Merck Serono, and Sanofi/Regeneron, and his institution has received research funding from Bristol-Myers Squibb and EMD Serono. The remaining authors state no conflict of interest.

ACKNOWLEDGMENTS

The study was supported by the Independent Research Fund Denmark (DFF Sapere Aude grant number 4004-00422) and the European Research Council, European Union (StG 677268 NextDART). The study is partly funded through the National Institutes of Health, United States/National Cancer Institute Cancer Center Support Grants P30 CA015704 and P01 CA225517. We would like to thank all patients with cancer who contributed samples; all collaborators taking part in the study; Bente Rotbøl, Tripti Tamhane, Anna Gyllenberg Burkal, Anni Flarup Løye, and Julien Candrian for excellent technical assistance; and all current and former members of the SRH research group for scientific discussions.

AUTHOR CONTRIBUTIONS

Conceptualization: UKH, RL, PTS, SRH; Formal Analysis: UKH, RL; Funding Acquisition: PN, SRH; Investigation: UKH, RL; Methodology: UKH, RL, DI, MV; Project Administration: SRH; Resources: DI, CC, MV, JCB, PTS, PN, SRH; Supervision: PN, SRH; Validation: UKH, RL; Visualization: UKH, RL; Writing - Original Draft Preparation: UKH, SRH; Writing - Review and Editing: UKH, CC, JCB, PTS, PN, SRH

Ulla Kring Hansen[1], Rikke Lyngaa[1], Dafina Ibrani[2,3], Candice Church[2,3], Monique Verhaegen[4], Andrzej Antoni Dlugosz[4,5], Jürgen Christian Becker[6,7,8,9], Per thor Straten[10], Paul Nghiem[2,3,11,12] and Sine Reker Hadrup[1,*]

[1]Department of Health Technology, Technical University of Denmark, Lyngby, Denmark; [2]Division of Dermatology, Department of Medicine, University of Washington, Seattle, Washington, USA; [3]Department of Laboratory Medicine & Pathology, University of Washington, Seattle, Washington, USA; [4]Department of Dermatology, University of Michigan, Ann Arbor, Michigan, USA; [5]Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, Michigan, USA; [6]Department of Translational Skin Cancer Research, University Hospital Essen, Essen, Germany; [7]German Cancer Consortium (DKTK), Essen, Germany; [8]German Cancer Research Center (DKFZ), Heidelberg, Germany; [9]Department of Dermatology, University Hospital Essen, Essen, Germany; [10]National Center for Cancer Immune Therapy, Copenhagen University Hospital, Herlev, Denmark; [11]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; and [12]Seattle Cancer Care Alliance Center, Seattle, Washington, USA

[*]Corresponding author e-mail: sirha@dtu.dk

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2021.06.027.

REFERENCES

Bichakjian CK, Olencki T, Aasi SZ, Alam M, Andersen JS, Blitzblau R, et al. merkel cell carcinoma, version 1.2018, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2018;16:742–74.

D'Angelo SP, Russell J, Lebbé C, Chmielowski B, Gambichler T, Grob JJ, et al. Efficacy and safety of first-line avelumab treatment in patients with stage IV metastatic Merkel cell carcinoma a preplanned interim analysis of a clinical trial. JAMA Oncol 2018;9:e180077.

Feng H, Shuda M, Chang Y, Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 2008;319:1096–100.

Harms KL, Healy MA, Nghiem P, Sober AJ, Johnson TM, Bichakjian CK, et al. Analysis of prognostic factors from 9387 Merkel cell carcinoma cases forms the basis for the new 8th edition AJCC staging system. Ann Surg Oncol 2016;23:3564–71.

Iyer JG, Afanasiev OK, McClurkan C, Paulson K, Nagase K, Jing L, et al. Merkel cell polyomavirus-specific CD8$^+$ and CD4$^+$ T-cell responses identified in Merkel cell carcinomas and blood. Clin Cancer Res 2011;17:6671–80.

Jing L, Ott M, Church CD, Kulikauskas RM, Ibrani D, Iyer JG, et al. Prevalent and diverse intratumoral oncoprotein-specific CD8+ T cells within polyoma virus-driven Merkel cell carcinomas. Cancer Immunol Res 2020;8:648–59.

Lyngaa R, Pedersen NW, Schrama D, Thrue CA, Ibrani D, Met O, et al. T-cell responses to oncogenic Merkel cell polyomavirus proteins distinguish patients with Merkel cell carcinoma from healthy donors. Clin Cancer Res 2014;20:1768–78.

Miller NJ, Church CD, Dong L, Crispin D, Fitzgibbon MP, Lachance K, et al. Tumor-infiltrating Merkel cell polyomavirus-specific T cells are diverse and associated with improved patient survival. Cancer Immunol Res 2017;5:137–47.

Nghiem P, Kaufman HL, Bharmal M, Mahnke L, Phatak H, Becker JC. Systematic literature review of efficacy, safety and tolerability outcomes of chemotherapy regimens in patients with metastatic Merkel cell carcinoma. Future Oncol 2017;13:1263–79.

Nghiem PT, Bhatia S, Lipson EJ, Kudchadkar RR, Miller NJ, Annamalai L, et al. PD-1 blockade with pembrolizumab in advanced Merkel-cell carcinoma. N Engl J Med 2016;374:2542–52.

Paulson KG, Iyer JG, Simonson WT, Blom A, Thibodeau RM, Schmidt M, et al. CD8+ lymphocyte intratumoral infiltration as a stage-independent predictor of Merkel cell carcinoma survival: a population-based study. Am J Clin Pathol 2014;142:452–8.

Samimi M, Benlalam H, Aumond P, Gaboriaud P, Fradin D, Kervarrec T, et al. Viral and tumor antigen-specific CD8 T-cell responses in Merkel cell carcinoma. Cell Immunol 2019;344:103961.

Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, et al. T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. Proc Natl Acad Sci USA 2008;105:16272–7.

Sihto H, Joensuu H. Tumor-infiltrating lymphocytes and outcome in Merkel cell carcinoma, a virus-associated cancer. Oncoimmunology 2012;1:1420–1.

Topalian SL, Bhatia S, Amin A, Kudchadkar RR, Sharfman WH, Lebbé C, et al. Neoadjuvant nivolumab for patients with resectable Merkel cell carcinoma in the CheckMate 358 trial. J Clin Oncol 2020;38:2476–87.

A GWAS Finds Variants at 2p21 Associated with Self-Reported Sensitive Skin in the Han Chinese Population

Journal of Investigative Dermatology (2022) **142,** 243–247; doi: 10.1016/j.jid.2021.04.021

Accepted manuscript published online 21 May 2021; corrected proof published online 17 November 2021

TO THE EDITOR

Sensitive skin is defined by the occurrence of unpleasant sensations (e.g., stinging, burning, pain, pruritus, tingling) in response to various stimuli (e.g., cosmetics, temperature variation, emotion) that generally should not

Table 1. Summary of the Four Signal SNPs at 2p21 in GWAS

CHR	SNP	BP[1]	REF	ALT	EA	EAF	Discovery			Replication			Meta-Analysis		
							OR	95% CI	*P*-Value	OR	95% CI	*P*-Value	OR	95% CI	*P*-Value
2	rs17030203	43155982	T	G	G	0.503	1.45	(1.27–1.66)	4.17×10^{-8}	1.25	(1.02–1.53)	0.034	1.39	(1.24–1.55)	9.04×10^{-9}
2	rs17030206	43156974	T	C	C	0.505	1.45	(1.27–1.66)	3.82×10^{-8}	1.23	(1.00–1.51)	0.049	1.38	(1.23–1.54)	1.35×10^{-8}
2	rs57908981	43157164	C	A	A	0.504	1.46	(1.28–1.67)	3.11×10^{-8}	1.23	(1.00–1.51)	0.044	1.39	(1.24–1.55)	9.90×10^{-9}
2	rs72803822	43158175	G	A	A	0.502	1.47	(1.28–1.68)	2.28×10^{-8}	1.26	(1.03–1.56)	0.026	1.42	(1.25–1.56)	3.78×10^{-9}

Abbreviations: ALT, alternative allele; BP, base-pair; CHR, chromosome; CI, confidence interval; EA, effect allele; EAF, effect allele frequency; REF, reference allele.
[1]Positions are according to human reference hg19.

provoke such sensations (Brenaut et al., 2020; Misery et al., 2017). The face tends to be the most common site of skin sensitivity owing to its rich innervation and its exposure to multiple types of irritants. Facial sensitive skin affects around 55–57% of the Asian population but varies substantially among countries and studies (Kamide et al., 2013; Kim et al., 2018). Genetic factors probably play a role, supported by the significant ethnic differences in skin sensitivity (Jourdain et al., 2002) and relevant skin properties (e.g., permeability [Kompaore and Tsuruta, 1993] and transepidermal water loss [Machado et al., 2010]). However, the genetics of sensitive skin is largely unknown. In a clinical study, sensitive skin is usually measured by stinging test methods (Christensen and Kligman, 1996), but this is not quite feasible in population-based studies. Instead, self-reported questionnaires have been the validated tools for identifying individuals with sensitive skin (Brenaut et al., 2020). In this study, we used self-reported sensitive skin as the phenotype for our GWAS. A direct-to-consumer platform (WeGene) was used to collect questionnaires and DNA data, a strategy that has been successfully applied in previous GWASs of various phenotypes (Kang et al., 2020; Shaffer et al., 2017). The discovery cohort was collected through an online platform in February and March 2018 (n = 1,872, age = 29.0 ± 7.2 years, 63.68% females) (Supplementary Table S1). An independent replication cohort was collected online in the next 10 months after the discovery set (n = 817, age = 28.6 ± 8.1 years, 62.30% females). Sensitive skin was defined as a self-reported case/control trait. Saliva samples of subjects were collected and then genotyped on Affymetrix WeGene V1 Arrays (Santa Clara, CA). Imputation was performed by Eagle and Minimac4 using 1000 Genomes phase 3 data as the reference. After quality control, a total of 596,744 genotyped SNPs and 5,370,088 imputed SNPs were included in the GWAS (Supplementary Materials and Methods). We found no outlier in our samples from a principal component analysis (Supplementary Figure S1). This research was conducted with official approval from the Ethical Committee of WeGene, and all subjects provided electronic informed consent.

When asked, "Would you consider your facial skin sensitive, and prone to a sensation of redness, itching, tension, sting, or burning when exposed to external stimuli?," 55.56% of the subjects in the discovery cohort and 59.73% in the replication cohort reported "very or rather sensitive" (Supplementary Tables S1 and S2), similar to the prevalence (55–57%) reported in East Asian populations (Kamide et al., 2013; Kim et al., 2018). The prevalence of self-reported sensitive skin was significantly higher in females than in males (63.30% vs. 45.67%, $P = 9.98 \times 10^{-19}$), also consistent with findings from previous studies (Chen et al., 2020). Overall, the age of the individuals with self-reported sensitive skin was slightly younger than that of the control group ($P = 0.04$); however, there was no significant association after sex stratification ($P_{female} = 0.17$, $P_{male} = 0.15$) (Supplementary Figure S2).

We then performed a GWAS using logistic regression with an additive genetic model in the discovery cohort, including sex, age, and the first five genetic principal components as covariates. We identified a signal at 2p21 significantly associated with self-reported sensitive skin (rs72803822 as the lead SNP; OR = 1.47, confidence interval = 1.28–1.68; $P = 2.28 \times 10^{-8}$) (Table 1 and Supplementary Figure S3a). This locus was successfully replicated ($P = 0.03$) in the replication cohort (Table 1). In the meta-analysis of the discovery and replication cohorts, four SNPs in strong linkage disequilibrium ($r^2 = 0.95$~0.99) at 2p21 reached the genome-wide significance level (OR = 1.42, confidence interval = 1.36–1.48; $P = 3.78 \times 10^{-9}$) (Figure 1a and b and Supplementary Figure S3b). Analysis after sex stratification showed that the effect size of the associated SNPs did not differ between males and females (Supplementary Table S3). In a GWAS further incorporating dermatologic conditions associated with self-reported sensitive skin as covariates (e.g., eczema/dermatitis, eczema/dermatitis since childhood, and asthma/hay fever) (Supplementary Table S4), the locus remained genome-wide significant in a meta-analysis (OR = 1.42, confidence interval = 1.30–1.53; $P = 5.43 \times 10^{-9}$)

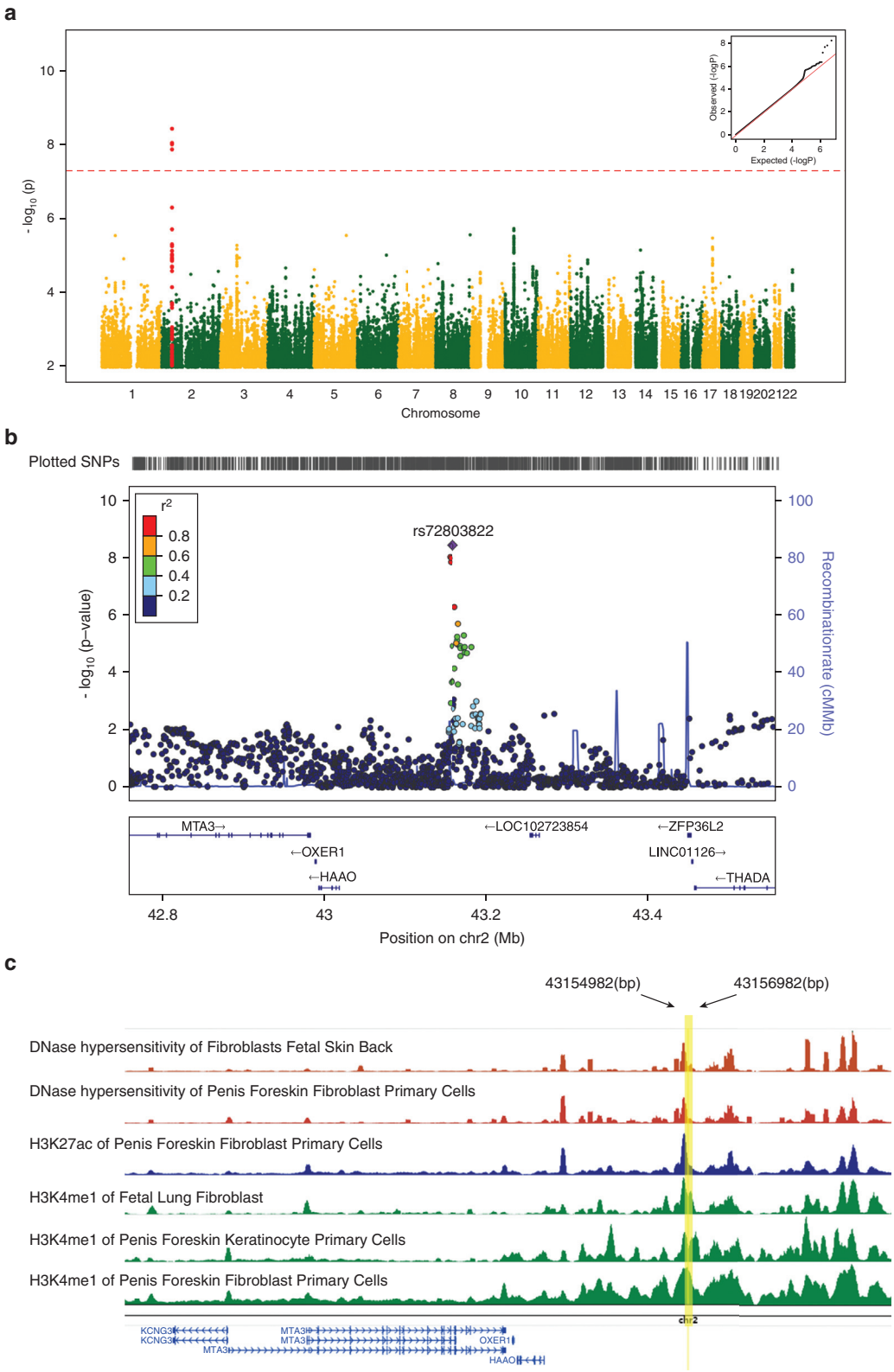


Figure 1. GWAS of self-reported sensitive skin identified a significant signal at 2p21. (**a**) Manhattan plot and quantile–quantile plot showing the results of the meta-analysis for GWASs of self-reported sensitive skin on discovery and replication sets. The red line corresponds to the genome-wide threshold ($P \leq 5 \times 10^{-8}$). SNPs within 5 kb to the signals were plotted in red. (**b**) Regional association plot for the significant region at 2p21. Increasing color intensities indicated increasing linkage disequilibrium (r^2) with rs72803822. (**c**) Epigenetic annotation at a region within 1 kb to rs17030203, 160 kb upstream of *MTA3*, and outlined in yellow. (**d**) Functionality scores from 3DSNP annotated the significant SNPs as TFBSs and enhancers. (**e**) The effect allele G on rs17030203 showed an increasing effect on more sensitive skin. (**f**) Geography of allele frequency on rs17030203. Chr, chromosome; Ctrl, control; Mb, megabase; SS, sensitive skin; TFBS, transcriptionfactor–binding site.

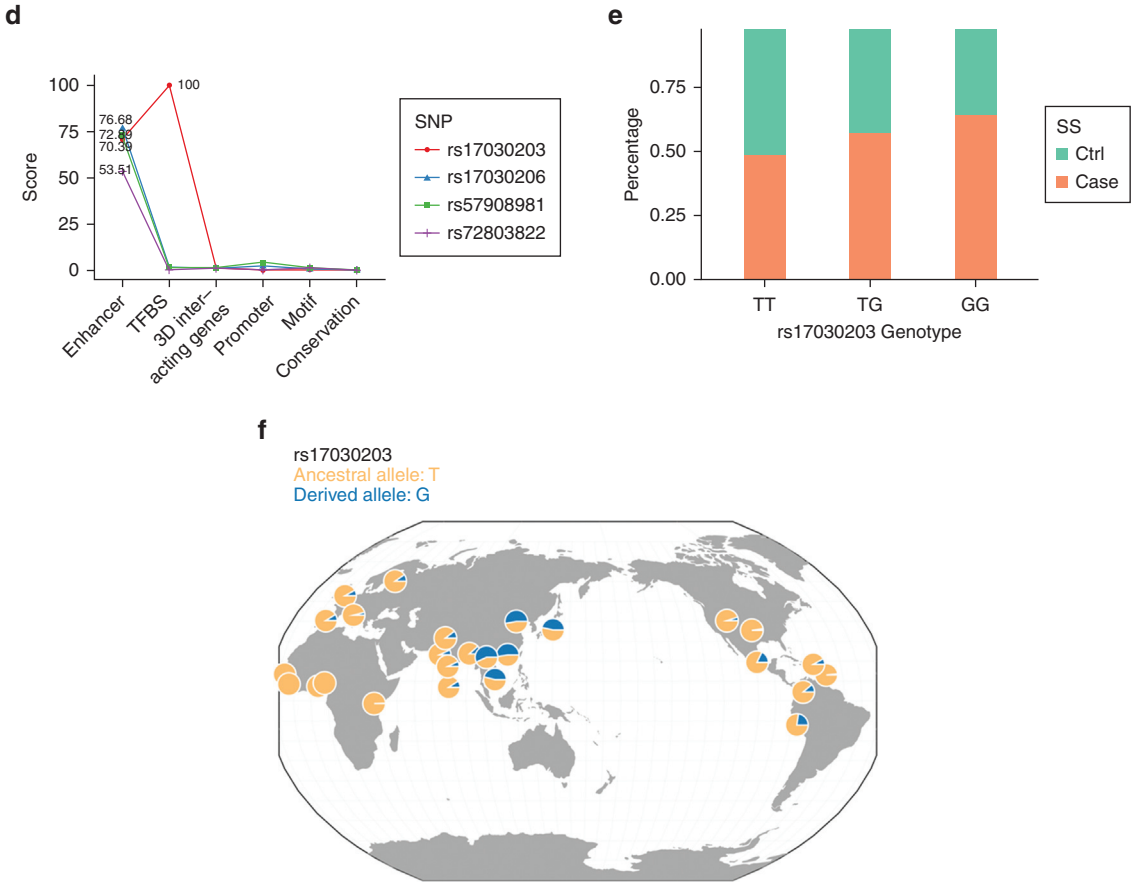


Figure 1. Continued.

(Supplementary Figure S3c), indicating that these findings were not affected by the skin disease history. None of the above SNPs has ever been reported to be associated with any dermatologic conditions in the GWAS catalog.

All four SNPs are located in the intergenic region between *ZFP36L2* and *MTA3* genes. Interestingly, rs17030203, one of the significant SNPs, is an expression quantitative trait locus of *MTA3* in the blood (z-score = 3.07, $P = 0.002$ [Westra et al., 2013]; data from the other tissues are not significant in Genotype-Tissue Expression). Reportedly, *MTA3* plays a role in the maintenance of the normal epithelial architecture by regulating E-cadherin levels (Fujita et al., 2003). E-cadherin and its encoding gene *CDH1*, essential in the maintenance of epithelium integrity and keratinocyte differentiation, was recently found to be upregulated in sensitive skin samples (Kim et al., 2014). Querying the signal locus against the 3DSNP database, we found high functionality scores on transcription factor–binding sites and enhancers (53.51~100) (Figure 1d). In addition, on the basis of the encyclopedia of DNA elements and Roadmap database, the locus exhibited distinct signatures of active enhancer epigenetic markers such as H3K4me1 histone modifications in fibroblasts and keratinocytes (Figure 1c). This evidence thus revealed that the signal locus has a potential regulatory function. Taken together, we speculate that SNP rs17030203 regulates the expression of the *MTA3* gene in an enhancer/transcription factor binding sites–dependent manner, and thus regulating E-cadherin level to affect skin sensitivity.

Individuals with the derived allele (G) of rs17030203 have a higher frequency of self-reported sensitive skin, with an increased prevalence of approximately 7.6% per copy (Figure 1e). The derived allele of rs17030203 is of high frequencies in East Asian populations but mostly rare in other populations (Figure 1f). The other three significant SNPs in linkage disequilibrium also show similar patterns (Supplementary Figure S4 and Supplementary Table S5). These results were consistent with the report of comparatively higher skin sensitivity in East Asians than in Europeans (Aramaki et al., 2002). However, we found no signal for positive natural selection in the 2p21 region in East Asians, Europeans, or Africans (Supplementary Figure S5).

In conclusion, this GWAS on self-reported sensitive skin in the Han Chinese population, to our knowledge previously unreported, identified a risk locus at 2p21, and one of the lead SNPs, rs17030203, is an expression quantitative trait locus of *MTA3*. A limitation of the study is that the self-reported phenotype is likely composed of several endophenotypes, and the questionnaire-based approach makes the phenotype even rougher. Further investigations on the role of subjective perception on sensitive skin will be beneficial. On the other hand, genetic findings of this GWAS are likely robust. Our findings provide insights into a mechanism involving the disrupting of the maintaining epithelial architecture in the development of sensitive skin.

Data availability statement

The GWAS summary statistics were deposited in the publicly available National Omics Data Encyclopedia (http://www.biosino.org/node/) and are available under the project identification document OEP001547. Data usage shall be in full compliance with the Regulations on Management of Human Genetic Resources in China.

ORCIDs

Bingjie Li: https://orcid.org/0000-0002-4871-7444
Xiyang Cai: https://orcid.org/0000-0002-3868-5426
Lizhong Wang: https://orcid.org/0000-0002-3571-8771
Jiarui Li: https://orcid.org/0000-0003-0581-5850
Ying Zou: https://orcid.org/0000-0003-0849-5688
Gang Chen: https://orcid.org/0000-0002-5808-0432
Sijia Wang: https://orcid.org/0000-0001-6961-7867

CONFLICT OF INTEREST

LW and GC are employees of WeGene. The remaining authors state no other conflict of interest.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Project (grant number 2018YFC0910403 to SW), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDB38020400 to SW), and the Shanghai Municipal Science and Technology Major Project (grant number 2017SHZDZX01 to SW). We thank all WeGene users who consented to participate in the research. For data-related information, please correspond with GC (cg@wegene.com).

AUTHOR CONTRIBUTIONS

Conceptualization: BL, SW; Data Curation: BL, XC; Formal Analysis: BL, XC, LW; Funding Acquisition: SW; Investigation: BL, XC; Methodology: BL, XC, LW; Project Administration: BL, GC, SW; Resources: LW, GC, SW; Software: BL, XC, LW; Supervision: GC, SW; Visualization: BL, XC, LW, GC, SW; Writing - Original Draft Preparation: BL, XC; Writing - Review and Editing: JL, YZ, GC, SW

***Bingjie Li*[1,2,7], *Xiyang Cai*[2,7], *Lizhong Wang*[3,4,7], *Jiarui Li*[2], *Ying Zou*[1], *Gang Chen*[3,4,5] and *Sijia Wang*[2,6,*]**

[1]*Department of Skin and Cosmetic Research, Shanghai Skin Disease Hospital, Tongji University School of Medicine, Tongji University, Shanghai, China;* [2]*CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China;* [3]*Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha, China;* [4]*WeGene, Shenzhen Zaozhidao Technology, Shenzhen, China;* [5]*Shenzhen WeGene Clinical Laboratory, Shenzhen, China; and* [6]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China*
[7]*These authors contributed equally to this work*
[*]*Corresponding author e-mail: wangsijia@picb.ac.cn*

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2021.04.021.

REFERENCES

Aramaki J, Kawana S, Effendy I, Happle R, Löffler H. Differences of skin irritation between Japanese and European women. Br J Dermatol 2002;146:1052–6.

Brenaut E, Barnetche T, Le Gall-Ianotto C, Roudot AC, Misery L, Ficheux AS. Triggering factors in sensitive skin from the worldwide patients' point of view: a systematic literature review and meta-analysis. J Eur Acad Dermatol Venereol 2020;34:230–8.

Chen W, Dai R, Li L. The prevalence of self-declared sensitive skin: a systematic review and meta-analysis. J Eur Acad Dermatol Venereol 2020;34:1779–88.

Christensen M, Kligman AM. An improved procedure for conducting lactic acid stinging tests on facial skin. J Soc Cosmet Chem 1996;47:1–11.

Fujita N, Jaye DL, Kajita M, Geigerman C, Moreno CS, Wade PA. MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer. Cell 2003;113:207–19.

Jourdain R, de Lacharrière O, Bastien P, Maibach HI. Ethnic variations in self-perceived sensitive skin: epidemiological survey. Contact Dermatitis 2002;46:162–9.

Kamide R, Misery L, Perez-Cullell N, Sibaud V, Taïeb C. Sensitive skin evaluation in the Japanese population. J Dermatol 2013;40:177–81.

Kang K, Sun X, Wang L, Yao X, Tang S, Deng J, et al. Direct-to-consumer genetic testing in China and its role in GWAS discovery and replication. Quant Biol 2020:1–15.

Kim EJ, Lee DH, Kim YK, Kim MK, Kim JY, Lee MJ, et al. Decreased ATP synthesis and lower pH may lead to abnormal muscle contraction and skin sensitivity in human skin. J Dermatol Sci 2014;76:214–21.

Kim YR, Cheon HI, Misery L, Taieb C, Lee YW. Sensitive skin in Korean population: an epidemiological approach. Skin Res Technol 2018;24:229–34.

Kompaore F, Tsuruta H. In vivo differences between Asian, black and white in the stratum corneum barrier function. Int Arch Occup Environ Health 1993;65(Suppl. 1):S223–5.

Machado M, Hadgraft J, Lane ME. Assessment of the variation of skin barrier function with anatomic site, age, gender and ethnicity. Int J Cosmet Sci 2010;32:397–409.

Misery L, Ständer S, Szepietowski JC, Reich A, Wallengren J, Evers AW, et al. Definition of sensitive skin: an expert position paper from the Special Interest Group on Sensitive Skin of the International Forum for the Study of Itch. Acta Derm Venereol 2017;97:4–6.

Shaffer JR, Li J, Lee MK, Roosenboom J, Orlova E, Adhikari K, et al. Multiethnic GWAS reveals polygenic architecture of earlobe attachment. Am J Hum Genet 2017;101:913–24.

Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013;45:1238–43.

Transient Induction of Fever in the Imiquimod C57BL/6 Mouse Model of Psoriasis-Like Disease Involves IL-1 and IL-6 but Not IL-36

Journal of Investigative Dermatology (2022) **142,** 247–251; doi: 10.1016/j.jid.2021.05.028

TO THE EDITOR

Psoriasis embodies a number of related chronic inflammatory skin conditions. Plaque psoriasis is the most common form and affects 1–3% of the

Abbreviation: GPP, generalized pustular psoriasis

Accepted manuscript published online 17 July 2021; corrected proof published online 5 August 2021

SUPPLEMENTARY MATERIALS AND METHODS

Population and samples

All participants were recruited from consenting WeGene customers from Shenzhen Zaozhidao Technology, a direct-to-consumer genetic testing service provider. The discovery set was collected in February and March 2018 and included 1,872 volunteers (1,192 females and 680 males). The replication set was collected in the following 10 months and included 817 volunteers (509 females and 308 males). The sample summary is provided in Supplementary Table S1.

Phenotyping

Information on personal data and phenotypes for each participant was gathered through a self-reported questionnaire. Participants replied to the question "Would you consider your facial skin sensitive, and prone to a sensation of redness, itching, tension, sting, or burning when exposed to external stimuli?" In the following analysis, we defined individuals who answered very sensitive and rather sensitive as a sensitive group and those who answered slightly sensitive and not sensitive as a nonsensitive group (Kamide et al., 2013). Potentially relevant dermatologic conditions (e.g., eczema/dermatitis, eczema/dermatitis since childhood, and asthma/hay fever) were also queried in the survey. The initial survey questions were worded in Chinese and are shown in Supplementary Table S2.

Genotyping

Saliva samples for DNA extraction and genotyping were collected and then genotyped on the Affymetrix WeGene V1 Arrays (Santa Clara, CA), which covered 596,744 SNPs at the WeGene genotyping center, Shenzhen. To control for genotype quality, we used PLINK, version 1.9 (Purcell et al., 2007), to exclude individuals with >5% missing data, outlying heterozygous rate, or discordant sex information. Unrelated was filtered by pairwisely checking for all the samples, and those whose identities were identified by descent scores >0.125 were removed. Ancestry was assigned from self-reported surveys and further examined with principal component analysis. We also discarded SNPs with unbalanced call rates in case and controls, those with >2% missing data, those with a minor allele frequency <1%, or any SNPs that failed the Hardy Weinberg deviation test ($P < 1 \times 10^{-5}$). After pruning, 373,988 SNPs were retained for further analysis. We used Eagle (Loh et al., 2016) and Minimac4 (Fuchsberger et al., 2015) to impute the genotypes at nongenotyped SNPs with a 10-Mb chunk size, a 3-Mb step size, and 1000 Genomes phase 3 data (Genomes Project et al., 2015) as an imputation reference panel. X chromosome was also phased and imputed by Eagle and Minimac4 following the guidelines of Minimac4. The final imputed dataset contained genotypes for 5,370,088 autosomal SNPs and 183,199 X-chromosomal SNPs with an imputation quality score (Minimac Rsq) > 0.3, mean allele frequency > 1%, and missing rate < 2%.

Statistical analyses

Population stratification analysis. We first used PLINK, version 1.9 (Purcell et al., 2007), to carried out principal component analysis to correct for possible population stratification. We combined our datasets with samples of 102 Han Chinese in Beijing, China, 98 Utah residents with Northern and Western European ancestry from the CEPH collection, and 107 Yoruba in Ibadan, Nigeria from 1000 Genomes phase 3 data and selected 56,723 SNPs with linkage equilibrium $r^2 < 0.1$ for analysis to account for potential biases introduced by linkage disequilibrium structure. Han Chinese in Beijing, China, and both our discovery and replication set were clustered, and there were no outliers in our dataset (Supplementary Figure S1). The top five genetic principal components were later included in the GWAS model. The number of principal components used was based on the proportion of variance explained and the degree of genomic inflation (lambda).

Association test. Correlation analysis was performed using R software (version 4.0.1) between sensitive skin and sex by Pearson chi-square test and between sensitive skin and chronological age by Welch two-sample *t*-test. We found that reported case of self-reported sensitive skin was significantly higher in females than in males (63.30% vs. 45.67%, $P = 9.98 \times 10^{-19}$). Overall, the age of the individuals with self-reported sensitive skin was slightly younger than that of the control group ($P = 0.04$); however, there was no significant association between age and self-reported sensitive skin after stratified by sex ($P_{female} = 0.17$, $P_{male} = 0.15$). The distribution of the trait stratified by sex and age is shown in Supplementary Figure S2. Age and sex are included as covariates in the GWAS model as a convention. We also performed Pearson chi-square test between potentially relevant dermatologic conditions with self-reported sensitive skin and found significant associations (Supplementary Table S4).

We then performed a genome-wide association analysis for self-reported sensitive skin with logistic regression and an additive genetic model incorporating age, sex, and the first five genetic principal components. Associations were significant if the *P*-values passed the genome-wide significant threshold (Supplementary Figure S3a). We performed a meta-analysis with PLINK, including the discovery set and the replication set, applying the standard error–based method (Purcell et al., 2007). X chromosomes were also included in the GWAS and meta-analysis (Supplementary Figure S3d). For loci of interest, we performed linkage disequilibrium analysis using an online tool LocusZoom (Pruim et al., 2010). We also performed a conditional GWAS, controlling the top-associated SNP to identify any independent signals (Supplementary Figure S3b). We further performed another GWAS, including potentially relevant dermatologic conditions as covariances to control for skin health (Supplementary Figure S3c). In addition, GWAS and meta-analysis were repeated after sex stratification (Supplementary Table S3) or on SNP–sex interaction (Supplementary Figure S3e).

Tests for natural selection. To test for signals of natural selection, we used the Composite of Multiple Signals method (Grossman et al., 2010), which incorporates the scores on the basis of three distinct signatures of selection: long-range haplotypes, differentiated alleles, and high frequency–derived alleles. We tested for signals of positive natural selection with Composite of Multiple Signals scores at 2p21 region extracted from the genome-wide scores of Yoruba in Ibadan, Nigeria, Utah residents with Northern and Western European ancestry from the CEPH collection, and East Asia (Japanese in Tokyo, Japan + Han Chinese in Beijing, China) provided by the Broad Institute (https://www.broadinstitute.org/cms/cms-composite-multiple-signals). Results are presented in Supplementary Figure S5.

Functional annotation

We queried the signal locus against the 3DSNP database (Lu et al., 2017) for functionality scores, and we used HaploReg, version 4.1 (Ward and Kellis, 2012), to annotate the identified variants for their functional relevance and the University of California, Santa Cruz Genome Browser (Kent et al., 2002) on Human Feb. 2009 (GRCh37/hg19) Assembly for visualization (http://epigenomegateway.wustl.edu/legacy/). DNase hypersensitive site and Histone mark annotation (H3K27ac, H3K4me1) peaks for the human skin cells were obtained from the encyclopedia of DNA elements and Roadmap database (ENCODE Project Consortium et al., 2007). Results are presented in Figure 1c and d.

REFERENCES

ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007;447:799–816.

Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. Bioinformatics 2015;31:782–4.

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015;526:68–74.

Grossman SR, Shlyakhter I, Karlsson EK, Byrne EH, Morales S, Frieden G, et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. Science 2010;327:883–6.

Kamide R, Misery L, Perez-Cullell N, Sibaud V, Taïeb C. Sensitive skin evaluation in the Japanese population. J Dermatol 2013;40:177–81.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res 2002;12:996–1006.

Loh PR, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 2016;48:1443–8.

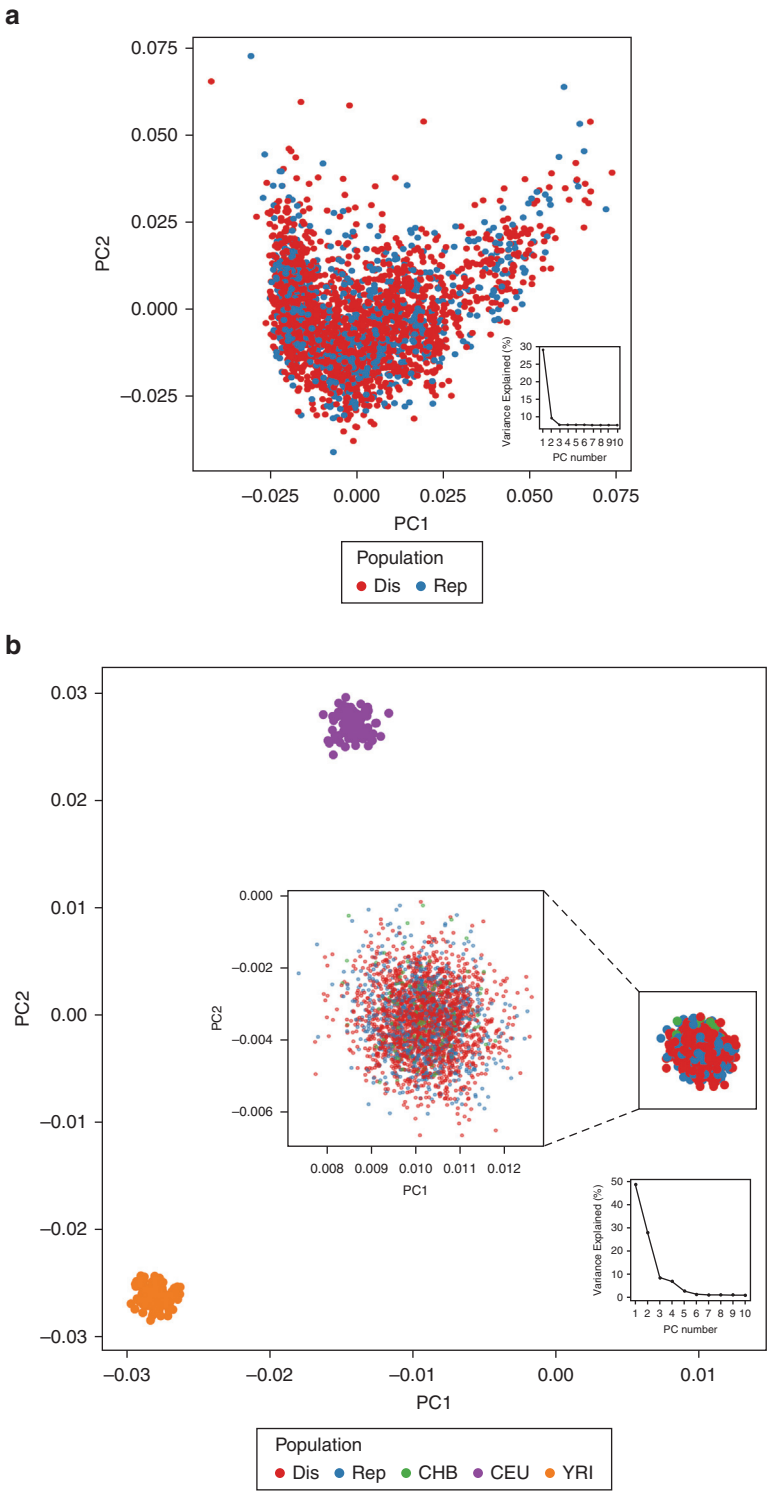
Lu Y, Quan C, Chen H, Bo X, Zhang C. 3DSNP: a database for linking human noncoding SNPs to their three-dimensional interacting genes. Nucleic Acids Res 2017;45:D643–9.

Marcus JH, Novembre J. Visualizing the geography of genetic variants. Bioinformatics 2017;33:594–5.

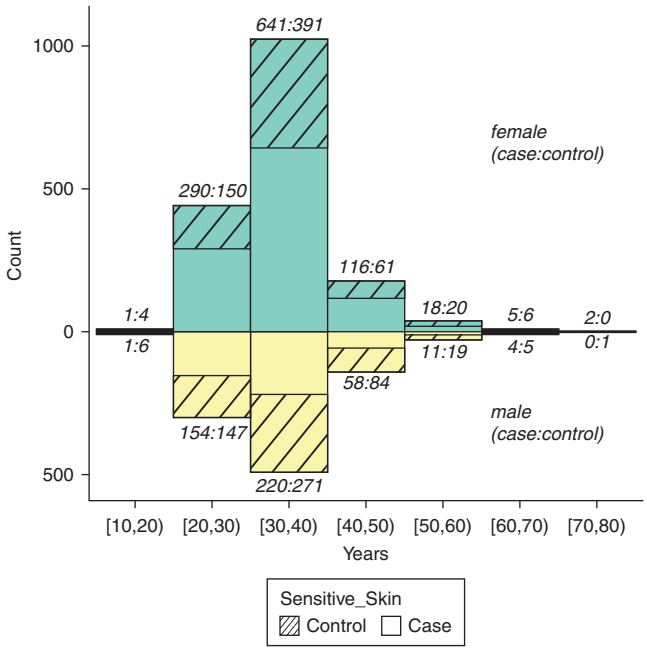
Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336–7.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.

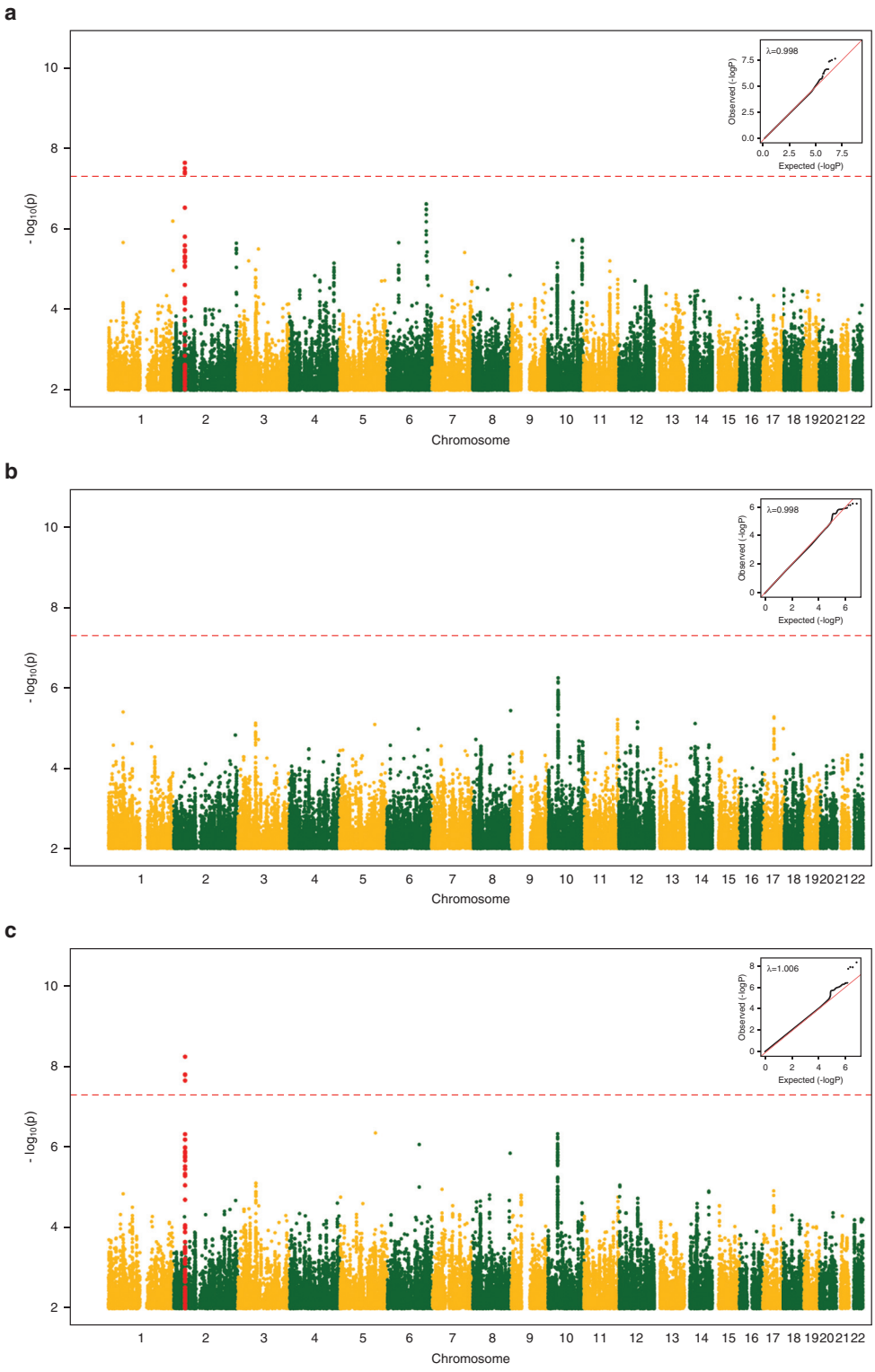
Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012;40:D930–4.



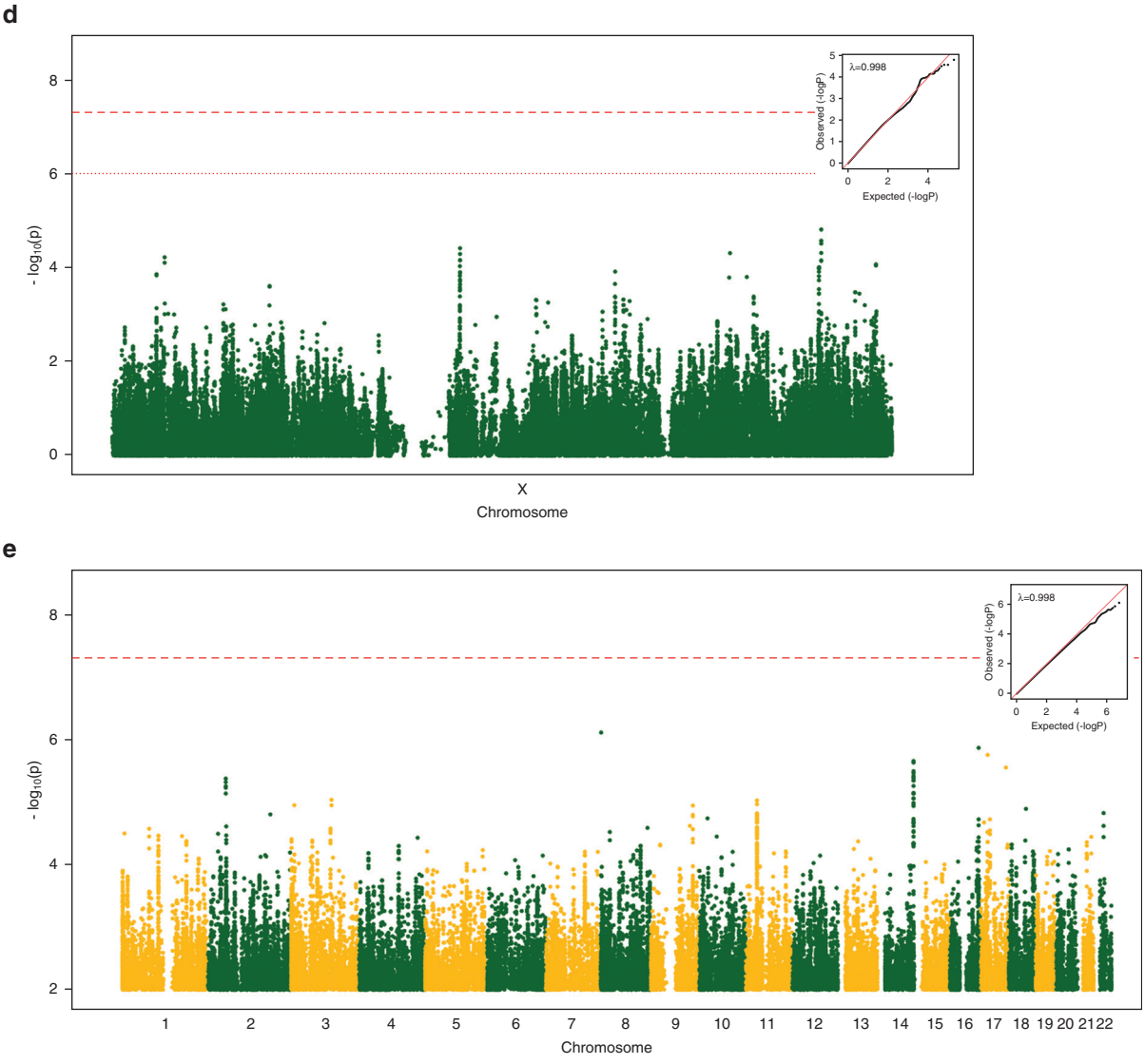
Supplementary Figure S1. Structural analysis of the samples from discovery and replication sets, combined with representative populations from the 1000 Genomes Project. (**a**) Principal component analysis for the discovery and replication set showed that the two populations cluster together. (**b**) Joint principal component analysis with three populations from 1000 Genomes phase 3 data. The zoomed-in lot again showed that WeGene cohort samples and Han Chinese samples are mostly in one cluster, which indicated that our samples are well overlapped with Han Chinese samples from 1000 Genomes phase 3 data and that there were no outliers in our samples. CEU, Utah residents with Northern and Western European ancestry from the CEPH collection; CHB, Han Chinese in Beijing, China; Dis, discovery; PC, principal component; Rep, replication; YRI, Yoruba in Ibadan, Nigeria.



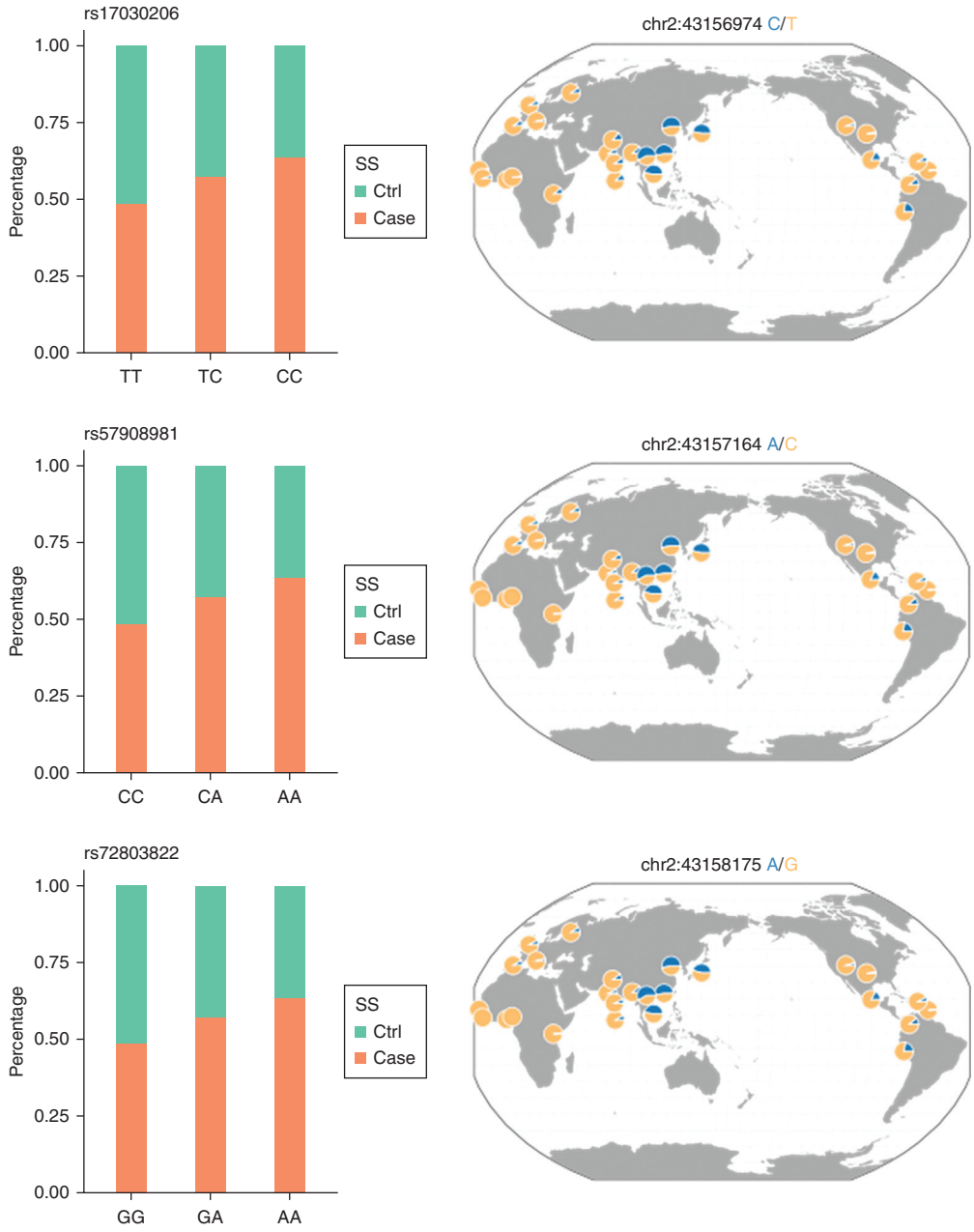
Supplementary Figure S2. Distribution of self-reported sensitive skin, stratified by sex and chronological age. The case of the trait was counted in every 10 years of age, and the count numbers were marked on the corresponding bars.



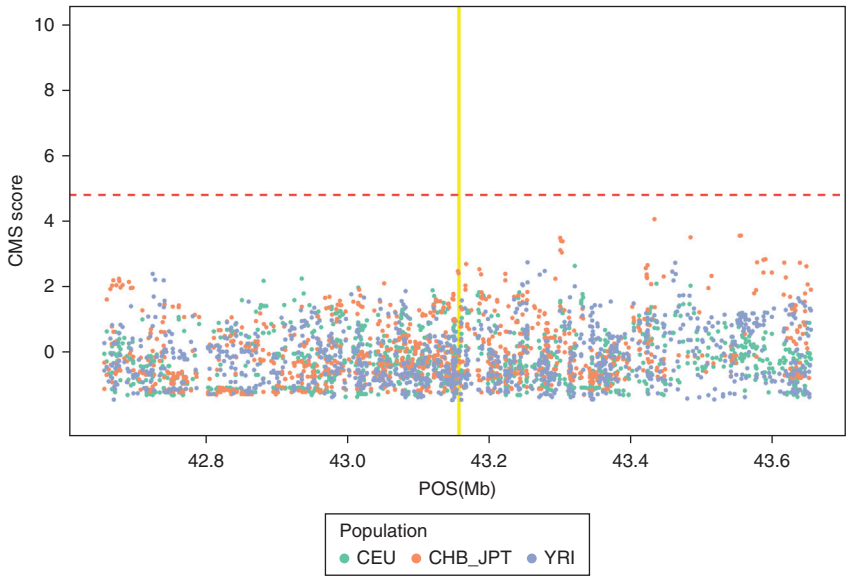
Supplementary Figure S3. Manhattan plot and quantile–quantile plot showing the results of the genome-wide scan for self-reported sensitive skin in the discovery set, the conditional model, and a GWAS-controlling health factors. The panels showed the Manhattan plot illustrating the results of (**a**) the GWAS in 1,872 samples adjusted for age, sex, and the first five genetic PCs; (**b**) the conditional GWAS adjusted for age, sex, the first five genetic PCs, and the top-associated SNP rs72803822; (**c**) the meta-analysis of discovery and replication set, adjusted for age, sex, the first five genetic PCs, and potentially relevant health factors, eczema/dermatitis, atopic dermatitis in childhood, and asthma/hay fever; and (**d**) the meta-analysis of the discovery and replication set, adjusted for age, sex, the first five genetic PCs on the X chromosome, and (**e**) on SNP–sex interaction. The red line corresponds to the threshold for genome-wide statistical significance ($P \leq 5 \times 10^{-8}$). SNPs that were close (<5 kb) to the genome-wide significant signals were plotted in red. The quantile–quantile plot (embedded) compared the empirical values on the vertical axis with simulated values on the horizontal axis. The x-axis and y-axis showed the expected *P*-values under null distribution and the observed *P*-values, respectively. The overall inflation of the observed versus expected distribution of association test statistics was reflected by lambda (λ). Here, $\lambda = 0.998$~1.006, showing no sign of confounding effects. PC, principal component.



Supplementary Figure S3. Continued.



Supplementary Figure S4. Effect and frequency of derived alleles on chromosome 2p21. The left panel showed the proportion of the trait against the genotypes of the three loci on region 2p21. The right panel showed the geography of allele frequency on the corresponding SNPs. The effect alleles were marked by blue. Allele frequencies were obtained from 1000 Genome data and visualized by the Geography of Genetic Variants Browser (Marcus and Novembre, 2017). Ctrl, control; SS, sensitive skin.



Supplementary Figure S5. Results of the test for natural selection on chromosome 2p21. CMS scores were plotted against physical distance for region 2p21 in CEU, CHB + JPT, and YRI populations, shown in green, orange, and blue, respectively. The LD block containing the four significantly associated SNPs was denoted by a yellow rectangle. No enrichment for high CMS scores (Top 0.1%) was found, indicating that the region was not subject to natural selection. CEU, Utah residents with Northern and Western European ancestry from the CEPH collection; CHB, Han Chinese in Beijing, China; CMS, Composite of Multiple Signal; LD, linkage disequilibrium; POS, position; JPT, Japanese in Tokyo, Japan; YRI, Yoruba in Ibadan, Nigeria.

Supplementary Table S1. Sample Summary of the Two Genome-Wide Studies in WeGene Cohort

Characteristics	Discovery	Replication
Subjects, n	1,872	817
Sex, n (%)		
Female	1,192 (63.68)	509 (62.30)
Male	680 (36.32)	308 (37.70)
Age, years, mean (SD)	29.03 (7.22)	28.55 (8.14)
Self-reported sensitive skin case, n (%)		
Very sensitive	270 (14.42)	152 (18.60)
Rather sensitive	770 (41.13)	336 (41.13)
Slightly sensitive	591 (31.57)	227 (27.78)
Not sensitive	241 (12.87)	102 (12.48)

Supplementary Table S2. Original Survey Questions in Chinese

请问您认为自己面部皮肤是否容易对外界刺激产生不适,如发红,瘙痒,紧绷,刺痛,灼烧感等等不舒服的感觉?:非常敏感/较为敏感/不太敏感/不敏感
请问您目前是否患有湿疹或皮炎:是/否
请问您是否从小就患有湿疹或皮炎(非后天传染所得):是/否
请问您是否患有哮喘或者过敏性鼻炎:是/否

Supplementary Table S3. Summary of Locus at 2p21 in Meta-Analysis after Sex Stratification

				Female (n = 1,701)			Male (n = 988)		
CHR	SNP	BP[1]	EA	OR	95% CI	*P-Value*	OR	95% CI	*P-Value*
2	43155982	rs17030203	G	1.40	(1.21−1.62)	3.68×10^{-6}	1.36	(1.14−1.63)	0.0008
2	43156974	rs17030206	C	1.40	(1.21−1.62)	4.49×10^{-6}	1.35	(1.12−1.62)	0.0012
2	43157164	rs57908981	A	1.40	(1.21−1.62)	3.74×10^{-6}	1.35	(1.13−1.61)	0.0010
2	43158175	rs72803822	A	1.41	(1.22−1.63)	2.64×10^{-6}	1.39	(1.16−1.67)	0.0005

Abbreviations: BP, base-pair; CHR, chromosome; CI, confidence interval; EA, effect allele.
[1]Base-pair positions are according to human reference hg19.

Supplementary Table S4. Breakdown of Dermatologic Conditions and Self-Reported Sensitive Skin

		Self-Reported Sensitive Skin			
		Discovery		Replication	
Dermatologic Conditions	Catalog	Case	Control	Case	Control
Eczema/dermatitis	Case	409	173	210	79
(present)	Control	631	659	278	250
	$P_{Chi\text{-}squared\ test}$	1.14×10^{17}		3.75×10^{-8}	
Atopic dermatitis in childhood	Case	189	62	91	29
	Control	851	770	397	300
	$P_{Chi\text{-}squared\ test}$	2.14×10^{-11}		1.49×10^{-4}	
Asthma/hay fever	Case	448	239	181	95
	Control	592	593	307	234
	$P_{Chi\text{-}squared\ test}$	2.11×10^{-10}		0.018	

Supplementary Table S5. Allele Frequencies of Locus at 2p21 across Populations

				Frequencies[2]			
CHR	BP[1]	SNP	EA	AFR	AMR	ASN	EUR
2	42928842	rs17030203	G	0.01	0.12	0.51	0.06
2	42929834	rs17030206	C	0.06	0.14	0.51	0.07
2	42930024	rs57908981	A	0.01	0.12	0.51	0.06
2	42931035	rs72803822	A	0.01	0.12	0.51	0.06

Abbreviations: AFR, African; AMR, Ad Mixed American; ASN, Asian; BP, base-pair; CHB, Han Chinese in Beijing, China; EA, effect allele; EUR, European; JPT, Japanese in Tokyo, Japan.
[1]Base-pair positions are according to human reference hg19.
[2]Allele frequencies were obtained from HaploReg, version 4.1.