Genetic Loci Associated with Nail Plate Morphology in East Asian Populations

Journal of Investigative Dermatology (2025) **145,** 969–972; doi:10.1016/j.jid.2024.08.035

TO THE EDITOR

The nail plate, a robust layer of hard keratin, serves as a critical protective barrier for the fingertips. Its morphology, influenced by intricate genetic determinants, presents considerable diversity across various populations and individuals. This morphological variation is not merely of cosmetic interest but is pivotal in the clinical context, particularly for disorders such as nail–patella syndrome characterized by distinct nail plate abnormalities (Khadka et al, 2023). The genetic architecture of the nail plate is largely unexplored, with significant implications for dermatology, anthropology, and population genetics. We report a GWAS to reveal the genetic variations shaping nail plate area, length, and width.

This study included 6795 individuals, mostly of Han ethnicity, with 5034 and 1231 from 2 Jidong cohorts (JD2019 and JD2022) and 530 from the cohort of National Survey of Physical Traits (Supplementary Table S1). Ethical approval was obtained from the Ethics Committees of Fudan University (14117), the Shanghai Institutes for Biological Sciences (ER-SIBS-261410), and the Shanghai Institute of Nutrition and Health (ER-SINH-262203) (Chen et al, 2024; Zhang et al, 2022). All participants provided written informed consent. Genomic principal component analysis indicated no significant population substratifications (Zhang et al, 2022).

We analyzed 3 distinct shape metrics—area (mm^2), length (mm), and width (mm) (Figure 1a)—of the nail plates on 4 fingers: the pinky (D5), ring (D4), middle (D3), and index (D2) fingers (Supplementary Figure S1). Our GitHub-accessible pipeline (https://github.com/MengxiangYou/NailNet) is based on the DeepLabv3+ model for automatically segmenting nail plates from hand photos (Fan et al, 2024[1]). Distributions were statistically not normal (Shapiro–Wilk test, false discovery rate $P < .05$), despite their visually normal appearance. An inverse normal transformation was applied for subsequent analyses (Supplementary Figure S2). All phenotypes were positively correlated with age (Pearson's $r > 0.06$) and body mass index ($r > 0.21$) (Supplementary Figure S3). Males showed larger nail plate area, length, and width than females (*t*-test, $P < .05$) (Supplementary Figure S3).

Individual GWASs were independently conducted on nail plate area, length, and width across 4 fingers in each cohort and then meta-analyzed across 3 cohorts, including covariates chip batch, sex, age, body mass index, and genomic principal components (Supplementary Figure S4). Individual GWASs and meta-analyses did not show inflation in test statistics ($\lambda < 1.02$). We applied the C-GWAS method (Xiong et al, 2022) to integrate the meta-analysis results, which revealed 5 study-wide significant loci ($P <$ 5e-8) (Figure 1b): 1q41 (*LYPLAL1_AS1*, rs2605107; $P =$ 4.94e-10), 2q31.1 (*OLA1*, rs10497417; $P =$ 2.91e-08), 2q33.1 (*FTCDNL1*, rs4672566; $P =$ 1.79e-08), 5q32 (*SPINK6*, rs12186491; $P =$ 1.14e-10), and 16q12.2 (*RPGRIP1L*, rs4784319, $P =$ 9.88e-09) (Table 1).

Our post-GWAS analyses and literature review provide robust evidence for the functional roles of loci 1q41, 5q32, and 16q12.2 in nail formation. The 1q41 locus shows genome-wide significant associations with the plate areas of the D5, D4, and D2 nails as well as the width of the D4 nail ($P <$ 5e-8) (Supplementary Figure S5 and Supplementary Table S2). The predominant SNP at this locus, rs2605107, located within an intron of *LYPLAL1_AS1*, functions as expression and splicing quantitative trait loci for *LYPLAL1_AS1* in skin tissues. The derived C allele frequency in our cohort (0.23) closely matches that of the Korean Genome Project (0.22) but differs from frequencies in African (0.07) and European (0.32) populations in the 1000 Genomes Project (Supplementary Figure S6). *LYPLAL1-AS1* is highly expressed in the skin, playing a significant role in processes such as epidermal cell differentiation, and is involved in structural formations such as desmosomes (Yang et al, 2022). Deletions in *LYPLAL1-AS1* was previously associated with moderate nail dysplasia (Filges et al, 2010). In UK Biobank, this locus showed significant associations with traits such as male-pattern baldness and skin cancer, underscoring its critical role in hair growth and skin cell proliferation (Supplementary Table S3).

At 5q32, the missense variant rs12186491 in *SPINK6* shows a significant association with the nail plate areas of D4 and D3 as well as the lengths of D4, D3, and D2. This association reaches genome-wide significance in single-trait GWAS ($P <$ 5e-8) (Supplementary Figure S7, Figure 1c, and Supplementary Table S4). The presence of the C allele is linked to larger nail plate areas and lengths, a pattern that is consistently observed across the 3 cohorts (Figure 1e and Supplementary Figure S8). The allele frequency of the C allele is 0.34 in our study, which is comparable with its frequency in East Asian populations (0.34) but differs from those observed in African (0.62) and European (0.08) populations (Figure 1d). With a combined annotation-dependent depletion score of 21.4, indicative of its missense

Accepted manuscript published online 9 October 2024; corrected proof published online 29 October 2024

[1] Fan Y, You M, Ge J, Zhai G, Wang S. Segmentation and feature extraction of fingernail plate and lunula based on deep learning. bioRxiv, https://www.biorxiv.org/content/10.1101/2024.07.26.605289v1; 2024 (accessed July 2024).

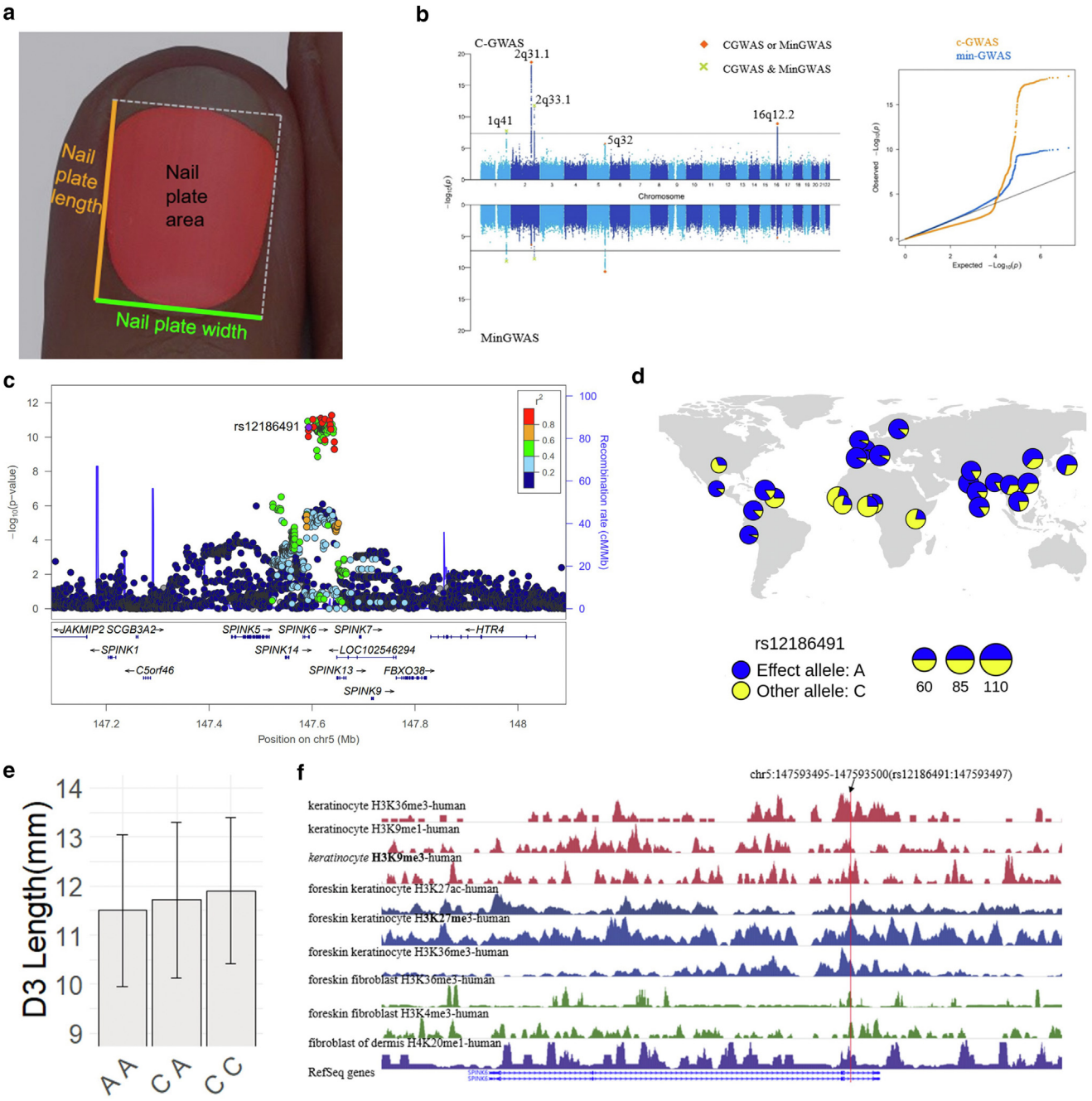


Figure 1. A genome-wide association scan on nail plate area, nail length, and nail width. (**a**) Phenotypic characteristics of nail plate area, nail length, and nail width (detailed definitions are provided in Supplementary Figure S1). (**b**) C-GWAS identified a total of 5 significant loci within the study scope, with 2 significant in both C-GWAS and MinGWAS, 2 significant only in MinGWAS, and 1 significant only in C-GWAS. $-\text{Log}_{10}$ values are plotted against the physical positions of all SNPs and their associated *P*-values (on the basis of genome assembly GRCh38.p13), with the black solid line indicating the genome-wide significance threshold ($P = 5 \times 10^{-8}$). (**c**) Regional association plot for the 5q32 signal in the meta-analysis of nail plate length in D3, where increasing color intensities represent higher levels of linkage disequilibrium (r2) with rs12186491. (**d**) Geographical distribution of allele frequency on rs12186491. (**e**) Functional relationship between the mean nail plate length of D3 in both sexes and the genotype of rs12186491 in the discovery cohort (JD2019, comprising 4465 adults, including 2136 males and 2329 females). Vertical bars correspond to the SEM. (**f**) Epigenetic annotation status of rs12186491, with the region where rs12186491 is located annotated in red.

nature, rs12186491 likely plays a functional role. Further investigations indicated that the region encompassing rs12186491 features active enhancer marks in keratinocytes, such as H3K27ac histone modification (Figure 1f), hinting at an enhancer-dependent regulatory mechanism for *SPINK6* expression. *SPINK6*, acting as a serine protease inhibitor, shows substantially increased expression in nail-specific keratinocytes compared with that in normal skin (Kim et al, 2021). Serving as a specific inhibitor for kallikrein-related peptidases, *SPINK6* is pivotal in maintaining skin barrier homeostasis and has demonstrated efficacy in inhibiting skin desquamation in vitro (Fischer et al, 2014). *SPINK6* expression was found to double during the fingertip regeneration process

Table 1. SNPs Significantly Associated with Nail Plate Area, Length, and Width

Cytoband	Lead SNP	Mapped Gene	Meta		JD2019[1]		JD2022[1]		NSPT[1]		CGWAS[2]	MinGWAS[2]			EA Frequency		
			β	*P*-Value	β	*P*-Value	β	*P*-Value	β	*P*-Value	*P*-Value	*P*-Value	EA	OA	JD2019	JD2022	NSPT
1q41	rs2605107	*LYPLAL1-AS1*	−0.12	2.25E-09	−0.16	4.94E-10	0.00	9.48E-01	−0.16	9.00E-02	9.50E-07	2.30E-08	C	A	0.224	0.23	0.232
5q32	rs12186491	*SPINK6*	0.09	2.29E-08	0.15	1.14E-10	0.10	1.80E-02	0.05	3.65E-01	8.96E-06	3.38E-10	C	A	0.202	0.206	0.185
2q33.1	rs4672566	*FTCDNL1*	−0.13	5.08E-09	−0.15	1.79E-08	−0.14	3.00E-03	−0.01	8.56E-01	1.14E-10	1.36E-08	C	A	0.203	0.205	0.184
2q31.1	rs10497417	*OLA1*	0.11	2.91E-08	0.11	1.62E-06	0.11	1.39E-02	0.10	1.51E-01	9.86E-19	3.17E-07	A	G	0.339	0.346	0.347
16q12.2	rs4784319	*RPGRIP1L*	−0.06	6.29E-06	−0.04	1.94E-02	−0.10	6.06E-04	−0.15	9.93E-04	9.88E-09	6.44E-05	A	T	0.477	0.471	0.474

Abbreviations: EA, effect allele; NSPT, National Survey of Physical Traits; OA, other allele.
[1]JD2019: The Jidong cohort collected in 2019 (Chen et al, 2024; Zhang et al, 2022); JD2022: The Jidong cohort collected in 2022; and NSPT: The cohort of NSPT (Chen et al, 2024; Zhang et al, 2022).
[2]C-GWAS: A method for combining GWAS summary statistics of multiple potentially correlated traits (Xiong et al, 2022) and MinGWAS: The minimal *P*-values of multiple single-trait GWASs.

(Maan et al, 2021). These insights suggest *SPINK6*'s regulatory influence on nail size, mediated through its impact on formation and positioning of the nail matrix.

The lead SNP rs4784319 (16q12.2) is situated in 3′ untranslated region of *RPGRIP1L*. The T allele was associated with reduced nail plate area and length for fingers D2−5, albeit with an increased width (Supplementary Figure S8). The T allele frequency (0.66) contrasts with those of East Asian and European populations (0.58 each) and is notably higher than that in African populations (0.28) (Supplementary Figure S6). *RPGRIP1L*, highly expressed in skin tissues, is vital for the structural integrity of desmosomes among keratinocytes. This gene is also crucial in hair follicle development and is linked to skin conditions such as epidermolysis bullosa (Choi et al, 2019). Analysis in UK Biobank reveals a significant association between this locus and skin and hair color (Supplementary Table S3).

The functional involvement of the remaining 2 loci, 2q31 and 2q33.1, in nail development is less clear. At 2q31.1, the lead SNP rs10497417 resides within an intron of *OLA1* (Supplementary Figure S9) and acts as an expression quantitative trait locus for *OLA1* in musculoskeletal tissues. In our dataset, the derived G allele frequency was 0.79, slightly differing from those of East Asian (0.76), African (0.82), and European (0.85) populations (Supplementary Figure S6). It was linked to reduced dimensions in D3 and D2 nails (Supplementary Figure S8). Analysis in UK Biobank reveals a significant association of *OLA1* with male-pattern baldness (Supplementary Table S3). At 2q33.1, the lead SNP rs72924480 is located within an intron of *FTCDNL1*(Supplementary Figure S10), holding a combined annotation-dependent depletion score of 16.98. The derived allele C of rs4672566, present at a frequency of 0.20 in our sample, varies across populations: 0.13 in East Asians, 0.01 in Africans, and 0.06 in Europeans (Supplementary Figure S6). Carriers of this allele showed decreased nail length and nail plate area (Supplementary Figure S8).

In summary, this study presents, to our knowledge, a previously unreported GWAS of nail plate morphology, identifying 5 significant genetic loci influencing nail plate area, length, and width. These loci are associated with broader dermatological and hair-related traits, highlighting their potential roles in skin and nail biology. Although our study's focus on a genetically homogenous Chinese Han cohort may limit the identification of genetic factors relevant to other populations, future research with larger and more diverse samples is needed to explore the full range of genetic influences on nail size.

DATA AVAILABILITY STATEMENT

The GWAS summary statistics are available from the National Omics Data Encyclopedia (http://www.biosino.org/node/) under the project identification document OEP005227. Data usage must be in full compliance with the Regulations on Management of Human Genetic Resources in China. Individual genotype and phenotype data cannot be shared owing to restrictions of the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (ER-SIBS-261410); the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences (ER-SINH-262203); and the Ethics Committee of Human Genetic Resources of School of Life Sciences, Fudan University (Shanghai, China) (14117) on privacy concerns. Other relevant data supporting the key findings of this study are available within the letter and supplementary materials or from the corresponding author on reasonable request.

KEYWORDS

Deep learning; GWASs; Nail plate shape; SPINK6

ORCIDs

Jieyu Ge: http://orcid.org/0000-0002-1070-2087
Mengxiang You: http://orcid.org/0009-0004-9518-8399
Yu Fan: http://orcid.org/0009-0009-0957-5541
Yong Zhou: http://orcid.org/0000-0001-5221-8026
Li Jin: http://orcid.org/0000-0001-9201-2321
Guangtao Zhai: http://orcid.org/0000-0001-8165-9322
Fan Liu: http://orcid.org/0000-0001-9241-8161
Sijia Wang: http://orcid.org/0000-0001-6961-7867

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant number 32325013), the CAS Project for Young Scientists in Basic Research, the Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDB38020400 to SW), Shanghai Science and Technology Commission Excellent Academic Leaders Program (22XD1424700), and the Human Phenome Data Center of Fudan University.

AUTHOR CONTRIBUTIONS

Conceptualization: SW; Data Curation: JG, MY; Methodology: JG, MY, SW, FL; Data Curation: SW, LJ, YZ, JG, MY; Formal Analysis: MY, JG; Visualization: SW, MY, JG, YF, GZ; Funding Acquisition: SW; Supervision: SW, FL; Writing - Original Draft Preparation: MY, JG, SW, FL; Writing - Review and Editing: JG, MY, YF, YZ, LJ, GZ, FL, SW

***Jieyu Ge*[1,9], *Mengxiang You*[1,9], *Yu Fan*[2], *Yong Zhou*[3], *Li Jin*[4,5,6], *Guangtao Zhai*[2], *Fan Liu*[7] and *Sijia Wang*[1,8,*]**

[1]CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China; [2]School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, People's Republic of China; [3]Clinical Research Institute, Shanghai General Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; [4]State Key Laboratory of Genetic Engineering, Human Phenome Institute, Zhangjiang Fudan International Innovation Center, Fudan University, Shanghai, China; [5]Fudan-Taizhou Institute of Health Sciences, Taizhou, China; [6]Research Unit of Dissecting the Population Genetics and Developing New Technologies for Treatment and Prevention of Skin Phenotypes and Dermatological Diseases (2019RU058), Chinese Academy of Medical Sciences, Shanghai, China; [7]Department of Forensic Sciences, College of Criminal Justice, Naif Arab University for Security Sciences, Riyadh, Saudi Arabia; and [8]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[9]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.2024.08.035.

REFERENCES

Chen W, Li J, Xu K, Luo J, Wang M, Liu Y, et al. Genetic variants associated with sweat gland phenotypes in 6210 Han Chinese individuals. J Invest Dermatol 2024;144:2096–2100.e8.

Choi YJ, Laclef C, Yang N, Andreu-Cervera A, Lewis J, Mao X, et al. RPGRIP1L is required for stabilizing epidermal keratinocyte adhesion through regulating desmoglein endocytosis. PLoS Genet 2019;15:e1007914.

Filges I, Röthlisberger B, Boesch N, Weber P, Wenzel F, Huber AR, et al. Interstitial deletion 1q42 in a patient with agenesis of corpus callosum: phenotype-genotype comparison to the 1q41q42 microdeletion suggests a contiguous 1q4 syndrome. Am J Med Genet A 2010;152A:987–93.

Fischer J, Wu Z, Kantyka T, Sperrhacke M, Dimitrieva O, Koblyakova Y, et al. Characterization of Spink6 in mouse skin: the conserved inhibitor of kallikrein-related peptidases is reduced by barrier injury. J Invest Dermatol 2014;134:1305–12.

Khadka N, Kc P, Sharma S, Sharma S, Rayamajhi S. Dilated cardiomyopathy and nail-patella syndrome: a case report. Cureus 2023;15:e48969.

Kim HJ, Shim JH, Park JH, Shin HT, Shim JS, Jang KT, et al. Single-cell RNA sequencing of human nail unit defines RSPO4 onychofibroblasts and SPINK6 nail epithelium. Commun Biol 2021;4:692.

Maan ZN, Rinkevich Y, Barrera J, Chen K, Henn D, Foster D, et al. Epidermal-derived hedgehog signaling drives mesenchymal proliferation during digit tip regeneration. J Clin Med 2021;10:4261.

Xiong Z, Gao X, Chen Y, Feng Z, Pan S, Lu H, et al. Combining genome-wide association studies highlight novel loci involved in human facial variation. Nat Commun 2022;13:7832.

Yang Y, Liu S, He C, Lv T, Zeng L, Zhang F, et al. LncRNA LYPLAL1-AS1 rejuvenates human adipose-derived mesenchymal stem cell senescence via transcriptional MIR-LET7B inactivation. Cell Biosci 2022;12:45.

Zhang M, Wu S, Du S, Qian W, Chen J, Qiao L, et al. Genetic variants underlying differences in facial morphology in East Asian and European populations. Nat Genet 2022;54:403–11.

Estimation of Body Mass Index from 3-Dimensional Total Body Photography

Journal of Investigative Dermatology (2025) **145,** 972–975; doi:10.1016/j.jid.2024.06.1294

TO THE EDITOR

Obesity is considered a major preventable risk factor for cancer (Lauby-Secretan et al, 2016; Reilly and Saltiel, 2017). Recent evidence suggests that obesity, measured through the body mass index (BMI) coefficient, is a relevant marker for melanoma incidence and prognostic features, including Breslow thickness, tumor ulceration, and mitotic rate (Stenehjem et al, 2018). In response to this evidence, the International Agency for Research on Cancer has called for further research on the association between melanoma and obesity (Lauby-Secretan et al, 2016).

The increased use of 3-dimensional total body photography (3D-TBP) systems for melanoma surveillance offers spatial geometric data that could facilitate automated anthropomorphic measurements to estimate BMI. Similar methods using spatial data from purpose-built laser body scanners are reported to calculate BMI with high accuracy compared with in-person measurements (Koepke et al, 2017). It is previously untested whether anthropomorphic measurements derived from 3D-TBP offer a similar degree of accuracy. In this study, we validate the anthropomorphic measurements extracted from 3D-TBP to estimate BMI and compare with actual BMI.

A total 125 consecutive participants were recruited from the Australian Centre of Excellence in Melanoma Imaging and Diagnosis Cohort Study with written, informed consent (registration: ACTRN12619001706167) (Koh et al, 2023). Participants were imaged using the Vectra whole-body scanner (Vectra WB360, Canfield Scientific, Parsippany, NJ) to capture 3D-TBP. The Vectra System uses images from 92 cameras to digitally construct 3-dimensional avatars of the participant (Figure 1). Weight (using Seca 762 Mechanical Personal Scale) and height were measured at the imaging appointments. BMI was

Abbreviations: 3D-TBP, 3-dimensional total body photography; BMI, body mass index; LoA, limit of agreement; TBP, total body photography

Accepted manuscript published online 26 September 2024; corrected proof published online 28 October 2024

SUPPLEMENTARY MATERIALS AND METHODS

The JD2019 and JD2022 cohorts

Population statistics for hand image collection are shown in Supplementary Table S1. The Jidong cohort is a community-based, long-term observational cohort study aimed at assessing risk factors related to health. Baseline data were collected from 2013 to 2014 at the Workers' Hospital of the Jidong Oilfield Branch of China National Petroleum Corporation. The study obtained approval from the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (ER-SIBS-261410) and the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences (ER-SINH-262203). There were no duplicate individuals between the JD2019 cohort and the JD2022 cohort. Both cohorts consisted of natural populations, meaning that individual recruitment did not rely on any specific disease status. All participants provided written informed consent. Hand photos and blood samples were collected concurrently at the Workers' Hospital (Chen et al, 2024; Zhang et al, 2022).

During the quality control process for cohorts JD2019 and JD2022, individuals were excluded on the basis of several criteria. These included (i) improper nail plate presentation or blurry images; (ii) signs of onychomycosis or other fungal infections; (iii) presence of nail polish; and (iv) physical nail abnormalities such as injuries, unclean nails, or deformities affecting the fingers. In cohort JD2019, 62 individuals were excluded for improper nail plate presentation or blurry images, 7 were excluded for fungal infections, 83 were excluded for nail polish presence, and 2 were excluded for physical nail abnormalities. In cohort JD2022, 190 individuals were excluded for improper nail plate presentation or blurry images, 4 were excluded for fungal infections, 154 were excluded for nail polish presence, and 8 were excluded for physical nail abnormalities. After quality control, the JD2019 cohort comprised 4465 adults (2136 males and 2329 females, aged 18−89 years, the average age was 46.82 ± 13.34 years), and the JD2022 cohort comprised 1231 adults (610 males and 621 females, aged 21−79 years, the average age was 41.86 ± 9.94 years).

The National Survey of Physical Traits cohort

The National Survey on Physical Traits (NSPT) conducted sample collection in Zhengzhou, Henan in 2017. The NSPT cohort is a population-based prospective cohort study aimed at exploring environmental and genetic factors related to physical traits and diseases. The NSPT cohort study was conducted with formal approval from the Ethics Committee of Human Genetic Resources of School of Life Sciences, Fudan University, Shanghai (14117) and the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (ER-SIBS-261410). All voluntary participants in this project provided written informed consent and obtained approval from the ethics committee (Chen et al, 2024; Zhang et al, 2022).

This cohort also consisted of natural populations, meaning that individual recruitment did not rely on any specific disease status. In cohort NSPT, 207 individuals were excluded for improper nail plate presentation or blurry images, 27 were excluded for fungal infections, 44 were excluded for nail polish presence, and 45 were excluded for physical nail abnormalities. After quality control, the NSPT cohort used in this study included 530 adults (179 males and 351 females, aged 19−71 years, the average age was 43.97 ± 13.39 years).

Phenotyping

The hand image data in this study were collected using a digital SLR camera. At the sampling site, we set up a small photo studio (dimensions: 50 ×60 ×50 cm), equipped with stable light-emitting diode lighting, a black background paper, and a color calibration card with a scale ruler (ColorChecker Passport, X-Rite). There is an opening at the top, and the lens of the digital SLR camera shoots vertically downward to capture the hands of the volunteers inside the photo studio (Supplementary Figure S1a). Each volunteer sequentially places their left-hand dorsum, right-hand dorsum, and thumbs flat in the photo studio, while the staff takes several photos (Supplementary Figure S1b). The scale ruler on the color calibration card is used to restore the real shape size.

In our study, the nail plate area is defined as the area of the entire nail plate, including the lunula, but excluding the free edge and surrounding skin. It is not calculated as the product of length and width. This definition is consistent with previous studies (De Berker et al, 2007; Kumuda and Dinesh, 2015). For illustration, the nail plate area is depicted as the red-colored region in Figure 1a. Supplementary Figure S1b also demonstrates the nail plate area, represented by the number of white pixels.

To efficiently and accurately extract nail plate area, length, and width, this study proposes a novel deep learning framework called NailNet, based on the DeepLabv3+ model (Fan et al, 2024[2]), capable of accurately segmenting the nail plate and lunula in a large number of nail photos. Our proposed NailNet adopts a similar encoder−decoder structure as DeepLabv3+ but incorporates the following improvements: to prevent overfitting and reduce computational costs, we replace the ResNet101 backbone of DeepLabv3+ with a ResNet50 backbone; we utilize depth-wise separable convolutions. Furthermore, to alleviate the model's burden, we replace the convolutional layers in the Atrous Spatial Pyramid Pooling module with depth-wise separable convolutions; to reduce feature discrepancy, we design a feature pyramid structure to integrate some high-level information into low-level feature maps. Ultimately, NailNet achieves an Intersection over Union of 0.9529 and accuracy of 0.9725 for nail plate segmentation (Fan et al, 2024[1]).

Technically, we first use NailNet to segment the nail plate area from whole-hand photos. The length and width of the nail plate are then defined by the dimensions of the minimum bounding rectangle of this segmented area.

[2] Fan Y, You M, Ge J, Zhai G, Wang S. Segmentation and feature extraction of fingernail plate and lunula based on deep learning. bioRxiv, https://www.biorxiv.org/content/10.1101/2024.07.26.605289v1; 2024 (accessed July 2024).

Before conducting GWAS, further quality control steps were performed: individuals with missing identity information and nail photos with inaccurate predictions by NailNet. After quality control of the nail photos, the distribution of the nail plate area, length, and width in the 3 cohorts did not follow a normal distribution. Therefore, we applied an inverse normal transformation to the nail plate area, length, and width, making their mean 0 and SD 1, in accordance with the requirements of the fastGWA mixed linear model for phenotype distribution (Speliotes et al, 2010) (Supplementary Figure S2).

Genotyping

For all the JD2019, JD2022, and NSPT cohorts, genomic DNA was extracted from blood samples using the MagPure Blood DNA KF Kit. All samples were genotyped using the Illumina Infinium Global Screening Array (Illumina) consisting of about 710,000 SNPs. We implemented exclusion criteria for quality control using PLINK, version 1.9 (Chang et al, 2015). In detail, we excluded subjects with >5% missing data, the duplicated subjects, and subjects that failed the X-chromosome sex concordance check or had ethnic information incompatible with their genetic information. We excluded SNPs that had >2% missing data, SNPs with a minor allele frequency <1%, and the ones that failed the Hardy–Weinberg deviation test ($P < 1 \times 10^{-5}$). Imputation was performed using the 1000 Genomes Project phase 3 as the reference panel. The chip genotype data were firstly phased using SHAPEIT3 (O'Connell et al, 2016), and IMPUTE2 was then used to impute genotypes (Howie et al, 2009). SNPs with an imputation quality score (INFO) <0.6, minor allele frequency <0.01, or a missing rate >0.01 were eliminated from further analyses. Finally, 8,039,700 SNPs passed quality control and were used for further analyses (Wang et al, 2022).

Statistical analysis

Heritability. To estimate the genetic variance underlying nail plate area, length, and width in our study population, we employed the restricted maximum likelihood method. We incorporated the effects of all autosomal loci from chip data as random effects and adjusted for the top 10 genomic principal components, sex, age, and body mass index as covariates. Subsequently, we fitted linear mixed models to the data to estimate the narrow-sense heritability of nail plate area, length, and width. Considering the impact of sample size, we combined the JD2019 and JD2022 cohorts and used the GCTA (genome-wide complex trait analysis) software to estimate heritability on the basis of genome-wide SNPs. The overall heritability of nail plate area, length, and width ranged from 0.20 to 0.44 (Supplementary TableS5)

Sex-specific GWAS. In our GWAS, we adjusted for the potential confounding effect of sex by including sex as a covariate in our mixed linear model, as described in the main text. To further examine sex-specific differences, we additionally conducted sex-stratified GWAS. The results show that *LYPLAL1-AS1* (1q41) and *SPINK6* (5q32) remain significant in both female-only and male-only GWAS, with no new significant loci emerging (Supplementary Table S6).

Functional annotation. In this study, we utilized various websites such as dbSNP (Sherry et al, 2001); HaploReg, version 4.1 (Ward and Kellis, 2012); and 3DSNP (Lu et al, 2017) to search for functional annotation information of significant loci. We also employed databases such as GeneCard (Safran et al, 2010) and GTEx (GTEx Consortium et al, 2017) to annotate the functionality, expression, and pathways of candidate genes. Finally, we visualized chromatin modifications, chromatin-binding states, region conservativeness, and nearby genes using the WashU EpiGenome Browser website (Bult et al, 2019) (Figure 1f). For gene-level phenotype–genotype associations, we conducted a phenome-wide association study using GWASATLAS (https://atlas.ctglab.nl/PheWAS), which includes 4756 GWAS results from the UK Biobank database covering 3302 unique traits (Watanabe et al, 2019).

SUPPLEMENTARY REFERENCES

Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE. Mouse Genome Database Group. Mouse genome database (MGD) 2019. Nucleic Acids Res 2019;47:D801–6.

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.

Chen W, Li J, Xu K, Luo J, Wang M, Liu Y, et al. Genetic variants associated with sweat gland phenotypes in 6210 Han Chinese individuals. J Invest Dermatol 2024;144:2096–2100.e8.

De Berker DA, André J, Baran R. Nail biology and nail science. Int J Cosmet Sci 2007;29:241–75.

GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, et al. Genetic effects on gene expression across human tissues. Nature 2017;550:204–13.

Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.

Kumuda NS, Dinesh MS. Human fingernail segmentation International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). Mandya, India: IEEE;2015; 2015. p. 270–5. https://ieeexplore.ieee.org/abstract/document/7499025. (accessed June 27, 2016).

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.

O'Connell J, Sharp K, Shrine N, Wain L, Hall I, Tobin M, et al. Haplotype estimation for biobank-scale data sets. Nat Genet 2016;48:817–20.

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards version 3: the human gene integrator. Database (Oxford) 2010;2010:baq020.

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001;29:308–11.

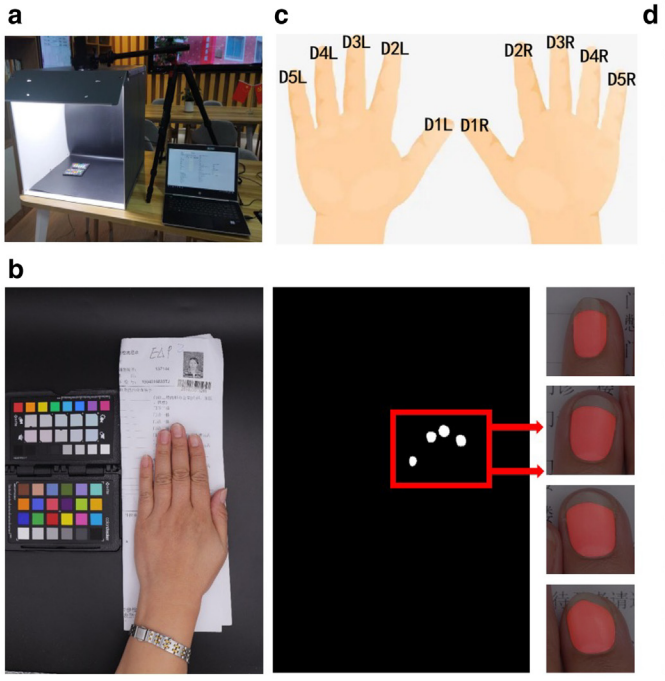
Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010;42:937–48.

Wang F, Luo Q, Chen Y, Liu Y, Xu K, Adhikari K, et al. A genome-wide scan on individual typology angle found variants at SLC24A2 associated with skin color variation in Chinese populations. J Invest Dermatol 2022;142:1223–1227.e14.

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.

Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet 2019;51:1339–48.

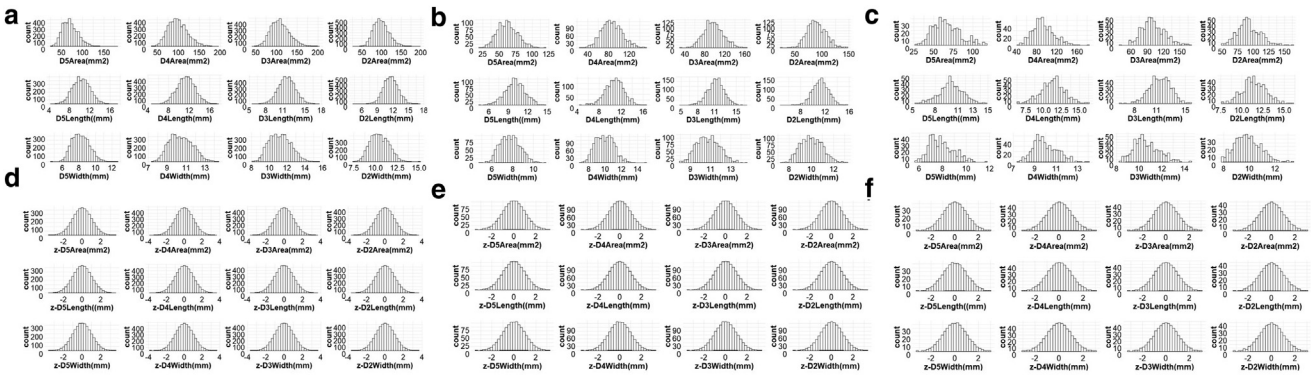
Zhang M, Wu S, Du S, Qian W, Chen J, Qiao L, et al. Genetic variants underlying differences in facial morphology in East Asian and European populations. Nat Genet 2022;54:403–11.



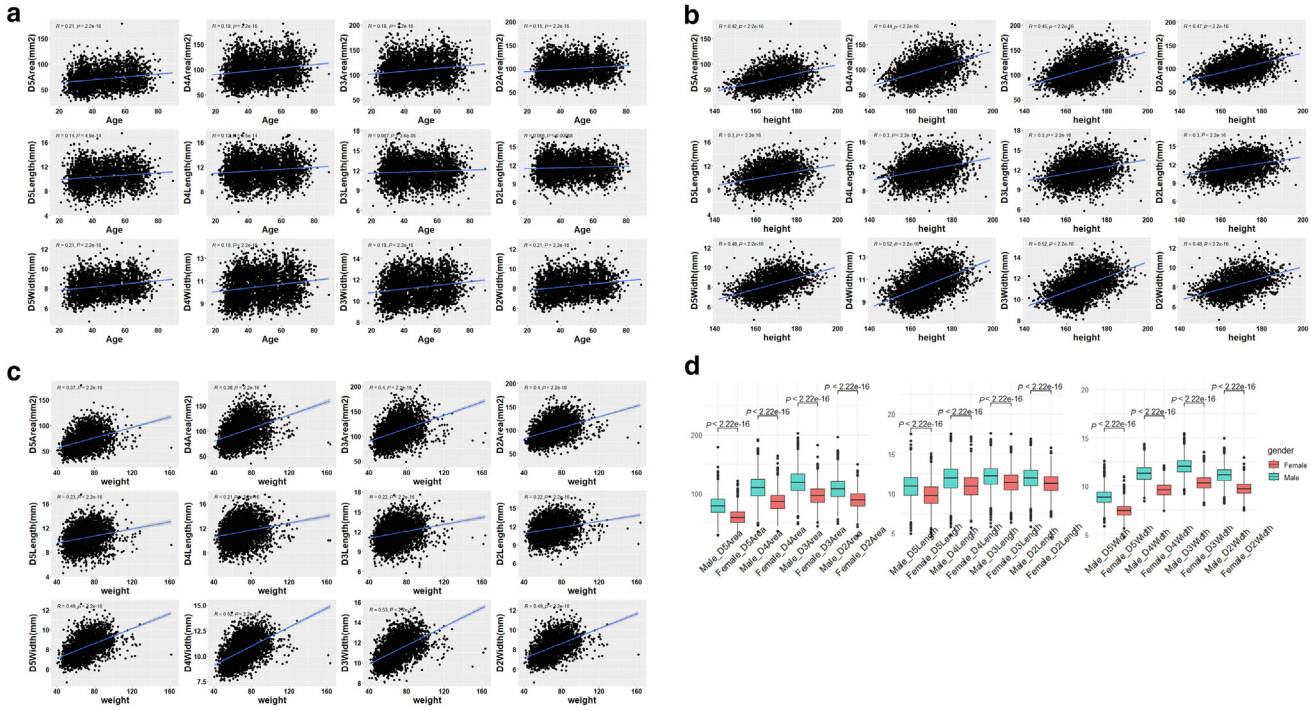
d

Phenotype name	Phenotypic definition
D5 Area	(D5L area + D5R area) ÷ 2
D4 Area	(D4L area + D4R area) ÷ 2
D3 Area	(D3L area + D3R area) ÷ 2
D2 Area	(D2L area + D2R area) ÷ 2
D5 Length	(D5L length+ D5R length) ÷ 2
D4 Length	(D4L length + D4R length) ÷ 2
D3 Length	(D3L length + D3R length) ÷ 2
D2 Length	(D2L length + D2R length) ÷ 2
D5 Width	(D5L width+ D5R width) ÷ 2
D4 Width	(D4L width + D4R width) ÷ 2
D3 Width	(D3L width + D3R width) ÷ 2
D2 Width	(D2L width + D2R width) ÷ 2

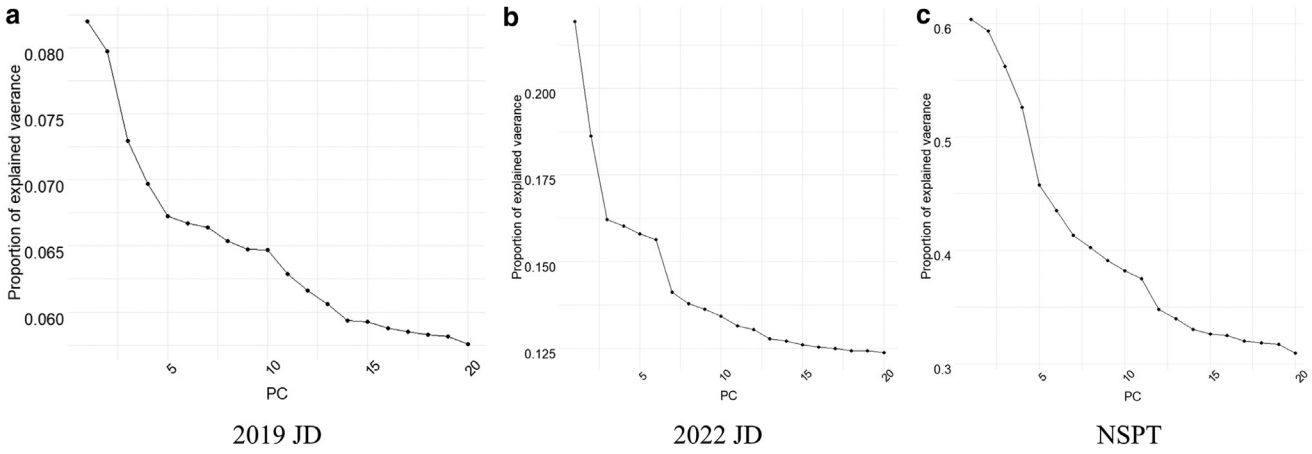
Supplementary Figure S1. Definition of nail plate area, length, and width. (**a**) Equipment for hand image collection: a small photography studio (dimensions: 50 × 60 ×50 cm) equipped with stable LED lighting, black background paper, and a color calibration card (ColorChecker Passport, X-Rite). There is an opening at the top, and the lens of the digital SLR camera shoots vertically downward to capture the hands of volunteers inside the photography studio. (**b**) Utilizing NailNet to segment the nail plate and extract nail plate area from hand images. In our study, the nail plate area is defined as the area of the entire nail plate, including the lunula but excluding the free edge and surrounding skin. It is not calculated as the product of length and width. This definition is consistent with previous studies (De Berker et al, 2007; Kumuda and Dinesh, 2015). Nail plate area is represented by the number of white pixels in this image. We first use NailNet to segment the nail plate area from whole-hand photos. The length and width of the nail plate are then defined by the dimensions of the minimum bounding rectangle of this segmented area. More information about NailNet, which focuses on the color and shape properties of the nail plate, can be found in our article (Fan et al, 2024[1]). (**c**) Finger naming convention: L (left) for the left hand, R (right) for the right hand, pinky (D5), ring finger (D4), middle finger (D3), index finger (D2), thumb (D1). (**d**) Nail plate area, length, and width are averaged for both left and right hands across D5, D4, D3, and D2. LED, light-emitting diode.



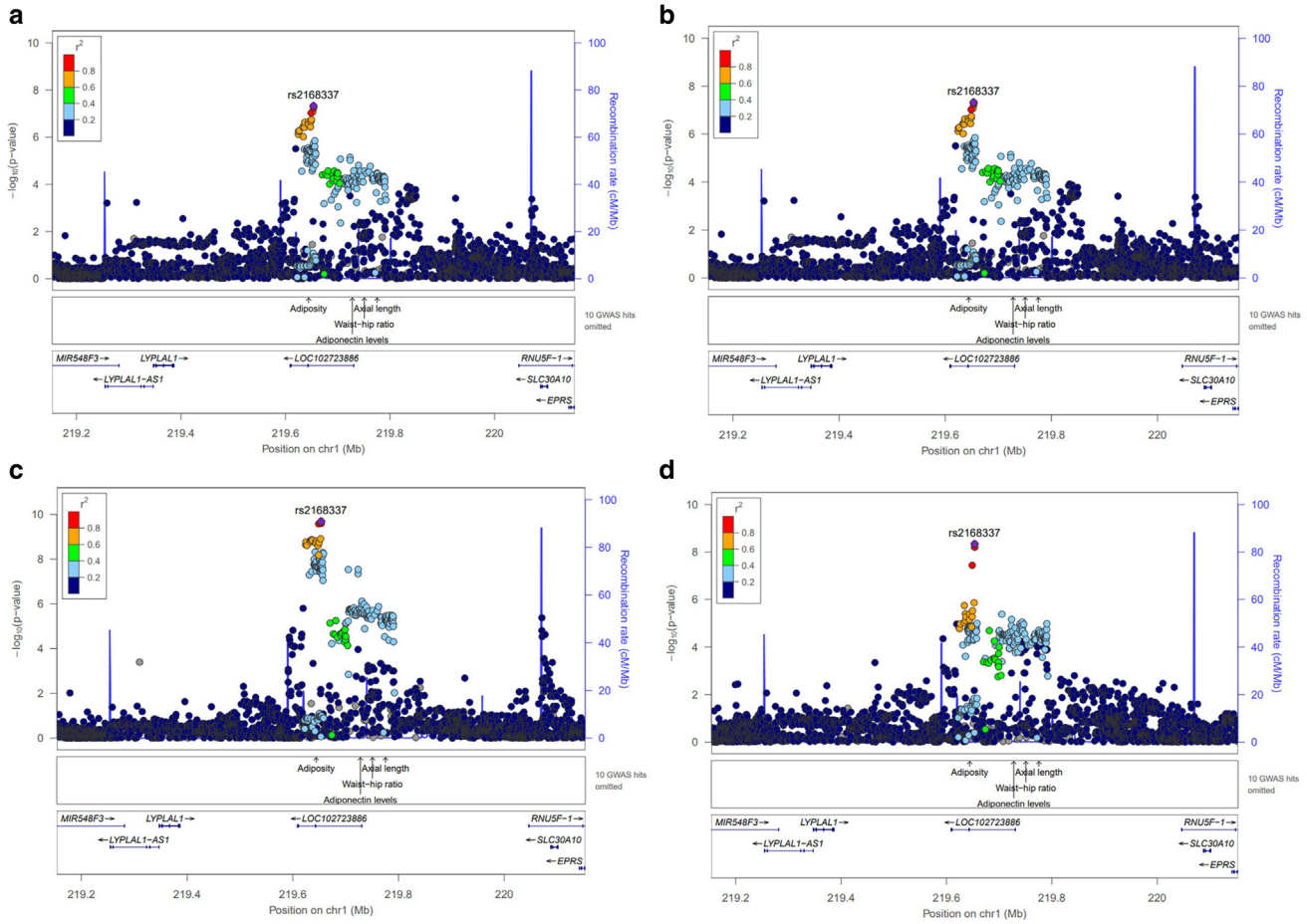
Supplementary Figure S2. Histograms of 12 nail plate area, length, and width for the JD2019, JD2022, and NSPT cohorts. (**a**–**c**) Nail plate area, length, and width do not follow a normal distribution in the JD2019, JD2022, and NSPT cohorts. (**d**–**f**) Histograms of nail plate area, length, and width after inverse normal transformation for the JD2019, JD2022, and NSPT cohorts, respectively. NSPT, National Survey of Physical Traits.



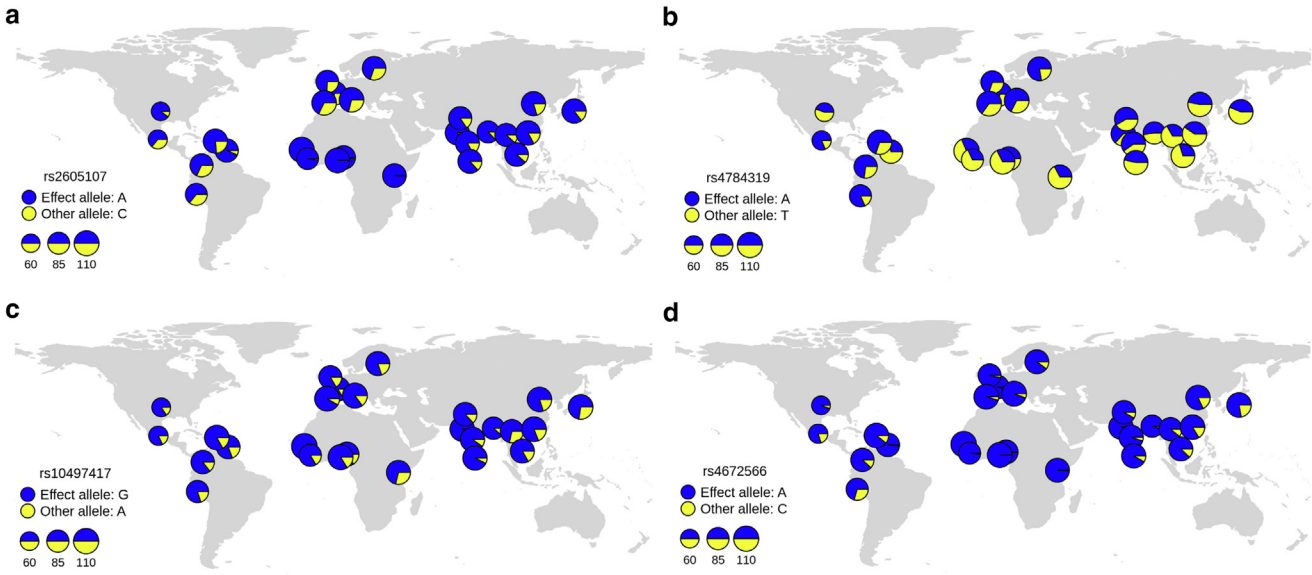
Supplementary Figure S3. Effects of age, height, weight, and sex on nail plate area, length, and width. (**a**–**d**) The effects of age, height, weight, and sex on nail plate area, length, and width, which exhibit a significant positive correlation with age, height, and weight. Males demonstrate significantly larger nail plate area, length, and width than females.



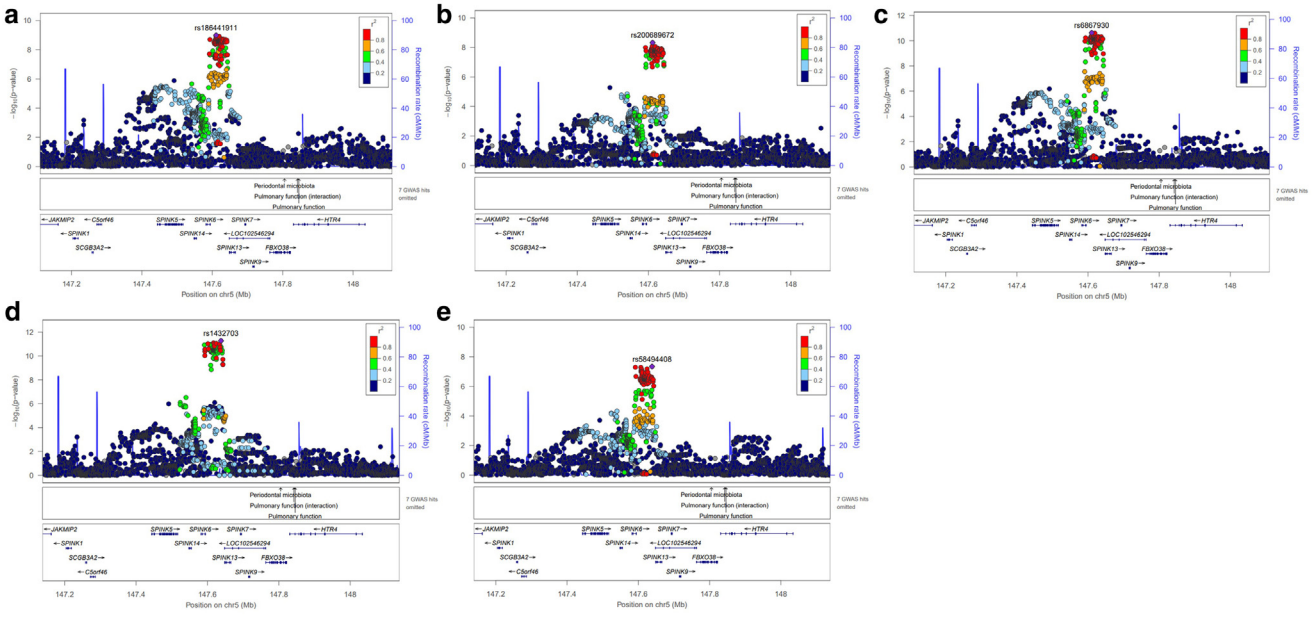
Supplementary Figure S4. Population structure of the JD2019, JD2022, and NSPT cohorts based on PC analysis. (**a**–**c**) The population structure of the JD2019, JD2022, and NSPT cohorts, respectively, using a scatterplot. The top 10 PCs of the genetic data were selected as covariates. NSPT, National Survey of Physical Traits; PC, principal component.



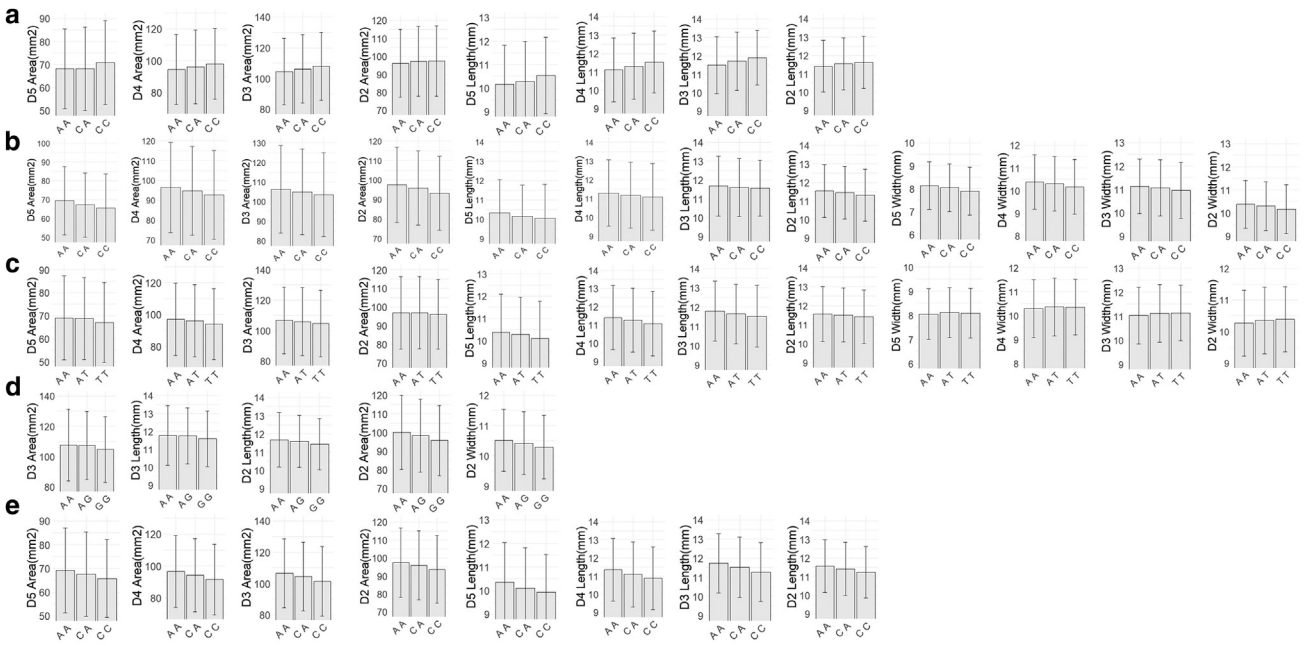
Supplementary Figure S5. The regional plots of the 1q41 signal are depicted. (**a**–**d**) The regional signals for D5 area, D4 area, D2 area, and D4 width, respectively, using LocusZoom.



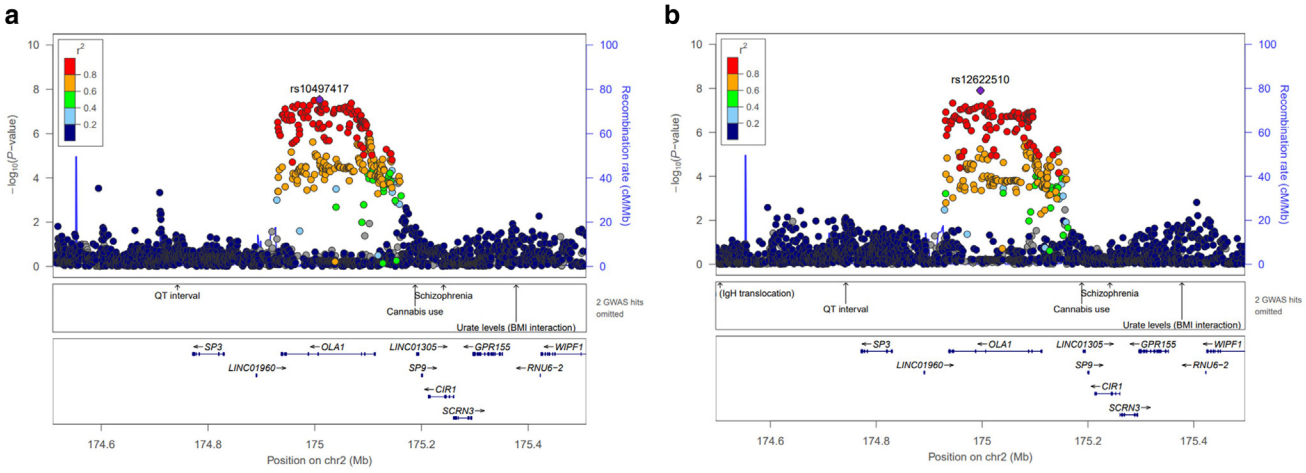
Supplementary Figure S6. Discrepancy between nail plate area, length, and width and allele frequencies. (**a**–**d**) Respectively showing the geographic distribution of minor allele frequencies for rs2605107, rs4784319, rs10497417, and rs4672566 across 2504 multiethnic individuals from the 1000 Genomes Project. Ancestral alleles are depicted in deep blue, whereas derived alleles are shown in yellow.



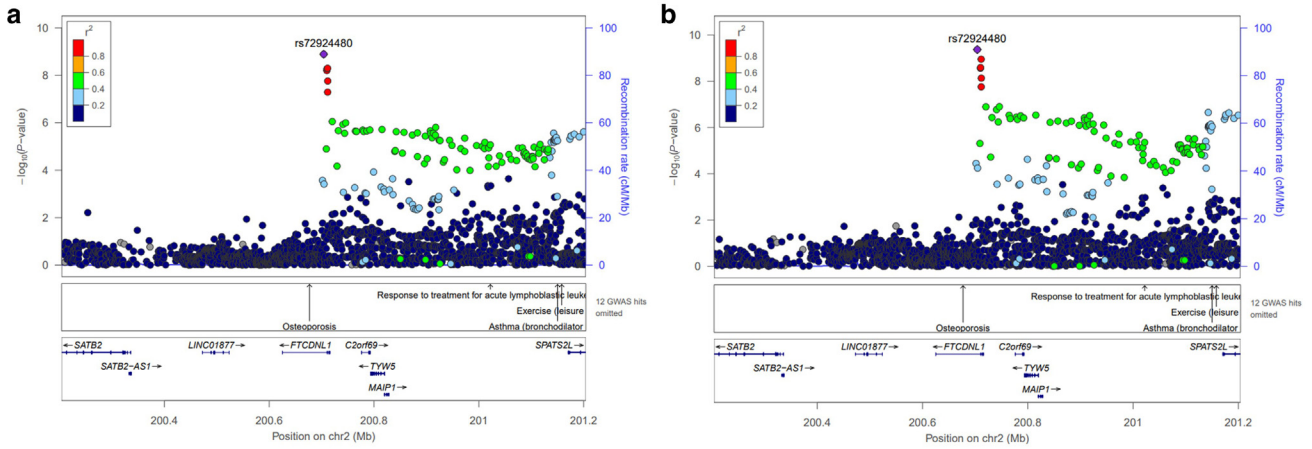
Supplementary Figure S7. The regional plots of the 5q32 signal are presented. (**a**–**e**) Depict the regional signals for D4 area, D3 area, D4 length, D3 length, and D2 length, respectively, using LocusZoom.



Supplementary Figure S8. The functional relationship between nail plate area, length, and width and genotypes. (**a**) Individuals with the C allele of rs12186491 exhibit larger nail length and area. (**b**) The C allele of rs2605107 shows a trend of reducing nail length, width, and area. (**c**) Individuals with the A allele of rs4784319 display increased nail plate area and length for D2−5 and decreased width. (**d**) Individuals carrying the G allele of rs10497417 show decreased nail length, width, and area for D3 as well as decreased area and width for D2. (**e**) Individuals with the C allele of rs4672566 demonstrate decreased nail length and area. Vertical bars represent the SEM.



Supplementary Figure S9. The regional plots for the 2q31.1 signal are shown. (**a, b**) Respectively display the regional signals for D2 area and D2 width using LocusZoom.



Supplementary Figure S10. The regional plots for the 2q33.1 signal are presented. (**a, b**) Respectively show the regional signals for D4 length and D3 length using LocusZoom.

Supplementary Table S1. The Phenotype–Genotype Association Study Results for Candidate Genes from the GWASATLAS Database[1]

Gene	Atlas ID	PMID	y	Domain	Trait	*P*-Value	n
LYPLAL1	4352	30573740	2018	Dermatological	Male-pattern baldness (BOLT LMM noninfinitesimal mixed model)	1.14E-07	205327
	4351	30573740	2018	Dermatological	Male-pattern baldness (BOLT LMM infinitesimal mixed model)	9.27E-07	205327
	3658	31427789	2019	Neoplasms	Cancer register—type of cancer: ICD10: C44 other and unspecified malignant neoplasm of the skin	1.18E-04	54696
RPGRIP1L	3274	31427789	2019	Dermatological	Ease of skin tanning	2.34E-08	378364
	3498	31427789	2019	Dermatological	Hair color (natural, before graying): dark brown	4.16E-05	385603
	3273	31427789	2019	Dermatological	Skin color	1.07E-04	381433
OLA1	4352	30573740	2018	Dermatological	Male-pattern baldness (BOLT LMM noninfinitesimal mixed model)	3.40E-06	205327
	4351	30573740	2018	Dermatological	Male-pattern baldness (BOLT LMM infinitesimal mixed model)	6.93E-06	205327

Abbreviations: ICD10, International Classification of Diseases, Tenth Revision; ID, identification; LMM, linear mixed model.
[1]The phenotypes significantly associated with skin and its appendages in genes *LYPLAL1*, *LYPLAL1*, and *OLA1*. Bonferroni corrected *P*-values = 1.25e-4, 1.14e-4, 1.15e-4, respectively.

Supplementary Table S2. Associations of Variants on 5q32 with Nail Plate Area, Length, and Width

	Lead SNP: rs12186491		Top SNPs (within ±250 bp)			LD with rs12186491	
Phenotype	β	*P*-Value	Name	β	*P*-Value	R2	D′
D5 area	0.0859	2.55E-06	rs10651520	0.1082	2.35E-06	0.84	0.95
D2 area	0.0699	3.57E-05	rs2195927	0.0989	2.54E-06	0.72	1
D5 length	0.1068	4.53E-07	rs10651520	0.1392	1.59E-07	0.84	0.95
D5 width	0.0234	.1325	rs13158727	0.0226	.1261	0.67	1
D4 width	0.016	.2285	rs12186491	0.016	.2285	1	1
D3 width	0.0045	.7396	rs12186491	0.0045	.7396	1	1
D2 width	0.0077	.596	rs12186491	0.0077	.596	1	1

Supplementary Table S3. The Population Cohort Statistical Information from Hand Image Collection[1]

Group	Cohort	Sample Size	Male, n (%)	Age, y (mean ± SD)	Height, cm (mean ± SD)	Weight, kg (mean ± SD)
Discovery	JD2019	4465	2136 (47.84%)	46.82 ± 13.34	167.90 ± 7.96	69.27 ± 13.19
Validation	JD2022	1231	610 (49.55%)	41.86 ± 9.94	167.98 ± 8.26	69.35 ± 13.25
Validation	NSPT	530	179 (33.77%)	43.97 ± 13.39	164.54 ± 11.90	79.79 ± 5.26

Abbreviation: NSPT, National Survey of Physical Traits.

[1]All 3 cohorts consisted of natural populations, meaning that individual recruitment did not rely on any specific disease status. During the quality control process for cohorts JD2019, JD2022, and NSPT, individuals were excluded on the basis of several criteria. These included (i) improper nail plate presentation or blurry images; (ii) signs of onychomycosis or other fungal infections; (iii) presence of nail polish; and (iv) physical nail abnormalities such as injuries, unclean nails, or deformities affecting the fingers.

Supplementary Table S4. Associations of Variants on 1q41 with Nail Plate Area, Length, and Width

	Lead SNP: rs2605107		Top SNPs (within ±250 bp)			LD with rs2605107	
Phenotype	β	*P*-Value	Name	β	*P*-Value	R2	D′
D3 area	0.0781	3.79E-05	rs2605107	0.0781	3.79E-05	1.00	1.00
D5 length	0.1105	3.10E-06	rs2605107	0.1105	3.10E-06	1.00	1.00
D4 length	0.0707	1.10E-03	rs2820438	0.0761	9.78E-05	0.63	0.98
D3 length	0.0698	1.50E-03	rs2605107	0.0698	1.50E-03	1.00	1.00
D2 length	0.0948	1.94E-05	rs2820438	0.089	8.33E-06	0.63	0.98
D5 width	0.0897	2.87E-07	rs1563353	−0.0898	2.64E-07	0.97	0.99
D3 width	0.0583	1.43E-04	rs2605097	0.0609	1.42E-04	0.96	0.99
D2 width	0.0735	7.49E-06	rs5781122	0.0785	1.58E-06	0.81	1.00

Supplementary Table S5. The Heritability of Nail Plate Area, Length, and Width

Heritability	D5 Area	D4 Area	D3 Area	D2 Area	D5 Length	D4 Length	D3 Length	D2 Length	D5 Width	D4 Width	D3 Width	D2 Width
h^2	0.2826	0.4162	0.4404	0.3198	0.2043	0.3841	0.3873	0.2942	0.2540	0.3686	0.4412	0.2941
SE	0.1158	0.0918	0.0923	0.0929	0.1172	0.0921	0.0920	0.0925	0.1172	0.0924	0.0939	0.0942
P-value	6.25E-03	1.96E-06	5.50E-07	1.93E-04	4.01E-02	1.19E-05	7.65E-06	5.10E-04	1.44E-02	3.03E-05	1.41E-06	7.86E-04
n	3020	3908	3872	3791	3020	3908	3872	3791	3020	3908	3872	3791

Abbreviation: SE, standard error.

Supplementary Table S6. *LYPLAL1-AS1* (1q41) and *SPINK6* (5q32) Remain Significant when Conducting GWAS in Female or Male Separately

Phenotype	Sex	Cytoband	Mapped Gene	Top SNP	EA	OA	n	EA Frequency	β	SE	*P*-Value
D2 width	Female	1q41	*LYPLAL1-AS1*	rs139092604	G	A	1833	0.484	0.140	0.023	1.34E-09
D3 length	Female	5q32	*SPINK6*	rs1159204	A	G	1843	0.215	0.235	0.036	7.82E-11
D4 length	Male	1q41	*LYPLAL1-AS1*	rs191126199	G	A	1941	0.048	0.425	0.074	9.21E-09
D5 area	Male	1q41	*LYPLAL1-AS1*	rs4453077	C	T	1553	0.175	0.219	0.039	2.59E-08
D5 length	Male	1q41	*LYPLAL1-AS1*	rs4453077	C	T	1553	0.175	0.280	0.046	1.27E-09

Abbreviations: EA, effect allele; OE, other allele; SE, standard error.
No other new significant loci emerged.