Limb development genes underlie variation in human fingerprint patterns

Jinxi Li,[1,2,27] James D. Glover,[3,27] Haiguo Zhang,[4,27] Meifang Peng,[2,4,27] Jingze Tan,[4] Chandana Basu Mallick,[3,5] Dan Hou,[6] Yajun Yang,[4] Sijie Wu,[1,2] Yu Liu,[2] Qianqian Peng,[2] Shijie C. Zheng,[2] Edie I. Crosse,[7] Alexander Medvinsky,[7] Richard A. Anderson,[8] Helen Brown,[3] Ziyu Yuan,[9] Shen Zhou[10], Yanqing Xu[11], John P. Kemp,[12] Yvonne Y.W. Ho,[13] Danuta Z. Loesch,[14] Lizhong Wang,[15] Yingxiang Li,[15] Senwei Tang,[15] Xiaoli Wu,[15] Robin G. Walters,[16] Kuang Lin,[16] Ruogu Meng,[17] Jun Lv,[18] Jonathan M. Chernus,[19] Katherine Neiswanger,[20] Eleanor Feingold,[19] David M. Evans,[12] Sarah E. Medland,[13] Nicholas G. Martin,[13] Seth M. Weinberg,[19,20,21] Mary L. Marazita,[19,20,22,23] Gang Chen,[15] Zhengming Chen,[16] Yong Zhou,[24] Michael Cheeseman,[3] Lan Wang,[6] Li Jin,[1,2,25,28,*] Denis J. Headon,[3,28,*] Sijia Wang[2,26,28,29*]

[1] State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, and Human Phenome Institute, Fudan University, Shanghai 200438, P.R. 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 200031, P.R. China

[3] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK

[4] Ministry of Education Key Laboratory of Contemporary Anthropology, Department of Anthropology and Human Genetics, School of Life Sciences, Fudan University, Shanghai 200438, P.R. China

[5] Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, India

[6] Chinese Academy of Sciences Key Laboratory of Tissue Microenvironment and Tumor, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, P.R. China

[7] Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK.

[8] MRC Centre for Reproductive Health, Queens Medical Research Institute, University of Edinburgh, Edinburgh, UK

[9] Fudan-Taizhou Institute of Health Sciences, Taizhou, Jiangsu 225326, P.R. China

[10] Shanghai Foreign Language School, Shanghai 200083, P.R. China

[11] Forest Ridge School of the Sacred Heart, Bellevue, WA 98006, USA

[12] University of Queensland Diamantina Institute, University of Queensland, Brisbane, Australia

[13] QIMR Berghofer Medical Rese Institute, Brisbane, Australia
[14] Psychology Department, La Trobe University, Melbourne, Australia
[15] WeGene, Shenzhen, Guangdong 518040, P.R. China
[16] Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[17] Center for Data Science in Health and Medicine, Peking University, Beijing 100191, P.R. China
[18] Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing 100191, P.R. China
[19] Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA 15261, USA
[20] Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, University of Pittsburgh, Pittsburgh, PA, 15219
[21] Department of Anthropology, University of Pittsburgh, Pittsburgh, PA, 15260
[22] Clinical and Translational Science, University of Pittsburgh, Pittsburgh, PA 15261
[23] Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, 15261
[24] Clinical Research Institute, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, P.R. China
[25] 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 200438, P.R. China
[26] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, P.R. China
[27] These authors contributed equally
[28] Senior authors; these authors contributed equally
[29] Lead Contact
*Correspondence: lijin@fudan.edu.cn (L.J.), denis.headon@roslin.ed.ac.uk (D.J.H.), wangsijia@picb.ac.cn (S.W.)

Summary

Fingerprints are of longstanding practical and cultural interest, but little is known about the mechanisms that underlie their variation. Using genome-wide scans in Han Chinese cohorts, we identified 18 loci associated with fingerprint type across the digits, including a genetic basis for the long-recognized “pattern-block” correlations among the middle three digits. In particular, we identified a variant near *EVI1* that alters enhancer regulatory activity and established a role for EVI1 in dermatoglyph patterning in mice. Dynamic EVI1 expression during human development supports its role in shaping the limbs and digits, rather than influencing skin patterning directly. Trans-ethnic meta-analysis identified 43 fingerprint-associated loci, with nearby genes being strongly enriched for general limb development

pathways. We also found that fingerprint patterns were genetically correlated with hand proportions. Taken together, these findings support the key role of limb development genes in influencing the outcome of fingerprint patterning.

Keywords

Fingerprint pattern, genetics, genome-wide association study, trans-ethnic meta-analysis, limb development, *EVI1*

Introduction

Dermatoglyphs are parallel ridge formations present on the skin of the palms and fingers of the hands and on the soles and toes of the feet (Cummins and Midlo, 1926). On the fingertips these regular patterns of ridges and furrows form fingerprints of three principal pattern types: arch, loop, and whorl (Figure 1A). Though fingerprints probably evolved to aid grasping (André et al., 2010; Yum et al., 2020) and for sensing of surface textures (Medland et al., 2007; Scheibert et al., 2009), since the 19th Century the fingerprint has been widely used for personal identification as the patterns are unique to every individual, present from birth, and do not change over the lifespan (Galton, 1892).

Dermatoglyphic patterns on fingers begin to develop on the digit tips after the 10th week of gestation, forming on the skin overlying the swollen and regressing volar pads. By the 14th week, the configuration of the future fingerprint pattern (arch, loop or whorl) is defined at the epidermal-dermal junction by the primary ridges (Babler, 1991; Okajima, 1975). Several mechanisms have been proposed to explain the generation of these repeated patterns of epidermal ridges, including theories based on the resolution of mechanical strain on the epidermis through buckling (Kücken, 2007), the arrangement of ridge configurations according to a template set by blood vessels or nerves (Hirsch and Schweichel, 1973), and the operation of reaction-diffusion signaling processes (Garzón-Alvarado and Ramírez Martinez, 2011). However, the biological mechanisms underlying the generation of dermatoglyph patterns and the overall fingerprint configuration remain largely unknown.

Previous studies have reported population differences (Zhang et al., 2010) and considerable heritability of fingerprint pattern types (h^2=0.3-0.8) (Karmakar et al., 2011; Machado et al., 2010; Sengupta and Karmakar, 2004). A recent genome-wide association scan (GWAS) with moderate sample size discovered several loci associated with fingerprint patterns in a European-ancestry cohort, but none of the loci had previously ascribed functions in either limb or skin development (Ho et al., 2016), yielding little insight into the underlying biological mechanisms. In the current study, by performing large-scale GWAS in Han Chinese populations, as well as trans-ethnic meta-analysis of more than 23,000 individuals, we identified numerous novel loci underlying the systematic variation in human fingerprint patterns, implicating genes with important roles in

embryonic limb development as the principal determinants of heritable fingerprint variation. We functionally validated the top signal near *EVI1* as altering enhancer regulatory function, established the importance of the EVI1 protein in dermatoglyph patterning in mouse models, and assessed EVI1 expression across limb development to dermatoglyph formation in human fetal tissues. We further found evidence of a shared genetic basis between fingerprint type and hand proportions. Our findings highlight the importance of limb development genes and processes in defining human fingerprint patterns.

Results

1. Genome-wide scans identify 18 genomic regions associated with fingerprint patterns across all digits in Han Chinese cohorts

In the discovery phase, we conducted genome-wide association scans (GWAS) on fingerprint patterns on all ten digits (D1-5L, D1-5R) in 9,909 individuals from three Han Chinese cohorts: the Taizhou Longitudinal Study (TZL, n=2,961), the National Survey of Physical Traits (NSPT, n=2,679), and the Jidong cohort study (JD, n=4,269) (see also Table S1 for the details of cohorts). The GWAS on the ordinal phenotypes (coded as 0, 1, and 2 for arch, loop, and whorl, respectively; Figure 1A; STAR Methods) identified 18 genome-wide significant signals after adjusting for multiple testing ($P_{adj}<1.67\times10^{-8}$; Figure 1B), while the GWAS based on the binary phenotypes (non-whorl or whorl) showed similar results (Table S2). The majority of signals (17 out of 18) were replicated in at least one of the two validation cohorts: the China Kadoorie Biobank (CKB, n=1,785) and the WeGene cohort (WeGene, n=2,152), with consistent allelic effects across all the cohorts (Table 1). The narrow-sense heritability (estimated using GCTA software; see STAR Methods) of fingerprint patterns on each digit was between 0.295-0.432 in discovery cohorts. The 18 most significant SNPs explained 3.06-5.56% of the phenotypic variance (R^2; Table S3). We further conducted GWAS on 66 different derived phenotypes ($P_{adj}<3.57\times10^{-9}$, e.g. binary, nominal, ordinal, and quantitative phenotypes), and found most of the significant associations overlapped with the 18 aforementioned genomic signals, suggesting the simple ordinal phenotypes cover most of the phenotypic information related to variation of fingerprint pattern (Table S4).

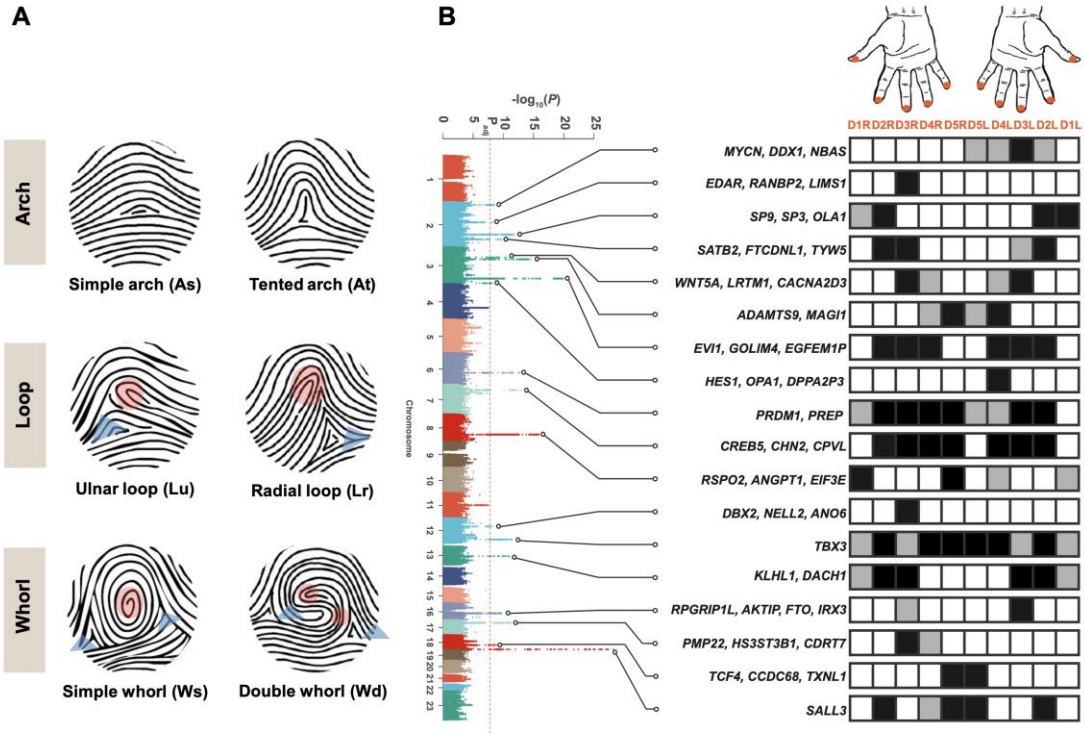

Figure 1. Genetic variants associated with ordinal fingerprint patterns (arch, loop and whorl) in Han Chinese (N=9,909)

(A) Pattern-types of fingerprints according to the number of triradii/deltas (triangles) and cores (circles) (STAR methods). There are three main types: arch, loop and whorl. Each main group contains two sub-types according to the steepness, direction of ridges and the variable core.

(B) Genome-wide scans of the ordinal arch-loop-whorl phenotype identify 18 genomic regions associated with fingerprint patterns. The red line indicates the threshold for genome-wide significance after adjusting for the effective number of independent phenotypes ($P_{adj}<1.67\times10^{-8}$; STAR Methods). Detailed patterns of adjusted association significance across different fingers are indicated by black squares for corresponding digits for significant associations ($P_{adj}<1.67\times10^{-8}$), and grey squares for marginal associations ($P_{adj}<3.33\times10^{-6}$). Notable genes are indicated for each locus (see Table 1 for selection criteria). Abbreviation: D1-5L/R = digit 1 to 5 of left or right hand. See also Table S2 and S4.

Table 1. GWAS signals for fingerprint pattern type in Han Chinese discovery and replication populations

Locus	Top SNPs	Alt/Ref allele	EAF[a]	Notable genes[b]	Associated digits[c]	Discovery-TZL+NSPT+JD (N=9909)[d]		Replication1-CKB (N=1785)[d]				Replication2-WeGene (N=2152)[d]			
						β	P^e (Top digit)	β^f	P^f	β^g	P^g	β^f	P^f	β^g	P^g
2p24.3	rs16862838	T/C	0.197	*MYCN(G/L),DDX1(G/GE/R/L),NBAS(L)*	D2L,**D3L,**D4L,D5L	-0.106	2.92×10^{-9}(D3L)	-0.047	3.2×10^{-2}(L3)	-0.042	0.083	-0.086	3.91×10^{-5}(L1)	/	/
2q13	rs371242548	C/A	0.054	*EDAR(G),RANBP2(G),LIMS1(L)*	**D3R**	-0.184	6.11×10^{-9}(D3R)	-0.104	7.8×10^{-3}(L3)	-0.018	0.63	-0.11	1.46×10^{-3}(L1)	/	/
2q31.1	rs11460049	AT/A	0.201	*SP9(L),SP3(G/L),OLA1(G/GE/R/L)*	**D1L**/R,**D2L/R**	-0.127	1.32×10^{-12}(D2R)	-0.104	4.6×10^{-5}(L2)	-0.08	3.6×10^{-3}	--	--	--	--
2q33.1	rs4673509	C/A	0.184	*SATB2(G/L),FTCDNL1/FONG(G/R/L),TYW5(L)*	**D2L/R,D3**L/**R**	0.118	1.74×10^{-10}(D2L)	0.073	8.50×10^{-3}(L1)	/	/	0.05	1.64×10^{-2}(L3)	0.023	0.331
3p14.3	rs358075	T/C	0.048	*WNT5A(G/L),LRTM1(G/L),CACNA2D3(GE/R/L)*	**D3L/R**,D4L/R	-0.222	2.99×10^{-11}(D3L)	-0.126	8.10×10^{-3}(L1)	/	/	-0.141	4.51×10^{-4}(L1)	/	/
3p14.1	rs17072351	G/A	0.225	*ADAMTS9(G/L),MAGI1(G/L)*	**D4L**/R,**D5**L/**R**	-0.133	7.42×10^{-15}(D4L)	-0.057	9.40×10^{-3}(L1)	/	/	-0.053	3.30×10^{-3}(L2)	-0.048	1.53×10^{-2}
3q26.2	rs6444832	A/G	0.459	*EVI1(G/R/L),GOLIM4(G),EGFEM1P(L)*	**D2L/R,D3L/R,D4L/R**	-0.133	1.31×10^{-20}(D3L)	-0.074	2.10×10^{-4}(L1)	/	/	-0.048	6.80×10^{-3}(L2)	-0.032	0.055
3q29	rs80252354	T/C	0.031	*HES1(G/L),OPA1(G/L),DPPA2P3(L)*	**D4L**	-0.238	7.36×10^{-9}(D4L)	0.094	0.068	/	/	0.115	0.125	/	/
6q21	rs28700026	C/T	0.254	*PRDM1(G/L), PREP(G/R/L)*	D1R,**D2L/R,D3L/R,D4L/R,D5**L/**R**	-0.12	2.57×10^{-13}(D2R)	-0.053	5.2×10^{-3}(L3)	-0.033	0.18	-0.043	4.09×10^{-2}(L2)	-0.019	0.355
7p14.3	rs2075127	C/T	0.284	*CREB5(G/L),CHN2(G/L),CPVL(GE/R/L)*	**D2L/R,D3L/R,D4L/R, D5R**	0.118	9.18×10^{-14}(D3R)	0.038	4.3×10^{-2}(L2)	0.036	0.063	0.058	2.32×10^{-3}(L2)	0.035	4.89×10^{-2}
8q23.1	rs1494910	C/T	0.161	*RSPO2(G/R/L),ANGPT1(G/L),EIF3E(L)*	**D1L/R**,D4L,**D5R**	0.215	1.30×10^{-16}(D1R)	0.142	1.70×10^{-7}(L2)	0.113	8.60×10^{-6}	0.085	1.37×10^{-4}(L3)	0.033	0.155
12q12	rs2731043	C/T	0.182	*DBX2(G/GE/R/L),NELL2(G/L),ANO6(L)*	**D3R**	0.107	8.06×10^{-9}(D3R)	0.051	2.30×10^{-2}(L1)	/	/	0.028	0.184	/	/
12q24.21	rs7957733	G/C	0.28	*TBX3(G/R/L)*	D1L/R,**D2L/R,**D3L/R,**D4L/R**, **D5L/R**	0.111	3.78×10^{-12}(D5L)	0.054	7.8×10^{-3}(L3)	0.032	0.082	0.075	4.40×10^{-5}(L2)	0.068	4.88×10^{-5}
13q21.33	rs11618603	G/C	0.139	*KLHL1(G/L),DACH1(R/L)*	D1L/R,**D2L/R,D3L/R**	0.14	1.24×10^{-11}(D3L)	0.102	2.5×10^{-4}(L3)	0.02	0.47	0.096	3.57×10^{-4}(L2)	0.058	2.71×10^{-2}
16q12.2	rs5005161	C/T	0.425	*RPGRIP1L(G/GE/R/L),AKTIP(G/L),FTO(L),IRX3(G)*	**D3L**/R	0.092	1.93×10^{-10}(D3L)	0.05	1.5×10^{-3}(L3)	0.01	0.59	0.063	1.95×10^{-4}(L1)	/	/
17p12	rs7208722	G/A	0.299	*PMP22(G),HS3ST3B1(G),CDRT7(R/L)*	**D3R**,D4R	0.108	4.63×10^{-12}(D3R)	0.073	1.5×10^{-3}(L3)	0.031	0.096	0.045	9.00×10^{-3}(L1)	/	/
18q21.2	rs17089876	C/T	0.441	*TCF4(G/GE/R/L),CCDC68(G),TXNL1(G)*	**D5L/R**	0.087	1.46×10^{-9}(D5L)	0.033	0.057	/	/	0.039	1.55×10^{-2}(L3)	0.027	0.068
18q23	rs2004773	A/G	0.382	*SALL3(G/R/L)*	**D2L/R**,D4R,**D5L/R**	0.164	6.14×10^{-29}(D5R)	0.059	7.1×10^{-4}(L1)	/	/	0.101	1.89×10^{-10}(L1)	/	/

[a]The effect or alternative (Alt) allele frequency of the discovery populations.

[b]Notable genes are indicated as follows: 1) the two nearest genes within 1000 kb of the most significantly associated SNP annotated by GREAT (G), which uses the subset of the UCSC Known Genes; 2) the nearest gene mapped by GENCODE (GE) or RefSeq (R); 3) Protein-coding genes within 1000 kb of the most significantly associated SNP in regional LocusZoom plot (L). Underlining indicates that the best-associated SNP is located within the gene.

[c]Genome-wide significant level ($P_{adj}<1.67\times10^{-8}$, bold font) or suggestive level ($P_{adj}<3.33\times10^{-6}$) after multiple-testing adjustment.

[d]The sample sizes vary in GWAS on different phenotypes of digit: N=5415-9909 for discovery cohort (fingerprint patterns on digit 1 are not available in JD cohort), N=1634-1785 for replication cohort 1 (CKB), and N=2138-2152 for replication cohort 2 (WeGene).
[e]The associations between the top SNPs and the fingerprint pattern of the most significant digit (i.e. top digit, as indicated in parentheses).
[f]The signal was replicated at different levels of association, as following: the most significant replication is exactly the association between the top SNP and the top digit (L1); the most significant replication is the association between the top SNP and one of the other associated digits, while the association between the top SNP and the top digit is also significant (L2) or not significant (L3). "--" not available (INDEL polymorphisms are not available in the WeGene cohort).
[g]The associations between the Top SNPs and Top digits in replication cohorts. The "/" indicate that the associations have the same effect size and p value as the results of the two columns in front.
Abbreviation: Alt=alternative, ref=reference, EAF=Effect or Alt Allele Frequency, TZL=cohort from Taizhou Longitudinal Study, NSPT=cohort from National Survey of Physical Traits Project, JD=cohort from Jidong of Hebei Province, CKB=cohort from China Kadorie Biobank, WeGene=cohort from WeGene company. D1-5L/R=digit 1 to 5 of left or right hand.
See also Table S1.

2. The *EVI1* locus contributes to the "pattern-block" correlation between the middle three digits on both hands

Fingerprint pattern types on contralateral (left and right) digit pairs are highly correlated, also with strong genetic correlations among them (0.64-1, Figure 2A). In particular, the patterns on the middle three digits are more correlated with one another than are the patterns on the other digits, reflecting the long-recognized "pattern-block" (also known as "pattern influence"; (Martin et al., 1982; Nagy and Pap, 2005)) phenomenon. We found that using all SNPs across the genome, genetic correlations among the middle three digits on both hands (0.93-1) are significantly higher than those among all ten digits (Figure 2B). To model the common elements underlying such correlations among the "pattern-block", we adopted a partial least-square path model (PLSPM, STAR Methods) to extract the composite phenotype from the middle three digits on both hands. GWAS on this composite phenotype found 12 signals, 11 of which had been detected in the GWAS on the ordinal phenotypes; a novel signal at 5p12 was also discovered (Figure 2C, Table S3). This composite phenotype showed strengthened narrow-sense heritability (h^2_{CP} =0.524±0.039). Its phenotypic variance could be genetically explained by the top SNPs of the 12 signals with higher proportion (6.24%) than single phenotype (3.16-4.88%) (R^2; Table S3), suggesting a shared genetic basis with the individual digit phenotypes from which the trait was extracted. The top signal at 3q26.2 was also effectively strengthened (P=6.41×10^{-22}, while P values were between 1.31×10^{-20}-2.93×10^{-11} for single phenotypes).

The top signal at 3q26.2 is located ~100 kb downstream of the *EVI1* (Ecotropic Viral Integration Site 1) gene, also termed *MECOM* (*MDS1* and *EVI1* complex locus) (Figure 2D). Fine-mapping of the 3q26.2 association-enhanced composite phenotype using PAINTOR with functional annotation (STAR methods) found that the 99% credible set contained two SNPs: rs7646897 (posterior probability=0.383) and rs7623083 (posterior probability=0.617) (Figure 2D top). Interrogation of ENCODE and REMC databases revealed that the 3q26.2 signal region showed distinct active enhancer signatures in a range of human cells types (Figure 2D bottom). Further Hi-C data showed the SNP rs7646897 region and *EVI1* gene are located in the same Topologically Associating Domain (TAD), while other nearby genes are located outside the boundary of this domain (Figure S1A). There is also chromatin interaction between promoter region of *EVI1* gene and enhancer harboring SNP rs7646897 (Figure S1B). To verify this potential enhancer activity on the *EVI1* gene, we performed luciferase reporter assays (STAR methods) on rs7646897 and rs7623083 in HEK293T cells, and detected allele specific differences in enhancer activity only for rs7646897 (P=7.00×10^{-4}; Figure 2E, 2F and S2A-H). Further assays showed that the alternate SNPs at rs7646897 modulated the expression of *EVI1* (P=7.50×10^{-3}), but not the promoters of the closet up- and downstream genes *GOLIM4* (P=0.364) and *TERC* (P=0.778) (Figure S2I). Independent experimental repeats showed similar results (Figure S2J).

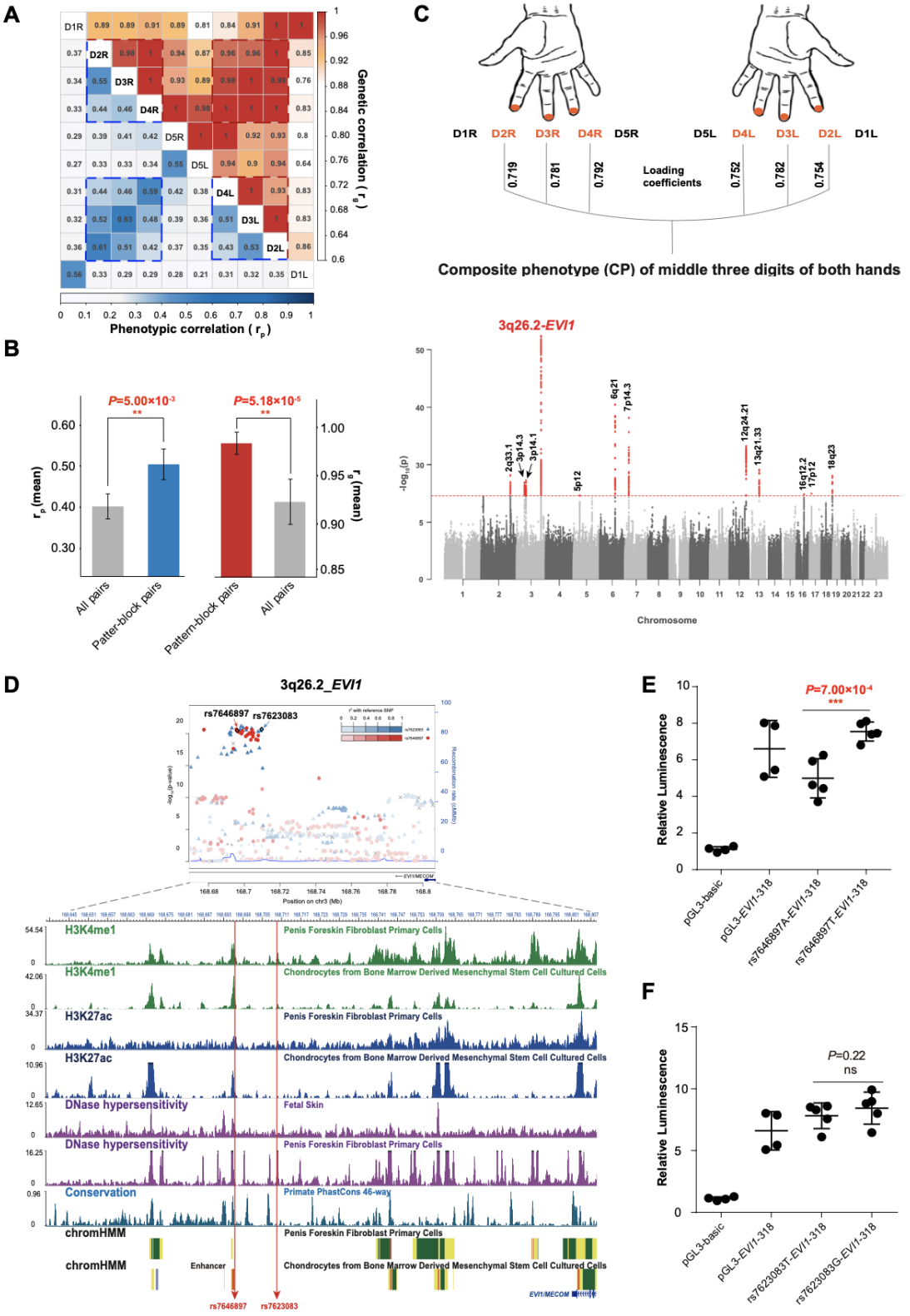

Figure 2. Genetic basis of the middle three digit "pattern-block" phenomenon, with top signal near *EVI1*

(A) "Pattern block" of the middle three digits on both hands revealed by pair-wise phenotypic correlation (blue) and genetic correlation (red) among the ten digits (N=9,909). The dashed box indicates high correlations between the same digits of both hands and neighboring digits. The

correlations from high to low were represented by both color and correlation coefficients (r) in the figure.

(B) The correlations of fingerprint patterns between the middle three digits on both hands (Pattern-block pairs) are higher than the correlations of all random pairs of the ten digits (All pairs).

(C) Genome-wide scan on the composite phenotype extracted from the fingerprint pattern of the middle three digits on both hands. The loading coefficients of the composite phenotype on the six correlated variables are between 0.719 and 0.792.

(D) Fine mapping of signals at 3q26.2: LocusZoom plot of SNPs at the 3q26.2 region (top) and mapping of epigenetic marks H3K4me1, H3K27ac, DNase hypersensitivity and conservation analysis at the same region, based on ENCODE and RMEC project data. SNP rs7646897 and rs7623083, indicated by red lines, are in a region that exhibits distinct active enhancer signatures defined by epigenetic marks, such as H3K4me1 (green), H3K27ac (blue) histone modifications and DNase hypersensitivity (purple), and with enhancer function by chromatin state assay (yellow box) in fibroblast primary cells and in chondrocytes. The phastCon score indicates the evolutionarily conservation in primates.

(E) Luciferase reporter assays on candidate regulatory elements carrying alternate alleles at SNPs rs7646897 and (F) rs7623083 in HEK293T cells. pGL3-basic is a negative control plasmid lacking enhancer activity, pGL3-318 is a positive control derived from the *EVI1* promoter region. Symbols indicate significance in t test (*P<0.05, **<0.01, ***<0.001). See also Figure S1, Figure S2 and Table S3.

3. *Evi1* mutation alters mouse dermatoglyph patterns and is expressed through early limb development stages

The regulatory SNP rs7646897 is intergenic, lying 100 kb from the *EVI1* gene. As regulatory elements can exert effects at large physical distances, we assessed the importance of the *EVI1*-encoded protein itself on dermatoglyph pattern formation using mouse models. Based on their location on the ventral side of digits, their formation prior to birth (Figure S3A, S3B), and their parallel arrangement of ridges carrying the pores of sweat glands, we analyzed the transverse digital ridges in the mouse as the closest model phenotype of human dermatoglyphs (see Supplementary Information for further justifications).

We analyzed dermatoglyph patterns in 21 day old $Evi1^{Jbo/+}$ heterozygous mutants (encoding EVI1p.Asn763Ile) and wild type littermate digits 2, 3 and 4 (Figures 3A, 3B, and 3C; Table S5). Homozygous $Evi1^{Jbo/Jbo}$ embryos die between midgestation and birth, while heterozygotes display a small spur on digit 5 (Figure 3A) (Parkinson et al., 2006), and slightly decreased digit length (Figure S4A). Using a mixed ordinal logistic regression model, we found reduced frequency of continuous ridges in all mutant digits ($P_{digit2}=0.02$; $P_{digit3}=0.0005$; $P_{digit4}<0.0001$), with digits 3 and 4 also carrying more discontinuous ridges ($P_{digit3}=0.03$; $P_{digit4}<0.003$; Figure 3C). These results demonstrate directly that EVI1 itself is a modulator of dermatoglyph patterns.

We assessed *Evi1* expression in intact embryonic mouse limbs by whole mount in situ hybridization, identifying high expression broadly in the autopod at E12.5, then becoming restricted to the distal regions of all emerging digits (Figure 3D). RNAscope in situ hybridization permits transcript detection at later stages on sectioned tissue. This approach identified broad *Evi1* expression throughout limb mesenchyme at E11.5, becoming restricted to the distal end of the limb at E13.5, and excluded from the cartilage elements of the digits by E15.5, reducing further by E17.5 (Figure 3E, S4B). Quantification of *Evi1* expression in embryonic mouse limbs (Figure 3F), agreed with the findings by in situ hybridization, finding a steep decline across the stages of limb bud outgrowth at embryonic day (E) 11.5, digit emergence (E13.5), digit outgrowth and definition of dermatoglyph patterns (E15.5 and E17.5). In the distal digit *Evi1* expression was detected in the same cells as *Prrx1*, also named *Prx1*, a marker of limb bud mesenchyme (Chesterman et al., 2001; Nohno et al., 1993) (Figure S4B).

We determined EVI1 expression during human embryonic development by immunofluorescence, finding broad expression in the mesenchyme of the early outgrowing limb bud at Carnegie Stage 17 (CS17, approximately 6 weeks estimated gestation age (EGA)), in contrast to low or absent expression in the trunk (Figure 3G). By 10 weeks EGA, EVI1 was prominently expressed in the mesenchyme of the distal ends of the digits, notably under the volar pads, the sites of later fingerprint formation (Figure 3H). By 13 weeks EGA expression is largely lost, being present only at the periphery of the distal phalange (Figure 3I), and at 16 weeks EGA, as dermatoglyphs are emerging as a periodically corrugated epidermis with extending sweat gland primordia, EVI1 expression is greatly reduced and it is not detected in these structures (Figure 3J). To further refine the cell population expressing *EVI1*, we assessed its coexpression with general limb bud mesenchyme marker *PRRX1*. At week 10 expression is detected in the distal digit and the deeper parts of the distal volar pad, in the same cells population as expresses *PRRX1* (Figure 3K). Expression of *EVI1* in human development, as in mouse (Figure 3E), is similar on dorsal and ventral sides of the digit, though broader on the ventral side as it extends into the lower volar pad. As observed for the protein, *EVI1* mRNA is not detected in emerging fingerprint ridges at week 16 EGA (Figure 3L). Thus, in human development EVI1 is expressed broadly in mesenchyme during limb growth and then under the digit pad upon which the dermatoglyph pattern will form, but expression is not associated with the epithelial folding of the fingerprint itself.

Early mouse and chicken limbs grow through proliferation from their distal end (Towers and Tickle, 2009) and we find high levels of cell proliferation, marked by Ki67 immunofluorescence, in 10 week EGA digits in the distal tip mesenchyme (Figure 3M). This region is a major site of EVI1 expression and as the principal function of this transcription factor is to promote proliferation (Hoyt et al., 1997), it is likely that altered *EVI1* regulation changes patterns of cell growth and thus the length and shape of the distal limb and digits.

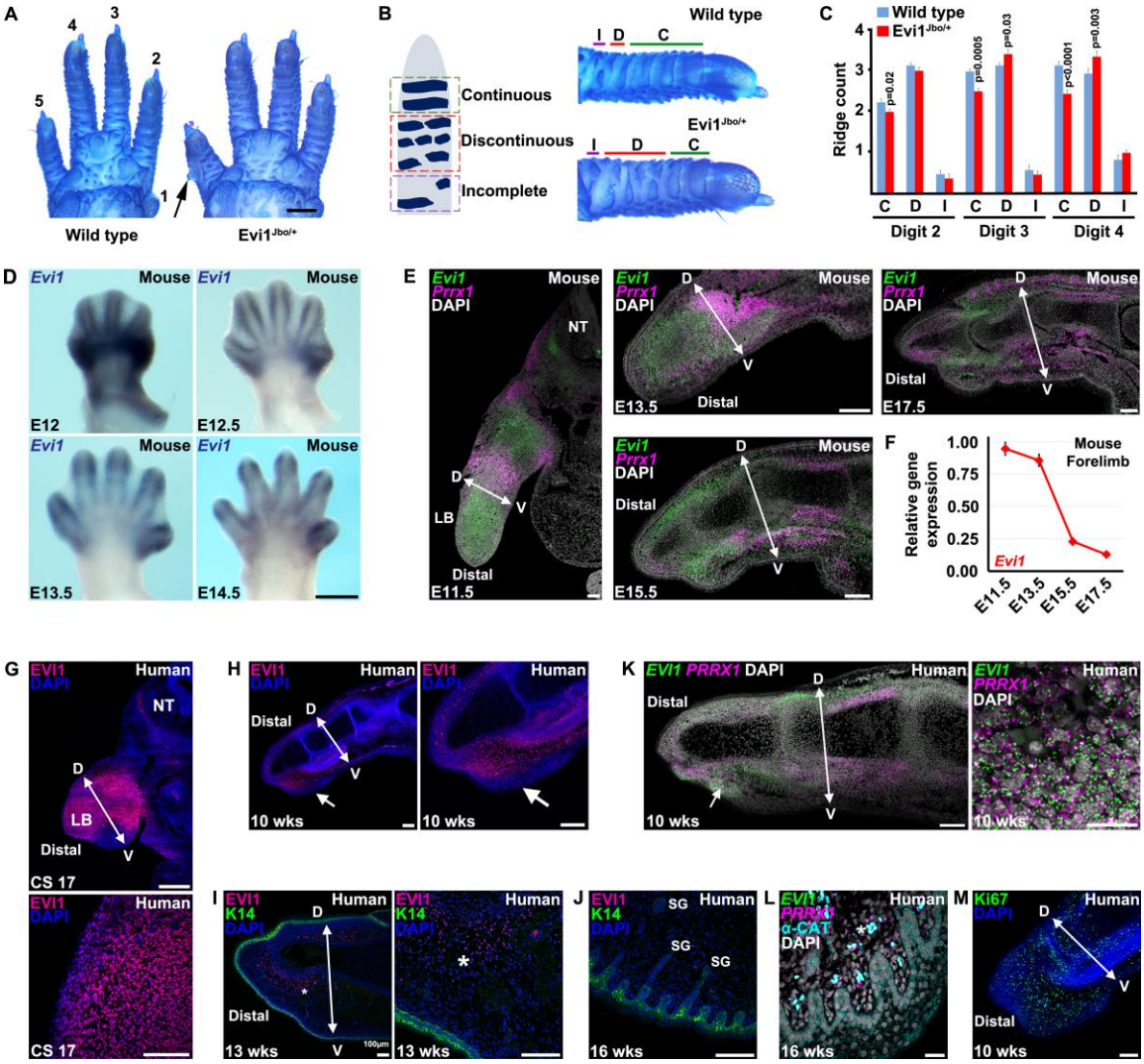

Figure 3. EVI1 in dermatoglyph patterning and limb development. A) Palmar dermal surface of toluidine blue stained paws from wild type and *Evi1*$^{Jbo/+}$ adult mice showing dermatoglyph arrangement. Arrow indicates spur on the mutant digit 5 (D5). B) Schematic depicting transverse ridge categories on mouse digits and ventral surface of D4 of wild type and *Evi1*$^{Jbo/+}$. Regions carrying continuous: C, discontinuous: D, and incomplete: I ridges are indicated. C) Quantification of digit ridge pattern in wild type and *Evi1*$^{Jbo/+}$ mutants. Continuous ridges are reduced on all mutant digits, while D3 and D4 carry more discontinuous ridges. D) Wholemount *in situ* hybridization detecting *Evi1* expression in mouse embryonic forelimbs. Ventral view. E) RNAscope *in situ* hybridization detecting *Evi1* and the limb mesenchyme marker *Prrx1* transcripts in mouse embryonic limb and digits between E11.5 and E17.5. F) Quantitative RT-PCR determination of *Evi1* expression in mouse forelimb at E11.5 (whole limb bud), E13.5, E15.5 and E17.5 (autopod only). (G-J) Immunofluorescence detecting EVI1 expression in human embryonic tissue. G) Transverse section of CS17 embryo (~6 week EGA) shows nuclear expression in mesenchymal cells of the limb bud (LB, magnified in lower panel). The neural tube (NT) indicates the dorsal midline. H) Longitudinal section of 10 week EGA digit, arrow indicates the raised volar pad across which fingerprints form. I) 13 week EGA digit and J) 16 week EGA digit detecting EVI1 and epithelial

marker K14. Dotted line indicates dermal-epidermal junction. SG: eccrine sweat gland. K, L) RNAscope in situ hybridization detecting *EVI1* and *PRRX1* transcripts in sectioned K) 10 week EGA and L) 16 week EGA human digit, with α-catenin immunofluorescence. Individual cells co-express *EVI1* and *PRRX1*. Asterisks indicates autofluorescent blood cells. M) Detection of proliferative cell marker Ki67 in 10 week EGA digit. Dorsal (D) and ventral (V) axes are annotated: Nuclei are stained with DAPI. Scale bars: A=1 mm; D, G upper = 500 μm; E, G lower, H, I, J, K left, M=100 μm; K right, L=20 μm. Error bars indicate S.E.M. See also Figure S3, Figure S4 and Table S5.

4. Trans-ethnic meta-analysis reveals a fingerprint pattern associated gene set enriched for limb development functions

To more fully understand the genetic architecture of fingerprint patterns, we performed a meta-analysis of both East Asian (EAS)-ancestry cohorts (TZL, NSPT, JD, CKB and WeGene), and European (EUR)-ancestry cohorts, including the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort study, the Queensland Institute of Medical Research (QIMR) twin studies, and the Pittsburgh cohort (Pittsburgh) using the binary phenotypes (non-whorl or whorl) on all ten digits (Table S1). This large-scale meta-analysis of more than 23,000 individuals identified 43 independent signals mapping to 105 notable genes ($P_{adj}<1.67\times10^{-8}$; Figure 4A, Table S6, see also Figure S5 for detailed LocusZoom plots), which explained 4.6%-7.9% of the phenotypic variance by polygenic risk score (STAR Methods). Of these 43 signals, 4 signals were independently identified in both EAS-ancestry and EUR-ancestry cohorts; 27 were only identified in EAS-ancestry cohorts, 9 of which were also nominally significant in EUR-ancestry cohorts; 2 were only identified in EUR-ancestry cohorts, both of which were also nominally significant in EAS-ancestry cohorts; 10 were identified only after trans-ethnic meta-analysis (Figure 4B; Table S6). Among the 18 signals that were only identified in EAS-ancestry and did not reach nominal significance in EUR-ancestry cohorts, 9 were either not available or with low minor allele frequency (MAF<0.04, while the other 9 all showed suggestive level of association in EUR-ancestry cohorts ($6.61\times10^{-6}<P<5.69\times10^{-3}$), with exactly the same effect direction as in EAS-ancestry cohorts (Table S6). Three signals were genome-wide significant in either the EAS or EUR-ancestry cohorts, but were only nominally significant in the trans-ethnic meta-analysis (Table S6). These results largely indicated that the fingerprint related genes are generally the same in European and East Asian ancestry populations, with some differences that are likely explained by differing allelic effect sizes or frequencies among populations (Figure S6).

Genomic enrichment analysis using GREAT V4.0.4 (McLean et al., 2010); see STAR Methods; Table S6) found that the fingerprint-associated signals were significantly enriched for embryonic-development and morphogenesis-related Gene Ontology biological processes including “limb development”, “Embryonic limb morphogenesis”, and “limb morphogenesis” (-log10(*P*)>8; Figure 4C); while epithelial and skin-related pathways including “morphogenesis of an epithelium” and

171 “epithelial tube morphogenesis” did not reach the significance threshold after correction for
172 multiple-testing (-log10(P_{thr})=5.42, Figure 4C). Similar patterns of enrichment for limb rather than
173 skin related terms were observed when using human phenotype annotations and mouse morphology
174 terms for the analysis (Figure 4C). Enrichment was also observed for expression in mouse limb
175 tissues at Theiler Stage 17 (E10.5) and 19 (E11.5) when the forelimbs are divided into two regions
176 indicating proximal limb and autopod, but not in the developing autopod epithelium (Figure 4C).
177 Network analysis using STRING v11 (Szklarczyk et al., 2019) linked fingerprint-associated
178 GWAS genes and proteins by known functions and regulatory interactions (see STAR Methods),
179 showing a network centered around WNT5A and IRX3 (Figure 4D), suggesting relevant
180 developmental pathways. Interestingly, both WNT5A and IRX3 were among the GWAS signals of
181 the composite phenotype of middle three fingers (Figure 2C). Eight proteins from 6 out of the 12
182 composite phenotype-associated signals appeared within two steps of the centered proteins with
183 higher degrees in the network (WNT5A and IRX3), suggesting that these developmental pathways
184 are relevant to the biological mechanism underlying the “pattern-block” phenomenon. Another
185 network was centered on NOTCH signaling (NOTCH1, though not itself a GWAS hit), responsible
186 for maintaining mesenchymal progenitors of the limb in an undifferentiated and proliferative state
187 during its development (Dong et al., 2010).

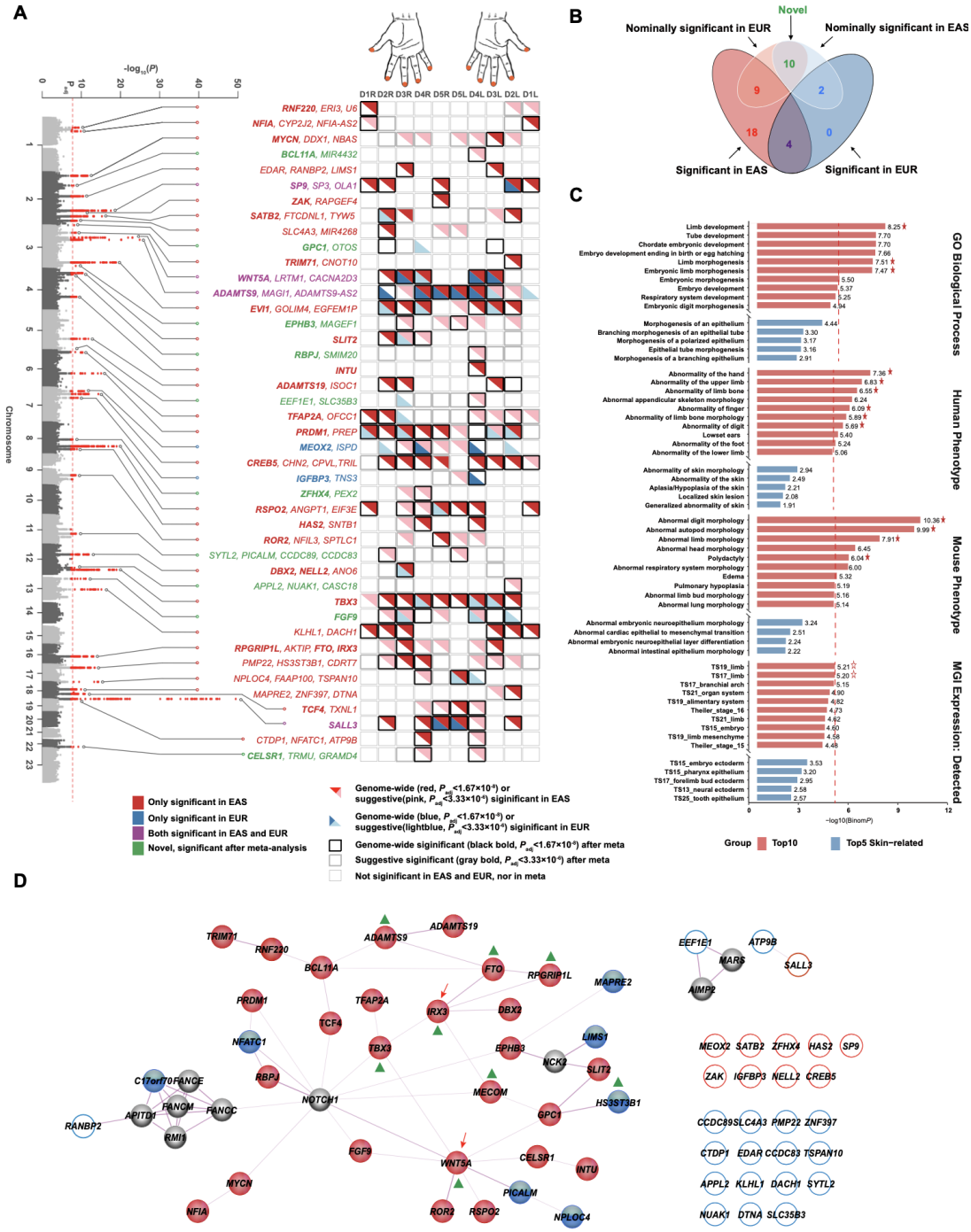

Figure 4. A meta-analysis of fingerprint patterns showing signals enriched in limb development

(A) A Manhattan plot showing the results of the meta-analyses combining GWAS of East Asian (EAS)-ancestry (TZL, NSPT, JD, CKB and WeGene) and European (EUR)-ancestry cohorts (ALSPAC, QIMR and Pittsburgh) across all ten digits (D1L/R were unavailable in JD and ALSPAC). There were 43 signals associated with fingerprint patterns of at least one digit (P_{adj}<1.67×10^{-8}; Table S6), with gene names in different colors: purple indicating significant in both EAS and EUR; red and blue indicating only significant in EAS and EUR, respectively; green indicating not significant in either EAS or EUR, but only significant after the meta-analysis combining both. Bold genes

showed associations with limb phenotypes abnormalities (Table 2). The block map on the right represented the digits corresponding to the signals on the left. Red and blue triangles indicate significance in EAS and EUR, respectively, while dark and light colors represented signals that reached the adjusted genome-wide significant (P_{adj}<1.67×10^{-8}) and suggestive levels (P_{adj}<3.33×10^{-6}), respectively. Bold frame indicated genome-wide significant (black) or suggestive (gray) significant after combined meta-analyses.

(B) Venn diagram summarizing fingerprint-associated signals corresponding to Figure 4A.

(C) Enrichment of annotations across ontologies for the 43 fingerprint-associated signals. The red asterisk indicates limb-relevant terms that genes are significantly enriched in after Bonferroni correction (the red dotted lines). Only the top 10 terms ranked after enrichment analysis and top 5 epithelial/skin-related terms are shown.

(D) Fingerprint pattern-associated proteins and their interactions. The nodes represent proteins and the links represent the existence of protein-protein or regulatory interactions. Edge thickness was proportional to the weight of the edge (assigned with respect to STRING score). Filled nodes indicate proteins involved in the interaction network, while empty nodes indicate proteins that are independent of the network. The two nodes indicated by red arrows (WNT5A and IRX3) represent the centered, highly connected proteins. Red nodes denote proteins reported to be involved in limb development, while blue nodes have not. Grey nodes are extended additional nodes to restrict the number of direct interactions with input nodes to 10 in the current network. Green triangles indicate notable genes associated with composite phenotypes (Figure 2C; Table S3). See also Table S6 and Figure S5 and S6.

Table 2. Functional annotation for notable genes (subset)

Genes	① Syndrome & phenotype description when mutated, in human (OMIM number) or ② mouse ③ Expression site in embryonic limb (mouse, unless stated)	Genes	① Syndrome & phenotype description when mutated, in human (OMIM number) or ② mouse ③ Expression site in embryonic limb (mouse, unless stated)
1p34.1-*RNF220*	③ Anterior limb bud(Ma et al., 2019)	6q21-*PRDM1*	② Mutants lack posterior digits, whisker development; ③ Limb bud posterior mesenchyme, including zone of polarizing activity (Robertson et al., 2007)
1p31.3-*NFIA*	① Brain malformations with or without urinary tract defects (613735); ③ Distal limb bud (Chaudhry et al., 1997)	7p21.2-*MEOX2*	② Mutants have selectively reduced limb musculature; ③ Limb bud myoblasts (Mankoo et al., 1999)
[a]2p24.3-*MYCN*	① Feingold syndrome; syndactyly and reduced middle phalanges (164280); ③ Proliferating cells of distal limb bud mesenchyme (Ota et al., 2007)	7p14.3-*CREB5*	③ Interdigital condensing mesenchyme (Lehoczky et al., 2004)
2p16.1-*BCL11A*	① Dias-Logan syndrome; Intellectual development disorder with persistent fetal hemoglobin (61711); ③ Mesenchyme of early limb bud then autopod, anterior and posterior margin of proximal limb, subsequently interdigital(Yamamoto et al., 2019)	7p12.3-*IGFBP3*	③ Interdigital region (van Kleffens et al., 1998)
[b]2q13-*EDAR*	① Hypohidrotic ectodermal dysplasia; skin appendages absent, aberrant dermatoglyphs (224900)	8q21.13-*ZFHX4*	③ Conserved human enhancer drives reporter expression in mouse limb buds (Ali et al., 2016)
2q31.1-*SP9*	③ Apical ectodermal ridge, regulated by FGF10 (Kawakami et al., 2004)	[a]8q23.1-*RSPO2*	① Tetraamelia syndrome 2; absence of limbs (618021); ③ Apical ectodermal ridge (Szenker-Ravi et al., 2018)

[a]2q31.1-*MAP3K20/ZAK*	① Split-foot malformation; mesoaxial polydactyly, nail duplications (616890); ③ Entire early limb bud (Spielmann et al., 2016)	8q24.12-*HAS2*	② Short limbs with phalange duplication and misplaced interphalangeal joints; ③ Distal limb bud mesenchyme, regulated by SHH (Liu et al., 2013)
[a]2q33.1-*SATB2*	① Glass syndrome; digit anomalies, sparse hair (612313); ③ Apical ectodermal ridge (Sheehan-Rooney et al., 2010)	[a]9q22.31-*ROR2*	① Robinow syndrome; limb shortening, including brachydactyly (268310); ③ Distal limb mesenchyme (Matsuda et al., 2001)
2q37.3-*GPC1*	③ Limb bud mesenchyme (chicken) (Saad et al., 2017)	[a]12q12-*DBX2*	① Deletion at this locus alters hand size, digit morphology and causes retention of fetal digital pads (Carlsen et al., 2015); ③ Genes at this locus are coordinately expressed in distal limb mesenchyme and embryonic digits(Beccari et al., 2021)
3p22.3-*TRIM71 (LIN41)*	③ Distal limb and digit mesenchyme (mouse and chicken) (Lancman et al., 2005)	[a]12q12-*NELL2*	
[a]3p14.3-*WNT5A*	① Robinow syndrome; limb shortening, brachydactyly (180700); ③ Apical ectodermal ridge and progress zone (Yamaguchi et al., 1999)	[a]12q24.21-*TBX3*	① Ulnar-mammary syndrome; posterior digits reduced or absent (181450); ③ Anterior and posterior mesenchyme of embryonic limb bud and apical ectodermal ridge (Gibson-Brown et al., 1996)
3p14.1-*ADAMTS9*	② Limb specific deletion causes syndactyl; ③ Broadly in early limb mesenchyme, subsequently digit perichondrium (McCulloch et al., 2009)	[a]13q12.11-*FGF9*	① Multiple Synostoses Syndrome 3; impaired interphalange joint formation, broadened thumbs (612961); ③ Apical ectodermal ridge (Mariani et al., 2008)
[a]3q26.2-*EVI1*	① RUSAT2; radioulnar synostosis, digit defects (616738); ③ Limb bud and digit pad mesenchyme (Human, Figure 3G-I)	[a]16q12.2-*RPGRIP1L*	① Meckel syndrome; polydactyly (611561); ③ Protein located at primary cilium of embryonic limb mesenchymal cells; regulates SHH signaling (Gerhardt et al., 2015)
3q27.1-*EPHB3*	③ Limb bud epithelium and nonchondrogenic mesenchyme (Compagni et al., 2003)	[c]16q12.2-*FTO*	① Growth retardation, developmental delay and facial dysmorphism; brachydactyly and cutis marmorata (612938)
4p15.31-*SLIT2*	③ Interdigital mesenchyme and digit lateral margins (Holmes et al., 1998)	16q12.2-*IRX3*	② Smaller limb when mutation combined with Irx5 mutation; ③ Proximal-anterior limb bud, interacts with SHH (Li et al., 2014)
[a]4p15.2-*RBPJ*	① Adams-Oliver syndrome; scalp and distal limb defects (short distal phalanges) (614814)	[b]18q12.1-*MAPRE2*	① Congenital symmetric circumferential skin creases; excess skin leading to ringed creases, principally on limbs (616734)
[a]4q28.1-*INTU*	① Short-rib thoracic dysplasia 20 with polydactyly (617925)	[a]18q21.2-*TCF4*	① Pitt-Hopkins syndrome; persistent fetal digital pads (610954); ③ Peridigital mesenchyme at distal digit tips (Cho and Dressler, 1998)
5q23.3-*ADAMTS19*	③ Posterior-proximal limb bud, regulated by SHH signalling (Lewandowski et al., 2015)	18q23-*SALL3*	② Lack of digit development when mutated with SALL1; ③ Distal posterior mesenchyme, regulated by SHH (Kawakami et al., 2009)
[a]6p24.3-*TFAP2A*	① Branchiooculofacial syndrome, incompletely penetrant polydactyly (113620); ③ Limb bud ectoderm and distal mesenchyme (Feng et al., 2008)	22q13.31-*CELSR1*	③ Early limb bud with distal bias (Shima et al., 2002)

[a]Abnormalities on limb phenotype when gene mutated in human.
[b]Abnormalities on skin or skin appendage phenotypes when gene mutated in human.
[c]Abnormalities on both skin and limb phenotypes when gene mutated in human.

5. Fingerprint patterns are significantly associated with hand proportions, showing strong genetic correlations with fingerprint pattern type

To further test the hypothesis that embryonic and fetal limb development may influence fingerprint pattern types, we examined the correlation between fingerprint patterns and limb-related phenotypes (i.e. hand phenotypes). We measured hand phenotypes (see Figure 5A for hand and digit length; Figure S7 for distal phalanx length and other hand phenotypes) in NSPT and JD cohorts. We found broad associations between hand phenotypes and fingerprint patterns (e.g. frequency of whorls) (Table S7). Higher frequency of whorl patterns was associated with longer little finger (D5) relative to hand length (ratio of D5 length to hand length (DHR5); β=0.15, P=2.42×10^{-5}, Figure 5B;

Table S7). This was clearly seen in the direct comparison of DHR5 in individuals with whorl and non-whorl fingerprint patterns on different digits, particularly on D5 (P_{D5L}=1.74×10^{-13}, P_{D5R}=5.85×10^{-13}, Figure 5C). Individuals with whorl pattern on D5 of both hands (56.83±4.34 mm) had on average a 1.32 mm longer D5 than individuals with no whorl pattern on D5 (55.51±4.52 mm). These phenotypic correlations also showed a strong genetic basis, as significant genetic correlations were found between DHR5 and fingerprint patterns on different digits, especially on D5 ($r_{g\text{-}D5L}$=0.31, P=3.89×10^{-16}, $r_{g\text{-}D5R}$=0.22, P=1.26×10^{-11}, Figure 5D; Table S8). Indeed, we found that fingerprint-associated signals (2p24.3-*MYCN*, 3p14.3-*WNT5A*, 12q24.21-*TBX3* and 18q23-*SALL3*) were also associated with the length of D5 or DHR5 (Table S6). Apart from the digit length, we found that the distal phalanx of D2 and D3 (ratio of D2 or D3 distal phalanx length to hand length (DPHR2 or DPHR3)) were also associated with whorl patterns (β_{DPHR2}=-0.31, P=9.13×10^{-18}, β_{DPHR3}=-0.24, P=2.02×10^{-11}; Table S7), and with fingerprint-associated signals (2q31.1-*SP9*, 8q23.1-*RSPO2*, etc.; Table S6). Together, these results provided further evidence that fingerprint patterns are strongly influenced by the process of limb development.

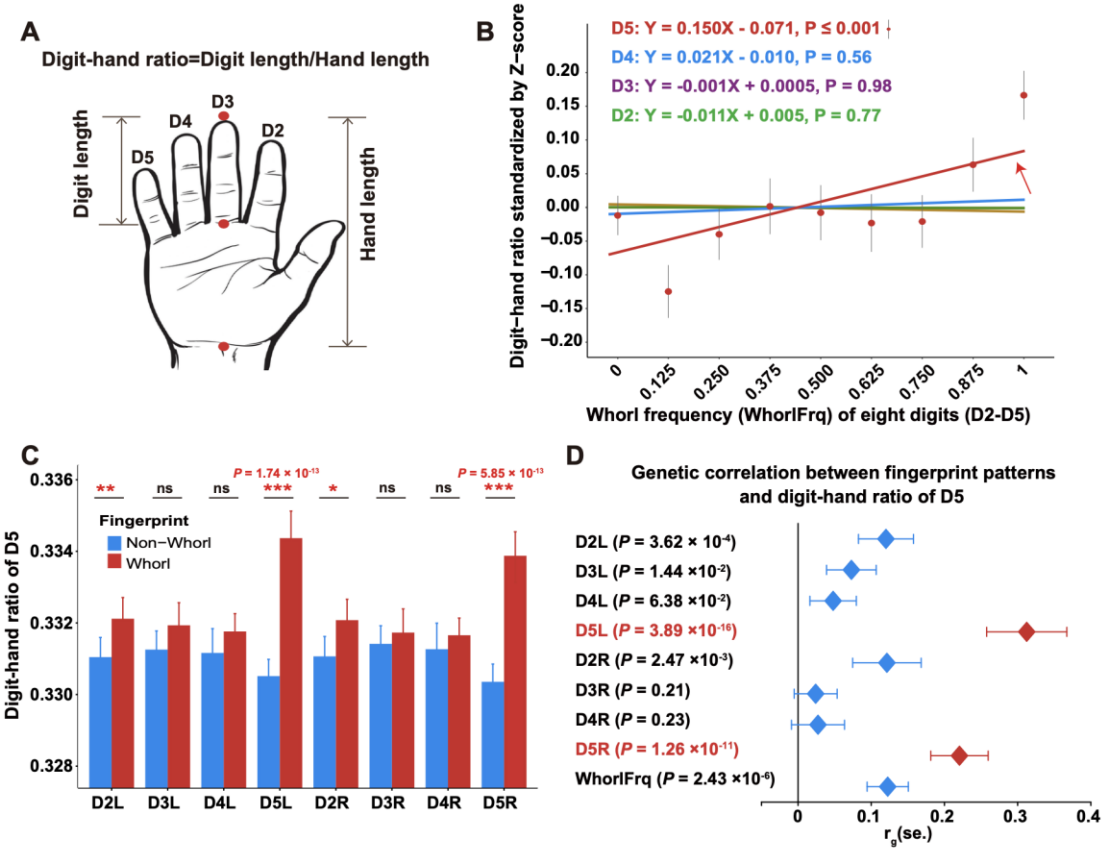

Figure 5. Association between fingerprint patterns and hand phenotypes (N=6,318)

(A) Diagrammed human hand with measured phenotypes, including hand and digit length. The digit-hand ratio (DHR) is the ratio of digit length and hand length.

(B) The association between the whorl frequency of eight digits (D2-D5) and the DHR of each digit. We used Z-score to standardize the mean DHR of left and right hands. Red dots indicate the average values and short black lines the standard deviation for each group. The arrow indicates the linear

regression passes the significance test.
(C) Bar plot of fingerprint patterns of each digit (D2-D5) and the mean DHR of D5. Error bars indicate S.E.M. *P<0.05, **<0.01, ***<0.001.
(D) Genetic correlations between fingerprint patterns and the mean DHR of D5. Estimates and tests were performed using the bivariate GREML of GCTA software. Error bars indicate S.E.M. See also Figure S7 and Table S7 and S8.

Discussion

By leveraging a large-scale GWAS, we identified many fingerprint pattern-associated genetic variants and inferred their associated biological processes. As with other complex traits, the phenotypic variance explained by the 43 top genetic signals from meta-analysis using polygenic risk scores was small, and together with the inherent stochasticity of skin patterning processes (Painter et al., 2012), it is clear that this work will not permit prediction of fingerprint pattern from an individual's genotype. This work shows that the enormous diversity of fingerprint patterns is at a basic level influenced by the dynamics and shaping of the underlying limb structure.

There are various ways to quantify fingerprint phenotypes (see Table S4). Interestingly, we found that GWAS of the ordinal phenotype (0, 1, and 2 for arch, loop, and whorl, respectively) provided the most significant and the greatest number of signals, suggesting this ordinal phenotype might better reflect the underlying genetic mechanisms of pattern formation. It has been proposed that the morphology (i.e. height, shape, size) of volar pads (Bonnevie, 1924; Wertheim and Maceo, 2002) and the growth stresses on the pad surface (Kücken, 2007) play important roles to influence pattern types. In particular, while ridges forming on high volar pads typically conform to the whorl-type pattern, low volar pads produce arch-type patterns; and asymmetric and intermediate height volar pads often form loop-type patterns (Babler, 1987; Mulvihill and Smith, 1969). This proposed link between height of volar pads and type of fingerprint pattern is consistent with the empirical justification of the ordinal phenotype (regarding loop as an intermediate phenotype between arch and whorl) in our study.

Of the genes identified in the 43 meta-GWAS signals, only *EDAR* had previously been implicated in dermatoglyph formation, with loss of EDAR activity producing highly abnormal fingerprints as part of the rare condition hypohidrotic ectodermal dysplasia (Kargul et al., 2001; Verbov, 1970). We found that in *Edar* mutant mice, the transverse digital ridges are profoundly aberrant, though, as in humans, limb and digit structure appears normal (Figure S3C). EDAR signaling in the surface ectoderm is thus likely to influence dermatoglyph patterns directly, in a manner similar to this pathway's role in defining hair follicle spatial arrangement (Mou et al., 2006). Collectively, however, the set of genes implicated from our GWAS represents a strong signal of limb growth regulation, particularly of the distal limb, evidenced by their

developmental expression patterns and the anatomical phenotypes induced by their mutation in humans and mice.

Evi1 mutation in mice alters the transverse digital ridge pattern, establishing the EVI1 protein as a determinant of dermatoglyph pattern. During human development, we did not detect EVI1 expression in the epidermal ridges that form the dermatoglyph; rather, expression was prominent at an earlier stage in the distal ends of the digits and under the transiently raised pads at the fingertips on which the fingerprints later form. These volar pads, built on a mesenchymal core, have been postulated to be key determinants of dermatoglyph characteristics as their variable shapes in the fetus are reported to be correlated with different dermatoglyph types (Babler, 1991). EVI1 promotes cell proliferation during development (Hoyt et al., 1997), and its expression in the proliferative mesenchyme of the distal limb suggests that it may modulate the shape and size of these pads and distal digit elongation by altering cell production. Volar pads are present but undergoing regression across the period of fingerprint establishment, though mutation at fingerprint pattern associated genes *TCF4* or the *DBX2* locus are associated with retention of fetal pads into adulthood, indicating a role for these genes in pad growth and regression (Table 2). *EVI1* is also a proto-oncogene, with increased expression driving acute myeloid leukemia and pediatric mixed lineage leukemia through suppression of cellular differentiation and apoptosis in hematopoietic lineages (Glass et al., 2014). Fingerprint pattern type has been reported to be associated with incidence of leukemia (Menser and Purvis-Smith, 1969; Rathee, 2014; Rosner, 1969), with the functional *EVI1* variants we identify potentially explaining some of this phenotypic association. Further clinical studies are needed to formally test this potential pleiotropic effect of *EVI1*.

Proximal-distal limb growth is driven by proliferation of distal mesenchyme in response to signals from its overlying apical ectodermal ridge (AER) (Towers and Tickle, 2009). Apically produced WNT5A, and its receptor ROR2 (Mikels and Nusse, 2006), RSPO2, and the downstream regulator TCF4, play roles in β-catenin transcriptional function, which is finely balanced to achieve appropriate AER maintenance and limb outgrowth (Hill et al., 2006). Our network analysis of fingerprint pattern associated genes extends this canonical WNT/β-catenin cluster to the planar cell polarity pathway, arising from the dual roles of WNT5A and ROR2, and connecting to INTU and CELSR1 as cytoplasmic effector and plasma membrane cell polarity factors. Planar cell polarity is crucial for coordinating directed limb outgrowth (Gao and Yang, 2013) and in particular regulating the formation of distal skeletal elements, consistent with the brachydactyly caused by *ROR2* or *WNT5A* mutation in Robinow syndrome and polydactyly caused by *INTU* mutation (Table 2).

A second node of network connectivity centers on IRX3, expressed in the proximal limb bud and antagonistically regulated by SHH, a key morphogen imparting anterior-posterior

polarity to the limb (Li et al., 2014). SHH, emanating from the zone of polarizing activity (ZPA) located in the posterior autopod mesenchyme, coordinates growth and digit identity across the handplate, resulting in altered digit number when SHH signaling is modulated (Towers and Tickle, 2009). RPGRIP1L and ADAMTS9 regulate SHH signal reception, by contributing to the formation and functioning of cilia, which serve as a signal-receiving projection from the cell surface, with mutations in *RPGRIP1L* causing formation of extra digits (polydactyly) (Arts et al., 2007; Delous et al., 2007; Zeng et al., 2010), while *ADAMTS9* mutation causes fusion of adjacent digits (syndactyly) (Dubail et al., 2014; Nandadasa et al., 2019). SHH signaling regulates the expression of other fingerprint pattern associated genes, including the *HAS2* (Liu et al., 2013), *SALL3* (Kawakami et al., 2009) and *DBX2* (Pierani et al., 1999) genes, further implicating the embryonic limb's anterior-posterior specification system in fingerprint type determination.

The influence of both the anterior-posterior and the proximal-distal limb development systems on dermatoglyph type is also supported by our assessment of their correlation with hand and digit lengths. The proximal-distal length proportions of hand, digit and phalanx are established during the fetal period, at the stage of fingerprint determination (Hamrick, 2001; Rao et al., 2019). We found the strongest phenotypic associations between fingerprint type and the proportions of the anterior-most and posterior-most fingers, substantiating a role for the anterior-posterior system in their selection. Moreover, 3p21.1-*WNT5A* ($p=6.93\times10^{-5}$) and 18q23-*SALL3* ($p=5.07\times10^{-4}$) *loci, as well as 3q26.2-EVI1* ($p=7.62\times10^{-3}$) which was similar to the mouse finding (Figure S4A), were associated with both fingerprint type and length of digit 5 (Table S6). This is consistent with demonstrated roles for SALL3 as a key regulator of autopod development (Kawakami et al., 2009) and WNT5A promoting distal limb and digit outgrowth during embryogenesis (Yamaguchi et al., 1999). Unlike the length proportions of the mature hand, which reflect their configuration from the stage at which fingerprint patterns were being established (Rao et al., 2019), other embryonic structures, such as the volar pads, are transient and their variation is unlikely to be detectable in adult limb proportions. In conclusion, the limb development genes uncovered in this study and the correlation of hand proportions with dermatoglyph types demonstrate the key role of embryonic limb growth processes in defining the intricate surface patterning of the human fingerprint.

Limitations of the study

We performed functional analysis on the most prominent signal of all for the composite phenotype, adjacent to *EVI1*, finding that the best supported SNP at this locus lies within an enhancer, with its alternate alleles displaying different regulatory activities. We mapped enhancer activities from functional datasets using 12 epithelial or mesenchymal and cell types

similar to those of the developing limb (See STAR Methods). However, this work could be extended and supported by generating functional datasets for limb cells at various developmental stages. Since the entire region of the 3q26.2 signal covering all the associated SNPs appeared in the credible sets using various methods, this region (chr3:168448858-168948263) is in the same TAD as the *EVI1* gene (chr3:168600000-169520000) (Figure S1A), and *EVI1* is strongly supported as the most likely causal gene. For our reporter assays we used the HEK293T cell line, widely used in biology to assay enhancer activity and for other applications, and used here because mesenchymal cell lines from early limb formation do not exist. However, being embryonic cells that express developmental regulators, including many HOX genes, brings HEK293T cells closer to embryonic human limb mesenchyme than cell lines derived from adult or from other species.

Acknowledgments

We would like to thank the participants of the TZL, NSPT, JD, CKB, WeGene, ALSPAC, QIMR, and Pittsburgh study who consented to participate in research, and the whole TZL, NSPT, and JD teams, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We also thank F.W. for her help with Figure 1A. This project was funded by the following grants and contracts: National Key Research and Development Project 2018YFC0910403 (S.W.); Shanghai Municipal Science and Technology Major Project 2017SHZDZX01 (L.J., S.W.); Max Planck-CAS Paul Gerson Unna Independent Research Group Leadership Award to S.W.; National Natural Science Foundation of China Grant 31521003 (L.J.); China Postdoctoral Science Foundation Grant 2019M651351 (J.L.); National Science & Technology Basic Research Project 2015FY111700 (L.J.); CAMS Innovation Fund for Medical Sciences 2019-I2M-5-066 (L.J.); National Natural Science Foundation of China 81622003 (L.W.); National Natural Science Foundation of China Grant 81373082 (J.L.); The 111 Project B13016 (L.J.); National Natural Science Foundation of China Grant 31771325 (J.T.); Major Research Program of National Natural Science Foundation of China 91731303 (J.T.); The National Institute for Dental and Craniofacial Research (NIDCR): R01-DE016148 (M.L.M and S.M.W) and R21-DE016930 (M.L.M); BBSRC: BBS/E/D/10002071 and BB/T007788/1 (D.J.H.); MRC: G1100357 (R.A.A) and MR/N022556/1 (MRC Centre for Reproductive Health); European Commission, Marie Sklodowska-Curie Actions Individual fellowship 706429 (C.B.M); Australian Research Council A7960034, A79906588, A79801419, DP0212016, and DP0343921 (S.E.M); Australian National Health and Medical Research Council (NHMRC) 241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496739, 552485, and 552498 (S.E.M); S.E.M. is supported by NHMRC SRF 1103623. The UK Medical Research Council and the Wellcome Trust (grant reference 102215/2/13/2) and the University of Bristol; Medical Research Council MC_UU_12013/4 (D.M.E.); Medical Research Council MC_UU_12013/3 (N.J.T.);

Australian Research Council Future Fellowship FT130101709 (D.M.E.).

Author Contributions

S.W., D.J.H., and L.J. conceived the project and provided main resources. For the discovery GWAS: H.Z., J.L., J.T, Y.Y., S.Z., J.Z., H.Y., and W.Z. performed data and sample collections. H.Z., J.L., S.Z., and Y.X. contributed to generate the fingerprint and hand phenotype data. S.W., Y.L., and Q.P. contributed to generate the SNP array data. J.L. and M.P. performed computational analysis.
For the replication and meta-analysis:
J.L., R.M., R.G.W., K.L., and Z.C. performed data analysis and provided summary statistics results and relevant descriptions for CKB cohort; L.W., Y.L., S.T., X.W., and G.C. for WeGene cohort; D.M.E. and J.P.K. for ALSPAC; S.E.M., N.G.M., Yvonne Y. W. Ho, and D.Z.L. for QIMR; J.M.C., K.N., E.F., S.M.W., and M.L.M. for Pittsburg cohort. J.L. and M.P. performed the overall analysis.
For the functional experiments:
D.J.H., J.D.G., C.B.M., and M.C. performed the experiments on mouse models and D.J.H, J.D.G, A.M., R.A.A. and E.I.C. did the gene and protein expression. HB aided in data analysis from mouse models. D.H. and L.W. performed the luciferase reporter assay.
J.L., D.J.H., and S.W. wrote the manuscript with input from J.M.C., E.F., S.M.W., M.L. M., M.C., S.E.M. and other co-authors.

Declaration of Interests

L.W., Y.L. S.T., X.W. and G.C. are employees of WeGene Inc. The other authors declare no competing interests.

Figure Titles and Legends

Figure 1. Genetic variants associated with ordinal fingerprint patterns (arch, loop and whorl) in Han Chinese (N=9,909)

(A) Pattern-types of fingerprints according to the number of triradii/deltas (triangles) and cores (circles) (STAR methods). There are three main types: arch, loop and whorl. Each main group contains two sub-types according to the steepness, direction of ridges and the variable core.
(B) Genome-wide scans of the ordinal arch-loop-whorl phenotype identify 18 genomic regions associated with fingerprint patterns. The red line indicates the threshold for genome-wide significance after adjusting for the effective number of independent phenotypes ($P_{adj}<1.67\times10^{-8}$; STAR Methods). Detailed patterns of adjusted association significance across different fingers are indicated by black squares for corresponding digits for significant associations ($P_{adj}<1.67\times10^{-8}$), and grey squares for marginal associations ($P_{adj}<3.33\times10^{-6}$). Notable genes are indicated for each locus

(see Table 1 for selection criteria). Abbreviation: D1-5L/R = digit 1 to 5 of left or right hand. See also Table S2 and S4.

Figure 2. Genetic basis of the middle three digits "pattern-block" phenomenon, with top signal near *EVI1*

(A) "Pattern block" of the middle three digits on both hands revealed by pair-wise phenotypic correlation (blue) and genetic correlation (red) among the ten digits (N=9,909). The dashed box indicates high correlations between the same digits of both hands and neighboring digits. The correlations from high to low were represented by both color and correlation coefficients (r) in the figure.

(B) The correlations of fingerprint patterns between the middle three digits on both hands (Pattern-block pairs) are higher than the correlations of all random pairs of the ten digits (All pairs).

(C) Genome-wide scan on the composite phenotype extracted from the fingerprint pattern of the middle three digits on both hands. The loading coefficients of the composite phenotype on the six correlated variables are between 0.719 and 0.792.

(D) Fine mapping of signals at 3q26.2: LocusZoom plot of SNPs at the 3q26.2 region (top) and mapping of epigenetic marks H3K4me1, H3K27ac, DNase hypersensitivity and conservation analysis at the same region, based on ENCODE and RMEC project data. SNP rs7646897 and rs7623083, indicated by red lines, are in a region that exhibits distinct active enhancer signatures defined by epigenetic marks, such as H3K4me1 (green), H3K27ac (blue) histone modifications and DNase hypersensitivity (purple), and with enhancer function by chromatin state assay (yellow box) in fibroblast primary cells and in chondrocytes. The phastCon score indicates the evolutionarily conservation in primates.

(E) Luciferase reporter assays on candidate regulatory elements carrying alternate alleles at SNPs rs7646897 and (F) rs7623083 in HEK293T cells. pGL3-basic is a negative control plasmid lacking enhancer activity, pGL3-318 is a positive control derived from the *EVI1* promoter region. Symbols indicate significance in t test (*P<0.05, **<0.01, ***<0.001). See also Figure S1, Figure S2 and Table S3.

Figure 3. EVI1 in dermatoglyph patterning and limb development. A) Palmar dermal surface of toluidine blue stained paws from wild type and $Evi1^{Jbo/+}$ adult mice showing dermatoglyph arrangement. Arrow indicates spur on the mutant digit 5 (D5). B) Schematic depicting transverse ridge categories on mouse digits and ventral surface of D4 of wild type and $Evi1^{Jbo/+}$. Regions carrying continuous: C, discontinuous: D, and incomplete: I ridges are indicated. C) Quantification of digit ridge pattern in wild type and $Evi1^{Jbo/+}$ mutants. Continuous ridges are reduced on all mutant digits, while D3 and D4 carry more discontinuous ridges. D) Wholemount *in situ* hybridization detecting *Evi1* expression in mouse embryonic forelimbs. Ventral view. E) RNAscope *in situ* hybridization detecting *Evi1* and the limb mesenchyme marker *Prrx1* transcripts in mouse

embryonic limb and digits between E11.5 and E17.5. F) Quantitative RT-PCR determination of *Evi1* expression in mouse forelimb at E11.5 (whole limb bud), E13.5, E15.5 and E17.5 (autopod only). (G-J) Immunofluorescence detecting EVI1 expression in human embryonic tissue. G) Transverse section of CS17 embryo (~6 week EGA) shows nuclear expression in mesenchymal cells of the limb bud (LB, magnified in lower panel). The neural tube (NT) indicates the dorsal midline. H) Longitudinal section of 10 week EGA digit, arrow indicates the raised volar pad across which fingerprints form. I) 13 week EGA digit and J) 16 week EGA digit detecting EVI1 and epithelial marker K14. Dotted line indicates dermal-epidermal junction. SG: eccrine sweat gland. K, L) RNAscope in situ hybridization detecting *EVI1* and *PRRX1* transcripts in sectioned K) 10 week EGA and L) 16 week EGA human digit, with α-catenin immunofluorescence. Individual cells co-express *EVI1* and *PRRX1*. Asterisks indicates autofluorescent blood cells. M) Detection of proliferative cell marker Ki67 in 10 week EGA digit. Dorsal (D) and ventral (V) axes are annotated: Nuclei are stained with DAPI. Scale bars: A=1 mm; D, G upper = 500 μm; E, G lower, H, I, J, K left, M=100 μm; K right, L=20 μm. Error bars indicate S.E.M. See also Figure S3, Figure S4 and Table S5.

Figure 4. A meta-analysis of fingerprint patterns showing signals enriched in limb development

(A) A Manhattan plot showing the results of the meta-analyses combining GWAS of East Asian (EAS)-ancestry (TZL, NSPT, JD, CKB and WeGene) and European (EUR)-ancestry cohorts (ALSPAC, QIMR and Pittsburgh) across all ten digits (D1L/R were unavailable in JD and ALSPAC). There were 43 signals associated with fingerprint patterns of at least one digit ($P_{adj}<1.67\times10^{-8}$; Table S6), with gene names in different colors: purple indicating significant in both EAS and EUR; red and blue indicating only significant in EAS and EUR, respectively; green indicating not significant in either EAS or EUR, but only significant after the meta-analysis combining both. Bold genes showed associations with limb phenotypes abnormalities (Table 2). The block map on the right represented the digits corresponding to the signals on the left. Red and blue triangles indicate significance in EAS and EUR, respectively, while dark and light colors represented signals that reached the adjusted genome-wide significant ($P_{adj}<1.67\times10^{-8}$) and suggestive levels ($P_{adj}<3.33\times10^{-6}$), respectively. Bold frame indicated genome-wide significant (black) or suggestive (gray) significant after combined meta-analyses.

(B) Venn diagram summarizing the fingerprint-associated signals corresponding to Figure 4A.

(C) Enrichment of annotations across ontologies for fingerprint-associated 43 signals. The red asterisk indicates limb-relevant terms that genes are significantly enriched in after Bonferroni correction (the red dotted lines). Only the top 10 terms ranked after enrichment analysis and top 5 epithelial/skin-related terms are shown.

(D) Fingerprint pattern-associated proteins and their interactions. The nodes represent proteins and the links represent the existence of protein-protein or regulatory interactions. Edge thickness

was proportional to the weight of the edge (assigned with respect to STRING score). Filled nodes indicate proteins involved in the interaction network, while empty nodes indicate proteins that were independent of the network. The two nodes indicated by red arrows (WNT5A and IRX3) represent the centered, highly connected proteins. Red nodes denote proteins reported to be involved in limb development, while blue nodes have not. Grey nodes are extended additional nodes to restrict the number of direct interactions with input nodes to 10 in the current network. Green triangles indicate notable genes associated with composite phenotypes (Figure 2C; Table S3). See also Table S6 and Figure S4 and S5.

Figure 5. Association between fingerprint patterns and hand phenotypes (N=6,318)

(A) Diagrammed human hand with measured phenotypes, including hand and digit length. The digit-hand ratio (DHR) is the ratio of digit length and hand length.

(B) The association between the whorl frequency of eight digits (D2-D5) and the DHR of each digit. We used Z-score to standardize the mean DHR of left and right hands. Red dots indicate the average values and short black lines the standard deviation for each group.

(C) Bar plot of fingerprint patterns of each digit (D2-D5) and the mean DHR of D5. Error bars indicate S.E.M. *P<0.05, **<0.01, ***<0.001.

(D) Genetic correlations between fingerprint patterns and the mean DHR of D5. Estimates and tests were performed using the bivariate GREML of GCTA software. Error bars indicate S.E.M. See also Figure S7 and Table S7 and S8.

Tables with Titles and Legends

Table 1. GWAS signals for fingerprint pattern type in Han Chinese discovery and replication populations

[a]The effect or alternative (Alt) allele frequency of the discovery populations.

[b]Notable genes are indicated as follows: 1) the two nearest genes within 1000 kb of the most significantly associated SNP annotated by GREAT (G), which uses the subset of the UCSC Known Genes; 2) the nearest gene mapped by GENCODE (GE) or RefSeq (R); 3) Protein-coding genes within 1000 kb of the most significantly associated SNP in regional LocusZoom plot (L). Underlining indicates that the best-associated SNP is located within the gene.

[c]Genome-wide significant level (P_{adj}<1.67×10^{-8}, bold font) or suggestive level (P_{adj}<3.33×10^{-6}) after multiple-testing adjustment.

[d]The sample sizes vary in GWAS on different phenotypes of digit: N=5415-9909 for discovery cohort (fingerprint patterns on digit 1 are not available in JD cohort), N=1634-1785 for replication cohort 1 (CKB), and N=2138-2152 for replication cohort 2 (WeGene).

[e]The associations between the top SNPs and the fingerprint pattern of the most significant digit (i.e. top digit, as indicated in parentheses).

[f] The signal was replicated at different levels of association, as following: the most significant

replication is exactly the association between the top SNP and the top digit (L1); the most significant replication is the association between the top SNP and one of the other associated digits, while the association between the top SNP and the top digit is also significant (L2) or not significant (L3). "--" not available (INDEL polymorphisms are not available in the WeGene cohort).

[g]The associations between the Top SNPs and Top digits in replication cohorts. The "/" indicate that the associations have the same effect size and p value as the results of the two columns in front.

Abbreviation: Alt=alternative, ref=reference, EAF=Effect or Alt Allele Frequency, TZL=cohort from Taizhou Longitudinal Study, NSPT=cohort from National Survey of Physical Traits Project, JD=cohort from Jidong of Hebei Province, CKB=cohort from China Kadorie Biobank, WeGene=cohort from WeGene company. D1-5L/R=digit 1 to 5 of left or right hand.

Table 2. Functional annotation for notable genes (subset)

[a]Abnormalities on limb phenotype when gene mutated in human.

[b]Abnormalities on skin or skin appendage phenotypes when gene mutated in human.

[c]Abnormalities on both skin and limb phenotypes when gene mutated in human.

STAR★Methods

Key Resources Table

Resource Availability

Lead Contact

Further information and requests for resources should be directed to Sijia Wang (wangsijia@picb.ac.cn), Li Jin (lijin@fudan.edu.cn) and Denis Headon (denis.headon@roslin.ed.ac.uk).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

All raw chip array data of discovery cohorts generated during this study are available at NODE (http://www.biosino.org/node). The accession number for the data reported in this paper is NODE: OEP000198. This study did not generate any unique code. All software is freely or commercially available and is listed in the STAR Methods description and Key Resources Table.

Experimental Model and Subject Details

Human subjects

There are in total 23,966 samples from 8 independent cohorts (Table S1). In discovery stage,

the Taizhou longitudinal (TZL) cohort includes 2961 adults (1059 males and 1902 females, aged 31-81 years); the National Survey of Physical Traits (NSPT) cohort comprises 2679 individuals from three different regions of China (1045 males and 1634 females, aged 18-83 years); the Jidong (JD) cohort includes 4269 adults (2104 males and 2165 females, aged 20-82 years). Two cohorts for replication: the Chinese Kadoorie biobank (CKB) includes 1785 adults (596 males and 1189 females, aged 36-55 years), and the WeGene cohort comprises 2152 unrelated individuals (954 males and 1198 females, aged 6-68 years). In trans-ethnic meta-analyses, the Pittsburgh cohort includes 1480 volunteers (690 males and 790 females, aged 0-86 years), the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort includes 5339 individuals, and the Queensland Institute of Medical Research (QIMR) cohort includes 3301 individuals. All participants provided written informed consent, and all study protocols were approved by the institutional review boards of the pertinent research institutions. For detailed information about study populations, donor enrollment, blood extraction, specification of fingerprint patterns and hand traits review the Methods Details.

Cell lines and cultures

Luciferase reporter plasmids were transfected into HEK293T cells. HEK293T cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% Penicillin-Streptomycin (Gibco) and incubated at 37℃ in 5% CO2. Plasmids construction and cell culture are indicated in the Methods Details.

Mice

FVB, *Tabby* (Eda^{Ta}) and *downless* mutants ($Edar^{dlJ/dlJ}$ and $Edar^{dlJ/+}$) mice were bred at the Roslin Institute. *Junbo* mice were bred in both the Roslin Institute and MRC, Harwell centres. *Junbo* mice are congenic on a C3H/HeH genetic background (European Mouse Mutant Archive; EM: 00091) and maintained by crossing *Junbo* heterozygote males with C3H/HeH females. They bear a mutation in the transcription factor *Evi1/Mecom* and are therefore referred as $Evi1^{Jbo}$. This mouse model with altered EVI1 function through amino acid substitution ensures that Evi1 itself is the modified gene. *downless* mice (*Edar* p.E379K) were maintained by intercrossing $Edar^{dlJ/+}$ mice. Homozygous $Edar^{dlJ/dlJ}$ represents a loss of function mutation, with mice having a sparse hair coat and hairless tail, whereas $Edar^{dlJ/+}$ mice have an appearance identical to wild-type. Both *Eda* and *Edar* mutant strains were maintained on the FVB genetic background. The effect of *Evi1* mutation was studied at the age of P21 (postnatal 21 days) and for other mouse samples the age is specified. 16 $Evi1^{Jbo}$ and 10 wild type mice were used (Table S5). Mice were killed by cervical dislocation and DNA from tail tips were used for determining the genotypes. Forelimbs were dissected at the distal end of the zeugopod. $Evi1^{Jbo/+}$ mutants are characterized by the presence of an extra spur on digit 5, present in either left, right or both the forelimbs. The study was performed under UK Home Office license and approved by the Roslin Institute Animal Welfare and Ethical Review Body. Mouse embryonic tissue for in situ hybridization and qRT-PCR was obtained by timed mating of FVB/N male and female mice. Noon on the day of copulation plug detection denoted embryonic day 0.5.

Fetal tissue collection

Fetal tissue samples, no gender requirement, used in immunofluorescence were obtained after elective medical termination of pregnancy from the Royal Infirmary of Edinburgh, UK with informed consent (approved by the Lothian Research Ethics Committee, Ref: 08/S1101/1). All were morphologically normal and gestational age was determined according to Carnegie Stages (CS) of human development for embryos <10 weeks or by ultrasound >10 weeks gestation.

Methods Details

Ethics statement

All participants provided written informed consent, and all study protocols were approved by the institutional review boards of the pertinent research institutions. The Taizhou Longitudinal Study (TZL) was approved by the Ethics Committee of Human Genetic Resources at the Shanghai institute of life Sciences, Chinese Academy of Sciences (ER-SIBS-261410). The Jidong cohort (JD) was approved by the Ethics Committee of Human Genetic Resources at the Shanghai Institute of Life Sciences, Chinese Academy of Sciences (ER-SIBS-261410-A1801). The National Survey of Physical Traits (NSPT) is the sub project of The National Science & Technology Basic Research Project which was approved by the Ethics Committee of Human Genetic Resources of School of Life Sciences, Fudan University, Shanghai (14117). The CKB ethics approval was obtained from the Institutional Review Board (IRB) at the Peking University (IRB00001052-13055). Participants of WeGene cohort provided informed consent and participated in the research online, under a protocol approved by the Ethical Committee of WeGene. The ethical approval of ALSPAC birth cohort study was obtained from the ALSPAC Ethics and Law committee and the Local Research Ethics Committees. The QIMR study has been approved by QIMR Berghofer Human Research Ethics Committee (P193 & P455). The Pittsburgh cohort ethics approval was obtained locally at each recruitment site, including the University of Pittsburgh, which served as the coordinating center for this project (IRB0405013). Written informed consent was granted for each participant before enrollment in the study. We confirm that our study is compliant with the Guidance of the Ministry of Science and Technology (MOST) for the Review and Approval of Human Genetic Resources. Ethical approval for analysis of human fetal tissue was obtained from Lothian Research Ethics Committee (study code LREC 08/S1101/1), with informed written consent.

Study population and design

The overall design in current study is shown in Figure S7. This study is based on data from eight independent cohorts (Table S1). Study participants in discovery stage were from three cohorts: 1) Taizhou, Jiangsu Province, as part of the Taizhou Longitudinal Study (TZL)(Wang et al., 2009). In total, 2961 Han Chinese individuals (including 1059 males and 1902 females) who were aged 31-81 years were enrolled in 2014; 2) four sub-cohorts collected from three different regions of

China in different years: 15HanTZ, 17HanZZ, 18HanNN, and 19HanTZ. These four sub-cohorts from The National Survey of Physical Traits (NSPT) were part of the National Science & Technology Basic Research Project. Totally the NSPT cohort consisted of 1045 males and 1634 females, aged 18-83 years; 3) 2104 males and 2165 females, aged 20-82 years from Jidong of Hebei Province (JD). The summary statistics of fingerprint patterns for replication analysis were from two other independent Chinese Han cohorts consisting of: 1) 596 males and 1189 females, aged 36-55 years, as part of the China Kadoorie Biobank (CKB); 2) WeGene cohort of 2152 unrelated participants (954 males and 1198 females) who received the Personal Genome Service of WeGene, aged 6-68 recruited primarily online including organic posts on WeGene's Wechat social media channels and website. The GWAS summary statistics of three European-ancestry populations (ALSPAC, QIMR and Pittsburgh cohorts) also were included in the large-scale meta-analysis stage. Data from the ALSPAC and QIMR studies have been described in detail in a previous study (Ho et al., 2016). The Pittsburgh cohort comprised 1480 participants (690 males and 790 females) who were age 0-86 years. These individuals were recruited from a number of international sites as part of the larger Pittsburgh Orofacial Cleft Study (Weinberg et al., 2006), a collaborative effort to investigate the genetics of orofacial clefts. Participants include individuals with orofacial clefts, their unaffected relatives, and individuals from control families without a history of clefting. Recruitment of the participants in this study took place in the United States (N = 630), Hungary (N = 678), Spain (N = 117), and Argentina (N = 55).

Specification of fingerprint patterns and hand traits

Fingerprints were collected using rolled ink prints on paper, or using an electronic fingerprint scanner in TZL (Greenbit DactyScan26) and NSPT (DactyScan40i) cohorts. The images of the palmar surface of the hands were collected using an electronic scanner (EPSONScanV370) in NSPT and JD cohorts. As the full patterns of the thumbs were not clearly visible for JD samples, we excluded D1L and D1R digits from analyses in JD study. Fingerprint patterns were visually categorized by two investigators according to the number of triradii/delta (triangles) and core (circles) (Cummins, 1969; Holt and Penrose, 1968): a) Arch pattern which has 0 triradii and 0 core with gentle (Simple Arch, As) or steep (Tented Arch, At) lines; b) Loop pattern which has 1 triradius and 1 core. Its ridge opens away from the triradius towards the radial or ulnar side and are sub classified as either Radial Loop (Lr) or Ulnar Loop (Lu), respectively; c) Whorl pattern which has 2 triradii and 1 (Simple Whorl, Ws) or 2 (Double Whorl, Wd) cores because of the highly variable inside. For highly correlated phenotypes, we extracted their composite phenotype by a partial least square path model using the "plspm" package in R. The bootstrap confidence interval test (bootstrap resampling times = 1000, significance level = 0.05) was applied to test the significance of each path coefficient. Further 66 derived phenotypes (e.g. binary, nominal, ordinal, and quantitative phenotypes) review Table S4. Hand traits were measured by manually calibrating feature points on palmar images which were collected using electronic scanner (EPSONScanV370) in NSPT and JD

cohorts. The digit length, distal phalanx length, palm length and palm width were calculated by coordinates of landmarks using MATLAB_R2019a. Further derived ratio phenotypes (e.g., the ratio of digit length to hand length (DHR), the ratio of distal phalanx to hand length (DPHR)) are detailed in Table S8.

In CKB cohort, fingerprints were collected using an electronic fingerprint scanner (Hanlintongxin PU-JY500U) and categorized fingerprint patterns as both ordinal and binary phenotypes. Fingerprint patterns were obtained twice by manual work (two trained investigators). When two results were inconsistent, a third independent investigator would make final judgments. In WeGene cohort, the same judgments criteria were given to the participants who replied with their self-reported fingerprint patterns both ordinal and binary phenotypes. Within the ALSPAC sample (Boyd et al., 2013; Fraser et al., 2013), pattern type for each digit was scored from photocopies of the palmar surface of the hands, which were collected for the purpose of measuring digit ratio (Medland et al., 2010). Pattern type was manually coded into arch, loop, and whorl. Arches were not analyzed in ALSPAC study. Thumbs were excluded from analyses because the full patterns were not clearly visible. After initial quality control analyses, 8 variables were included in the study: the presence of whorls across all digits, except D1L and D1R digits. Please note the study website contains details of data available through a fully searchable data dictionary, http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/). In QIMR cohort, fingerprints were collected using rolled ink prints on paper, and or using an electronic rolled fingerprint scanner (Smiths Heimann Biometrics ACCO1394) (Medland et al., 2007). The fingerprint patterns were then manually coded. Arches were also not included in analysis. Presence of whorls across D4 and D5 were unavailable for QIMR adult sample. So these four digits have smaller sample size than others. For the Pittsburgh cohort, fingerprints for all 10 digits were collected on paper using standard ink-based methods. Three raters independently classified each pattern as an arch, loop, or whorl. A few patterns (0.14%, from 157 people) could not be easily classified and were treated as missing data. In the discovery (TZL, NSPT, JD) and replication GWAS (CKB, WeGene), we considered fingerprint pattern as an ordinal phenotype according to the ordinal number of triradii. The same ordinal phenotype was previously used to estimate heritability (Arrieta et al., 1991; Reed et al., 1975). In the trans-ethnic meta-analyses, the binary phenotype according to the absence or presence of whorl type was used for comparability reason across the cohorts.

Genotyping, quality control, and imputation

Genotyping was performed in separately in the eight cohorts (Table S1). For TZL cohort, DNA was extracted from peripheral blood samples using GENEray^{TM} DNA extraction kit and genotyped along with HapMap Phase 1-3 and 1000 Genome Project samples for 776,213 SNPs on the Illumina HumanOmniZhongHua-8 chip. Genetic data cleaning and quality control was done using PLINK 1.9 (Purcell et al., 2007). In brief, samples were interrogated for sex, chromosomal aberrations, relatedness, and genotype call rate (>5% excluded). SNPs were interrogated for call rate (>5%

excluded), discordance, Mendelian errors, deviations from Hardy-Weinberg equilibrium (HWE; $P<1\times10^{-5}$) and differences in minor allele frequencies (MAF<1% excluded) and heterozygosity. The chip genotype data were firstly phased using SHAPEIT (Delaneau et al., 2011). IMPUTE2 (Howie et al., 2009) was then used to impute unobserved variants using the 1000 Genomes Project (1KGP) Phase 3 (International HapMap et al., 2010) as the reference. SNPs with an imputation quality scores (INFO) less than 0.8, MAF less than 1% or a missing rate of more than 2% of genotypes were eliminated from further analyses. Finally, a total of 7,057,720 SNPs from 2961 individuals were passed quality control and were used for further analyses.

For NSPT and JD 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 that analyzes over 710,000 SNPs. It is a fully custom array designed by WeGene (https://www.wegene.com/). Genetic data cleaning and quality control was done using PLINK v1.9 (Purcell et al., 2007). We excluded subjects with more than 5% missing data, duplicated subjects, and subject samples that failed the X-chromosome gender concordance check. We excluded SNPs that had >2% missing data, MAF <1%, or a deviation from Hardy-Weinberg equilibrium ($P<1\times10^{-5}$), leaving 707,146 SNPs from 4269 individuals in JD cohort and 2679 individuals in NSPT cohort for further analyses. Imputation of unobserved variants was performed using haplotypes from the 1000 Genomes Project Phase 3 as the reference. The chip genotype data was firstly phased using SHAPEIT. IMPUTE2 was then used to impute genotypes. SNPs with an imputation quality scores (INFO) less than 0.6, MAF <1% or a missing rate >1% of genotypes were eliminated from further analyses. Finally, a total of 8,039,700 SNPs were passed quality control and were used for further analyses.

The genotyping of the CKB cohort was performed using custom-designed 800K-SNP Affymetrix Axiom arrays (Axiom_CKB_1 and Axiom_CKB_2). Genetic data cleaning and QC procedures utilised PLINK v1.9. Subjects of non-East Asian ancestry, mismatch with reported gender, with genotype heterozygosity > 3×SDs from the mean and X/Y aneuploidy were excluded. The selected SNPs are of > 98% call rate and with less than 20% allele frequency difference from the 1KGP East Asian population. SNPs with HWE $P< 1\times10^{-6}$ were manually examined. The hard-genotyped data were prephased using SHAPEIT3 and imputed using IMPUTE4. SNPs with 0 MAF in the 1KGP East Asian population were ignored.

For WeGene cohort, DNA extraction and genotyping were performed on saliva samples. Genotyping was performed on the Affymetrix WeGene V1 Arrays covering 596,744 SNPs at the WeGene genotyping centre, Shenzhen. Quality control (QC) was performed in PLINK V1.07. Ancestry was assigned using self-reported ancestry obtained from customer surveys and further checked with principal components analysis. Unrelated filtering was done by checking pair wisely for all the samples and where identity by descent (IBD) scores of >0.125 (3rd-degree relative) were identified with one from each such pair removed. Individuals with discordant sex information were removed. The individuals with genotype call rate of <95% and outlying heterozygous rate were

excluded. To minimize the influence of bias, the following SNP were discarded: (1) sites with unbalanced call rate in case and controls, (2) sites that failed the Hardy-Weinberg Equilibrium test (P <1×10^{-5}), and (3) MAF <1%. The 1000 Genomes Project Phase 3 data was used as imputation reference panel. Phasing and imputation on the autosomes were carried out using SHAPEIT2 and IMPUTE2. Post-imputation filtering were done by removing SNPs with imputation quality scores (INFO) less than 0.5, MAF less than 1% or missing rate more than 2%. The combined application of these filters left us with a data set of 7,124,171 SNPs and these data set was used for further analyses.

ALSPAC participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by the Wellcome Trust Sanger Institute, Cambridge, UK and the Laboratory Corporation of America, Burlington, NC, US. Genotype data were cleaned using standard thresholds (SNPs excluded if MAF< 1%, call rate <95% and P-value from an exact test of Hardy-Weinberg equilibrium <5×10^{-7}). Individual samples were excluded on the basis of incorrect sex assignment, minimal or excessive heterozygosity, and high levels of missingness or cryptic relatedness. We combined child's genotypes with cleaned genome-wide SNP data from 9,048 ALSPAC mothers (Fatemifar et al., 2013) and removed subjects due to potential sample mismatches. ALSPAC samples were imputed using the Hapmap2 r22.36 CEU reference. Single nucleotide polymorphisms (SNPs) that had a MAF >0.01 and could be imputed with confidence (R^2 > 0.3) were used in these analyses.

Participants in the QIMR cohort were genotyped on the Illumina Human610-Quad SNP chip. These samples were genotyped in the context of a larger genome-wide association project that resulted in the genotyping of 28,028 individuals59 using the Illumina 317, 370, 610, 660, Core+Exome, PsychChip, Omni2.5 and OmniExpress SNP chips which included data from twins, their siblings and their parents. Genotype data were screened for genotyping quality (< 0.7), SNP and individual call rates (<0.95), HWE failure (P <1×10^{-6}) and MAF (< 0.01). As these samples were genotyped in the context of a larger project, the data were integrated with the larger QIMR genotype project and the data were checked for pedigree, sex and Mendelian errors and for non-European ancestry. QIMR samples were imputed using the Hapmap2 r22.36 CEU reference. Single nucleotide polymorphisms (SNPs) that had a minor allele frequency >0.01 and could be imputed with confidence (R^2 > 0.3) were used in these analyses.

For the Pittsburgh cohort, DNA samples primarily from blood or saliva were assayed for 557,577 SNP genotypes (including 15,890 SNPs of custom content) with the Illumina HumanCore+Exome platform at the Center for Inherited Disease Research. The University of Washington Genetics Coordinating Center performed cleaning and quality control analysis, which included investigation of genetic sex, chromosomal anomalies, relatedness, call rate, and batch effects. 455,449 SNPs initially passed filters that included call rate, deviation from Hardy-Weinberg equilibrium, difference by sex in allele frequency or heterozygosity, discordance in duplicate samples, and Mendelian errors in controls. Genotypes for 34.4 million additional SNPs were

imputed to the 1000 Genomes Project Phase 3 worldwide reference panel using IMPUTE2 after phasing with SHAPEIT. Imputed SNPs were filtered out for low info score (< 0.5), evidence of extreme deviation from HWE, and low MAF (<5%). Individual genotypes with low probability (< 0.9) were treated as missing. After filtering, about 6.7 million SNPs were available for GWAS.

Population stratification analysis

We corrected the effects of possible population stratification within cohort using EIGENSTRAT (Price et al., 2006) utility from the EIGENSOFT package. 102,284 SNPs in low linkage equilibrium (r^2<0.2) were selected for analysis. TZL cohort was compared with YRI, CHB, and CEU from 1000 Genomes Phase 3 (International HapMap et al., 2010) and principal component (PC) analysis did not find any outliers. Here, we also adjusted top 4 PCs to further avoid inflation from different genetic background. The similar criteria were used in NSPT and JD cohorts and adjusted top 5 PCs for further analyses and without any inflation.

PCs of CKB (and 1KGP) subjects were obtained using the PLINK implementation of the GCTA algorithm (Yang et al., 2011). After the removal of high-LD regions, 142,165 SNPs with low linkage equilibrium (r^2 <0.2) were selected for the PCA of CKB. We used the combination of CKB and 1KGP for the detection of CKB subjects of non-East Asian ancestry. Eight PCs were employed for the further association studies.

For WeGene cohort, we selected 308 unrelated samples from the YRI, CEU and CHB populations (1KGP Phase 3), and then chose SNPs using following criteria: (1) MAF$\geqslant$0.05 and HWE P>10^{-6}, in each of the populations YRI, CEU and CHB, (2) pairwise $r^2\leqslant$0.1 to exclude SNPs in high LD (calculated using PLINK indep-pairwise function with a step window of size 1000 bp), (3) remove C/G and A/T SNPs to avoid unresolvable strand mismatches. With the remaining 38,144 SNPs, we computed PCA for the combined samples. Ten PCs were employed in further association studies.

For Pittsburgh cohort, participants comprised a subset from a large, multiethnic study. PCs of ancestry were calculated for all participants in the larger study and then projected for the subset included here. Because there were no outlier, association studies were performed without adjusted genetic PCs.

As there was no evidence of systematic inflation in the ALSPAC dataset (λ = 1.007–1.034), results were not corrected for ancestry informative PCs.

For QIMR cohorts, EigenSoft (version 6.0.1) was used to perform principal component analysis. The QIMR data were combined with Genome-EUTWIN and HapMap Phase 3 populations. The EUTWIN and HapMap populations were used to produce the internal axes, and PCA coordinates were calculated on those axes for those populations and the QIMR participants. Individuals more than 6 standard deviations from the PC1 or PC2 means of European populations were excluded and we included the first 4 PCs as covariates to further avoid inflation from and residual population related effects in QIMR cohort.

Association analyses

Genome-wise association analyses were conducted separately in TZL, NSPT, JD, CKB, ALSPAC, QIMR, Pittsburgh, and WeGene cohorts and sex was adjusted for in all analyses (Table S1). Initial genome-wide association analyses on ordinal phenotype (coded as 0, 1, and 2 for arch, loop, and whorl, respectively) were performed in PLINK 1.9 (Purcell et al., 2007), using multiple linear regression model of additive allelic effects with additional 4 genetic PCs, 5 genetic PCs, and 5 genetic PCs as covariates in TZL, NSPT, and JD cohorts, respectively. Then we conducted genome-wide association analyses on different variety and derived phenotypes (e.g. binary, nominal, ordinal, and quantitative phenotypes), using a linear or logistic regression model also incorporating sex and 4 genetic PCs as covariates. All Manhattan plots and quantile-quantile plots were created using *qqman* package in R software (Turner, 2014). Quantile-quantile plots were used for all association tests to assess systematic inflation from population stratification or other systematic causes of bias. The genomic control factor λ in all tests did not showed any sign of inflation (<1.03, Table S1). The narrow-sense heritability of fingerprint patterns was estimated using GCTA and then estimated the contribution of SNPs to phenotypic variance of ordinal fingerprint patterns using linear regression model (R^2). GWAS within in CKB were performed using BOLT-LMM v.2.3.2 with the linear mixture model. The covariates used are gender, regions, genotyping array versions and the 8 genetic PCs. For WeGene cohort, genome-wide association analyses were performed with PLINK 1.07 combined with top 10 principal components (PC1-10) generated from GCTA. A genomic inflation factor was generated on the basis of the $\chi 2$-values obtained from PLINK results using R programming (<1.021). Genome-wide analyses were conducted for each digit in each cohort using merlin-offline (QIMR) or Mach2dat (ALSPAC) (Abecasis et al., 2002). Both studies adjusted for sex, with the former also adjusting for 4 genetic PCs. Information in details have been described in a previous study (Ho et al., 2016). For Pittsburgh cohort, all association analyses were performed with the linear mixed model software EMMAX (Kang et al., 2010), which tests the additive effect of each allele while using the subjects' kinship matrix to account for relatedness and ancestry. Sex was included as a covariate. For each of the ten fingers, both an ordinal phenotype (coded as 0, 1, and 2 for arch, loop, and whorl, respectively) and a binary phenotype (presence vs. absence of a whorl) were analyzed. Strong evidence of inflation was not observed (< 1.018).

Meta-analyses

To avoid batch effects from different SNP chip products, meta-analyses were performed in the discovery stage. Then to maximize the statistical power to detect associated genetic variants of small effect, we conducted trans-ethnic meta-analyses for TZL, NSPT, JD, WeGene, CKB, ALSPAC, QIMR, and Pittsburgh cohorts based on the summary results of binary phenotypes (presence vs. absence of a whorl). GWAS results based on binary phenotype for each study were combined via sample-size-weighted fixed-effects analysis using METAL (Willer et al., 2010). In addition, the

heterogeneity of the associations across the different cohorts was assessed by the I^2 and Cochran's Q statistics as reported by METAL. For SNPs with significant heterogeneity, a random effects model was applied for meta-analysis using METASOFT (Han and Eskin, 2011).

Multiple-testing corrections

Given the burden of multiple comparisons, a strict significance threshold of $P < 5\times10^{-8}$ was used to declare 'genome-wide significance', which corresponds to a Bonferroni correction for 1 million independent tests. Given that we tested fingerprint patterns variation as one of many derived phenotypes separately, the multiple-comparisons burden was magnified. Therefore, we also determined a more stringent threshold for declaring 'study-wide significance' (Claes et al., 2018) corresponding to an additional adjustment for the effective number of independent test (Li and Ji, 2005). The eigenvalues of pairwise multivariate corrections of 10 ordinal/binary phenotypes on ten fingers determined a total of 3 effective independent tests (the number of eigenvalues greater than 1 is the number of independent tests) in the discovery cohorts with available phenotype data. Therefore, the study-wide significance threshold was determined to be 1.67×10^{-8} (i.e., $5\times10^{-8}/3$). The same threshold applied to meta-analyses when binary phenotype was used. In addition, when 66 different derived phenotypes were further tested, they were also not strictly independent to each other and the eigenvalues of pairwise multivariate corrections of these phenotypes determined 14 effective independent tests. Therefore, the "study-wide significance" threshold for the association tests of the 66 different derived phenotypes was determined to be 3.57×10^{-9} (i.e., $5\times10^{-8}/14$).

Fine-mapping credible set analysis

First, we created a 3q26.2 signal (GWA SNPs) space including SNPs with $P<5\times10^{-8}$ associated with composite phenotype. The online tool HaploReg (V4.1) was used to explore the function, chromatin states and the nearest genes for each SNP in the signal space. These SNPs were located in non-coding region and some of them showed potential regulatory function annotated in the reference epigenomes of 127 human tissues and cell types obtained from the NIH Roadmap Epigenomics Mapping Consortium (Bernstein et al., 2010). Then we performed fine mapping analysis to detect the potential causal variants for a 500 kb genomic interval flanking the top SNP (250 kb upstream and 250 kb downstream) of 3q26.2 locus using PAINTOR (Kichaev et al., 2014). For each SNP within the 500-kb window, we calculated the posterior probability of driving the association, and then constructed 99% credibility set. We created credibility sets by using combined TZL, NSPT, and JD data sets. We also integrated the linkage disequilibrium information and functional annotation data including seven highlighted epigenomic marks (H3K4me3, H3K4me1, H3K36me3, H3K27me3, H3K9me3, H3K27ac and H3K9ac) for 12 epithelial or mesenchymal cell types (from ectoderm and mesoderm, possibly fingerprint-relevant types) of 127 human tissues and cell types above. Each annotation data was enter the model independently according to the suggested

pipeline. The top one (with the highest sum log bayes factors) was selected for further analyses to compute trait-specific posterior probabilities for causality. A 99% credible set was then constructed by variants (going down the sorted list by posterior probability) whose cumulative posterior probability of representing the causal variant at each locus exceeded 0.99.

Gene mapping, functional annotation and genomic enrichment analyses

The top SNPs found at association loci were used to query the evidence of the candidate genes based on physical distance, biological pathways, the tissue location of expression and whether other traits affected by the mutations in these genes in NCBI (Sherry et al., 2001), UCSC genome browser (Fujita et al., 2010) and Ensemble genome browser 89 (Aken et al., 2016). We used HaploReg V4.1(Ward and Kellis, 2011), an integrative database browser combined of histone modification (ChIP-seq tracks) and ChIA-PET (Chromatin Interaction) from ENCODE and Roadmap Epigenomics Project, eQTLs from GTEx, and conserved regions from GERP and Phastcons to identify more regulatory annotations of genetic variants. 3D genome browser (*www.3dgenome.org*) were used to visualize the chromatin interaction of genome (Wang et al., 2018). Further, we performed genomic enrichment analyses by using the Genomic Regions Enrichment of Annotation Tool (GREAT, abbreviated as G) to identify whether the mapped genes (Table S6) nearby both coding and non-coding genomic regions within 1000kb of top SNPs were enriched with relevant annotations across GO biological processes, human or mouse morphology, and gene expression (McLean et al., 2010). Gene regulatory domains utilized for region annotation were defined as the two nearest genes, and extended up to 1000 kb to the nearest gene's Transcription Start Site ('Two nearest genes' option). The Gene Ontology (GO) Biological Process had 13,145 terms (significant threshold after Bonferroni correction 3.80×10^{-6}) (Ashburner et al., 2000), Human Phenotypes 6,672 terms (7.49×10^{-6}), Mouse Phenotypes 9,554 terms (5.23×10^{-6}) (Dickinson et al., 2016), and gene expression from the MGI (Mouse Genome Informatics) database 9,337 terms (5.35×10^{-6}) (Eppig et al., 2017). Enrichment was tested against the whole human genome (hg19) using standard parameters.

Protein network

To build a meaningful network, the genes (Table S6) which showed limb related by literature for each locus and all notable genes if none of them has been shown to be limb-related for that locus were used as input into the STRING (Szklarczyk et al., 2019) (http://string-db.org/) with specifying Homo sapiens in organism. STRING tool was used to construct a PPI (protein-protein interaction) network with an interacting confidence scores of > 0.4 designated as the cutoff and limiting the number of interactions that directly connect with input by setting the 1st shell to 10. STRING is one of the largest databases of known and predicted protein-protein interactions. In STRING, the functional associations are derived from four sources: genomic context, high-throughput

experiments, conserved coexpression, and previous knowledge. The network was produced in STRING and recolored in Cytoscape 3.8.1 (Kohl et al., 2011) (http://apps.cytoscape.org/), which is an open source platform for visualizing complex networks. cytoHubba plugin in Cytoscape was used to extract the top 2 hub genes (not include extended additional ones) from the PPI network based on maximal clique centrality (MCC) algorithm.

Gene-phenotype associations

Genes and their associated disorders in human were taken from Online Mendelian Inheritance in Man (OMIM) database. Gene name was used to search OMIM and if associated with a Mendelian condition then the description of phenotypic condition and OMIM code were collected. Literature was searched for embryonic expression and mouse phenotypes using Google Scholar. Gene names were searched together with digit, finger, limb, skin, or dermatoglyph. Embryonic limb expression data were summarized for the table if expression was determined using a spatially explicit method (RNA in situ hybridization, immunodetection, or in vivo reporter) only.

Genetic correlation estimation

The pairwise genetic correlation (r_g) values among fingerprint patterns on ten digits in discovery cohorts (genotype available) and genetic correlation between fingerprint and hand traits in NSPT and JD cohorts were calculated using bivariate GCTA-GREML (Lee et al., 2012; Yang et al., 2011) where phenotypes are measured in the same sample. This approach estimates the extent to which genetic similarities correlate with phenotypic similarities. Then we calculated the average r_g values of overall pairs (C_{10}^2) and the middle three digits pairs (C_6^2). The significance test between genetic correlations values of overall pairs and the middle three digits (D2, D3, and D4) pairs was conducted using t test in R.

Polygenic Risk Score (PRS) Analysis

We calculated the polygenic risk score for fingerprint patterns of ten digits by PRSice-2 software (Choi and O'Reilly, 2019). The PRS is a method which is the sum of the trait-related SNPs at multiple genetic loci and weighted according to the effect size estimated by genome-wide association studies. In this study, the "base" data was obtained from GWAS summary statistics of fingerprint-associated 43 SNPs of meta-analysis. The "target" data was individual-level genotype (PLINK format) and ordinal or binary fingerprint phenotypes from three EAS cohorts (TZL, NSPT and JD). Sex was included as a covariate, for p-value thresholds using a lower bound of p = 0.0001, an upper bound of = 0.5 and an increment of 5e-05.

Correlation between fingerprint patterns and hand traits

We used linear regression to analyze the correlation between digit or distal phalanx length and fingerprints on eight digits (statistical significance level at $P < 0.05$). Wilcoxon tests were performed to compare the difference of digit five (D5) length and distal phalanx length of digit two (DP2) and three (DP3) between whorl types and non-whorl types fingerprint (significant threshold at $P <$ 0.00083 (i.e. 0.05/60) after FDR correction). To evaluate the performance of 43 independent genomic signals significantly associated with fingerprints in the GWAS results of hand phenotypes, we use the p value, 0.0042 (i.e. 0.05/12), adjusted by eigen as the threshold. All statistical analysis was done with R.

Plasmids construction, cell culture and luciferase assay

The promoter region of *EVI1* gene was amplified from genomic DNA of an unspecified participant in TZL cohort using the primers EVI1 pro-318-F (5'-ATCGAGATCTAAAGTCTGGGCGATGTG-3') and EVI1 pro-318-R (5'-ATCGAAGCTTAAACCGACGGACAGAGACA -3'). The 303 bp fragment was cloned into pGL3-promoter vector digested with Bgl II and Hind III, resulting in the replacement of the SV40 promoter with the human *EVI1* promoter, yielding the *EVI1* promoter vector pGL3-318. The human TERC promoter fragment was subcloned in the luciferase reporter pGL3-promoter with Bgl II/Nco I, which resulted in the replacement of SV40 promoter with the 867 bp TERC promoter, namely pGL3-TERC. The construction of TERC promoter is based on the previous study (Zhao et al., 2005) which has been proven to be functional. The human GOLIM4 promoter fragment was subcloned in the luciferase reporter pGL3-promoter with Sac I and Mlu I, which resulted in the replacement of SV40 promoter with the 1021 bp GOLIM4 promoter (upstream of transcription start site), namely pGL3-GOLIM4.

The genomic fragment that contains the SNP rs7646897 was amplified from, genomic DNA of an unspecified participant, who was verified as homozygous at rs7646897 by Sanger sequencing using the primers rs7646897-cloning-F (5'-ATCGACGCGTACTGCCATCTCAAGACTAAGC-3') and rs7646897-cloning-R (5'-AGCTAGATCTCATCCTGCACATGTACCTCTG-3'), then the 1469 bp fragment was cloned into pGL3-promoter vector. The same approach was applied with SNP rs7623083. The genomic fragment containing the rs7623083 SNP was amplified using the primers rs7623083-cloning-F (5'-ATCGACGCGTGAGATGACCCCAAAGGATGGG-3') and rs7623083-cloning-R (5'-AGCTCTCGAGGTCACTGCCTTAATAGCTCCCC-3'), then the 1415 bp fragment was cloned into pGL3-promoter vector with restriction enzymes MluI and Bgl II. SNP rs7646897 is at the position of 786bp in its fragment and SNP rs7623083 is at the position of 752bp in its fragment. Individual mutations were incorporated using site-directed mutagenesis (Yeasen). The inserts in each construct were verified by Sanger sequencing. Luciferase reporter plasmids were transfected into HEK293T cells (200ng), using Lipofectamine2000 (Invitrogen) according to

manufacturer's instructions. HEK293T cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% Penicillin-Streptomycin (Gibco) and incubated at 37℃ in 5% CO2. Firefly luciferase expression was normalized to values from co-transfected Renilla luciferase plasmid (10 ng pRL-TK). Cells were harvested 48 h after transfection. Luminescence activity was measured with a Lumat LB 9508 Single Tube Luminometer. Data were represented at least three independent experiments.

Visualization and measurement of dermal structure of mouse forelimbs

Mouse forelimbs were collected and fixed in 4% formaldehyde for 72 hours. The fixed limbs were treated with 5% KOH at 37°C overnight to remove the epidermis. Limbs were washed using 1X PBS and tubes containing the forelimb vortexed to remove residual epidermis. Each forelimb was dipped in 0.05% Toluidine Blue solution for 45 seconds, placed immediately in 1X PBS and visualized using Stereo microscope (Olympus SZX10) and images recorded. Images were used for measurement of digit length using ImageJ.

Categorization of transverse digital ridges

We observed four different categories of transverse digital ridges and they were defined as follows (Figure 3B-C). Pads running continuously across the digit were referred as continuous (c), those running discontinuously across the digit along a single line were referred to as discontinuous (d), and pads originating only at one side of the digit were coded as incomplete (i). Scoring was done for transverse ridges of the middle three digits (D2, D3 and D4) for each forelimb. Both right (R) and left (L) forelimb was scored for each individual mouse. In total, 4 litters were analyzed (Table S5). Two pups of FVB mice were collected at postnatal day 0 (P0), P4, P8 and P12 to study the development of transverse digital ridge patterning (Figure S2B).

Statistical analyses on transverse digital ridge patterns

The effect of genotype on three transverse digit ridge patterns (continuous(c), discontinuous (d), and incomplete (i)/half (h)) was analyzed using a mixed ordinal logistic regression model fitting: genotype, digit, side (L/R), digit × genotype and digit × side interaction were included as fixed effects and litter and line were fitted as random effects. All the statistical analyses were performed using statistical package SAS.

Gene expression analyses

For quantitative RT-PCR, total RNA was isolated from mouse embryonic limb buds using the RNEasy micro kit (Qiagen) following disruption of the tissue in RLT buffer (Qiagen) using a handheld homogeniser. cDNA was synthesized from 1 µg of total RNA using random primers and Superscript III reverse transcriptase (Life Technologies). cDNA was diluted 20-fold and 3 µl used as a template for each qRT-PCR using the Universal SYBR Green Master Mix (Life Technologies), according to manufacturer's instructions. qRT-PCRs were performed using a Stratagene MX 3000p with primer annealing temperatures of 60°C for 40 cycles. Reactions were performed in triplicate, with at least three biological replicates used to determine each data point. Relative expression levels

were determined from cDNA dilution standard curves, and normalized to *Gapdh* values. Sequences of oligonucleotides used as primers were:

Evi1 F: 5'-GCTATGATCAGCACAACCTTGTTG-3'

Evi1 R: 5'-TGTCTGCGACTACTCGGTAGAATATC-3'

Gapdh F: 5'-CGTATTGGGCGCCTGGTCAC-3'

Gapdh R: 5'-ATGATGACCCTTTTGGCTCC-3'

For *Evi1* whole mount in situ hybridization, mouse embryos were fixed in 4% PFA overnight at 4°C. Samples were dehydrated into 100% methanol, bleached in 5:1 MeOH: 30% H_2O_2, rehydrated, treated with 20 µg/ml proteinase K then fixed in 4% PFA, 0.2% glutaraldehyde. Samples were hybridized with a digoxigenin-labeled riboprobe synthesized from expressed sequence tag clone IMAGp998F2411055Q (Source Bioscience) at 60°C overnight in 50% formamide, 5 X saline sodium citrate (SSC), 1% sodium dodecyl sulfate (SDS), 50 µg/mL heparin, and 50 µg/mL yeast RNA. Samples were washed to remove unbound probe and signal detected by incubating with alkaline phosphatase conjugated anti-digoxigenin antibody (Sigma Aldrich, 1:1000) overnight at 4ºC followed by a subsequent 5-bromo-4-chloro-3'-indolylphosphate/nitro-blue-tetrazolium (BCIP/NBT) (Sigma) color reaction.

RNA in situ hybridization

Formaldehyde-fixed forelimb or digit samples embedded in paraffin were sectioned at 6 µm and processed using the RNAscope Multiplex fluorescent reagent kit according to manufacturer's instructions. Probes were species-specific *Evi1*(*Mecom)* and *Prrx1* (see STAR table). Digits 1 and 5 were excluded from analyses. Positive and negative (dapB) control probes were used for both mouse and human sections. After *in situ* hybridisation sections were blocked (5% goat serum/TBST) and incubated overnight at 4ºC with primary antibodies (see STAR table) in blocking buffer. Detection was with fluorescent secondary antibodies diluted in blocking buffer, followed by counterstaining with DAPI and mounting in Prolong Gold. Sections were imaged using a Zeiss LSM 880 confocal microscope.

Immunofluorescence

Fetal tissue samples were fixed overnight at 4°C in 10% neutral buffered formalin, or, for CS14 and CS17 specimens, were embedded in OCT for cryosectioning. 10, 13 and 16 week estimated gestational age samples were embedded in paraffin. Digits 1 and 5 were not included in analyses. For immunofluorescence, cryosections were incubated in PBS for 30 minutes at 37°C to remove gelatin; paraffin sections were dewaxed, rehydrated and antigen retrieved in sodium citrate (pH 6) using Bio-retriever2000 (Aptum Bio, Southampton, UK). Tissue sections were permeabilized in 0.1% Triton X-100/PBS, washed with PBS and treated with TrueBlack (Biotium). Sections were blocked in 5% goat serum/PBS then incubated at 4°C with rabbit anti-EVI1 antibody (1:200, Cell

Signalling Technologies #2593). Sections were washed with PBS, incubated with goat anti-rabbit IgG Alexafluor 546 (1:500 in blocking buffer, ThermoFisher Scientific #A-11035) at RT and washed with PBS, then counterstained with DAPI (Sigma Aldrich) and mounted in Prolong Gold (ThermoFisher Scientific). Fluorescent images were collected using a Zeiss LSM710 inverted confocal microscope.

Quantification and Statistical Analysis

Details on statistical tests, n numbers, and significance cutoffs can be found in the figures or legends. When relevant, further details are found in the method details for the specific measurement.

Supplemental Information

Transverse digital ridges as a murine model of human dermatoglyphs

We examined the structures on the ventral side of the mature mouse forepaw to assess candidate structures for modelling the human dermatoglyph ridges. Chemical removal of the epidermis followed by staining with toluidine blue reveals the arrangement of the uppermost dermal structures. In mouse this approach leads to visualization of different types of dermal structures in the forepaw (Tsugane and Yasuda, 1995). The contact surface of the foot sole has large pads, which have been termed cobblestone, interdigital, or carpal pads. The digits themselves have a series of transverse parallel ridges, some complete, some interrupted and some partial. The digit tips have a single pad with very fine reticulated network of dermal ridges under the epidermis (Tsugane and Yasuda, 1995), which differs from the other dermal structures detected in that they are not present in the topology of the skin surface, which is smooth. This provides a range of candidate structures for comparison to human dermatoglyph ridges.

As human fingerprint patterns are established before birth, we explored the developmental origin of dermatoglyphs in mouse embryos. We found that the transverse digit ridges are apparent prior to birth, at embryonic day 17.5 (E17.5), as a set of stripes visualized in the WNT reporter TCF/LEF::H2B-GFP line (Ferrer-Vaquer et al., 2010) (Figure S2A). Thus the ridge patterns in mouse digits emerge prenatally and prior to physical functioning of the digits, as do the dermatoglyph ridges in human. Newborn pups at postnatal day 1 (P1) already have both continuous and discontinuous ridges, which become more distinct by toluidine blue staining in the subsequent stages (Figure S2B).

In humans, a profound distortion of dermatoglyph pattern is a characteristic of the condition hypohidrotic ectodermal dysplasia, such that ridges are reduced and distorted, with fingerprint patterns sometimes indecipherable (Kargul et al., 2001; Verbov, 1970). This condition is caused by mutations in the *EDA*, *EDAR* or *EDARADD* genes, and also causes reduction or absence of hairs, teeth and a number of types of glands. We investigated dermal structures on the forelimb of Eda^{Tabby}

and *Edar*^{downlessJ} mutant mice, these being models of human hypohidrotic ectodermal dysplasia (Headon and Overbeek, 1999; Kowalczyk-Quintas and Schneider, 2014; Monreal et al., 1999; Srivastava et al., 1997). The cobblestone pads of the palmar surface and the fine patterns at the digit tips were not notably altered in these mutants when compared to control mice, but the transverse digital ridges were markedly distorted and diminished (Figures S2C). This supports the transverse digital ridges as the closest identifiable homologs of human dermatoglyphs, due to their shared and selective requirement for EDA function for their normal development.

Based on their location on ventral side of digits, prenatal formation, parallel periodic arrangement transverse to the long axis of the digit, presence of Merkel cells from early in development (data not shown), carrying of eccrine sweat gland pores, and selective disruption of their pattern by EDA or EDAR pathway mutations, we conclude that transverse digital ridges in mouse serve as the best model of human dermatoglyphs. These transverse ridges in mouse cover only approximately the proximal two-thirds of the digit, with the distal portion having a smooth surface. Thus these transverse ridges are akin to the ridges running across the proximal two phalanges on the human digit, rather than the more complex patterns of the human digit tip.

Supplemental Figures Titles and Legends

Figure S1. Visualizing chromatin interaction between EVI1 gene, Enhancer and rs7646897 region, related to Figure 2

(A) Visualization of genome structure by Hi-C data as a heatmap indicated that the rs7646897 region and EVI1 are located within the same TAD surveyed in 3D genome browser.

(B) Combined with three methods implemented in 3D Genome browser: Virtual 4C (Circular chromosomal conformation capture, top) that surveys for one-vs-many interactions in the genome, DNase I Hypersensitivity Site linkage (middle) that detects distal-proximal DHSs pairs, and ChIA-PET (bottom) that detects long-range interactions between genomic regions, chromatin interactions were identified between the promoter region of EVI1 gene and enhancer harboring SNP rs7646897 (primarily supported by the DHS-linkage data).

Figure S2. Independent luciferase reporter assays on candidate regulatory elements carrying alternate alleles at SNPs rs7646897 in HEK293T cells.

(A-D) Four independent luciferase reporter assays to test the changes of rs7646897 modulating the expression of *EVI1*. Three out of four experiments confirmed the regulatory activity.

(E-H) Four independent luciferase reporter assays to test the changes of rs7623083 modulating the expression of *EVI1*. The effect of rs7623083 on modulating expression *EVI1* isinconsistent.

(I-J) Two independent luciferase reporter assays to test whether SNP rs7646897 modulates the expression of other closest up- and downstream genes *GOLIM4* and *TERC*.

pGL3-basic is a negative control plasmid lacking enhancer activity, pGL3-EVI1, pGL3-GOLIM4

and pGL3-TERC are the positive controls derived from EVI1, GOLIM4 and TERC promoter region, respectively. Symbols indicate significance in t test (*P<0.05, **<0.01, ***<0.001).

Figure S3. Development of digit ridge pattern in mouse, Related to Figure 3

(A) Embryonic origin of digit ridge pattern in mouse. Ventral views of forelimbs of mice carrying a TCF/Lef::H2B-GFP WNT pathway reporter gene at embryonic day 17.5 (E17.5) and newborn (P1). Transverse digital ridges (examples indicated by arrows) are apparent from E17.5. Scale bar = 500 μm. (B) Development of mouse digital ridge patterning through different time points starting from P0 (at birth) followed by post natal days P4, P8 and P12. (C) Altered digit ridge patterns in *EDA* pathway mutant mice. Palmar dermal surface of toluidine blue stained right forepaws from wild type, Eda^{Ta}, $Edar^{dlJ/dlJ}$ and heterozygous $Edar^{dlJ/+}$ mice at postnatal day 21, showing footpad and digit pad types.

Figure S4. *Evi1*^{Jbo} effect on digit proportions and coexpression with *Prrx1*, Related to Figure 3

(A) Digit length in wild type (blue) and $Evi1^{Jbo/+}$ heterozygous (red) adult mice. Digits are shorter in mutant animals. (B) RNAscope *in situ* hybridization detecting *Evi1* and the limb mesenchyme marker *Prrx1* transcripts in mouse E13.5 embryonic limb. Right panels are higher magnification views of area indicated by the arrow on the left panel. Dotted line demarcates epithelium, as defined by immunofluorescent detection of KERATIN14 (K14). *Evi1* and *Prrx1* are coexpressed in individual mesenchymal cells, with *Evi1* also expressed in prominently in mesenchymal cells condensing to form cartilage of the digits, as in human limb development. No expression is detected in epithelium. Nuclei are counterstained with DAPI. Scale bar = 100 μm.

Figure S5. Regional plots of 43 loci associated with fingerprint patterns observed in the meta-analysis across all eight cohorts, Related to Figure 4

Regional association and linkage disequilibrium plot are shown for each locus around the lead SNP (purple diamond) and the color of the remaining markers reflects the linkage disequilibrium (r^2) with the lead SNP. The recombination rate (right-side y axis) is plotted in blue and is based on the ASN 1000 genome population for Han Chinese. Vertical bars on the bottom represent the exons for each gene available from the hg19 UCSC Genome Brower.

Figure S6. Effect sizes (regression coefficients) for the derived allele at top SNPs in the genome regions associated with fingerprint patterns, Related to Figure 4

Estimates obtained in each cohort are shown as blue boxes. Box size is proportional to allele frequency of reference allele. Horizontal bars indicate confidence intervals representing 2×standard errors. Intervals that include zero (that is, non-significant effects) intersect the dashed vertical line. The estimates obtained in the Asian cohorts are shown in red boxes and those obtained in the European cohorts in blue boxes. Diamonds represent the beta and the error bars indicate the 95%

confidence interval.

Figure S7. Association between fingerprint patterns and hand phenotypes, Related to Figure 5

(A) Diagrammed human hand with measured phenotypes, including hand and distal phalanx length. The distal phalanx-hand ratio (DPHR) is the ratio of distal phalanx length to hand length.

(B) The association between the whorl frequency on eight digits and the DPHR of each digit. We used Z-score to standardize the mean DPHR of left and right hands. Solid dots indicated the average values and black short lines were standard deviation for each group.

(C-D) Bar plot of fingerprint patterns of each digit (D2-D5) and the mean DPHR of D2 or D3. Error bars indicate S.E.M. * P<0.05, ** <0.01, *** <0.001;

(E-F) Genetic correlations between fingerprint patterns and the mean DPHR of D2 or D3. Estimates and tests were performed using the bivariate GREML of GCTA software. Error bars indicate S.E.M.

Figure S8. Flow chart of analyses conducted in the study of fingerprint patterns.

Flowchart depicting strategy for the association analysis and functional validation, as well as numbers of samples and SNPs by stage. The yellow diamonds represent bioinformatics tools and software, the green box indicates the main results of the analysis, and the purple box (asterisk) summarizes the conclusions of this study. Abbreviation: TZL = cohort from Taizhou Longitudinal Study, NSPT = cohort from National Survey of Physical Traits Project, JD = cohort from Jidong of Hebei Province, CKB = cohort from China Kadorie Biobank, WeGene = cohort from WeGene company, ALSPAC = The Avon Longitudinal Study of Parents and Children cohort, QIMR = The Queensland Institute of Medical Research cohort.

Supplemental Tables with Titles and Legends

Table S1. Characteristics of cohorts and summary of imputation methods and quality control filters, Related to Table 1

sex difference test P *. <0.05; **. <0.01; ***. <0.001

"--"not available

Table S2. The performance of signals in Han Chinese discovery and replication populations based on ordinal and binary fingerprint patterns, Related to Figure 1

NA, no TopSNP found.

[a]The effect or alternative (Alt) allele frequency of the discovery population.

[b]The two nearest genes within 1000 kb of the most significantly associated SNP annotated by GREAT (http://great.stanford.edu/public/html/index.php), which uses the subset of the UCSC Known Genes.

[c]Bold font means this SNP association was replicated (P<0.05).

Abbreviation: FP_c3_Significant = The SNP only reached the genome-wide significance level ($P<5\times10^{-8}$) in the ordinal (arch, loop and whorl) fingerprint patterns;

FP_c2_Significant = The SNP only reached the genome-wide significance level ($P<5\times10^{-8}$) in the binary (nonwhorl & whorl) fingerprint patterns;

FP_both_Significant = The SNP reached the genome-wide significance level ($P<5\times10^{-8}$) in both binary and ordinal fingerprint patterns.

D1L/R = digit 1 of left or right hand; D2L/R = digit 2 of left or right hand; D3L/R = digit 3 of left or right hand;

D4L/R = digit 4 of left or right hand; D5L/R = digit 5 of left or right hand.

Table S3. Association results of the composite phenotype (extracted from the middle three digits of both hands), Related to Figure 2

[a]Notable genes indicate the same meaning as in Table 1.

[b]The effect or alternative (Alt) allele frequency of the discovery population.

[c]Narrow-sense heritability estimated by GCTA, and standard error in brackets

dExplained variance (R2) for total variance estimated using linear regression model, and R2 is under adjusted

Abbreviation: CP = composite phenotype extracted from the fingerprint pattern of middle three digits on both hands, D1L/R = digit 1 of left or right hand; D2L/R = digit 2 of left or right hand; D3L/R = digit 3 of left or right hand; D4L/R = digit 4 of left or right hand; D5L/R = digit 5 of left or right hand.

Table S4. The 66 additional phenotypes derived from fingerprint patterns used in this study, Related to Figure 1

[a]Genome-wide significance after multiple-testing correction (Padj <3.57×10-9, see STAR Methods)

[b]Novel signals associated with derived fingerprint pattern. CNTN6 is a glycosylphosphatidylinositol (GPI)-anchored neuronal membrane protein that functions as a cell adhesion molecule. It may play a role in the formation of axon connections in the developing nervous system. Participates in oligodendrocytes generation by acting as a ligand of NOTCH1. ADAMTS19 expressed in posterior-proximal limb bud. FGF18 played an important role in the regulation of cell proliferation, cell differentiation and cell migration and required for normal ossification and bone development. EPHB3 controled other aspects of development through regulation of cell migration and positioning, and expressed in limb bud epithelium and nonchondrogenic mesenchyme. Data from GeneCard database (https://www.genecards.org).

We found 202 significant associations for 66 phenotypes.Among them, 190 (94.1%) overlapped with the 18 aforementioned signals identified on ordinal arch-loop-whorl phenotypes.

"--" no significant signal was found.

Table S5. Transverse digit pad patterning in *Evi1*^{Jbo} mice, Related to Figure 3

c: continuous transverse digital ridge; Continuous transverse digital ridge =3

d: discontinuous transverse digital ridge; Discontinuous transverse digital ridge =2

h: half transverse digital ridge; Incomplete/half transverse digital ridge = 1

I: incomplete transverse digital ridge

Y: Presence of the extra digit on digit 5

N: Presence of the extra digit on digit 5 NOT SEEN

R: Right forepaw

L: Left forepaw

Table S6. The performance of 43 signals after meta-analysis of eight populations (the top SNP in each locus was displayed), Related to Figure 4

[a]Signals showed genome-wide significant association ($P_{adj}<1.67\times10^{-8}$) with fingerprint patterns in European (ALSPAC, QIMR, and Pittsburgh) populations or East Asian (TZL, NSPT, JD, CKB, and WeGene) populations and suggestive significant association ($P_{adj}<3.33\times10^{-6}$) in meta-analyses of eight cohorts.

[b]Population frequency of allele A1 in AFR (African), EAS (East Asian), and EUR (European) populations from 1000 Genomes Project phase 3.

[c]The associations between the top SNPs and the fingerprint pattern of the most significant digit (i.e. top digit, as indicated in parentheses).

[d]Top SNPs associated fingerprint pattern of digits passed genome-wide significant ($P_{adj}<1.67\times10^{-8}$, bold font) or suggestive ($P_{adj}<3.33\times10^{-6}$) level after multiple-testing adjustment. Meta-analysis on D1 were conducted in cohorts exclude JD and ALSPAC.

[e]Notable genes indicate the same meaning as in Table 1. Bold genes showed associations with limb phenotypes abnormalities (Table 2). Underline indicates that the SNP is located within the gene.

[f]The associations between the top SNPs and hand traits. Significant level ($P_{adj}<5.6\times10^{-3}$) after multiple-testing adjustment, see Table S7 for abbreviations of hand traits.

Abbreviation: TZL = cohort from Taizhou Longitudinal Study, NSPT = cohort from National Survey of Physical Traits Project, JD = cohort from Jidong of Hebei Province, CKB = cohort from China Kadorie Biobank, WeGene = cohort from WeGene company.

"--" no significant associations.

Table S7. Single-factor regression analysis between hand traits and whorl frequency of eight digits (except the thumbs) in the Chinese Han populations (N=6,318, JD+NSPT), Related to Figure 5

[a]Bold indicates below the significance level (P<0.05).

[b]The region contains JD cohort and three sub-groups (17HanZZ, 18HanNN, and 19HanTZ) from NSPT cohort.

Table S8. Genetic correlation (r_g) based on SNP between fingerprint and hand traits in the Chinese Han populations (N=6,318, JD + NSPT), Related to Figure 5

[a]Bold indicates below the significance level (P<0.05).

[b]Abbreviation: WhorlFrq=Whorl frequency of eight digits (except for the thumbs of both hands)

"--" not available (might be due to the small sample size).

References

Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet *30*, 97-101.

Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Fernandez Banet, J., Billis, K., García Girón, C., and Hourlier, T. (2016). The Ensembl gene annotation system. Database *2016*.

Ali, S., Amina, B., Anwar, S., Minhas, R., Parveen, N., Nawaz, U., Azam, S.S., and Abbasi, A.A. (2016). Genomic features of human limb specific enhancers. Genomics *108*, 143-150.

André, T., Lefèvre, P., and Thonnard, J.L. (2010). Fingertip moisture is optimally modulated during object manipulation. J Neurophysiol *103*, 402-408.

Arrieta, M., Salazar, L., Criado, B., Martinez, B., and Lostao, C. (1991). Twin study of digital dermatoglyphic traits: investigation of heritability. Am J Hum Biol *3*, 11-15.

Arts, H.H., Doherty, D., van Beersum, S.E., Parisi, M.A., Letteboer, S.J., Gorden, N.T., Peters, T.A., Marker, T., Voesenek, K., Kartono, A.*, et al.* (2007). Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome. Nat Genet *39*, 882-888.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T.*, et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.

Babler, W. (1987). Prenatal development of dermatoglyphic digital patterns: Associations with epidermal ridge, volar pad and bone morphology. Coll Antropol *11*, 297-303.

Babler, W.J. (1991). Embryologic development of epidermal ridges and their configurations. Birth Defects Orig Artic Ser *27*, 95-112.

Beccari, L., Jaquier, G., Lopez-Delisle, L., Rodriguez-Carballo, E., Mascrez, B., Gitto, S., Woltering, J., and Duboule, D. (2021). Dbx2 regulation in limbs suggests interTAD sharing of enhancers. Dev Dyn.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R.*, et al.* (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol *28*, 1045-1048.

Bonnevie, K. (1924). Studies on papillary patterns of human fingers. Journal of Genetics *15*, 1-111.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., and Davey, S.G. (2013). Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol *42*, 111-127.

Carlsen, E., Frengen, E., Fannemel, M., and Misceo, D. (2015). Haploinsufficiency of ANO6, NELL2 and DBX2 in a boy with intellectual disability and growth delay. Am J Med Genet A *167a*, 1890-1896.

Chaudhry, A.Z., Lyons, G.E., and Gronostajski, R.M. (1997). Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. Dev Dyn *208*, 313-325.

Chesterman, E.S., Gainey, G.D., Varn, A.C., Peterson, R.E., Jr., and Kern, M.J. (2001). Investigation of Prx1 protein expression provides evidence for conservation of cardiac-specific posttranscriptional regulation in vertebrates. Dev Dyn *222*, 459-470.

Cho, E.A., and Dressler, G.R. (1998). TCF-4 binds beta-catenin and is expressed in distinct regions of the embryonic brain and limbs. Mech Dev *77*, 9-18.

Choi, S.W., and O'Reilly, P.F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience *8*.

Claes, P., Roosenboom, J., White, J.D., Swigut, T., Sero, D., Li, J., Lee, M.K., Zaidi, A., Mattern, B.C., Liebowitz, C.*, et al.* (2018). Genome-wide mapping of global-to-local genetic effects on human facial shape. Nat Genet *50*, 414-423.

Compagni, A., Logan, M., Klein, R., and Adams, R.H. (2003). Control of skeletal patterning by ephrinB1-

EphB interactions. Dev Cell *5*, 217-230.
Cummins, H. (1969). The genetics of dermal ridges. Am J Hum Genet *21*, 516.
Cummins, H., and Midlo, C. (1926). Palmar and plantar epidermal ridge configurations (dermatoglyphics) in European‐Americans. Am J Phys Anthropol *9*, 471-502.
Delaneau, O., Marchini, J., and Zagury, J.F. (2011). A linear complexity phasing method for thousands of genomes. Nat Methods *9*, 179-181.
Delous, M., Baala, L., Salomon, R., Laclef, C., Vierkotten, J., Tory, K., Golzio, C., Lacoste, T., Besse, L., Ozilou, C.*, et al.* (2007). The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome. Nat Genet *39*, 875-881.
Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., and Adissu, H. (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508-514.
Dong, Y., Jesse, A.M., Kohn, A., Gunnell, L.M., Honjo, T., Zuscik, M.J., O'Keefe, R.J., and Hilton, M.J. (2010). RBPjkappa-dependent Notch signaling regulates mesenchymal progenitor cell proliferation and differentiation during skeletal development. Development *137*, 1461-1471.
Dubail, J., Aramaki-Hattori, N., Bader, H.L., Nelson, C.M., Katebi, N., Matuska, B., Olsen, B.R., and Apte, S.S. (2014). A new Adamts9 conditional mouse allele identifies its non-redundant role in interdigital web regression. Genesis *52*, 702-712.
Eppig, J.T., Smith, C.L., Blake, J.A., Ringwald, M., and Bult, C.J. (2017). Mouse Genome Informatics (MGI): Resources for Mining Mouse Genetic, Genomic, and Biological Data in Support of Primary and Translational Research (Springer New York).
Fatemifar, G., Hoggart, C.J., Paternoster, L., Kemp, J.P., Prokopenko, I., Horikoshi, M., Wright, V.J., Tobias, J.H., Richmond, S., Zhurov, A.I.*, et al.* (2013). Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. Hum Mol Genet *22*, 3807-3817.
Feng, W., Huang, J., Zhang, J., and Williams, T. (2008). Identification and analysis of a conserved Tcfap2a intronic enhancer element required for expression in facial and limb bud mesenchyme. Mol Cell Biol *28*, 315-325.
Ferrer-Vaquer, A., Piliszek, A., Tian, G., Aho, R.J., Dufort, D., and Hadjantonakis, A.K. (2010). A sensitive and bright single-cell resolution live imaging reporter of Wnt/ss-catenin signaling in the mouse. BMC Dev Biol *10*, 121.
Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A.*, et al.* (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. Int J Epidemiol *42*, 97-110.
Fujita, P.A., Rhead, B., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Cline, M.S., Goldman, M., Barber, G.P., Clawson, H., and Coelho, A. (2010). The UCSC genome browser database: update 2011. Nucleic Acids Res *39*, D876-D882.
Galton, F. (1892). Finger prints (Macmillan and Company).
Gao, B., and Yang, Y. (2013). Planar cell polarity in vertebrate limb morphogenesis. Curr Opin Genet Dev *23*, 438-444.
Garzón-Alvarado, D.A., and Ramírez Martinez, A.M. (2011). A biochemical hypothesis on the formation of fingerprints using a turing patterns approach. Theor Biol Med Model *8*, 24.
Gerhardt, C., Lier, J.M., Burmühl, S., Struchtrup, A., Deutschmann, K., Vetter, M., Leu, T., Reeg, S., Grune, T., and Rüther, U. (2015). The transition zone protein Rpgrip1l regulates proteasomal activity at the primary cilium. J Cell Biol *210*, 115-133.
Gibson-Brown, J.J., Agulnik, S.I., Chapman, D.L., Alexiou, M., Garvey, N., Silver, L.M., and Papaioannou, V.E. (1996). Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. Mech Dev *56*, 93-101.
Glass, C., Wilson, M., Gonzalez, R., Zhang, Y., and Perkins, A.S. (2014). The role of EVI1 in myeloid malignancies. Blood Cells Mol Dis *53*, 67-76.
Hamrick, M.W. (2001). Primate origins: evolutionary change in digital ray patterning and segmentation. J Hum Evol *40*, 339-351.
Han, B., and Eskin, E. (2011). Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet *88*, 586-598.
Headon, D.J., and Overbeek, P.A. (1999). Involvement of a novel Tnf receptor homologue in hair follicle induction. Nat Genet *22*, 370-374.
Hill, T.P., Taketo, M.M., Birchmeier, W., and Hartmann, C. (2006). Multiple roles of mesenchymal beta-

1413 catenin during murine limb patterning. Development *133*, 1219-1229.
1414 Hirsch, W., and Schweichel, J. (1973). Morphological evidence concerning the problem of skin ridge
1415 formation. J Intellect Disabil Res *17*, 58-72.
1416 Ho, Y.Y., Evans, D.M., Montgomery, G.W., Henders, A.K., Kemp, J.P., Timpson, N.J., St Pourcain, B., Heath,
1417 A.C., Madden, P.A., Loesch, D.Z*., et al.* (2016). Common Genetic Variants Influence Whorls in Fingerprint
1418 Patterns. J Invest Dermatol *136*, 859-862.
1419 Holmes, G.P., Negus, K., Burridge, L., Raman, S., Algar, E., Yamada, T., and Little, M.H. (1998). Distinct
1420 but overlapping expression patterns of two vertebrate slit homologs implies functional roles in CNS
1421 development and organogenesis. Mech Dev *79*, 57-72.
1422 Holt, S.B., and Penrose, L.S. (1968). The genetics of dermal ridges.
1423 Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method
1424 for the next generation of genome-wide association studies. PLoS Genet *5*, e1000529.
1425 Hoyt, P.R., Bartholomew, C., Davis, A.J., Yutzey, K., Gamer, L.W., Potter, S.S., Ihle, J.N., and Mucenski, M.L.
1426 (1997). The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme
1427 development. Mech Dev *65*, 55-70.
1428 International HapMap, C., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A.,
1429 Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F*., et al.* (2010). Integrating common and rare genetic
1430 variation in diverse human populations. Nature *467*, 52-58.
1431 Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010).
1432 Variance component model to account for sample structure in genome-wide association studies. Nat
1433 Genet *42*, 348-354.
1434 Kargul, B., Alcan, T., Kabalay, U., and Atasu, M. (2001). Hypohidrotic ectodermal dysplasia: dental, clinical,
1435 genetic and dermatoglyphic findings of three cases. The Journal of clinical pediatric dentistry *26*, 5-12.
1436 Karmakar, B., Malkin, I., and Kobyliansky, E. (2011). Inheritance of 18 quantitative dermatoglyphic traits
1437 based on Factors in MZ and DZ twins. Anthropol Anz *68*, 185-193.
1438 Kawakami, Y., Esteban, C.R., Matsui, T., Rodríguez-León, J., Kato, S., and Izpisúa Belmonte, J.C. (2004).
1439 Sp8 and Sp9, two closely related buttonhead-like transcription factors, regulate Fgf8 expression and
1440 limb outgrowth in vertebrate embryos. Development *131*, 4763-4774.
1441 Kawakami, Y., Uchiyama, Y., Rodriguez Esteban, C., Inenaga, T., Koyano-Nakagawa, N., Kawakami, H.,
1442 Marti, M., Kmita, M., Monaghan-Nichols, P., Nishinakamura, R*., et al.* (2009). Sall genes regulate region-
1443 specific morphogenesis in the mouse limb by modulating Hox activities. Development *136*, 585-594.
1444 Kichaev, G., Yang, W.Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A.L., Kraft, P., and Pasaniuc, B.
1445 (2014). Integrating Functional Data to Prioritize Causal Variants in Statistical Fine-Mapping Studies. PLoS
1446 Genet *10*, e1004722.
1447 Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: software for visualization and analysis of
1448 biological networks. In Data mining in proteomics (Springer), pp. 291-303.
1449 Kowalczyk-Quintas, C., and Schneider, P. (2014). Ectodysplasin A (EDA) - EDA receptor signalling and its
1450 pharmacological modulation. Cytokine Growth Factor Rev *25*, 195-203.
1451 Kücken, M. (2007). Models for fingerprint pattern formation. Forensic Sci Int *171*, 85-96.
1452 Lancman, J.J., Caruccio, N.C., Harfe, B.D., Pasquinelli, A.E., Schageman, J.J., Pertsemlidis, A., and Fallon,
1453 J.F. (2005). Analysis of the regulation of lin-41 during chick and mouse limb development. Dev Dyn *234*,
1454 948-960.
1455 Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., and Wray, N.R. (2012). Estimation of pleiotropy
1456 between complex diseases using single-nucleotide polymorphism-derived genomic relationships and
1457 restricted maximum likelihood. Bioinformatics *28*, 2540-2542.
1458 Lehoczky, J.A., Williams, M.E., and Innis, J.W. (2004). Conserved expression domains for genes upstream
1459 and within the HoxA and HoxD clusters suggests a long-range enhancer existed before cluster
1460 duplication. Evol Dev *6*, 423-430.
1461 Lewandowski, J.P., Du, F., Zhang, S., Powell, M.B., Falkenstein, K.N., Ji, H., and Vokes, S.A. (2015).
1462 Spatiotemporal regulation of GLI target genes in the mammalian limb bud. Dev Biol *406*, 92-103.
1463 Li, D., Sakuma, R., Vakili, N.A., Mo, R., Puviindran, V., Deimling, S., Zhang, X., Hopyan, S., and Hui, C.C.
1464 (2014). Formation of proximal and anterior limb skeleton requires early function of Irx3 and Irx5 and is
1465 negatively regulated by Shh signaling. Dev Cell *29*, 233-240.
1466 Li, J., and Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a
1467 correlation matrix. Heredity (Edinb) *95*, 221.
1468 Liu, J., Li, Q., Kuehn, M.R., Litingtung, Y., Vokes, S.A., and Chiang, C. (2013). Sonic hedgehog signaling
1469 directly targets Hyaluronic Acid Synthase 2, an essential regulator of phalangeal joint patterning. Dev

Biol *375*, 160-171.
Ma, P., Song, N.N., Li, Y., Zhang, Q., Zhang, L., Zhang, L., Kong, Q., Ma, L., Yang, X., Ren, B.*, et al.* (2019). Fine-Tuning of Shh/Gli Signaling Gradient by Non-proteolytic Ubiquitination during Neural Patterning. Cell Rep *28*, 541-553.e544.
Machado, J.F., Fernandes, P.R., Roquetti, R.W., and Fernandes Filho, J. (2010). Digital dermatoglyphic heritability differences as evidenced by a female twin study. Twin Research and Human Genetics *13*, 482-489.
Mankoo, B.S., Collins, N.S., Ashby, P., Grigorieva, E., Pevny, L.H., Candia, A., Wright, C.V., Rigby, P.W., and Pachnis, V. (1999). Mox2 is a component of the genetic hierarchy controlling limb muscle development. Nature *400*, 69-73.
Mariani, F.V., Ahn, C.P., and Martin, G.R. (2008). Genetic evidence that FGFs have an instructive role in limb proximal-distal patterning. Nature *453*, 401-405.
Martin, N.G., Eaves, L.J., and Loesch, D.Z. (1982). A genetical analysis of covariation between finger ridge counts. Ann Hum Biol *9*, 539-552.
Matsuda, T., Nomi, M., Ikeya, M., Kani, S., Oishi, I., Terashima, T., Takada, S., and Minami, Y. (2001). Expression of the receptor tyrosine kinase genes, Ror1 and Ror2, during mouse development. Mech Dev *105*, 153-156.
McCulloch, D.R., Nelson, C.M., Dixon, L.J., Silver, D.L., Wylie, J.D., Lindner, V., Sasaki, T., Cooley, M.A., Argraves, W.S., and Apte, S.S. (2009). ADAMTS metalloproteases generate active versican fragments that regulate interdigital web regression. Dev Cell *17*, 687-698.
McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.
Medland, S.E., Loesch, D.Z., Mdzewski, B., Zhu, G., Montgomery, G.W., and Martin, N.G. (2007). Linkage analysis of a model quantitative trait in humans: finger ridge count shows significant multivariate linkage to 5q14. 1. PLoS Genet *3*, e165.
Medland, S.E., Zayats, T., Glaser, B., Nyholt, D.R., Gordon, S.D., Wright, M.J., Montgomery, G.W., Campbell, M.J., Henders, A.K., Timpson, N.J.*, et al.* (2010). A variant in LIN28B is associated with 2D:4D finger-length ratio, a putative retrospective biomarker of prenatal testosterone exposure. Am J Hum Genet *86*, 519-525.
Menser, M.A., and Purvis-Smith, S.G. (1969). Dermatoglyphic defects in children with leukaemia. Lancet *1*, 1076-1078.
Mikels, A.J., and Nusse, R. (2006). Purified Wnt5a protein activates or inhibits beta-catenin-TCF signaling depending on receptor context. PLoS Biol *4*, e115.
Monreal, A.W., Ferguson, B.M., Headon, D.J., Street, S.L., Overbeek, P.A., and Zonana, J. (1999). Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia. Nat Genet *22*, 366-369.
Mou, C., Jackson, B., Schneider, P., Overbeek, P.A., and Headon, D.J. (2006). Generation of the primary hair follicle pattern. Proc Natl Acad Sci U S A *103*, 9075-9080.
Mulvihill, J.J., and Smith, D.W. (1969). The genesis of dermatoglyphics. The Journal of pediatrics *75*, 579-589.
Nagy, A., and Pap, M. (2005). Pattern influence on the fingers. HOMO-Journal of Comparative Human Biology *56*, 51-67.
Nandadasa, S., Kraft, C.M., Wang, L.W., O'Donnell, A., Patel, R., and Gee, H.Y. (2019). Secreted metalloproteases ADAMTS9 and ADAMTS20 have a non-canonical role in ciliary vesicle growth during ciliogenesis. *10*, 953.
Nohno, T., Koyama, E., Myokai, F., Taniguchi, S., Ohuchi, H., Saito, T., and Noji, S. (1993). A chicken homeobox gene related to Drosophila paired is predominantly expressed in the developing limb. Dev Biol *158*, 254-264.
Okajima, M. (1975). Development of dermal ridges in the fetus. J Med Genet *12*, 243-250.
Ota, S., Zhou, Z.Q., Keene, D.R., Knoepfler, P., and Hurlin, P.J. (2007). Activities of N-Myc in the developing limb link control of skeletal size with digit separation. Development *134*, 1583-1592.
Painter, K.J., Hunt, G.S., Wells, K.L., Johansson, J.A., and Headon, D.J. (2012). Towards an integrated experimental-theoretical approach for assessing the mechanistic basis of hair and feather morphogenesis. Interface Focus *2*, 433-450.
Parkinson, N., Hardisty-Hughes, R.E., Tateossian, H., Tsai, H.-T., Brooker, D., Morse, S., Lalane, Z., MacKenzie, F., Fray, M., and Glenister, P. (2006). Mutation at the Evi1 locus in Junbo mice causes susceptibility to otitis media. PLoS Genet *2*, e149.

Pierani, A., Brenner-Morton, S., Chiang, C., and Jessell, T.M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. Cell *97*, 903-915.
Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet *38*, 904-909.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J.*, et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559-575.
Rao, R., Gornbein, J., Afshar, Y., Platt, L.D., DeVore, G.R., and Krakow, D. (2019). A new biometric: In utero growth curves for metacarpal and phalangeal lengths reveal an embryonic patterning ratio. Prenat Diagn *39*, 200-208.
Rathee, R., Kamal, N., Kumar, A., Vashist, M., and Yadav, R. (2014). Dermatoglyphic Patterns of Acute Leukemia Patients. International Research Journal of Biological Sciences *3*, 90-93.
Reed, T., Sprague, F., Kang, K., Nance, W., and Christian, J. (1975). Genetic analysis of dermatoglyphic patterns in twins. Hum Hered *25*, 263-275.
Robertson, E.J., Charatsi, I., Joyner, C.J., Koonce, C.H., Morgan, M., Islam, A., Paterson, C., Lejsek, E., Arnold, S.J., Kallies, A.*, et al.* (2007). Blimp1 regulates development of the posterior forelimb, caudal pharyngeal arches, heart and sensory vibrissae in mice. Development *134*, 4335-4345.
Rosner, F. (1969). Dermatoglyphics of leukaemic children. The Lancet *294*, 272-273.
Saad, K., Theis, S., Otto, A., Luke, G., and Patel, K. (2017). Detailed expression profile of the six Glypicans and their modifying enzyme, Notum during chick limb and feather development. Gene *610*, 71-79.
Scheibert, J., Leurent, S., Prevost, A., and Debrégeas, G. (2009). The role of fingerprints in the coding of tactile information probed with a biomimetic sensor. Science *323*, 1503-1506.
Sengupta, M., and Karmakar, B. (2004). Mode of inheritance of finger dermatoglyphic traits among Vaidyas of West Bengal, India. Ann Hum Biol *31*, 526-540.
Sheehan-Rooney, K., Pálinkášová, B., Eberhart, J.K., and Dixon, M.J. (2010). A cross-species analysis of Satb2 expression suggests deep conservation across vertebrate lineages. Dev Dyn *239*, 3481-3491.
Sherry, S.T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res *29*, 308-311.
Shima, Y., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Chisaka, O., Takeichi, M., and Uemura, T. (2002). Differential expression of the seven-pass transmembrane cadherin genes Celsr1-3 and distribution of the Celsr2 protein during mouse development. Dev Dyn *223*, 321-332.
Spielmann, M., Kakar, N., Tayebi, N., Leettola, C., Nürnberg, G., Sowada, N., Lupiáñez, D.G., Harabula, I., Flöttmann, R., Horn, D.*, et al.* (2016). Exome sequencing and CRISPR/Cas genome editing identify mutations of ZAK as a cause of limb defects in humans and mice. Genome Res *26*, 183-191.
Srivastava, A.K., Pispa, J., Hartung, A.J., Du, Y., Ezer, S., Jenks, T., Shimada, T., Pekkanen, M., Mikkola, M.L., Ko, M.S.*, et al.* (1997). The Tabby phenotype is caused by mutation in a mouse homologue of the EDA gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. Proc Natl Acad Sci U S A *94*, 13069-13074.
Szenker-Ravi, E., Altunoglu, U., Leushacke, M., Bosso-Lefèvre, C., Khatoo, M., Thi Tran, H., Naert, T., Noelanders, R., Hajamohideen, A., Beneteau, C.*, et al.* (2018). RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6. Nature *557*, 564-569.
Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., and Bork, P. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res *47*, D607-D613.
Towers, M., and Tickle, C. (2009). Growing models of vertebrate limb development. Development *136*, 179-190.
Tsugane, M., and Yasuda, M. (1995). Dermatoglyphics on volar skin of mice: the normal pattern. The Anatomical Record *242*, 225-232.
Turner, S.D. (2014). qqman: an R package for visualizing GWAS results using QQ and manhattan plots. BioRxiv, 005165.
van Kleffens, M., Groffen, C., Rosato, R.R., van den Eijnde, S.M., van Neck, J.W., Lindenbergh-Kortleve, D.J., Zwarthoff, E.C., and Drop, S.L. (1998). mRNA expression patterns of the IGF system during mouse limb bud development, determined by whole mount in situ hybridization. Mol Cell Endocrinol *138*, 151-161.
Verbov, J. (1970). Hypohidrotic (or anhidrotic) ectodermal dysplasia--an appraisal of diagnostic methods.

Br J Dermatol *83*, 341-348.
Wang, X., Lu, M., Qian, J., Yang, Y., Li, S., Lu, D., Yu, S., Meng, W., Ye, W., and Jin, L. (2009). Rationales, design and recruitment of the Taizhou Longitudinal Study. BMC Public Health *9*, 223.
Wang, Y., Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., Li, D., Choudhary, M.N.K., Li, Y., Hu, M.*, et al.* (2018). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol *19*, 151.
Ward, L.D., and Kellis, M. (2011). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res *40*, D930-D934.
Weinberg, S.M., Neiswanger, K., Martin, R.A., Mooney, M.P., Kane, A.A., Wenger, S.L., Losee, J., Deleyiannis, F., Ma, L., De Salamanca, J.E.*, et al.* (2006). The Pittsburgh Oral-Facial Cleft study: expanding the cleft phenotype. Background and justification. Cleft Palate Craniofac J *43*, 7-20.
Wertheim, K., and Maceo, A. (2002). The critical stage of friction ridge and pattern formation. Journal of Forensic Identification *52*, 35.
Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190-2191.
Yamaguchi, T.P., Bradley, A., McMahon, A.P., and Jones, S. (1999). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. Development *126*, 1211-1223.
Yamamoto, S., Uchida, Y., Ohtani, T., Nozaki, E., Yin, C., Gotoh, Y., Yakushiji-Kaminatsui, N., Higashiyama, T., Suzuki, T., Takemoto, T.*, et al.* (2019). Hoxa13 regulates expression of common Hox target genes involved in cartilage development to coordinate the expansion of the autopodal anlage. Dev Growth Differ *61*, 228-251.
Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82.
Yum, S.M., Baek, I.K., Hong, D., Kim, J., Jung, K., Kim, S., Eom, K., Jang, J., Kim, S., Sattorov, M.*, et al.* (2020). Fingerprint ridges allow primates to regulate grip. Proc Natl Acad Sci U S A *117*, 31665-31673.
Zeng, H., Hoover, A.N., and Liu, A. (2010). PCP effector gene Inturned is an important regulator of cilia formation and embryonic development in mammals. Dev Biol *339*, 418-428.
Zhang, H.-G., Chen, Y.-F., Ding, M., Jin, L., Case, D.T., Jiao, Y.-P., Wang, X.-P., Bai, C.-X., Jin, G., and Yang, J.-M. (2010). Dermatoglyphics from all Chinese ethnic groups reveal geographic patterning. PLoS One *5*, e8783.
Zhao, J., Bilsland, A., Jackson, K., and Keith, W.N. (2005). MDM2 negatively regulates the human telomerase RNA gene promoter. BMC Cancer *5*, 6.