | 0.02        | 1.35 (1.19–1.54) | 10^{-6}   |
| rs6473227    | 8q  | ZBTB10| A/G  | 0.24        | 0.93 (0.91–0.96) | 10^{-6}   |
| rs7127637    | 11q | IL15RA| C/T  | 0.05        | 1.09 (1.06–1.13) | 10^{-6}   |
| rs14061755   | 17q | SFMBT1| C/T  | 0.04        | 0.99 (0.98–1.01) | 10^{-4}   |

* indicates that the variant is within 1 Mb of a gene.

In conclusion, we have identified ten new loci robustly associated with atopic dermatitis in Europeans (six of which also reach genome-wide significance in random-effects analysis across studies of all ancestry groups), bringing the total number of susceptibility loci to 31 (24 in Europeans), with evidence of secondary signals at four of these loci. Altogether, in the subset of European studies with clinically defined cases, the previously established and newly identified variants explain approximately 12.3% and 2.6% of the variance in liability, respectively (Supplementary Table 23). All newly identified susceptibility loci are related to (auto)immune regulation, in particular innate immune signaling and T cell activation and specification, and there appears to be a substantial genetic overlap with other inflammatory and autoimmune diseases. Although not detracting from the importance of maintaining the skin barrier in the prevention and treatment of atopic dermatitis, our findings lend support to new therapeutic approaches targeted at immune modulation.®

© 2015 Nature America, Inc. All rights reserved.
Lavinia Paternoster\textsuperscript{12,111}, Marie Standl\textsuperscript{3}, Johannes Waage\textsuperscript{4}, Hansjörg Baurecht\textsuperscript{5}, Melanie Hotze\textsuperscript{5}, David P Strachan\textsuperscript{6}, John A Curtin\textsuperscript{6}, Klaus Bonnelykke\textsuperscript{6}, Chao Tian\textsuperscript{7}, Atsushi Takahashi\textsuperscript{8}, Jorge Esparza-Gordillo\textsuperscript{10,11}, Alexandress Couto Alves\textsuperscript{12}, Jacob P Thyszen\textsuperscript{13}, Herman T den Dekker\textsuperscript{14–16}, Manuel A Ferreira\textsuperscript{17}, Elisabeth Altmaier\textsuperscript{18–20}, Patrick M A Sleiman\textsuperscript{21,22}, Feng Li Xiao\textsuperscript{23}, Juan R Gonzalez\textsuperscript{24–26}, Ingo Marenholz\textsuperscript{10,11}, Birgit Kalb\textsuperscript{10,27}, Maria Pino-Yanes\textsuperscript{28–30}, Cheng-Jian Xu\textsuperscript{31,32}, Lisbeth Carstensen\textsuperscript{33}, Maria M Groen-Blokhuys\textsuperscript{34}, Cristina Venturini\textsuperscript{35}, Craig E Pennell\textsuperscript{36}, Sheila J Barton\textsuperscript{37}, Albert M Levin\textsuperscript{38}, Ivan Curjic\textsuperscript{39,40}, Mariona Bustamante\textsuperscript{41–44}, Eskil Kreiner-Moller\textsuperscript{4}, Gabriele A Lockett\textsuperscript{42}, Jonas Bacelis\textsuperscript{43}, Supinda Bunyanavich\textsuperscript{44}, Rachel A Myers\textsuperscript{45}, Anja Matanovic\textsuperscript{10,11}, Ashish Kumar\textsuperscript{39,40,46,47}, Joyce Y Tung\textsuperscript{48}, Tomomitsu Hirota\textsuperscript{49}, Michiaki Kudo\textsuperscript{49}, Wendy L McArdle\textsuperscript{2}, A John Henderson\textsuperscript{2}, John P Kemp\textsuperscript{1,2,50}, Lie Zheng\textsuperscript{1,2}, George Davey Smith\textsuperscript{1,2}, Franz Rüschendorf\textsuperscript{10}, Anja Bauerfeind\textsuperscript{10}, Min Ae Lee-Kirsch\textsuperscript{51}, Andrea Arnold\textsuperscript{52}, Georg Homuth\textsuperscript{53}, Carsten O Schmidt\textsuperscript{54}, Elisabeth Mangold\textsuperscript{55}, Sven Cichon\textsuperscript{55–59}, Thomas Keil\textsuperscript{50,61}, Elke Rodriguez\textsuperscript{5}, Annette Peters\textsuperscript{5}, Andre Franke\textsuperscript{5}, Wolfgang Lieb\textsuperscript{5}, Natalija Novak\textsuperscript{5}, Regina Föster-Holst\textsuperscript{5}, Momoko Horikoshi\textsuperscript{6}, Juha Pekkanen\textsuperscript{6}, Sylvain Sebert\textsuperscript{6,8}, Lise L Husemoen\textsuperscript{70}, Niels Grarup\textsuperscript{71}, Johan C de Jongste\textsuperscript{14}, Fernando Rivadeneira\textsuperscript{15,16,72}, Albert Hofman\textsuperscript{73}, Vincent W V Jaddoe\textsuperscript{14–16}, Suzanne G M A Pasmans\textsuperscript{73}, Niels J Elbert\textsuperscript{14,73}, André G Uitterlinden\textsuperscript{15,72}, Guy B Marks\textsuperscript{74}, Philip J Thompson\textsuperscript{17,75,76}, Melanie C Matheson\textsuperscript{77}, Colin F Robertson\textsuperscript{78}, Australian Asthma Genetics Consortium (AAGC)\textsuperscript{79}, Janina S Ried\textsuperscript{80}, Jing Li\textsuperscript{81}, Xin Bo Zuo\textsuperscript{23}, Xiao Dong Zheng\textsuperscript{82}, Xiao Yang Yin\textsuperscript{23}, Liang Dan Sun\textsuperscript{83}, Mace A McEave\textsuperscript{84}, Graeme M O’Regan\textsuperscript{85}, Caomhín M R Fahy\textsuperscript{86,87}, Linda E Campbell\textsuperscript{88}, Milan Macek\textsuperscript{89}, Michael Kurek\textsuperscript{89}, Donglei Hu\textsuperscript{108}, Celeste Eng\textsuperscript{89}, Dirkje S Postma\textsuperscript{90}, Bjarke Feenstra\textsuperscript{91}, Frank Geller\textsuperscript{92}, Jouke Jan Hottenga\textsuperscript{93}, Christel Middeldorp\textsuperscript{94}, Pirro Hysi\textsuperscript{95}, Veronique Bataille\textsuperscript{96}, Tim Spector\textsuperscript{97}, Carla M T Tiesler\textsuperscript{98}, Elisabeth Thiering\textsuperscript{99}, Badri Pahukasahrasam\textsuperscript{100}, James J Yang\textsuperscript{101}, Medea Imboden\textsuperscript{102,103}, Scott Huntzmann\textsuperscript{104}, Natalia Vilor-Jeodaro\textsuperscript{104–106}, Caroline L Relton\textsuperscript{107}, Ronny Myhre\textsuperscript{108,109}, Wenche Nystad\textsuperscript{110}, Adriaan Custovic\textsuperscript{7}, Scott T Weiss\textsuperscript{111}, Deborah A Meyers\textsuperscript{112}, Cilla Soderhall\textsuperscript{113,114}, Erik Melén\textsuperscript{115,116}, Carole Ober\textsuperscript{117}, Benjamin A Raby\textsuperscript{118}, Angela Simpson\textsuperscript{119}, Bo Jacobsen\textsuperscript{120,121}, John W Holloway\textsuperscript{6,122}, Hans Bisgaard\textsuperscript{123}, Jordi Sunyer\textsuperscript{6,124}, Nicole M Probst-Hensch\textsuperscript{39,40}, L Keoki Williams\textsuperscript{87,88}, Keith M Godfrey\textsuperscript{37,99}, Carol A Wang\textsuperscript{89}, Dorret I Boomsma\textsuperscript{41,100}, Mads Melbye\textsuperscript{33,101,102}, Gerard H Koppelman\textsuperscript{103}, Deborah Jarvis\textsuperscript{104,105}, W H Irwin McLean\textsuperscript{106}, Alan D Irvine\textsuperscript{80–82},
University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children’s Hospital, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, the Netherlands. Respiratory Epidemiology, Occupational Medicine and Public Health, National Heart and Lung Institute, Imperial College London, London, UK. Department of Epidemiology and Biostatistics, Medical Research Council–Public Health England Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA. Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark. Department of Epidemiology and Biostatistics, Medical Research Council–Public Health England Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. Unit of Primary Care, Oulu University Hospital, Oulu, Finland. Department of Dermatology, Ninewells Hospital and Medical School, Dundee, UK. These authors contributed equally to this work. These authors jointly supervised this work. Correspondence should be addressed to L.P. (l.paternoster@bristol.ac.uk).
ONLINE METHODS

GWAS meta-analysis. We carried out genome-wide association analysis for atopic dermatitis case-control status in 26 individual studies (Supplementary Table 1), comprising a total of 21,399 cases and 95,464 controls. The majority of these studies included individuals of only European ancestry (22 studies; 18,900 cases, 84,166 controls). We also included one study of Japanese individuals (RIKEN; 1,472 cases, 7,966 controls), one study of African-American individuals (SAPPIRE; 422 cases, 844 controls), one study of Latin American individuals (GALA II; 300 cases, 1,592 controls) and one study with individuals of mixed non-European ancestry (Generation R; 305 cases, 896 controls). Details on informed consent and ethical approval procedures for the individual studies are included in the Supplementary Note.

Each cohort separately imputed their genetic data to the 1000 Genomes Project Phase 1 reference panel (with the majority using the March 2012 release; Supplementary Table 1) and carried out genome-wide association analysis across all imputed variants. Before meta-analysis, we restricted each study to only variants with MAF >1% and a moderate imputation quality score (Rsq >0.3 for variants imputed in MACH and proper info >0.4 for variants imputed in IMPUTE). For some cohorts, additional quality control filters were applied (the full methods for each study are available in the Supplementary Note).

Meta-analysis was conducted for Europeans only in GWAMA (using a fixed-effects model) and for all ancestry groups combined in MANTRA48. Rather than imposing a fixed- or random-effects model, MANTRA accounts for the heterogeneity of effects between ancestry groups by allowing the studies to cluster according to allele frequency profile (and, hence, population genetic similarity). To prevent very small European studies (with less precise estimates of allele frequencies) from having undue weight in our analysis, we fixed the Europeans to cluster together using the European fixed-effects results in the MANTRA analysis. Variants with $P < 5 \times 10^{-8}$ in the European analysis were considered to be associated, as were any additional variants with $\log_{10}(P) > 6.1$ (equivalent to $P < 5 \times 10^{-8}$) (ref. 61) in the MANTRA analysis. Each locus is represented in the corresponding results table by the variant with the strongest evidence for association. Heterogeneity was assessed using the $I^2$ statistic and Cochran’s Q test. Meta-analysis results were also stratified according to ancestry, method of case definition and age of onset to explore sources of heterogeneity.

For the epidermal differentiation complex region (where the FLG gene is located and which has previously shown complex association results), we repeated the association tests (across the region from 150.2 to 154.5 Mb on chromosome 1) conditioning on the four most common FLG variants (p.Arg501*, c.2282del4, p.Arg2447* and p.Ser3247*) in the individual studies where these data were available (ten studies, 20,384 individuals; Supplementary Table 12). These samples were subjected to meta-analysis to determine whether there were any remaining independent association signals in this region.

Identification of independent secondary signals at associated loci. To identify secondary independent signals at each of the other associated loci, we carried out conditional analysis of the European meta-analysis results using GCTA62, with ALSPAC 1000 Genomes Project imputation (restricted to variants with MAF >1% and imputation quality proper info score >0.8) serving as the reference. The regions tested were within 250 kb up- or downstream of the top hit at each locus. Locus-specific significance thresholds were estimated by first calculating the effective number of tests in each 500-kb region using Nyholt’s procedure43 and the 1000 Genomes Project reference data (European). For each locus, we estimated the new threshold for a locus-wide error rate of 5% by dividing $\alpha$ (0.05) by the corresponding number of effective tests in that region ($\alpha$ values are shown in Supplementary Table 11). For 4q27, we defined the region as lying within 500 kb of the top hit because a known hit was just less than 500 kb from the top SNP in our analysis. We conditioned each region on the top hit from our meta-analysis. Any variant that surpassed the locus-specific significance threshold was considered an independent secondary signal.

MAGENTA gene set enrichment analysis. We tested our meta-analysis results for enriched gene sets using MAGENTA33. This method assigns SNPs to genes on the basis of genomic distance (SNPs are assigned if within 110 kb upstream or 40 kb downstream of each gene) and generates gene-based summary $P$ values. Subsequently, genes are assigned to gene sets (using curated repositories including GO-data, Biocarta, PANTHER, KEGG, etc.), and each gene set is assigned a $P$ value by comparing gene summary $P$ values to a null model where SNPs are drawn randomly 10,000 times (normalizing for the number of SNPs genotyped in each gene) and controlling for FDR at $\alpha = 0.05$. Approximately 10,000 gene sets were tested. As MAGENTA requires a $P$ value as input and to take account of the differing effects between populations, we reanalyzed our meta-analysis of all studies using a random-effects model, to generate results to serve as input for the MAGENTA analysis. All genes in the human leukocyte antigen (HLA) region (chr 6: 29,710,331–33,150,000) were removed from the analysis. To identify additional variants of interest to take forward to replication, we examined any pathway with FDR <0.05. From these gene sets, we took forward to replication any additional locus with a genetic variant associated at $P < 1 \times 10^{-5}$.

Cross-phenotype comparisons. The National Human Genome Research Institute (NHGRI) GWAS catalog64 was mined for traits with reported associations at each of our genome-wide significant loci. To further investigate the genetic overlap between atopic dermatitis and autoimmune diseases, we took the genome-wide significant loci from recent GWAS of IBD55, psoriasis56, ankylosing spondylitis57, multiple sclerosis58, rheumatoid arthritis59 and type 1 diabetes60 and extracted the atopic dermatitis results for these variants from our European discovery GWAS, noting whether each variant was associated ($P < 0.05$) with atopic dermatitis (testing the enrichment of overlap using a two-sided binomial exact test) and carried the same or opposite directions of effect for the two traits (tested using a sign test).

Replication. The top SNPs from the 11 newly associated regions (log_{10}(P) > 6.1 or $P < 5 \times 10^{-8}$) were taken forward to replication, along with nine suggestively associated SNPs ($P < 1 \times 10^{-5}$) that were in genes highlighted in the MAGENTA analysis as good candidates. In addition, we included any variant with evidence of being a secondary independent signal in the associated loci. Replication consisted of 18 studies (32,059 cases and 228,628 controls) with genome-wide imputed data available or custom genotyping (Supplementary Table 1). The studies of European ancestry were combined in fixed-effects meta-analyses in GWAMA. We also carried out random-effects meta-analysis for the European studies to assess association for variants that showed evidence of heterogeneity ($P < 0.05$). Significant association in the replication phase was defined by a one-sided $P$ value meeting a Bonferroni-corrected threshold ($\alpha = 0.05/20 = 0.0025$). Random-effects meta-analyses of replication studies of all ancestry groups were also carried out, and forest plots were examined for evidence of population-specific effects. For replication of the three secondary signals, the secondary SNPs were tested for association after conditioning on the top SNP in each of the European cohorts. These results were then combined in fixed-effects meta-analyses.

Credible sets. To assemble a sensible list of variants at each locus for functional look-ups, we constructed credible sets34 that represented the SNPs most likely to be causal on the basis of statistical evidence from the MANTRA analysis (or from the European analysis for the three variants that appeared to be European specific). The European-only GWAS was repeated in MANTRA to generate the Bayes factors required for credible set analysis. Bayes factors were used to calculate posterior probabilities for all SNPs in each region ($\pm 1$ Mb from the top SNP); the minimum set of SNPs that had a cumulative posterior probability of at least 95% made up each credible set. These sets can be interpreted similarly to confidence intervals, assuming that the association signal at a locus can be attributed to a single causal variant (and that the true causal variant is included in the analysis and has been well imputed); there is a 95% probability that the 95% credible set contains that causal variant. Given that analysis based on 1000 Genomes Project imputation may not include the true causal variant or that each association may be driven by more than one causal variant, we do not expect these credible sets to necessarily contain the causal variants at the suggested probability of 95%. Nevertheless, they demonstrate which neighboring variants, in addition to the top SNP, also show strong association with atopic dermatitis, and we find them useful in
assessing the size of the regions of interest and for defining a set of variants to follow up. As the posterior probabilities for the MAGENTA-identified credible sets were extremely large (owing to the weaker signals at these loci), we instead carried out functional look-ups for all variants with $r^2 \geq 0.8$ with the top hit for these loci.

**Functional look-ups.** For variants identified as part of a credible set, we carried out look-ups in functional data resources as follows: (i) RegulomeDB and HaploReg were mined for evidence of coding or regulatory function (these resources collate annotations for example, coding variation, regulatory chromatin marks, DNase I hypersensitivity, protein binding and motif alteration) from the ENCODE Consortium, the US National Institutes of Health Roadmap Epigenomics Mapping Consortium and the literature over a wide range of tissues; (ii) cis eQTLs in skin or LCLs were identified in data from the MuTHER Consortium [67] (with variants considered to be eQTLs if the association with any transcript within 1 Mb had $P < 7 \times 10^{-4}$, corresponding to significance with Bonferroni correction for 36 loci and two tissues); (iii) implicated genes were identified with differential expression reported between uninvolved skin from cases and skin from controls [65] and between lesional and non-lesional skin in patients with atopic dermatitis in a study deposited in the Gene Expression Omnibus (GDS4444) (ref. 66) and (iv) mouse mutants of implicated genes were examined in the Mouse Genome Informatics (MGI) database.

Colocalization of atopic dermatitis GWAS signals and eQTLs in the MuTHER data was investigated using the R package coloc [67]. All SNPs within 100 kb of the lead atopic dermatitis SNP were included in the analysis, and we report the posterior probability of each pair of signals colocalizing in Supplementary Table 19.

To identify and visualize cell types implicated in atopic dermatitis pathogenesis, the tendency of disease-associated loci to fall in cell type–specific regulatory DNase I–hypersensitive sites (DHSs; a proxy for accessible and/or regulatory DNA) was calculated for the full range of $P$ values, essentially as described [68]. Enrichment was computed for 168 cell types and cell lines in the ENCODE and Roadmap repositories [66]. Duplicates and directly redundant cell types were removed before analysis. One-sided $P$ values for enrichment were calculated from an empirical null distribution of locus overlap for DHS regions, generated by 10,000 random permutations with overlap of an identical number of randomly selected loci as found at $P \leq 1 \times 10^{-10}$ with all DHS regions for all cell and tissue types.

**Variance in liability explained.** We estimated the proportion of variance in atopic dermatitis liability explained by the established and new variants in a subset of studies that had clinically diagnosed cases (GENEVA/KORAF/POPGEN, NCR/ADC, GENUFAD-SHIP1, GENUFAD-SHIP2, GENEVA (replication) and CECCS) using the method of So et al. [69].

---

60. Morris, A.P. Transethnic meta-analysis of genomewide association studies. *Genet. Epidemiol.* **35**, 809–822 (2011).
61. Wang, X. et al. Comparing methods for performing trans-ethnic meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **22**, 2303–2311 (2013).
62. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
63. Nyholt, D.R. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–769 (2004).
64. Welter, D. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42**, D1001–D1006 (2014).
65. Cole, C. et al. Filaggrin-stratified transcriptomic analysis of pediatric skin identifies mechanistic pathways in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* **134**, 82–91 (2014).
66. Tintle, S. et al. Reversal of atopic dermatitis with narrow-band UVB phototherapy and biomarkers for therapeutic response. *J. Allergy Clin. Immunol.* **128**, 583–593 (2011).
67. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
68. Maurano, M.T. et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
69. So, H.C., Li, M. & Sham, P.C. Uncovering the total heritability explained by all true susceptibility variants in a genome-wide association study. *Genet. Epidemiol.* **35**, 447–456 (2011).