GWASs Identify Genetic Loci Associated with Human Scalp Hair Whorl Direction

Journal of Investigative Dermatology (2023) **143,** 2065–2068; doi:10.1016/j.jid.2023.04.008

JID Open

TO THE EDITOR

A hair whorl is a patch of hair growing in a circular pattern around a point specified by hair follicle (HF) orientations (Ziering and Krenitsky, 2003). As an easily observed human trait, established between 10 and 16 weeks of gestation (Smith and Gong, 1973), scalp hair whorl pattern is typically defined by the whorl number (single or double whorl) and whorl direction (e.g., clockwise, counterclockwise, or diffuse) (Ziering and Krenitsky, 2003). Because atypical whorl patterns have been observed in patients with abnormal neurological development (Smith and Gong, 1973), understanding the genetic basis of whorl patterns may help to unravel important biological processes. Because the clockwise or counterclockwise direction of a single whorl is highly stable within an individual and varies considerably between individuals and populations (Ziering and Krenitsky, 2003), it is the most likely feature to yield insight into the embryology and development of the whorl.

In this study, we report a GWAS of whorl direction in 2,149 Chinese individuals from the National Survey of Physical Traits cohort, followed by replication in 1,950 Chinese individuals from the Taizhou Longitudinal Study cohort. The study was approved by the Ethics Committee of Fudan University (Shanghai, China) and the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (Shanghai, China). All individuals provided written informed consent. Because the small available sample size of other whorl patterns limits the statistical power of GWAS, only individuals with a single clockwise or counterclockwise whorl were included in this GWAS (n = 4,099) (Figure 1a and Supplementary Table S1). The proportion of individuals with each whorl direction in our study population was consistent with those of previous studies and did not differ between females and males (Ziering and Krenitsky, 2003). Complete sample characteristics are provided in Supplementary Table S1.

Using a linear mixed model (Supplementary Materials and Methods, and Supplementary Figure S1), a locus at 7p21.3 was identified for whorl direction in the National Survey of Physical Traits cohort (rs2024403 as the lead SNV, $P = 2.72 \times 10^{-8}$, OR = 1.08) (Supplementary Figure S2) and was replicated in the Taizhou Longitudinal Study cohort ($P_{rs2024403} = 4.61 \times 10^{-7}$, OR = 1.07) (Supplementary Figure S2). A meta-analysis of the two cohorts identified three additional signals (5q33.2, $P = 2.56 \times 10^{-10}$, OR = 1.06; 7q33, $P = 2.90 \times 10^{-9}$, OR = 1.21; and 14q32.13, $P = 7.69 \times 10^{-12}$, OR = 1.12) (Figure 1b).

We performed fine mapping to identify the putative causal variants for each signal and prioritized candidate genes by functional annotations using Combined Annotation-Dependent Depletion, DeepSEA, Genotype-Tissue Expression Portal, and 3DSNP (Table 1 and Supplementary Table S2). Combined Annotation-Dependent Depletion and DeepSEA scores both indicated that candidate variants were likely to have biological functions. For the most significant signal at 7p21.3, the putative causal variant rs246829 is located 252 kb downstream of *ARL4A* (Figure 1c) and is an expression quantitative trait locus of *ARL4A* in the skin. The 3DSNP further suggested that rs246829 is located within a set of transcription factor binding sites, indicating that this is a gene regulatory region. Thus, rs246829 may regulate the expression of *ARL4A* in the skin to influence whorl direction. Individuals with the risk allele C had a higher frequency of counterclockwise whorls (Figure 1d).

In single-cell sequencing data of mouse whisker HF from embryonic day 11.5 to embryonic day 17.0 (Morita et al., 2021), *ARL4A* had the greatest expression in the upper HF or infundibulum from the early hair bud stage (embryonic day 12) to the hair germ stage (embryonic day 11.5) (Figure 1e), the same period as when cell division angles are generated in the development of the HF epithelium (Morita et al., 2021). Because Arf small GTPases function in cell migration and actin cytoskeleton remodeling, ARL4A may influence cell polarity through the ELMO–DOCK180–Rac signaling pathway (Côté and Vuori, 2007; Patel et al., 2011) or by regulating CDC42 activation (Chiang et al., 2019; Etienne-Manneville, 2004). The global alignment of HF orientation in mouse embryogenesis is guided by planar cell polarity (Cetera et al., 2017). Thus, we hypothesize that ARL4A contributes to keratinocyte shape through the modulation of the actin cytoskeleton and through its integration with the planar cell polarity pathway (Henderson et al., 2018) during early skin and HF development, leading to variation of HF orientation and thus whorl direction. Modulation of different planar cell polarity components causes hair whorls to appear in the mouse coat, with alteration of different genes leading to patterns with opposite orientations (Cetera et al., 2017).

Signals at 14q32 and 5q33.2 discovered in the meta-analysis (Supplementary Figures S3–5) also implicated associated protein-coding genes nearby (Supplementary Table S2). The putative causal SNV at 14q32.13, rs179151, is an expression quantitative trait locus and at a promoter of *SYNE3*. Nesprin-3, encoded by *SYNE3*, is a component of the linker of

Abbreviation: HF, hair follicle

Corrected proof published online 9 August 2023

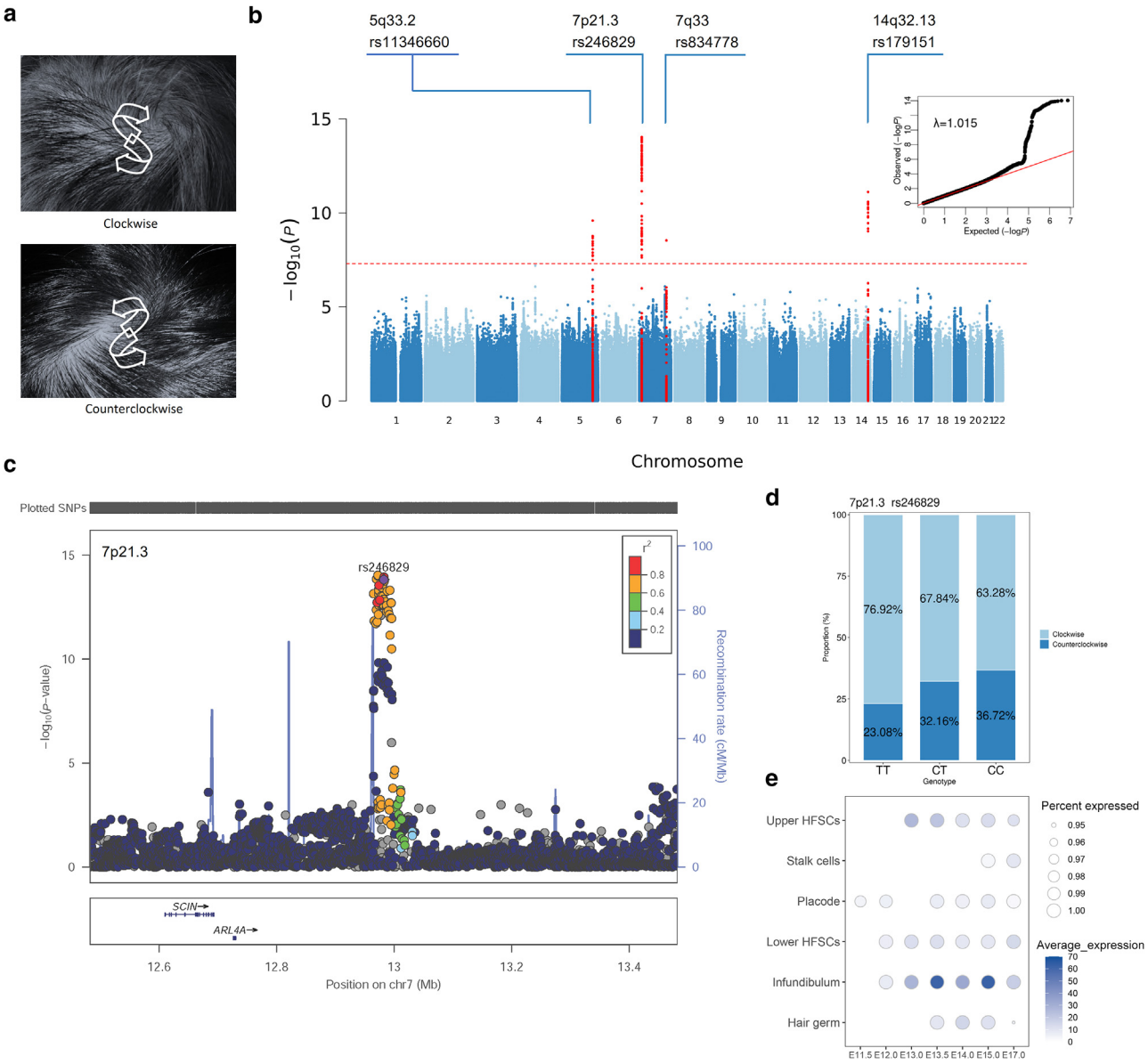


Figure 1. GWASs of hair whorl direction identified four significant signals. (**a**) Patterns of whorl direction used in the GWAS. (**b**) Manhattan plot and quantile−quantile plot from the discovery (NSPT) and replication (TZL) cohorts. The red line corresponds to the threshold for genome-wide significance ($P = 5 \times 10^{-8}$). (**c**) Regional association plot of the 7p21.3 signal in the meta-analysis. Color intensity indicates linkage disequilibrium (r^2) with the putative causal SNV. (**d**) The C allele of rs246829 shows a contribution to counterclockwise hair whorls. (**e**) Expression of *ARL4A* in different cell types of mouse whisker hair follicles from E11.5 to E17.0. E11.5, embryonic day 11.5; E17.0, embryonic day 17.0; HFSC, hair follicle stem cell; NSPT, National Survey of Physical Traits; TZL, Taizhou Longitudinal Study.

nucleoskeleton and cytoskeleton complex, which serves to connect nuclear structure with components of the cytoskeleton, including interaction with the actin cytoskeleton (Ketema et al., 2007) and which is involved in cell polarity and cortical development (Hakanen et al., 2019). Putative causal variant rs11346660 at 5q33.2 is located at an enhancer of *HAND1*, a basic helix-loop-helix transcription factor that is critical for craniofacial morphogenesis (Firulli et al., 2014). Unlike the *ARL4A* gene, expression levels of the *SYNE3* and *HAND1* genes are low during HF development (Supplementary Figure S6). Because the whorl appears in the position of cranial neural tube closure (Yamaguchi and Miura, 2013), it is possible that the arrangement and directionality of HFs on the scalp could be influenced by SYNE3 and HAND1 acting in cranial development rather than directly in the skin. The growth pattern of tissue underlying the skin may influence HF orientation through its mechanical effect in establishing planar cell polarity orientation in the epidermis (Aw et al., 2016).

Putative causal variant rs834778 at 7q33 is close to *MTPN*, which modulates the actin cytoskeleton (Bhattacharya et al., 2006). However, rs834778 had a low-risk allele frequency in our populations (<0.05), and possibly as a consequence, its association with whorl type is less well statistically supported.

Previous work proposed associations between hair whorl patterns and neurological diseases (Smith and Gong, 1973),

Table 1. Summary of the Putative Causal SNVs Associated with Whorl Direction

Locus	Putative Causal SNV	Mapped Gene	NSPT			TZL			Meta-Analysis			Non-Risk Allele	Risk Allele	Risk Allele Frequency		
			OR	95% CI	*P*-Value	OR	95% CI	*P*-Value	OR	95% CI	*P*-Value			NSPT	TZL	NSPT + TZL
7p21.3	rs246829	*ARL4A*	1.08	(1.05–1.11)	2.47×10^{-07}	1.08	(1.05–1.12)	1.41×10^{-08}	1.08	(1.06–1.10)	1.49×10^{-14}	C	T	0.447	0.44	0.45
14q32.13	rs179151	*SYNE3*	1.13	(1.08–1.18)	1.15×10^{-07}	1.1	(1.05–1.15)	9.36×10^{-05}	1.11	(1.08–1.15)	5.55×10^{-11}	G	A	0.113	0.102	0.107
5q33.2	rs11346660	*HAND1*	1.08	(1.04–1.10)	8.99×10^{-07}	1.06	(1.03–1.09)	6.54×10^{-05}	1.06	(1.04–1.09)	2.56×10^{-10}	G	GA	0.584	0.598	0.594
7q33	rs834778	*LOC105375523, MTPN*	1.17	(1.07–1.28)	5.42×10^{-04}	1.25	(1.14–1.36)	9.16×10^{-07}	1.21	(1.13–1.29)	2.90×10^{-09}	T	A	0.025	0.027	0.025

Abbreviations: CI, confidence interval; NSPT, National Survey of Physical Traits; TZL, Taizhou Longitudinal Study.

but we did not detect such associations by phenome-wide association studies (Supplementary Figure S7 and Supplementary Table S3). This does not support the hypothesis that hair whorl direction in the general population signifies a risk for neurological conditions. However, we did identify associations between *HAND1* variants and male-pattern baldness; although beyond influencing hair traits at the same anatomical site, the biological mechanism underlying this pleiotropy remains unclear.

In conclusion, this study reveals that hair whorl direction is a polygenic trait and identified four associated loci. The functions of the genes identified suggest that cell polarity and cytoskeletal structure in HF development and cranial neural tube closure and growth may influence HF orientation and contribute to whorl directionality. Our findings provide insights into hair whorl development, with important clues for future investigations on the potential mechanisms.

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 OEP003828. 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 Institutional Review Board restrictions 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.

ORCIDs

Junyu Luo: http://orcid.org/0000-0003-2206-1931
He Huang: http://orcid.org/0000-0002-8868-1431
Hui Qiao: http://orcid.org/0000-0002-0462-4844
Jingze Tan: http://orcid.org/0009-0000-4447-8653
Wenyan Chen: http://orcid.org/0000-0002-4821-9306
Manfei Zhang: http://orcid.org/0000-0003-2611-9440
Andrés Ruiz-Linares: http://orcid.org/0000-0001-8372-1011
Jiucun Wang: http://orcid.org/0000-0003-2765-0620
Yajun Yang: http://orcid.org/0000-0001-5713-0103
Jin Li: http://orcid.org/0000-0001-9201-2321
Denis J. Headon: http://orcid.org/0000-0002-0452-4480
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 Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDB38020400 to SW), CAS Project for Young Scientists in Basic Research (grant number YSBR-077), National Science and Technology Basic Research Project (grant number 2015FY111700), the National Natural Science Foundation of China (grants number to 32271186 to JT), and CAMS Innovation Fund for Medical Sciences (2019-I2M-5-066).

AUTHORS CONTRIBUTIONS

Conceptualization: JLu, SW, DJH; Data Curation: JLu, HH, HQ, WC, MZ; Formal Analysis: JLu, HH, HQ, WC, MZ; Funding Acquisition: SW, JT, JLi; Investigation: JLu, HH, HQ, JT; Methodology: JLu, HH, HQ; Project Administration: SW; Resources: SW, JT, JLi, ARL, JW, YY; Supervision: SW, DJH; Validation: JLu, HH; Visualization: JLu, HH, SW; Writing - Original Draft Preparation: JLu, HH; Writing - Review and Editing: JLu, HH, DJH, SW, HQ, MZ

***Junyu Luo*[1,11], *He Huang*[1,2,11], *Hui Qiao*[2,11], *Jingze Tan*[2], *Wenyan Chen*[1], *Manfei Zhang*[1,3], *Andrés Ruiz-Linares*[2,4,5], *Jiucun Wang*[6,7,8], *Yajun Yang*[2], *Li Jin*[6,7,8], *Denis J. Headon*[9] and *Sijia Wang*[1,10,*]**

[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]*Ministry of Education Key Laboratory of Contemporary Anthropology, Department of Anthropology and Human Genetics, School of Life Science, Fudan University, Shanghai, China;* [3]*Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Ministry of Education, Shanghai Jiao Tong University, Shanghai, China;* [4]*Aix-Marseille Université, Centre National de la Recherche Scientifique (CNRS), EFS, Anthropologie Bio-Culturelle, Droit, Ethique et Santé (ADES), Marseille, France;* [5]*Department of Genetics, Evolution and Environment, UCL Division of Biosciences, University College London, London, United Kingdom;* [6]*State Key Laboratory of Genetic Engineering, Human Phenome Institute, Zhangjiang Fudan International Innovation Center, Fudan University, Shanghai, China;* [7]*Fudan-Taizhou Institute of Health Sciences, Taizhou, China;* [8]*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;* [9]*The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, United Kingdom; and* [10]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China*

[11]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.2023.04.008.

REFERENCES

Aw WY, Heck BW, Joyce B, Devenport D. Transient tissue-scale deformation coordinates alignment of planar cell polarity junctions in the mammalian skin. Curr Biol 2016;26:2090–100.

Bhattacharya N, Ghosh S, Sept D, Cooper JA. Binding of myotrophin/V-1 to actin-capping protein: implications for how capping protein binds to the filament barbed end. J Biol Chem 2006;281:31021–30.

Cetera M, Leybova L, Woo FW, Deans M, Devenport D. Planar cell polarity-dependent and independent functions in the emergence of tissue-scale hair follicle patterns. Dev Biol 2017;428:188–203.

Chiang TS, Lin MC, Tsai MC, Chen CH, Jang LT, Lee F-JS. ADP-ribosylation factor-like 4A interacts with Robo1 to promote cell migration by regulating Cdc42 activation. Mol Biol Cell 2019;30:69–81.

Côté JF, Vuori K. GEF what? Dock180 and related proteins help Rac to polarize cells in new ways. Trends Cell Biol 2007;17:383–93.

Etienne-Manneville S. Cdc42-the centre of polarity. J Cell Sci 2004;117:1291–300.

Firulli BA, Fuchs RK, Vincentz JW, Clouthier DE, Firulli AB. Hand1 phosphoregulation within the distal arch neural crest is essential for craniofacial morphogenesis. Development 2014;141:3050–61.

Hakanen J, Ruiz-Reig N, Tissir F. Linking cell polarity to cortical development and malformations. Front Cell Neurosci 2019;13:244.

Henderson DJ, Long DA, Dean CH. Planar cell polarity in organ formation. Curr Opin Cell Biol 2018;55:96–103.

Ketema M, Wilhelmsen K, Kuikman I, Janssen H, Hodzic D, Sonnenberg A. Requirements for the localization of nesprin-3 at the nuclear envelope and its interaction with plectin. J Cell Sci 2007;120:3384–94.

Morita R, Sanzen N, Sasaki H, Hayashi T, Umeda M, Yoshimura M, et al. Tracing the origin of hair follicle stem cells. Nature 2021;594:547–52.

Patel M, Chiang TC, Tran V, Lee FJ, Côté JF. The Arf family GTPase Arl4A complexes with ELMO proteins to promote actin cytoskeleton remodeling and reveals a versatile Ras-binding domain in the ELMO proteins family. J Biol Chem 2011;286:38969–79.

Smith DW, Gong BT. Scalp hair patterning as a clue to early fetal brain development. J Pediatr 1973;83:374–80.

Yamaguchi Y, Miura M. How to form and close the brain: insight into the mechanism of cranial neural tube closure in mammals. Cell Mol Life Sci 2013;70:3171–86.

Ziering C, Krenitsky G. The Ziering whorl classification of scalp hair. Dermatol Surg 2003;29:817–21; discussion 821.

Plasma Vitamin D Is Not Associated with Moderate-to-Severe Psoriasis: Results from Danish General Population Studies

JID Open

Journal of Investigative Dermatology (2023) **143,** 2068–2071; doi:10.1016/j.jid.2023.04.004

TO THE EDITOR

In recent years, the association between low levels of vitamin D (measured as plasma level of 25-hydroxyvitamin D) and psoriasis has been debated. Besides playing a key role in calcium homeostasis and bone metabolism, vitamin D has several immunomodulatory actions, and suboptimal vitamin D status may be a risk factor for autoimmune diseases (Holick, 2007). Several case-control studies have found low levels of plasma vitamin D in patients with psoriasis compared with the levels in healthy controls; however, these findings have been in studies with small sample sizes (Lee and Song, 2018), and the findings could therefore be due to random effects.

To elucidate the association between vitamin D and psoriasis, we assessed the levels of plasma 25-hydroxyvitamin D in 35,973 individuals from the Danish general population using data from the Copenhagen City Heart Study and the Copenhagen General Population Study (Madsen et al., 2017). All participants gave written informed consent. Both studies were conducted according to the Declaration of Helsinki and were approved by the Danish Ethics Committees (KF-100.2039/91 and H-KF-01-144/01). Approximately 114,000 individuals from the two Copenhagen cohort studies were also genotyped for seven genetic variants biologically linked to and associated with plasma 25-hydroxyvitamin D (Çolak et al., 2021).

To avoid population stratification bias, we only included individuals of Danish or Scandinavian descent (Haycock et al., 2016). We collected data on psoriasis by linking participants to the Danish National Patient Registry, which contains diagnoses for all in-patients and out-patients in Danish hospitals since 1977 (Schmidt et al., 2015), registered according to the International Classification of Diseases. Patients with psoriasis were identified by the International Classification of Diseases, Eight Revision codes 696.09, 696.10, and 696.19 and International Classification of Diseases, Tenth Revision code L40, compatible with moderate-to-severe psoriasis (Ahlehoff et al., 2011) because these patients are referred to a dermatologist in a hospital setting and thus captured by the Danish National Patient Registry. Plasma samples for measurement of 25-hydroxyvitamin D were taken on the day of study enrollment. The concentrations of

Accepted manuscript published online 28 April 2023; corrected proof published online 20 June 2023

SUPPLEMENTARY MATERIALS AND METHODS

Study populations

The National Survey of Physical Traits cohort. The National Survey of Physical Traits (NSPT) cohort is a prospective cohort study to explore the genetic and environmental factors associated with physical traits and diseases in China. The NSPT cohort study was approved by the Ethics Committees of Fudan University (Shanghai, China) (14117). A total of 2,149 Han Chinese individuals recruited in 2015–2019 from Taizhou, Nanning, and Zhengzhou were included in this study. All individuals provided informed written consent.

The Taizhou longitudinal cohort. The Taizhou longitudinal (TZL) cohort is a long-term observational cohort study to explore the genetic and environmental risk factors for common and non-communicable diseases. This research program was under the approval of the Ethics Committee of Fudan University (Ethics Research Approval 85) and the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (ER-SIBS-261410) (Shanghai, China). A total of 1,950 Han Chinese recruited in Taizhou were included in this study. All individuals provided written informed consent.

Phenotyping

Hair whorl patterns were defined according to the Ziering whorl classification system (Ziering and Krenitsky, 2003). We collected all types of hair whorl directions: S pattern (a clockwise whorl), Z pattern (a counterclockwise whorl), SZ pattern (double whorls, one whorl in clockwise and the other whorl in counterclockwise direction), SS pattern (double whorls in clockwise directions), ZZ pattern (double whorls in counterclockwise directions), and diffuse pattern (Supplementary Table S1). In this study, we focused on the direction of a single hair whorl. Only individuals with the S pattern (clockwise) and Z pattern (counterclockwise) were included. In total, 2,149 individuals (1,071 men and 1,078 women) from the NSPT cohort and 1,950 individuals (762 men and 1,188 women) from the TZL cohort were included in this study. The terms clockwise and counterclockwise used in the main text specifically refer to the S and Z patterns defined by the Ziering classification system.

Genotyping

For the NSPT cohort, DNA was extracted from blood samples of individuals using MagPure Blood DNA KF Kits. Illumina Infinium Global Screening Array (Illumina, San Diego, CA) containing 707,180 SNPs designed by WeGene (https://www.wegene.com/) was used for genotyping.

For the TZL cohort, Generay DNA extraction kits were utilized to extract DNA from blood samples. DNA samples were genotyped by the Illumina HumanOmniZhongHua-8 chip (Illumina) containing 894,517 variants.

Imputations for the two cohorts were performed separately. Before imputation, samples with a genotype missing rate > 0.05 were excluded. Haplotypes were estimated by SHAPEIT, version 2.17 (Delaneau et al., 2011). Then, the samples were imputed by IMPUTE, version 2.3.2, using the 1000 Genomes Project Phase 3 reference panel (Howie et al., 2012). After imputation, variants with INFO score < 0.8 or certainty score < 0.9 were excluded. Quality control was performed by PLINK, version 1.9 (Purcell et al., 2007). Two datasets were filtered by variant missingness (–geno 0.05), minor allele frequency (–maf 0.02), and Hardy–Weinberg equilibrium ($P < 1\times10^{-6}$). In total, 2,131 individuals with 8,158,198 variants in the NSPT cohort and 1,947 individuals with 8,108,947 variants in the TZL cohort remained after filtering.

Statistical analysis

GWASs and meta-analysis. GWASs in the NSPT (n = 2,131) and TZL (n = 1,947) cohorts were conducted separately by GEMMA 0.92 (Zhou and Stephens, 2012) using additive allele effects and linear mixed models adjusted for the first two genomic principal components (Supplementary Figure S1). The linear mixed model implemented in GEMMA is suitable for both quantitative and binary traits and can more effectively correct for the relatedness of individuals, pedigree structures, and population structure compared with typical logistic regression models. To reduce the effects of confounding factors and to improve statistical power, we chose the linear mixed model in GEMMA over the logistic regression model, which is commonly used for binary traits. PLINK 1.9 (Purcell et al., 2007) was used to perform a meta-analysis of the two cohorts using an inverse variance fixed-effect model. GWAS results were visualized by the R package qqman. Regional linkage disequilibrium plots and association plots were generated by LocusZoom (Pruim et al., 2010).

Fine mapping. We performed fine mapping to identify putative causal variants for each significant locus identified in the meta-analysis. Fine mapping was conducted by PAINTOR (Kichaev et al., 2017) in a 1-Mb genomic window (500 kb upstream and downstream flanking the lead SNP).

Functional annotation. Combined Annotation-Dependent Depletion (Kircher et al., 2014), DeepSEA (Zhou and Troyanskaya, 2015), 3DSNP (version 2.0) (Quan et al., 2022), and Genotype-Tissue Expression (GTEx Consortium, 2015) were used to annotate the causal SNPs and prioritize the putative candidate genes.

To evaluate the expression levels of candidate causal genes during hair development, we used single-cell RNA sequencing data of mouse hair follicle development from Gene Expression Omnibus (GSE147372) (Morita et al., 2021). Gene expression levels, developmental stages, and cell types of candidate genes associated with whorl direction were obtained and visualized in bubble plots.

Phenome-wide association studies. Two databases were used to conduct phenome-wide association studies of candidate genes. We performed a gene-level phenome-wide association study in GWAS Atlas (https://atlas.ctglab.nl/PheWAS), which includes 4,756 GWASs from UK Biobank across 3,302 unique traits (Watanabe et al., 2019). We also performed SNP-level phenome-wide association study using Biobank Japan PheWeb (https://pheweb.jp), which contains 230 phenotypes.

SUPPLEMENTARY REFERENCES

GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648–60.

Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods 2011;9:179–81.

Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype

imputation in genome-wide association studies through pre-phasing. Nat Genet 2012;44:955–9.

Kichaev G, Roytman M, Johnson R, Eskin E, Lindström S, Kraft P, et al. Improved methods for multi-trait fine mapping of pleiotropic risk loci. Bioinformatics 2017;33:248–55.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–5.

Morita R, Sanzen N, Sasaki H, Hayashi T, Umeda M, Yoshimura M, et al. Tracing the origin of hair follicle stem cells. Nature 2021;594:547–52.

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 MAR, 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.

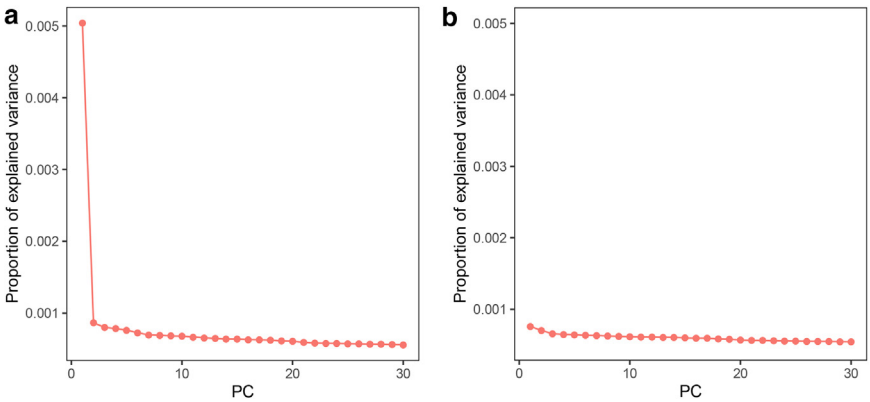
Quan C, Ping J, Lu H, Zhou G, Lu Y. 3DSNP 2.0: update and expansion of the noncoding genomic variant annotation database. Nucleic Acids Res 2022;50:D950–5.

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 [published correction appears in Nat Genet 2020;52:353] Nat Genet 2019;51:1339–48.

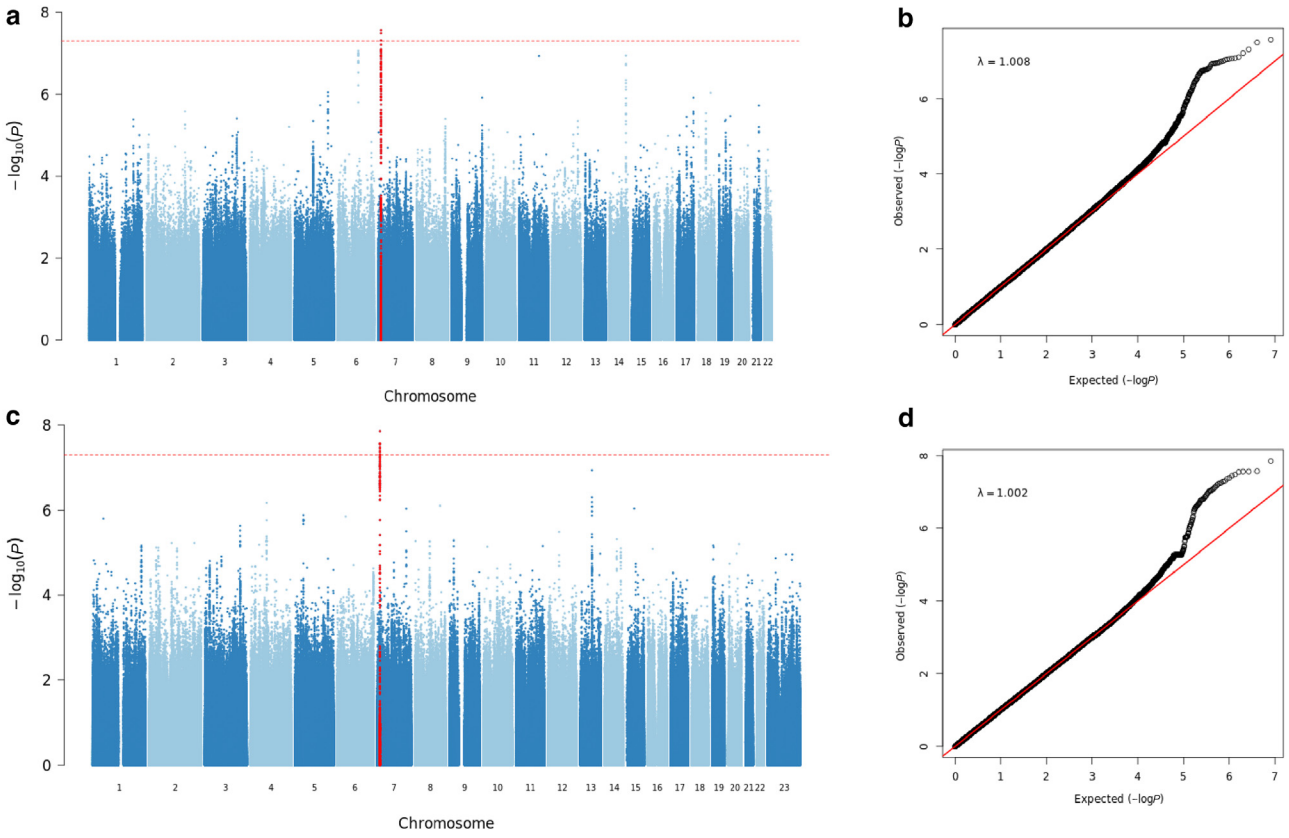
Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning–based sequence model. Nat Methods 2015;12:931–4.

Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 2012;44:821–4.

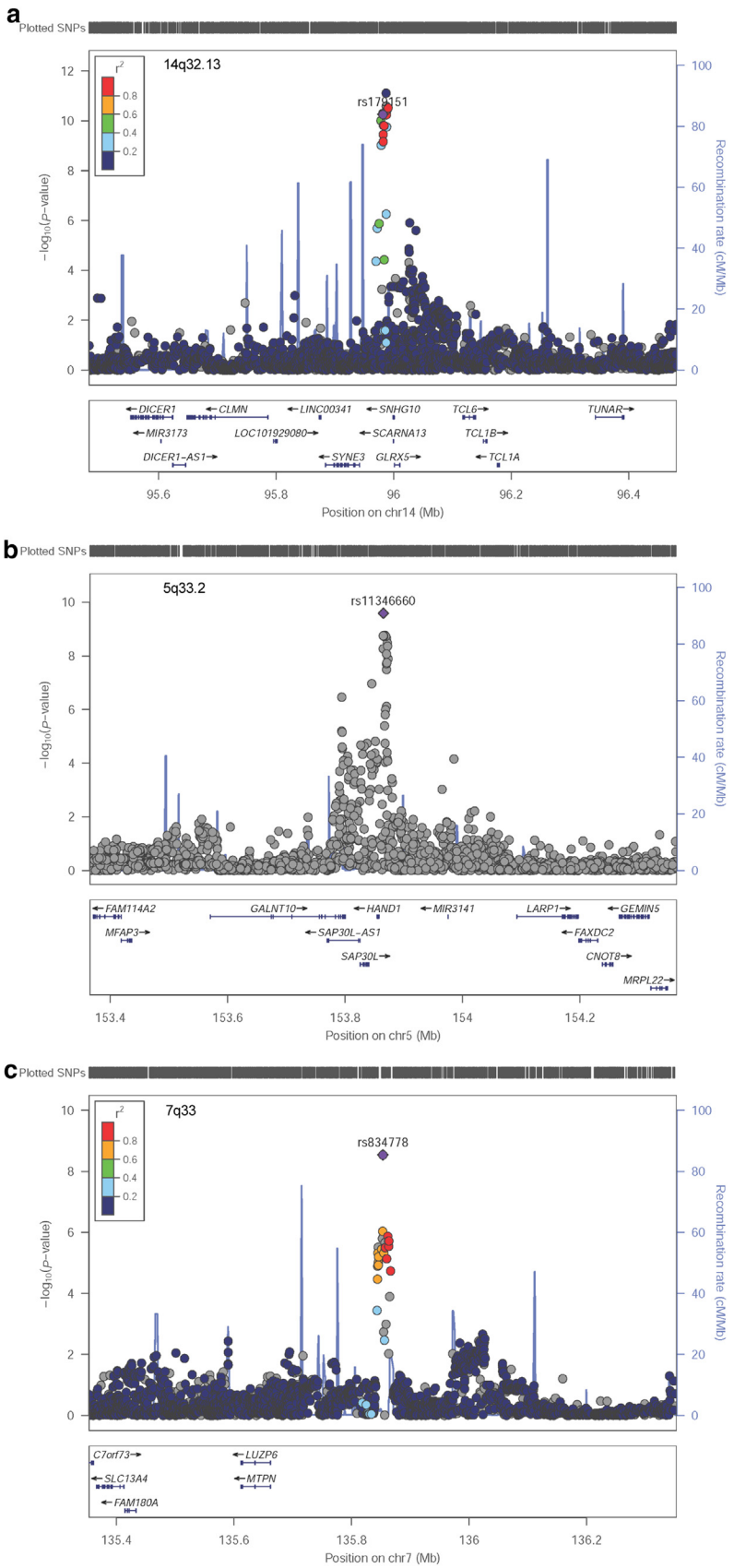
Ziering C, Krenitsky G. The Ziering whorl classification of scalp hair. Dermatol Surg 2003;29:817–21; discussion 821.



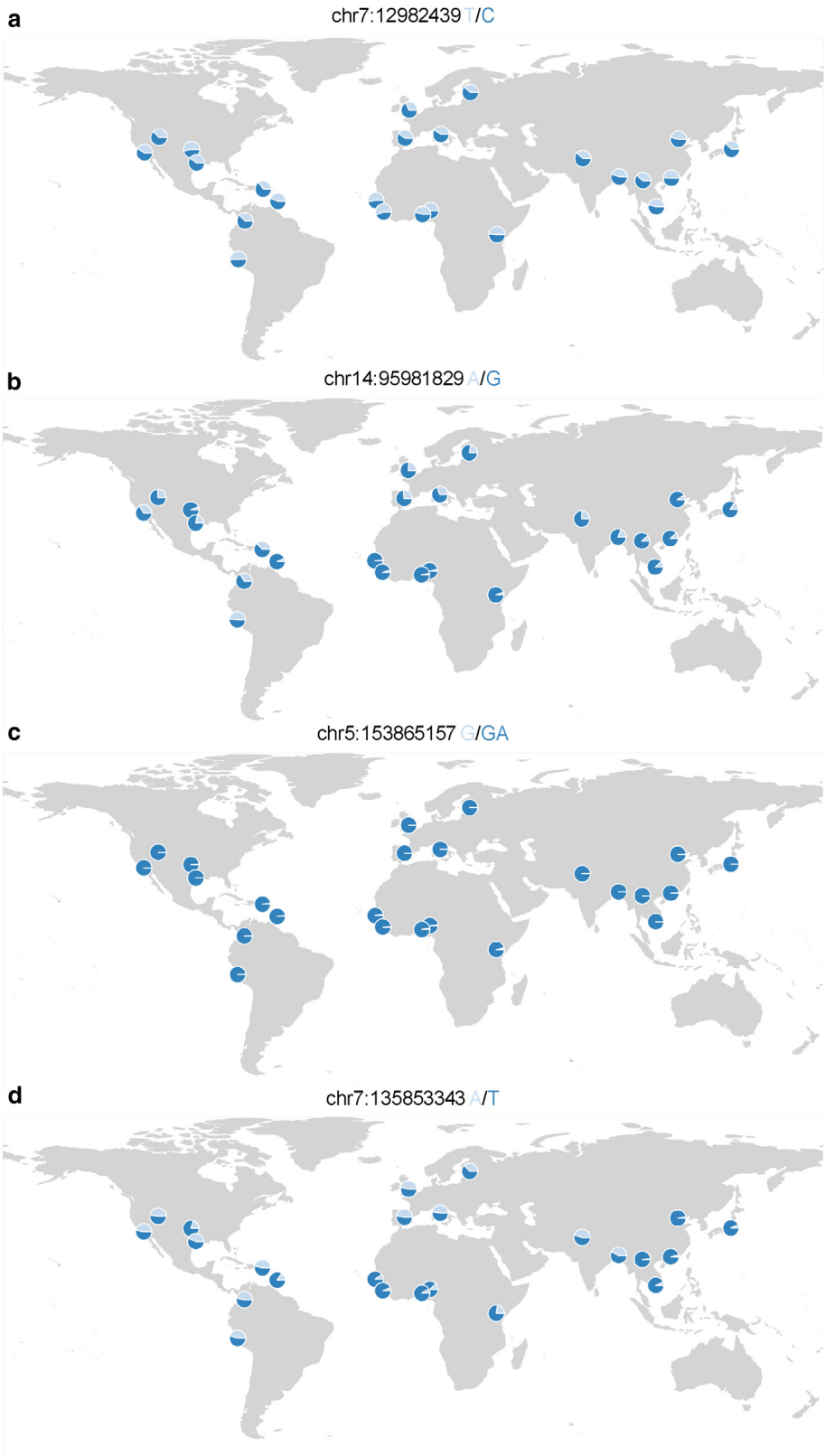
Supplementary Figure S1. Scree plots of population structure for the two cohorts. The first two PCs of the genetic data were selected to be used as covariates in the (**a**) NSPT cohort and the (**b**) TZL cohort. NSPT, National Survey of Physical Traits; PC, principal component; TZL, Taizhou Longitudinal Study.



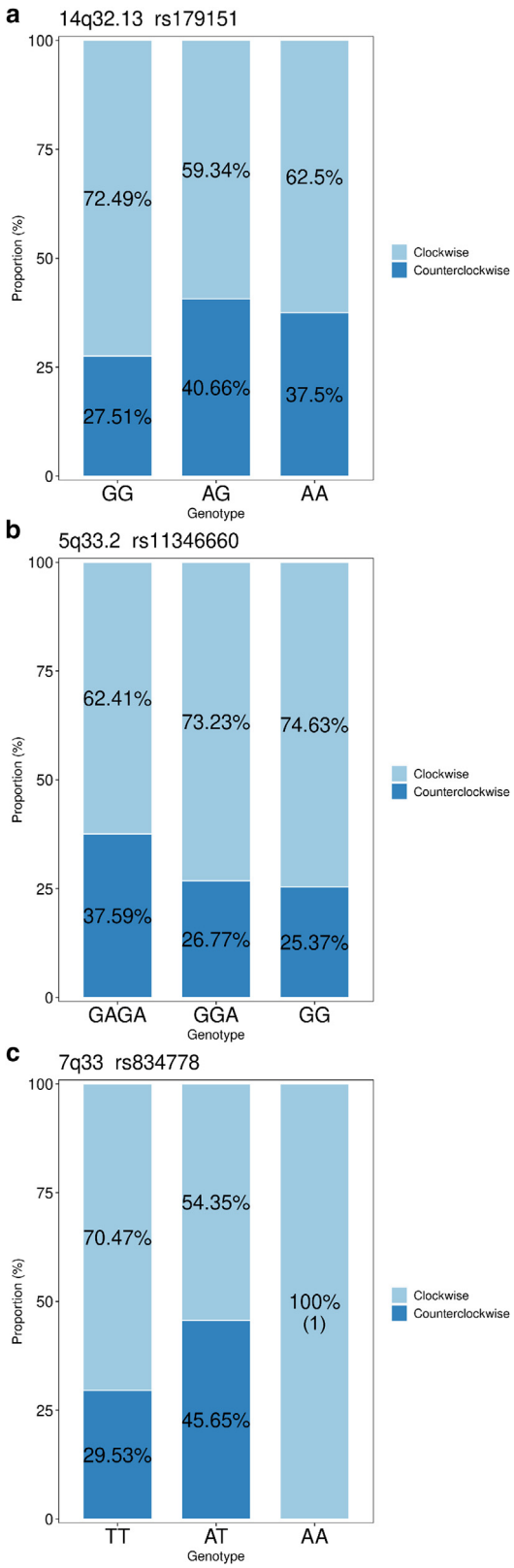
Supplementary Figure S2. Manhattan plots and quantile−quantile plots showing the results of the genome-wide scans for whorl direction in the NSPT and TZL cohorts. (**a**) GWAS in the discovery set (NSPT) adjusting the first two PCs; (**b**) the quantile−quantile plot for the discovery set with lambda (λ) = 1.008. (**c**) GWAS in the replication set (TZL) adjusting the first two PCs; (**d**) the quantile−quantile plot for the replication set with λ = 1.002. NSPT, National Survey of Physical Traits.



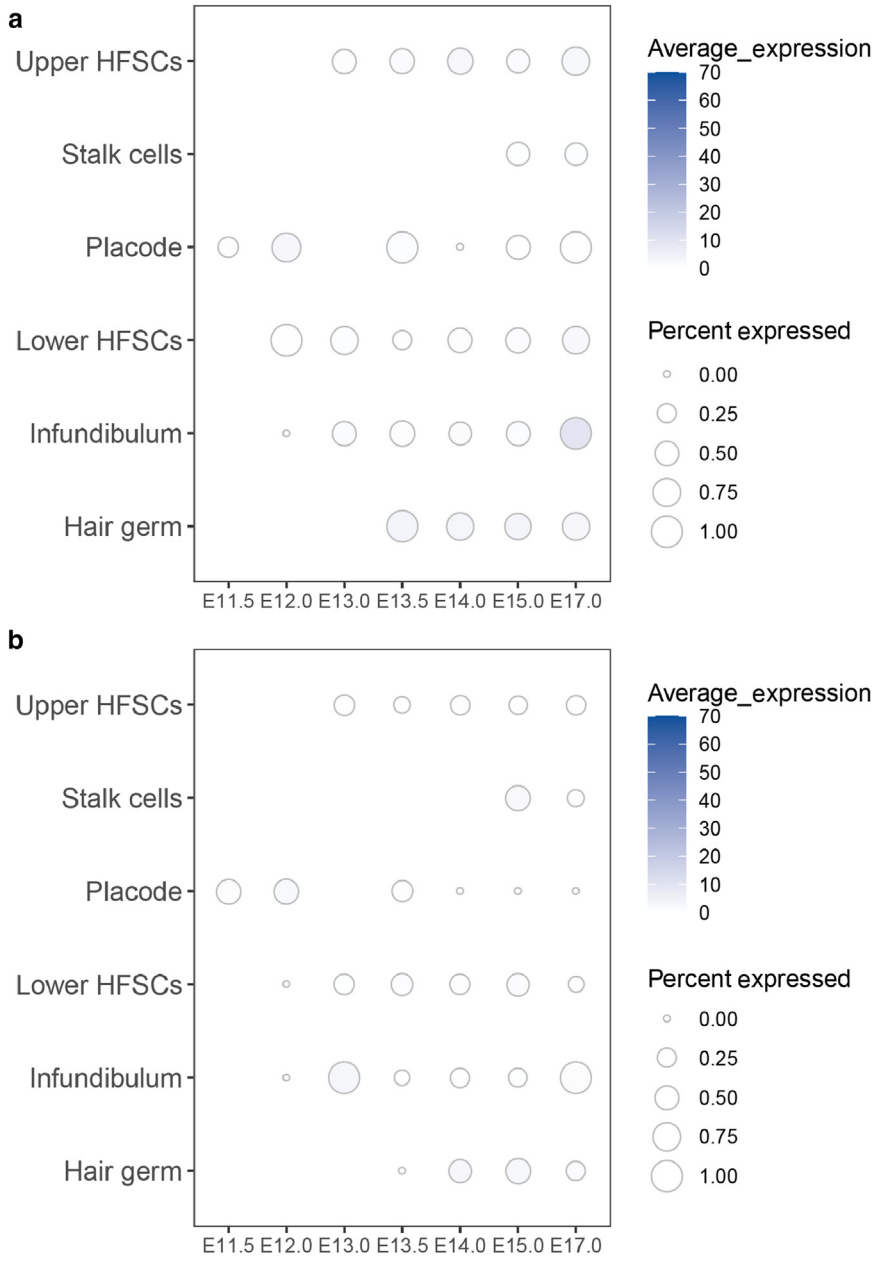
Supplementary Figure S3. Regional association maps. The significantly associated regions on chromosome (**a**) 14q32.13, (**b**) 5q33.2, and (**c**) 7q33 in the meta-analysis are depicted. Color intensity indicates linkage disequilibrium (r^2) with the candidate causal SNPs.



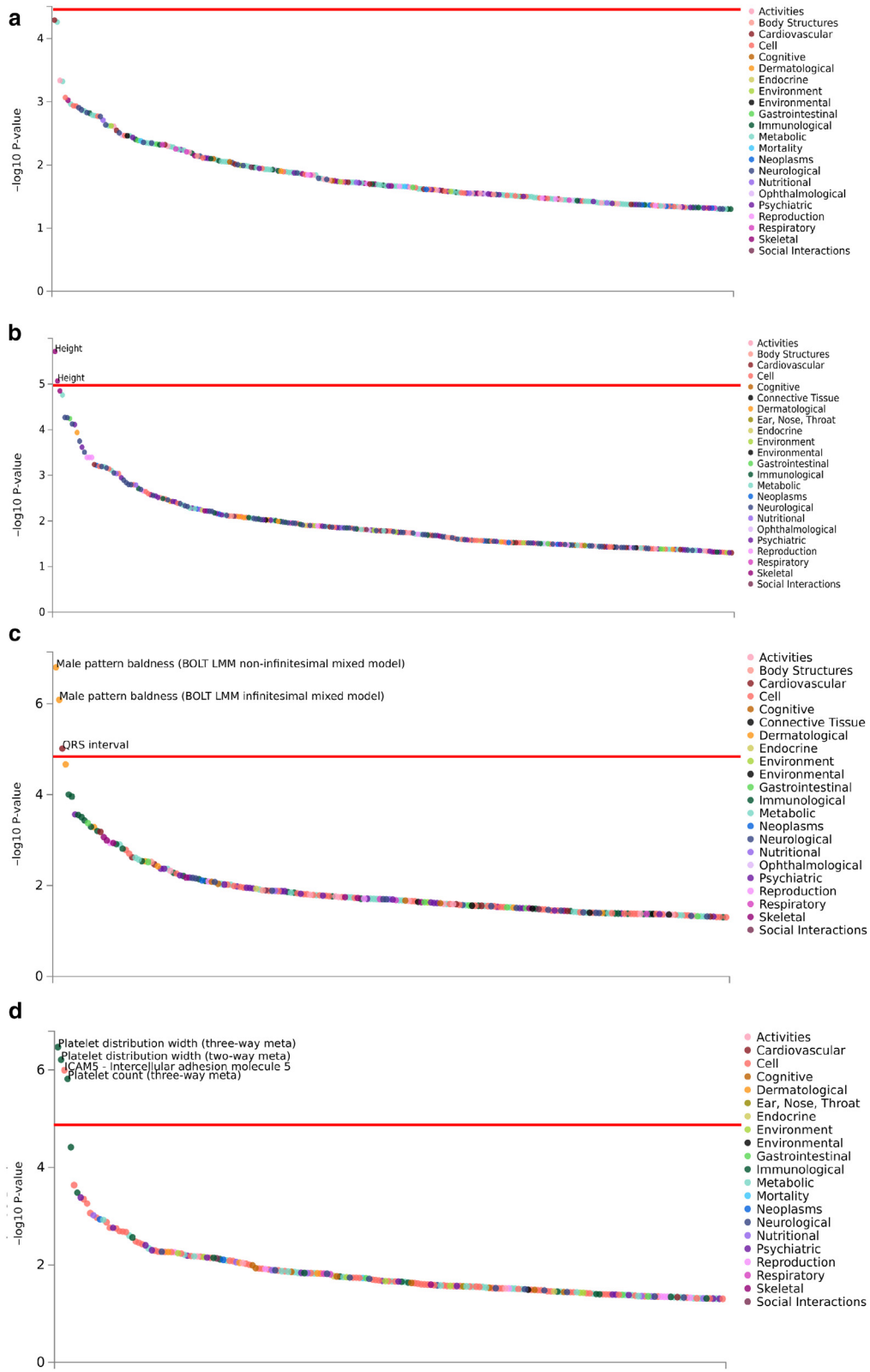
Supplementary Figure S4. Geographical distribution of associated SNP allele frequencies. Allele frequencies of candidate causal SNPs at (**a**) 7p21.3, (**b**) 14q32.13, (**c**) 5q33.2, and (**d**) 7q33 were obtained from the 1000 Genome Project and visualized by the R package rworldmap. Derived alleles are marked by orange, and ancestral alleles are marked by dark blue. Each pie chart represents the frequency of two tagged alleles. Chr, chromosome.



Supplementary Figure S5. Effect and frequency of derived alleles. SNPs on chromosome (**a**) 14q32.13, (**b**) 5q33.2, and (**c**) 7q33 showing the proportion of each whorl direction trait against the genotypes of the three SNPs, respectively. Only one person had the AA genotype of SNP rs834778.



Supplementary Figure S6. Expression pattern of candidate genes in developing whisker follicles of mice. (**a**) *SYNE3* expression pattern in distinct regions of the whisker hair follicle epithelium of mice at each embryonic stage. (**b**) *HAND1* expression pattern in whisker hair follicle epithelium of mice at each embryonic stage. HFSC, hair follicle stem cell.



Supplementary Figure S7. PheWAS results of candidate genes from the GWAS Atlas, showing other phenotypes associated with candidate genes identified in this study. (**a**) PheWAS result of *ARL4A*. (**b**) PheWAS result of *SYNE3*. (**c**) PheWAS result of *HAND1*. (**d**) PheWAS result of *MTPN*. The significance threshold is 1.05×10^{-5} with Bonferroni correction. Only statistically significant phenotypes are labeled in detail. PheWAS, phenome-wide association study.

Supplementary Table S1. Hair Whorl Trait Frequencies in Sample Populations

Cohort		Subjects, n	Male, n (%)	*P*-Value[1]
NSPT	C	1497	744 (49.70%)	0.8466
	CC	652	327 (50.15%)	
	C-CC	40	33 (82.50%)	
	C-C	7	5 (71.43%)	
	CC-CC	2	2 (100%)	
	D	175	98 (56.00%)	
TZL	C	1394	546 (38.24%)	0.9814
	CC	556	216 (37.96%)	
	C-CC	162	56 (34.57%)	
	C-C	205	102 (49.776%)	
	CC-CC	5	3 (60.00%)	
	C-D	1	1 (100%)	
	CC-D	4	3 (75.00%)	

Abbreviations: NSPT, National Survey of Physical Traits cohort; TZL, Taizhou Longitudinal cohort.
C denotes single clockwise hair whorl, CC denotes single counterclockwise hair whorl, C-CC denotes clockwise and counterclockwise double hair whorl, C-C denotes double clockwise hair whorl, CC-CC denotes double counterclockwise hair whorl, D denotes single diffuse hair whorl, C-D denotes clockwise and diffuse double hair whorl, and CC-D denotes counterclockwise and diffuse double hair whorl.
[1]*P*-value of Wilcoxon test for hair whorl type and sex.

Supplementary Table S2. Fine Mapping and Functional Annotation of Candidate Casual SNPs

CHR	SNP	PP[1]	CADD[2]	DeepSEA[3]	GTEx	3DSNP
7p21.3	rs246829	0.99	11.39	0.20	*ARL4A* eQTL in skin	3D interaction with *ARL4A* (HAP1) Enhancer in the skin (SKIN.NHEK)
14q32.13	rs179151	1	13.62	0.77	*SYNE3* and *RP11-1070N10.3* eQTL in skin	3D interaction with *SYNE, SNHG10, SCARNA13* (keratinocyte, NHEK) Enhancer in ESC (ESDR.H1.MSC) TFBS in H1-hESC
5q33.2	rs11346660	0.9998	0.347	0.69		3D interaction with *HAND1, MIR3141* (NHEK) Enhancer in ESC (ESC.H9)
7q33	rs834778	0.99997	2.827	0.29		Linear closest gene: *LOC105375523*

Abbreviations: 3D, three dimensional; CADD, Combined Annotation Dependent Depletion; CHR, chromosome; eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue Expression; NHEK, normal human epidermal keratinocyte; PP, posterior probability.
[1]PP in fine mapping conducted by PAINTOR.
[2]CADD: quantitatively prioritizes functional, deleterious, and disease-causal variants.
[3]DeepSEA: predicts genomic variant effects on transcription factor binding, DNase I hypersensitive sites, and histone marks in multiple human cell types.

Supplementary Table S3. PheWAS Results of Candidate Genes from BioBank Japan

Gene	Phenotype	Top SNP in Gene	*P*-Value[1]
ARL4A	Height	rs117253839	2.00×10^{-9}
	Blood urea nitrogen	7:12821117_A/G	3.60×10^{-6}
	Pollinosis	rs200882663	3.60×10^{-6}
SYNE3	Glaucoma	rs1187734	2.30×10^{-6}
	Iron deficiency anemia	rs142016308	3.20×10^{-6}
	Substance dependence	rs139961206	8.50×10^{-6}
	Goiter	rs118055876	8.60×10^{-6}
HAND1	Retinitis pigmentosa	rs376440284	1.30×10^{-5}
	Angina pectoris	rs201886344	1.60×10^{-5}
	Mean corpuscular hemoglobin	rs644716	2.60×10^{-5}
MTPN	Body mass index	rs1809626	6.40×10^{-8}
	Body weight	rs1809626	1.10×10^{-5}
	Sarcoidosis	rs182054768	5.30×10^{-5}

Abbreviation: PheWAS, phenome-wide association study.
[1] *P*-value of the most significant SNP in each gene. The threshold for statistical significance is 4.6×10^{-9} after Bonferroni correction.