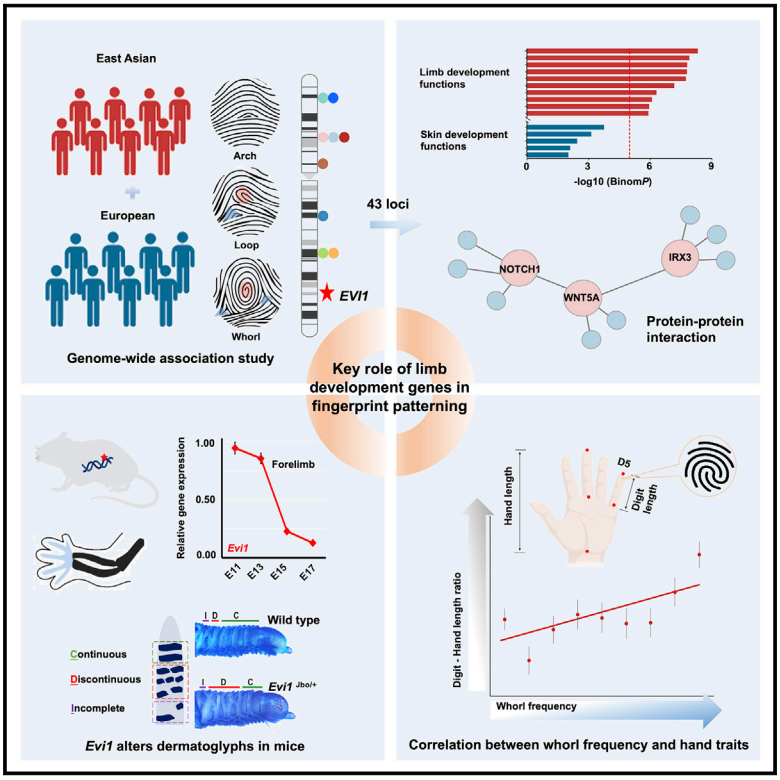
Cell

Limb development genes underlie variation in human fingerprint patterns

Graphical abstract



Highlights

- GWAS identifies variants associated with fingerprint type across all digits
- Fingerprint-associated genes are strongly enriched for limb development functions
- *Evi1* alters dermatoglyphs in mice by modulating limb rather than skin development
- Fingerprint patterns are genetically correlated with hand and finger proportions

Authors

Jinxi Li, James D. Glover, Haiguo Zhang, ..., Li Jin, Denis J. Headon, Sijia Wang

Correspondence

lijin@fudan.edu.cn (L.J.), denis.headon@roslin.ed.ac.uk (D.J.H.), wangsijia@picb.ac.cn (S.W.)

In brief

Genome-wide association scans in Han Chinese individuals and a trans-ethnic meta-analysis reveal genetic regions that are associated with specific fingerprint patterning; functional studies in mouse models confirm a role of *EVI1* in early limb development and ridge patterning.

Li et al., 2022, Cell *185*, 95–112
January 6, 2022
https://doi.org/10.1016/j.cell.2021.12.008

CellPress

Cell

CellPress
OPEN ACCESS

Article

Limb development genes underlie variation in human fingerprint patterns

Jinxi Li,[1,2,30] James D. Glover,[3,30] Haiguo Zhang,[4,30] Meifang Peng,[2,4,30,32] 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,[16,17] Kuang Lin,[16] Ruogu Meng,[18] Jun Lv,[19] Jonathan M. Chernus,[20] Katherine Neiswanger,[21] Eleanor Feingold,[20] David M. Evans,[12,22,23]

(Author list continued on next page)

[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, PRC
[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, PRC
[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, PRC
[5]Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh, 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, PRC
[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, PRC
[10]Shanghai Foreign Language School, Shanghai 200083, PRC
[11]Forest Ridge School of the Sacred Heart, Bellevue, WA 98006, USA
[12]University of Queensland Diamantina Institute, University of Queensland, Brisbane, QLD, Australia
[13]QIMR Berghofer Medical Rese Institute, Brisbane, QLD, Australia
[14]Psychology Department, La Trobe University, Melbourne, VIC, Australia
[15]WeGene, Shenzhen, Guangdong 518040, PRC

(Affiliations continued on next page)

SUMMARY

Fingerprints are of long-standing 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 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.

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). Although fingerprints probably evolved to aid grasping (André et al., 2010; Yum et al., 2020) and for sensing of surface textures (Loesch and Martin, 1984; Medland et al., 2007; Scheibert et al., 2009), since the 19th Century, the fingerprint has been widely used for personal identification because 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,

Sarah E. Medland,[13] Nicholas G. Martin,[13] Seth M. Weinberg,[20,21,24] Mary L. Marazita,[20,21,25,26] Gang Chen,[15] Zhengming Chen,[16,17] Yong Zhou,[27] Michael Cheeseman,[3] Lan Wang,[6] Li Jin,[1,2,28,31,*] Denis J. Headon,[3,31,*] and Sijia Wang[2,29,31,33,*]

[16]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[17]Medical Research Council Population Health Research Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[18]Center for Data Science in Health and Medicine, Peking University, Beijing 100191, PRC
[19]Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing 100191, PRC
[20]Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA 15261, USA
[21]Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, University of Pittsburgh, Pittsburgh, PA 15219, USA
[22]Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, Australia
[23]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[24]Department of Anthropology, University of Pittsburgh, Pittsburgh, PA 15260, USA
[25]Clinical and Translational Science, University of Pittsburgh, Pittsburgh, PA 15261, USA
[26]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15261, USA
[27]Clinical Research Institute, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, PRC
[28]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, PRC
[29]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, PRC
[30]These authors contributed equally
[31]Senior author
[32]Present address: The Core Laboratory in Medical Center of Clinical Research, Department of Molecular Diagnostics and Endocrinology, Shanghai Ninth People's Hospital, State Key Laboratory of Medical Genomics, Shanghai Jiao Tong University School of Medicine, Shanghai 200011, PRC
[33]Lead contact
*Correspondence: lijin@fudan.edu.cn (L.J.), denis.headon@roslin.ed.ac.uk (D.J.H.), wangsijia@picb.ac.cn (S.W.)
https://doi.org/10.1016/j.cell.2021.12.008

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; Penrose, 1965a), 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., 2010; 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 cohorts, 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 previously unreported 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

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

In the discovery phase, we conducted 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 ($p_{adj} < 1.67 \times 10^{-8}$) (Figure 1B), whereas the GWAS based on the binary phenotypes (non-whorl or whorl) showed similar results (Table S2). The majority of signals (17 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

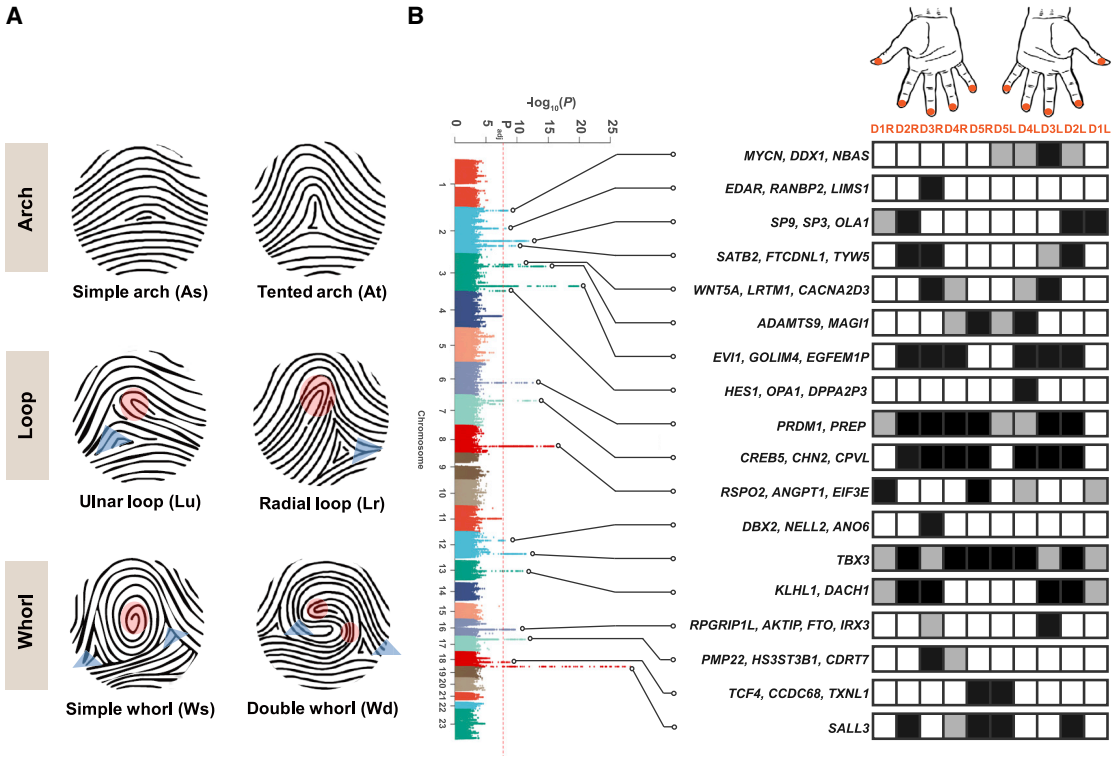


Figure 1. Genetic variants associated with ordinal fingerprint patterns (arch, loop, and whorl) in Han Chinese (n = 9,909)

(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 ($p_{adj} < 1.67 \times 10^{-8}$; STAR Methods). Detailed patterns of adjusted association significance across different fingers are indicated by black squares for corresponding digits for significant associations ($p_{adj} < 1.67 \times 10^{-8}$), and gray squares for marginal associations ($p_{adj} < 3.33 \times 10^{-6}$). Notable genes are indicated for each locus (see Table 1 for selection criteria). D1–5L/R, digit 1 to 5 of left or right hand.

See also Tables S2 and S3.

variance (R^2; Table S3). We further conducted GWAS on 66 different derived phenotypes ($p_{adj} < 3.57 \times 10^{-9}$, 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 S3).

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 previously unreported signal at 5p12 was also discovered (Figure 2C; Table S3). This composite phenotype showed strengthened narrow-sense heritability ($h^2_{CP} = 0.524 \pm 0.039$). 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 ($p = 6.41 \times 10^{-22}$, whereas 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 locus, which showed enhanced association with

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 = 9,909)[d]		Replication 1: CKB (n = 1785)[d]				Replication 2: 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	rs168 62838	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	rs371 242548	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	rs114 60049	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	rs46 73509	C/A	0.184	*SATB2(G/L)*, *FTCDNL1/FONG(G/R/L)*, *TYW5(L)*	D2L*/R*, D3L/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	rs35 8075	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	rs170 72351	G/A	0.225	*ADAMTS9(G/L)*, *MAGI1(G/L)*	D4L*/R, D5L/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	rs64 44832	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	rs802 52354	T/C	0.031	*HES1(G/L)*, *OPA1(G/L)*, *DPPA2P3(L)*	D4L*	−0.238	7.36 × 10^{-9} (D4L)	0.094	0.068	−0.065	0.22	0.115	0.125	0.008	0.917
6q21	rs287 00026	C/T	0.254	*PRDM1(G/L)*, *PREP(G/R/L)*	D1R,D2L*/R*, D3L*/R*, D4L/R*, D5L/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	rs20 75127	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	rs14 94910	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

(*Continued on next page*)

Table 1. ***Continued***

Locus	Top SNPs	Alt/Ref Allele	EAF[a]	Notable Genes[b]	Associated Digits[c]	Discovery: TZL+NSPT+JD (n = 9,909)[d] β	Discovery: TZL+NSPT+JD (n = 9,909)[d] p[e] (Top Digit)	Replication 1: CKB (n = 1785)[d] β[f]	Replication 1: CKB (n = 1785)[d] p[f]	Replication 1: CKB (n = 1785)[d] β[g]	Replication 1: CKB (n = 1785)[d] p[g]	Replication 2: WeGene (n = 2152)[d] β[f]	Replication 2: WeGene (n = 2152)[d] p[f]	Replication 2: WeGene (n = 2152)[d] β[g]	Replication 2: WeGene (n = 2152)[d] p[g]
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	rs728722	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	0.029	0.066	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)	/	/

Abbreviations: 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.

[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 1,000 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); and (3) protein-coding genes within 1,000 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 ($^*p_{adj} < 1.67 \times 10^{-8}$) or suggestive level ($p_{adj} < 3.33 \times 10^{-6}$) after multiple-testing adjustment.

[d]The sample sizes vary in GWAS on different phenotypes of digit: n = 5,415–9,909 for discovery cohort (fingerprint patterns on digit 1 are not available in JD cohort), n = 1,634–1,785 for replication cohort 1 (CKB), and n = 2,138–2,152 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 follows: 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, whereas 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. See also Table S1.

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 cell types (Figure 2D, bottom). Further Hi-C data showed that the variant containing region and the *EVI1* gene are located in the same topologically associating domain (TAD), whereas other nearby genes are located outside the boundary of this domain (Figure S1A). There is also chromatin interaction between the promoter region of the *EVI1* gene and anchoring SNP rs7646897 (Figure S1B). To verify this potential modulating activity on the *EVI1* gene, we performed luciferase reporter assays (STAR Methods; see also Data S1 and Methods S1 for the details of BLAST alignment for each fragment) on rs7646897 and rs7623083 in HEK293T cells and detected allele specific differences in modulating activity for rs7646897 ($p = 7.00 \times 10^{-4}$) (Figures 2E, S2A–S2D, S2I, and S2J). SNP rs7623083 showed no effect for three experiments and allele specific differences for two (Figures 2F and S2E–S2H). Further assays showed that the alternate SNPs at rs7646897 modulated the expression of *EVI1* ($p = 7.50 \times 10^{-3}$), 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).

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 (Figures S3A and 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 Methods S1 for further justifications).

We analyzed dermatoglyph patterns in 21 day old $Evi1^{Jbo/+}$ heterozygous mutant (encoding EVI1p.Asn763Ile) and wild-type littermate digits 2, 3, and 4 (Figures 3A–3C; Table S4). Homozygous $Evi1^{Jbo/Jbo}$ embryos die between midgestation and birth, whereas heterozygotes display a small spur on digit 5 (Figure 3A) (Parkinson et al., 2006), and slightly decreased digit length (Figure S3D). 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 embryonic day (E) 12.5, 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 (Figures 3E and S3E). 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 E11.5, 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 S3E).

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) (~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 phalanx (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 the 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 cell 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 because 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.

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)

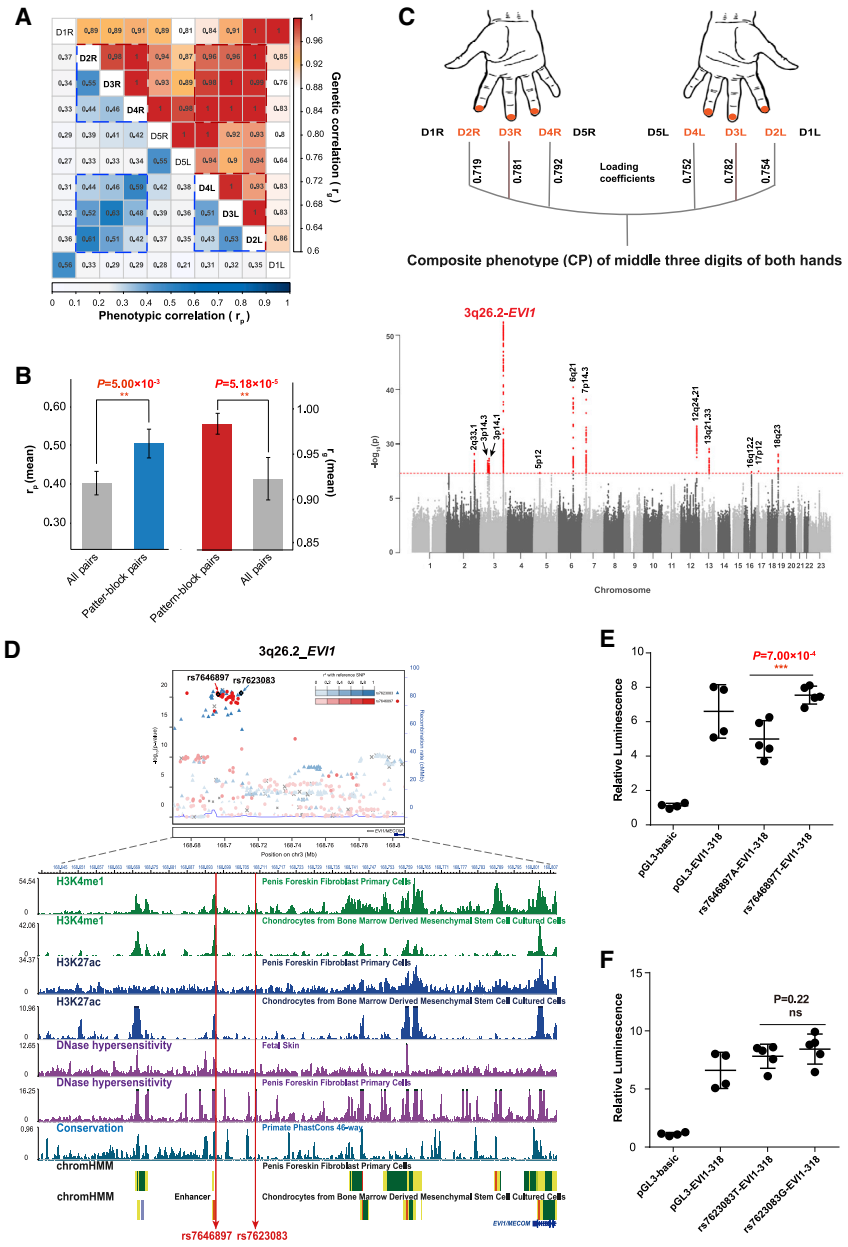


Figure 2. Genetic basis of the middle three digit "pattern-block" phenomenon, with top signal near *EVI1*

(A) "Pattern block" of the middle three digits on both hands revealed by pairwise 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 (Pruim et al., 2010) 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, area 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 and F) (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, and pGL3-318 is a positive control derived from the *EVI1* promoter region. Symbols indicate significance in t test (*p < 0.05, **p < 0.01, ***p < 0.001).

See also Figures S1 and S2, Table S3, Methods S1, and Data S1.

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 ($p_{adj} < 1.67 \times 10^{-8}$) (Figure 4A; Table S5; see also Data S1 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 S5). 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, whereas the other 9 all showed suggestive level of association in EUR-ancestry cohorts ($6.61 \times 10^{-6} < p < 5.69 \times 10^{-3}$), with exactly the same effect direction as in EAS-ancestry cohorts (Table S5). 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 S5). 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 S4).

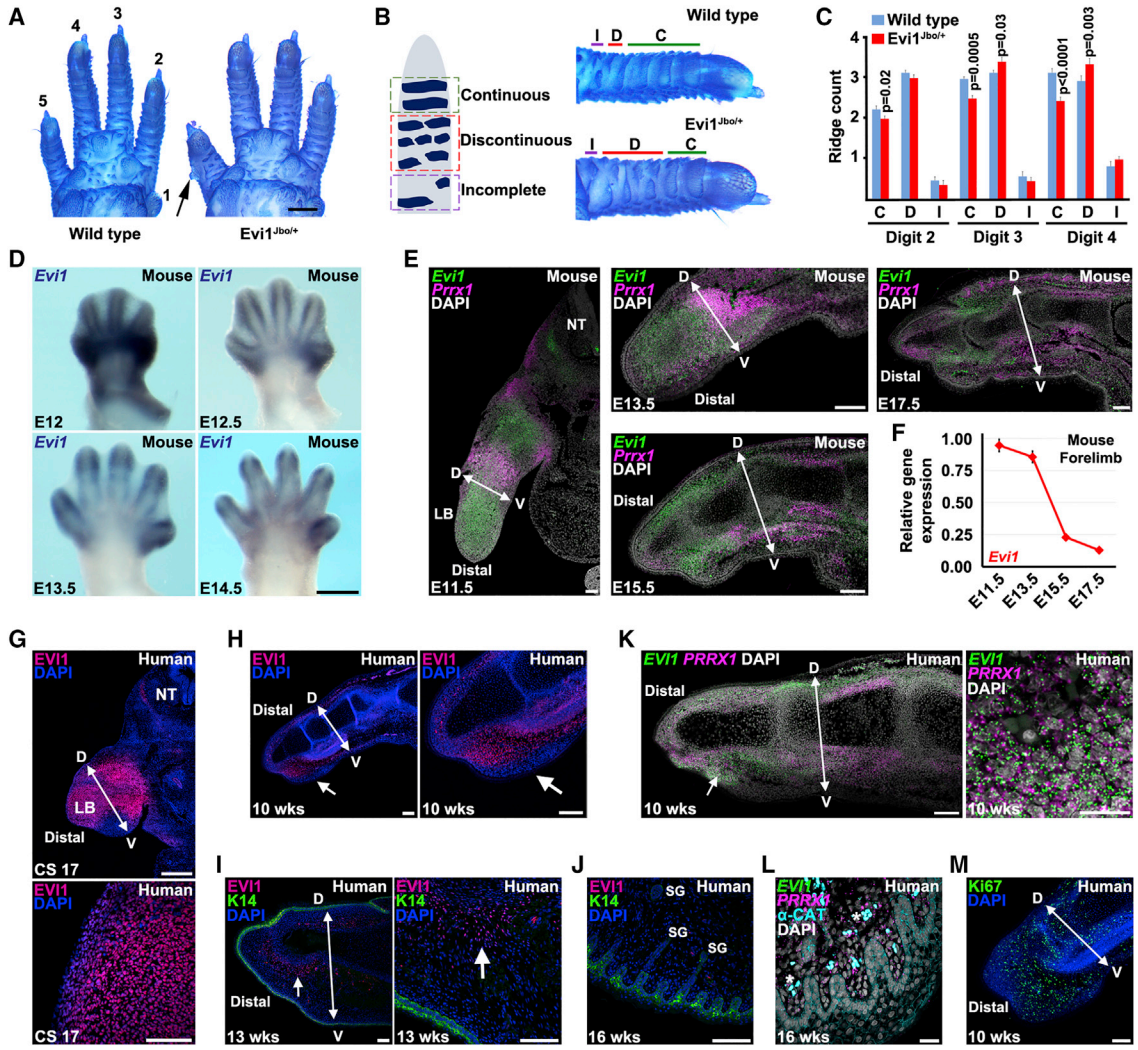


Figure 3. EVI1 in dermatoglyph patterning and limb development

(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*$^{Jbo/+}$. C, regions carrying continuous ridges are indicated; D, discontinuous ridges are indicated; I, incomplete ridges are indicated.

(C) Quantification of digit ridge pattern in wild-type and *Evi1*$^{Jbo/+}$ mutants. Continuous ridges are reduced on all mutant digits, whereas D3 and D4 carry more discontinuous ridges.

(D) Whole mount *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) qRT-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 (arrow in week 13) and epithelial marker K14. SG, eccrine sweat gland.

(K and 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* at 10 weeks, indicated by arrow and magnified in right panel. Asterisks indicate auto-fluorescent 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, 1 mm (A); 500 μm (D and G, upper); 100 μm (E, G, lower, H–J, K, left, and M); 20 μm (K, right, and L). Error bars indicate SEM.

See also Figure S3, Table S4, and Methods S1.

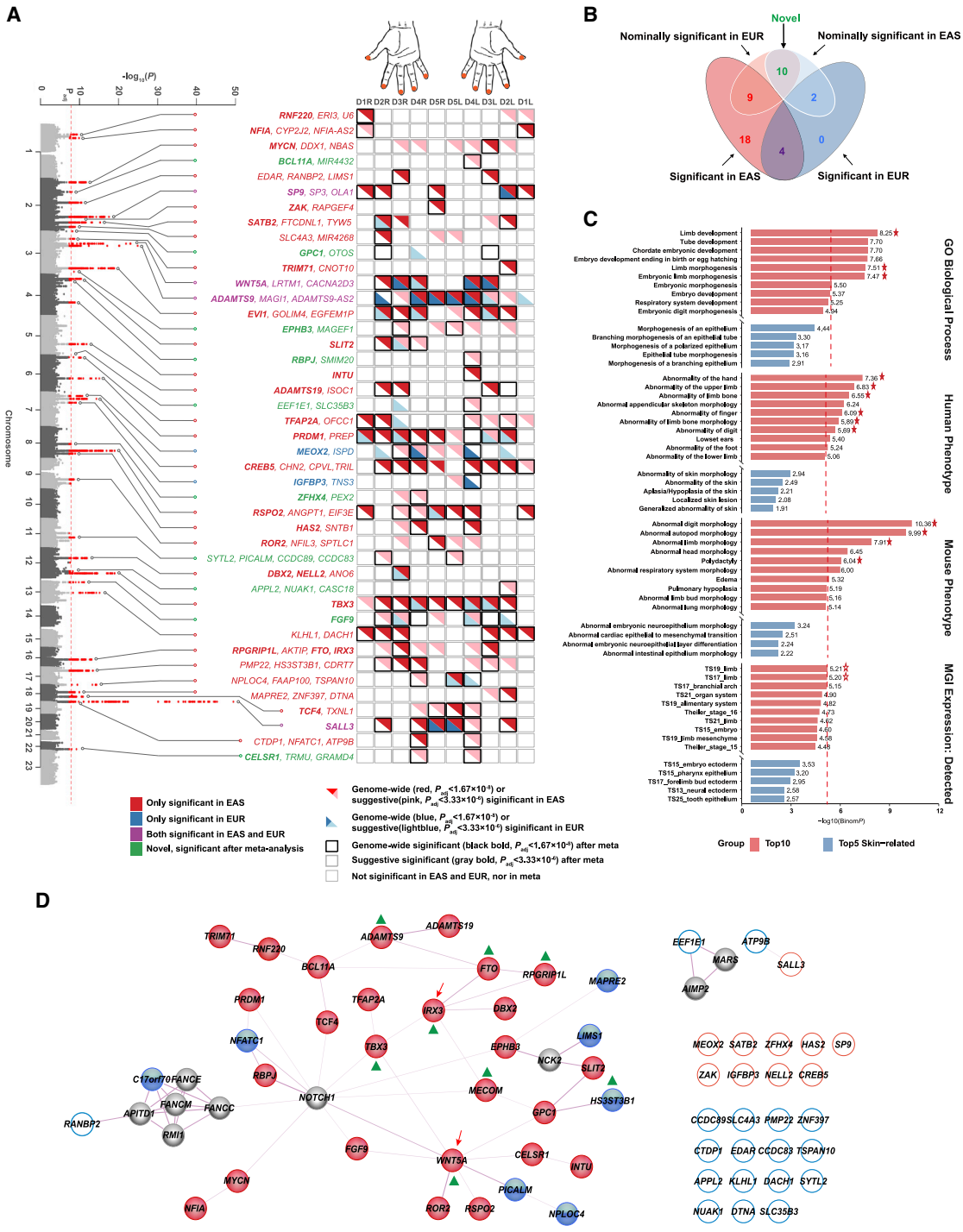


Figure 4. A meta-analysis of fingerprint patterns showing signals enriched in limb development

(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 ($p_{adj} < 1.67 \times 10^{-8}$; Table S5), 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; and 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, whereas dark and light colors represented

(legend continued on next page)

Genomic enrichment analysis using GREAT V4.0.4 (McLean et al., 2010) (Table S5; STAR Methods) 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); whereas epithelial and skin-related pathways including "morphogenesis of an epithelium" and "epithelial tube morphogenesis" did not reach the significance threshold after correction for multiple-testing (-log10(p_{thr}) = 5.42, Figure 4C). Similar patterns of enrichment for limb rather than skin-related terms were observed when using human phenotype annotations and mouse morphology terms for the analysis (Figure 4C). Enrichment was also observed for expression in mouse limb tissues at Theiler Stage 17 (E10.5) and 19 (E11.5) when the forelimbs are divided into two regions indicating proximal limb and autopod but not in the developing autopod epithelium (Figure 4C).

Network analysis using STRING v11 (Szklarczyk et al., 2019) linked fingerprint-associated GWAS genes and proteins by known functions and regulatory interactions (see STAR Methods), showing a network centered around WNT5A and IRX3 (Figure 4D), suggesting relevant developmental pathways. Interestingly, both WNT5A and IRX3 were among the GWAS signals of the composite phenotype of middle three fingers (Figure 2C). Eight proteins from 6 of the 12 composite phenotype-associated signals appeared within two steps of the centered proteins with higher degrees in the network (WNT5A and IRX3), suggesting that these developmental pathways are relevant to the biological mechanism underlying the "pattern-block" phenomenon. Another network was centered on NOTCH signaling (NOTCH1, although not itself a GWAS hit), responsible for maintaining mesenchymal progenitors of the limb in an undifferentiated and proliferative state during its development (Dong et al., 2010).

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 and Figure S5 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 S6). Higher frequency of whorl patterns was associated with longer little finger (D5) relative to hand length (ratio of D5 length to hand length [DHR5]; $\beta = 0.15$, $p = 2.42 \times 10^{-5}$) (Figure 5B; Table S6). 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 ($p_{D5L} = 1.74 \times 10^{-13}$, $p_{D5R} = 5.85 \times 10^{-13}$) (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 because significant genetic correlations were found between DHR5 and fingerprint patterns on different digits, especially on D5 ($r_{g\text{-}D5L} = 0.31$, $p = 3.89 \times 10^{-16}$, $r_{g\text{-}D5R} = 0.22$, $p = 1.26 \times 10^{-11}$) (Figure 5D; Table S6). 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 S5). 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 ($\beta_{DPHR2} = -0.31$, $p = 9.13 \times 10^{-18}$, $\beta_{DPHR3} = -0.24$, $p = 2.02 \times 10^{-11}$) (Table S6) and with fingerprint-associated signals (2q31.1-*SP9*, 8q23.1-*RSPO2*, etc.) (Table S5). Together, these results provided further evidence that fingerprint patterns are strongly influenced by the process of limb development.

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 S3). 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

signals that reached the adjusted genome-wide significant ($p_{adj} < 1.67 \times 10^{-8}$) and suggestive levels ($p_{adj} < 3.33 \times 10^{-6}$), 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 (A).

(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, whereas 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. Gray 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 Figure S4, Tables S5 and S6, and Data S1.

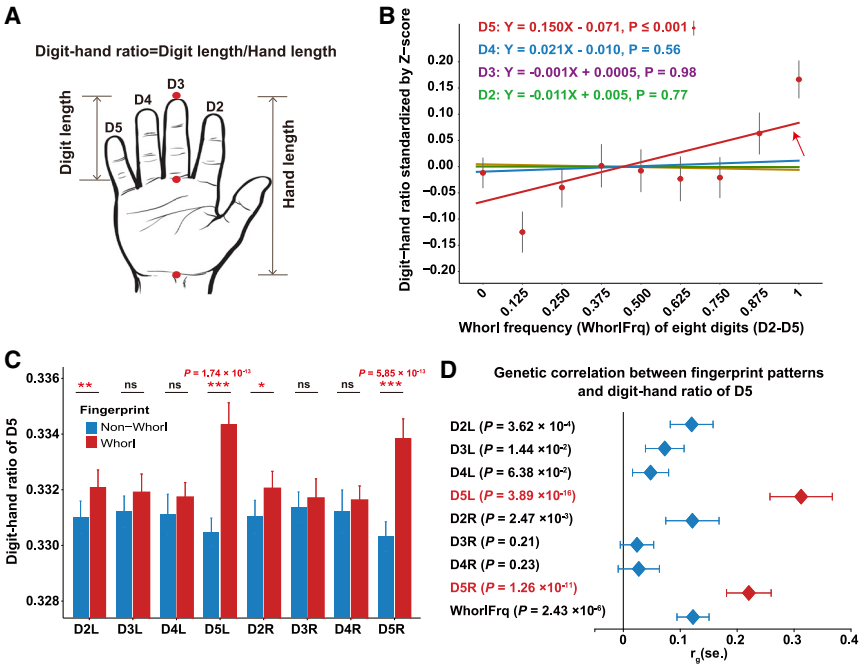


Figure 5. Association between fingerprint patterns and hand phenotypes (n = 6,318)

(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 SEM. *p < 0.05, **p < 0.01, ***p < 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 SEM.

See also Figure S5 and Table S6.

number of signals, suggesting that this ordinal phenotype might better reflect the underlying genetic mechanisms of pattern formation. It has been proposed that the morphology (i.e., height, shape, and size) of volar pads (Bonnevie, 1924; Penrose, 1965b; Wertheim and Maceo, 2002) and the growth stresses on the pad surface (Kücken, 2007) play important roles to influence pattern types. In particular, although 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; Penrose and Ohara, 1973). 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 (Kargül 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 because their variable shapes in the fetus are reported to be correlated with different dermatoglyph types (Babler, 1991; Loesch, 1983). 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 et al., 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

Table 2. Functional annotation for notable genes (subset)

Genes	① Syndrome and 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)
1p31.3-*NFIA*	① brain malformations with or without urinary tract defects (613735); ③ distal limb bud (Chaudhry et al., 1997)
2p24.3-*MYCN*[a]	① Feingold syndrome; syndactyly and reduced middle phalanges (164280); ③ proliferating cells of distal limb bud mesenchyme (Ota et al., 2007)
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)
2q13-*EDAR*[b]	① hypohidrotic ectodermal dysplasia; skin appendages absent, aberrant dermatoglyphs (224900)
2q31.1-*SP9*	③ apical ectodermal ridge, regulated by FGF10 (Kawakami et al., 2004)
2q31.1-*MAP3K20/ZAK*[a]	① split-foot malformation; mesoaxial polydactyly, nail duplications (616890); ③ entire early limb bud (Spielmann et al., 2016)
2q33.1-*SATB2*[a]	① Glass syndrome; digit anomalies, sparse hair (612313); ③ apical ectodermal ridge (Sheehan-Rooney et al., 2010)
2q37.3-*GPC1*	③ limb bud mesenchyme (chicken) (Saad et al., 2017)
3p22.3-*TRIM71 (LIN41)*	③ distal limb and digit mesenchyme (mouse and chicken) (Lancman et al., 2005)
3p14.3-*WNT5A*[a]	① Robinow syndrome; limb shortening, brachydactyly (180700); ③ apical ectodermal ridge and progress zone (Yamaguchi et al., 1999)
3p14.1-*ADAMTS9*	② limb-specific deletion causes syndactyl; ③ broadly in early limb mesenchyme, subsequently digit perichondrium (McCulloch et al., 2009)
3q26.2-*EVI1*[a]	① RUSAT2; radioulnar synostosis, digit defects (616738); ③ limb bud and digit pad mesenchyme (human, Figures 3G–3I)
3q27.1-*EPHB3*	③ limb bud epithelium and nonchondrogenic mesenchyme (Compagni et al., 2003)
4p15.31-*SLIT2*	③ interdigital mesenchyme and digit lateral margins (Holmes et al., 1998)
4p15.2-*RBPJ*[a]	① Adams-Oliver syndrome; scalp and distal limb defects (short distal phalanges) (614814)
4q28.1-*INTU*[a]	① short-rib thoracic dysplasia 20 with polydactyly (617925)
5q23.3-*ADAMTS19*	③ posterior-proximal limb bud, regulated by SHH signaling (Lewandowski et al., 2015)
6p24.3-*TFAP2A*[a]	① branchiooculofacial syndrome, incompletely penetrant polydactyly (113620); ③ limb bud ectoderm and distal mesenchyme (Feng et al., 2008)
6q21-*PRDM1*	② mutants lack posterior digits, whisker development; ③ limb bud posterior mesenchyme, including zone of polarizing activity (Robertson et al., 2007)
7p21.2-*MEOX2*	② mutants have selectively reduced limb musculature; ③ limb bud myoblasts (Mankoo et al., 1999)
7p14.3-*CREB5*	③ interdigital condensing mesenchyme (Lehoczky et al., 2004)
7p12.3-*IGFBP3*	③ interdigital region (van Kleffens et al., 1998)
8q21.13-*ZFHX4*	③ conserved human enhancer drives reporter expression in mouse limb buds (Ali et al., 2016)
8q23.1-*RSPO2*[a]	① tetraamelia syndrome 2; absence of limbs (618021); ③ apical ectodermal ridge (Szenker-Ravi et al., 2018)
8q24.12-*HAS2*	② short limbs with phalanx duplication and misplaced interphalangeal joints; ③ distal limb bud mesenchyme, regulated by SHH (Liu et al., 2013)
9q22.31-*ROR2*[a]	① Robinow syndrome; limb shortening, including brachydactyly (268310); ③ distal limb mesenchyme (Matsuda et al., 2001)
12q12-*DBX2*[a]	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)
12q12-*NELL2*[a]	③ distal limb and digit mesenchyme (mouse and chicken) (Lancman et al., 2005)
12q24.21-*TBX3*[a]	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)
13q12.11-*FGF9*[a]	① multiple synostoses syndrome 3; impaired interphalanx joint formation, broadened thumbs (612961); ③ apical ectodermal ridge (Mariani et al., 2008)
16q12.2-*RPGRIP1L*[a]	① Meckel syndrome; polydactyly (611561); ③ protein located at primary cilium of embryonic limb mesenchymal cells; regulates SHH signaling (Gerhardt et al., 2015)
16q12.2-*FTO*[c]	① growth retardation, developmental delay, and facial dysmorphism; brachydactyly and cutis marmorata (612938)

(*Continued on next page*)

Table 2. *Continued*

Genes	① Syndrome and Phenotype Description When Mutated, in Human (OMIM Number) or ② Mouse ③ Expression Site in Embryonic Limb (Mouse, Unless Stated)
16q12.2-*IRX3*	② smaller limb when mutation combined with Irx5 mutation; ③ proximal-anterior limb bud, interacts with SHH (Li et al., 2014)
18q12.1-*MAPRE2*[b]	① congenital symmetric circumferential skin creases; excess skin leading to ringed creases, principally on limbs (616734)
18q21.2-*TCF4*[a]	① Pitt-Hopkins syndrome; persistent fetal digital pads (610954); ③ peridigital mesenchyme at distal digit tips (Cho and Dressler, 1998)
18q23-*SALL3*	② lack of digit development when mutated with SALL1; ③ distal posterior mesenchyme, regulated by SHH (Kawakami et al., 2009)
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.

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), whereas *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* ($p = 6.93 \times 10^{-5}$) and 18q23-*SALL3* ($p = 5.07 \times 10^{-4}$) loci, as well as 3q26.2-*EVI1* ($p = 7.62 \times 10^{-3}$) which was similar to the mouse finding (Figure S3D), were associated with both fingerprint type and length of digit 5 (Table S5). 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 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. Because the entire region of the 3q26.2 signal covering all the associated SNPs appeared in the credible sets using various methods, this region (chromosome 3 [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 luciferase reporter assays, we used the HEK293T cell line, widely used to assay enhancer activity and for other applications, because they are easy to transfect and handle in the laboratory. In addition, the HEK293T cell line is used here because mesenchymal cell lines derived from early limb formation stages do not exist. From the results of luciferase reporter assays, we verified the regulatory effect of SNP rs7646897 on the *EVI1* gene. Although we have ruled out the

effect of this SNP on the expression of the two adjacent genes (*GOLIM4* and *TERC*), it is possible that SNP rs7646897 could potentially modulate the expression of other genes more distally. In addition, the present study lacks sufficient evidence to determine whether rs7623083 has a modulating effect on *EVI1* expression. Two experiments showed an effect and three did not. Furthermore, the plasmids containing the variants (rs7646897 or rs7623083) had repressing, activating or neutral effects on enhancer activity in different experimental trials, indicating significant variation in the reporter assay. Understanding the relative importance of these two closely linked SNPs on *EVI1* expression and dermatoglyph phenotype, and the detailed regulatory mechanisms at play, will thus require further experimental studies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
  - Cell lines and cultures
  - Mice
  - Fetal tissue collection
- METHOD DETAILS
  - Ethics statement
  - Study population and design
  - Specification of fingerprint patterns and hand traits
  - Genotyping, quality control, and imputation
  - Population stratification analysis
  - Association analyses
  - Meta-analyses
  - Multiple-testing corrections
  - Fine-mapping credible set analysis
  - Gene mapping, functional annotation and genomic enrichment analyses
  - Protein network
  - Gene-phenotype associations
  - Genetic correlation estimation
  - Polygenic Risk Score (PRS) Analysis
  - Correlation between fingerprint patterns and hand traits
  - Plasmids construction, cell culture and luciferase assay
  - Visualization and measurement of dermal structure of mouse forelimbs
  - Categorization of transverse digital ridges
  - Statistical analyses on transverse digital ridge patterns
  - Gene expression analyses
  - RNA *in situ* hybridization
  - Immunofluorescence
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2021.12.008.

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 include 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 and W.S. for her help with graphic abstract. We thank Dida Biotech for performing the Luciferase assay experiments. This project was funded by the "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDB38020400 to S. Wang); Science and Tecnology Commission of Shanghai Municipality; the Ministry of Science and Technology of the People's Republic of China; Shanghai Municipal Science and Technology Major Project (2017SHZDZX01 to L.J. and S. Wang); National Key Research and Development Project (2018YFC0910403 to S. Wang); Max Planck-CAS Paul Gerson Unna Independent Research Group Leadership Award (to S. Wang); National Natural Science Foundation of China (NSFC) (31521003 to L.J., 81622003 to L.W., 81373082 to J. Lv, and 31771325 to J.T.); China Postdoctoral Science Foundation (2019M651351 to J. Li); National Science & Technology Basic Research Project (2015FY111700 to L.J.); CAMS Innovation Fund for Medical Sciences (2019-I2M-5-066 to L.J.); The 111 Project (B13016 to L.J.); Major Research Program of National Natural Science Foundation of China (91731303 to J.T.); the National Institute for Dental and Craniofacial Research (NIDCR) (R01-DE016148 to M.L.M and S.M.W and R21-DE016930 to M.L.M); Biotechnology and Biological Sciences Research Council (BBS/E/D/10002071 and BB/T007788/1 to D.J.H.); MRC (G1100357 to R.A.A and MR/N022556/1 [MRC Centre for Reproductive Health]; European Commission; Marie Sklodowska-Curie Actions Individual fellowship 706429; Wellcome Trust/DBT India Alliance Early Career Fellowship (IA/E/18/1/504338 to C.B.M); Australian Research Council (A7960034, A79906588, A79801419, DP0212016, and DP0343921 to 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 to S.E.M); the UK Medical Research Council; the Wellcome Trust (102215/2/13/2); the University of Bristol; Medical Research Council (MC_UU_12013/4 to D.M.E. and MC_UU_12013/3 to N.J.T.); Australian Research Council Future Fellowship (FT130101709 to D.M.E.); the Hong Kong Kadoorie Charitable Foundation; UK Wellcome Trust (202922/Z/16/Z, 104085/Z/14/Z, and 088158/Z/09/Z); UK Medical Research Council (MC-PC-13049 and MC-PC-14135); and core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University from British Heart Foundation, UK MRC, and Cancer Research UK (to R.G.W. and Z.C.). S.E.M. is supported by NHMRC (SRF 1103623).

AUTHOR CONTRIBUTIONS

S. Wang, D.J.H., and L.J. conceived the project and provided main resources. For the discovery, GWAS, H.Z., J. Li, J.T., Y.Y., S.Z., J.Z., H.Y., and W.Z. performed data and sample collections. H.Z., J. Li, S.Z., and Y.X. contributed to generating the fingerprint and hand phenotype data. S. Wu, Y. Liu, and Q.P. contributed to generating the SNP array data. J. Li and M.P. performed computational analysis. For the replication and meta-analysis, J. Lv, R.M., R.G.W., K.L., and Z.C. performed data analysis and provided summary statistics results and relevant descriptions for the CKB cohort; Lizhong Wang, Y. Li, S.T., X.W., and G.C. performed data analysis and provided summary statistics results and relevant descriptions for the WeGene cohort; D.M.E. and J.P.K. performed data analysis and provided summary statistics results and relevant descriptions for the ALSPAC; S.E.M., N.G.M., Y.Y.W.H., and D.Z.L. performed data analysis and provided summary statistics results and relevant descriptions for the QIMR; and J.M.C., K.N., E.F., S.M.W., and M.L.M. performed data analysis and provided summary statistics results and relevant descriptions for the 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. H.B. aided in data analysis from mouse models. D.H. and Lan Wang performed the luciferase reporter assay. J. Li, D.J.H., and S. Wang 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

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

INCLUSION AND DIVERSITY

We worked to ensure gender balance in the recruitment of human subjects. We worked to ensure sex balance in the selection of non-human subjects. We worked to ensure ethnic or other types of diversity in the recruitment of human subjects. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

Received: April 16, 2021
Revised: October 20, 2021
Accepted: December 8, 2021
Published: January 6, 2022

SUPPORTING CITATIONS

The following references appear in the supplemental information: Headon and Overbeek (1999); Kowalczyk-Quintas and Schneider (2014); Monreal et al. (1999); Srivastava et al. (1997); Tsugane and Yasuda (1995).

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., Hourlier, T., et al. (2016). The Ensembl gene annotation system. Database (Oxford) *2016*, baw093.

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.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al.; International HapMap 3 Consortium (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58.

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

Arrieta, M.I., Salazar, L., Criado, B., Martinez, B., and Lostao, C.M. (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., Märker, 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.; The Gene Ontology Consortium (2000). Gene ontology: tool for the unification of biology. 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. *250*, 1280–1299.

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. J. Genet. *15*, 1–111.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., and Davey Smith, 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.Ø., 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*, giz082.

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.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., BBaymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res *46*, D794–D801.

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., Adissu, H., et al.; International Mouse Phenotyping Consortium; Jackson Laboratory; Infrastructure Nationale PHENOMIN, Institut Clinique de la Souris (ICS); Charles River Laboratories; MRC Harwell; Toronto Centre for Phenogenomics; Wellcome Trust Sanger Institute; RIKEN BioResource Center (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., Coelho, A., et al. (2011). 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.

GTEx Consortium. (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213.

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.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome research *22*, 1760–1774.

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-catenin during murine limb patterning. Development *133*, 1219–1229.

Hirsch, W., and Schweichel, J.U. (1973). Morphological evidence concerning the problem of skin ridge formation. J. Ment. Defic. Res. *17*, 58–72.

Ho, Y.Y.W., Evans, D.M., Montgomery, G.W., Henders, A.K., Kemp, J.P., Timpson, N.J., St Pourcain, B., Heath, A.C., Madden, P.A.F., Loesch, D.Z., et al. (2016). Common Genetic Variants Influence Whorls in Fingerprint Patterns. J. Invest. Dermatol. *136*, 859–862.

Holmes, G.P., Negus, K., Burridge, L., Raman, S., Algar, E., Yamada, T., and Little, M.H. (1998). Distinct but overlapping expression patterns of two vertebrate slit homologs implies functional roles in CNS development and organogenesis. Mech. Dev. *79*, 57–72.

Holt, S.B., and Penrose, L.S. (1968). The Genetics of Dermal Ridges (Charles C. Thomas).

Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. *5*, e1000529.

Hoyt, P.R., Bartholomew, C., Davis, A.J., Yutzey, K., Gamer, L.W., Potter, S.S., Ihle, J.N., and Mucenski, M.L. (1997). The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. Mech. Dev. *65*, 55–70.

The International HapMap Consortium (2003). The International HapMap Project. Nature *426*, 789–796.

International HapMap, Altshuler, C., Gibbs, D.M., Peltonen, R.A., Altshuler, L., Gibbs, D.M., Peltonen, R.A., Dermitzakis, L., Schaffner, E., et al. (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. *42*, 348–354.

Kargül, B., Alcan, T., Kabalay, U., and Atasu, M. (2001). Hypohidrotic ectodermal dysplasia: dental, clinical, genetic and dermatoglyphic findings of three cases. J. Clin. Pediatr. Dent. *26*, 5–12.

Karmakar, B., Malkin, I., and Kobyliansky, E. (2010). Inheritance of 18 quantitative dermatoglyphic traits based on factors in MZ and DZ twins. Anthropol. Anz. *68*, 185–193.

Kawakami, Y., Esteban, C.R., Matsui, T., Rodríguez-León, J., Kato, S., and Izpisúa Belmonte, J.C. (2004). Sp8 and Sp9, two closely related buttonhead-like transcription factors, regulate Fgf8 expression and limb outgrowth in vertebrate embryos. Development *131*, 4763–4774.

Kawakami, Y., Uchiyama, Y., Rodriguez Esteban, C., Inenaga, T., Koyano-Nakagawa, N., Kawakami, H., Marti, M., Kmita, M., Monaghan-Nichols, P., Nishinakamura, R., and Izpisua Belmonte, J.C. (2009). Sall genes regulate region-specific morphogenesis in the mouse limb by modulating Hox activities. Development *136*, 585–594.

Kichaev, G., Yang, W.Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A.L., Kraft, P., and Pasaniuc, B. (2014). Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. *10*, e1004722.

Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. In Data Mining in Proteomics (Springer), pp. 291–303.

Kowalczyk-Quintas, C., and Schneider, P. (2014). Ectodysplasin A (EDA) - EDA receptor signalling and its pharmacological modulation. Cytokine Growth Factor Rev. *25*, 195–203.

Kücken, M. (2007). Models for fingerprint pattern formation. Forensic Sci. Int. *171*, 85–96.

Lancman, J.J., Caruccio, N.C., Harfe, B.D., Pasquinelli, A.E., Schageman, J.J., Pertsemlidis, A., and Fallon, J.F. (2005). Analysis of the regulation of lin-41 during chick and mouse limb development. Dev. Dyn. *234*, 948–960.

Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., and Wray, N.R. (2012). Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540–2542.

Lehoczky, J.A., Williams, M.E., and Innis, J.W. (2004). Conserved expression domains for genes upstream and within the HoxA and HoxD clusters suggests a long-range enhancer existed before cluster duplication. Evol. Dev. *6*, 423–430.

Lewandowski, J.P., Du, F., Zhang, S., Powell, M.B., Falkenstein, K.N., Ji, H., and Vokes, S.A. (2015). Spatiotemporal regulation of GLI target genes in the mammalian limb bud. Dev. Biol. *406*, 92–103.

Li, J., and Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity *95*, 221–227.

Li, D., Sakuma, R., Vakili, N.A., Mo, R., Puviindran, V., Deimling, S., Zhang, X., Hopyan, S., and Hui, C.C. (2014). Formation of proximal and anterior limb skeleton requires early function of Irx3 and Irx5 and is negatively regulated by Shh signaling. Dev. Cell *29*, 233–240.

Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol *34*, 816–834.

Liu, J., Li, Q., Kuehn, M.R., Litingtung, Y., Vokes, S.A., and Chiang, C. (2013). Sonic hedgehog signaling directly targets Hyaluronic Acid Synthase 2, an essential regulator of phalangeal joint patterning. Dev. Biol. *375*, 160–171.

Loesch, D.Z. (1983). Quantitative Dermatoglyphics. Classification, Genetics, and Pathology (Oxford University Press).

Loesch, D.Z., and Martin, N.G. (1984). Finger ridge patterns and tactile sensitivity. Ann. Hum. Biol. *11*, 113–124.

Loh, P.R., Tucker, G., Bulik-Sullivan, B.K., Vilhjálmsson, B.J., Finucane, H.K., Salem, R.M., Chasman, D.I., Ridker, P.M., Neale, B.M., Berger, B., et al. (2015). Efficient Bayesian mixed-model analysis increases association power in large cohorts. . Nat Genet *47*, 284–290.

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.e4.

Machado, J.F., Fernandes, P.R., Roquetti, R.W., and Filho, J.F. (2010). Digital dermatoglyphic heritability differences as evidenced by a female twin study. Twin Res. Hum. Genet. *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*, 1736–1744.

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. USA *103*, 9075–9080.

Mulvihill, J.J., and Smith, D.W. (1969). The genesis of dermatoglyphics. J. Pediatr. *75*, 579–589.

Nagy, A.S., and Pap, M. (2005). Pattern influence on the fingers. Homo *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. Nat. Commun. *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., Glenister, P., et al. (2006). Mutation at the Evi1 locus in Junbo mice causes susceptibility to otitis media. PLoS Genet. *2*, e149.

Penrose, L.S. (1965a). Dermatoglyphic Topology. Nature *205*, 544–546.

Penrose, L.S. (1965b). On the geometry of loops and deltas. Ann. Hum. Genet. *29*, 207–211.

Penrose, L.S., and Ohara, P.T. (1973). The development of the epidermal ridges. J. Med. Genet. *10*, 201–208.

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.

Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics *26*, 2336–2337.

Pruitt, K.D., Brown, G.R., Hiatt, S.M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., Farrell, C.M., et al. (2014). RefSeq: an update on mammalian reference sequences. Nucleic Acids Res *42*, D756–D763.

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., and Sham, P.C. (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. Int. Res. J. Biol. Sci. *3*, 90–93.

Reed, T., Sprague, F.R., Kang, K.W., Nance, W.E., and Christian, J.C. (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. Lancet *2*, 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. USA *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., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47* (D1), D607–D613.

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.M., and Lauro, C. (2005). PLS path modeling. Comput Stat Data Anal *48*, 159–205.

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. Anat. Rec. *242*, 225–232.

Turner, S.D. (2014). qqman: an R package for visualizing GWAS results using QQ and manhattan plots. bioRxiv. https://doi.org/10.1101/005165.

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. (2012). 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. J. Forensic Identif. *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. Am. J. Hum. Genet. *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. USA *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., Yang, J.-M., et al. (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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Evi-1 (C50E12), Rabbit mAb (1:200)	Cell Signaling Technologies	Cat#2593, RRID:AB_2184098
Anti-Cytokeratin 14 antibody [LL002], Mouse mAb (1:200)	Abcam	Cat# ab7800, RRID:AB_306091
Anti-α E-catenin antibody (G-11), Mouse mAb (1:100)	Santa Cruz Biotechnology	Cat# sc9988, RRID:AB_626805
Recombinant Anti-Ki67 antibody [SP6], Rabbit mAb (1:200)	Abcam	Cat# ab16667, RRID:AB_302459
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 (1:500)	Thermo Fisher Scientific	Cat# A-11029, RRID:AB_2534088
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 546 (1:500)	Thermo Fisher Scientific	Cat#A-11035, RRID:AB_2534093
Anti-Digoxigenin-AP, Fab fragments from sheep (1:2000)	Sigma Aldrich	Cat# 11093274910, RRID:AB_2734716
Biological samples		
Human fetal tissue – digit and embryo samples	Royal Infirmary of Edinburgh (see STAR Methods for details)	N/A
Chemicals, peptides, and recombinant proteins		
4',6-Diamidino-2-phenylindole dihydrochloride (DAPI)	Sigma Aldrich	Cat# D9542
FastStart Universal Sybr Green Master (Rox)	Sigma Aldrich	Cat# 4913850001
SuperScript III Reverse Transcriptase	Thermo Fisher Scientific	Cat# 18080093
RNaseOUT	Thermo Fisher Scientific	Cat# 10777019
Nitrotetrazolium Blue Chloride	Sigma Aldrich	Cat# N6876
5-Bromo-4-chloro-3-indolyl phosphate p-toludine salt	Sigma Aldrich	Cat# B8503
TrueBlack	Biotium	Cat# 23007
RNAscope multiplex fluorescent reagent kit V2	ACD bio	Cat# 323100
Opal 570 fluorophore (1:1500)	Akoya biosciences	Cat# FP1488001KT
Opal 690 fluorophore (1:1500)	Akoya biosciences	Cat# FP1497001KT
Deposited data		
Raw and analyzed data: GWAS data from Chinese cohorts	This paper	NODE: OEP000198 http://www.biosino.org/node/project/detail/OEP000198
Raw and analyzed data: GWAS data from European cohorts	This paper	NODE: OEP000198 http://www.biosino.org/node/project/detail/OEP000198
Pittsburgh Orofacial Cleft Study	Weinberg et al., 2006	https://www.dental.pitt.edu/pittsburgh-orofacial-cleft-studies
HapMap Phase 1-3	Altshuler et al., 2010	https://www.genome.gov/10001688/international-hapmap-project
1000 Genomes Project Phase 3	Altshuler et al., 2010	https://www.internationalgenome.org/category/phase-3/
HAPMAP2 r22.36 CEU	Open source	ftp://share.sph.umich.edu/1000genomes/resources/CEU.r22.orig.tgz
NCBI dbSNP database	Sherry et al., 2001	http://www.ncbi.nlm.nih.gov/SNP
RefSeq database	Pruitt et al., 2014	https://www.ncbi.nlm.nih.gov/refseq/
UCSC genome browser	Fujita et al., 2011	http://genome.ucsc.edu
Ensemble genome browser 89	Aken et al., 2016	http://asia.ensembl.org/Homo_sapiens/Info/Index

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Roadmap Epigenomics Project	Bernstein et al., 2010	http://www.roadmapepigenomics.org
GENCODE v19	Harrow et al., 2012	https://www.gencodegenes.org/human/release_19.html
GTEx v8	GTEx Consortium, 2017	https://gtexportal.org/home/
ENCODE	Davis et al., 2018	https://www.encodeproject.org/
HaploReg v4.1	Ward and Kellis, 2012	https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php
STRING v11.0	Szklarczyk et al., 2019	https://string-db.org/
OMIM	Johns Hopkins School of Medicine	https://www.omim.org
Experimental models: organisms/strains		
Mouse: FVB/N	Charles River	Strain:207
Mouse: TCF/Lef::H2B-GFP (Ferrer-Vaquer et al., 2010)	Roslin Institute	N/A
Mouse: Tabby (*Eda*^{Ta})	Roslin Institute	N/A
Mouse: downless mutants (*Edar*$^{dlJ/dlJ}$ and *Edar*$^{dlJ/+}$)	Roslin Institute	N/A
Mouse: Junbo (C3H/HeH background)	European Mouse Mutant archive – Bred at Roslin Institute & Harwell	EM: 00091
Oligonucleotides		
Evi1 F: 5′-GCTATGATCAGCACAACCTTGTTG-3′	Thermo Fisher Scientific	N/A
Evi1 R: 5′-TGTCTGCGACTACTCGGTAGAATATC-3′	Thermo Fisher Scientific	N/A
Gapdh F: 5′-CGTATTGGGCGCCTGGTCAC-3′	Thermo Fisher Scientific	N/A
Gapdh R: 5′-ATGATGACCCTTTTGGCTCC-3′	Thermo Fisher Scientific	N/A
RNAscope Probe - Mm-Mecom - *Mus musculus* MDS1 and EVI1 complex locus (Mecom) transcript variant 1 mRNA	ACD bio	Cat# 432231
RNAscope Probe - Hs-MECOM - Homo sapiens MDS1 and EVI1 complex locus (MECOM) transcript variant 1 mRNA	ACD bio	Cat# 518021
RNAscope Probe - Mm-Prrx1-C2 - *Mus musculus* paired related homeobox 1 (Prrx1) transcript variant 1 mRNA	ACD bio	Cat# 485231-C2
RNAscope Probe - Hs-PRRX1-No-XMm-C2 - Homo sapiens paired related homeobox 1 (PRRX1) transcript variant pmx-1a mRNA	ACD bio	Cat# 511001-C2
RNAscope 3-plex Negative Control Probe	ACD bio	Cat# 320871
RNAscope 3-plex Positive Control Probe_Mm	ACD bio	Cat# 320881
RNAscope 3-plex Positive Control Probe_Hs	ACD bio	Cat# 320861
Recombinant DNA		
Mus musculus Mecom (Evi1) EST clone	Source Bioscience	IMAGp998F2411055Q
Software and algorithms		
MATLAB_R2019a	MathWorks	https://www.mathworks.com/products/matlab.html; RRID: SCR_013499
R v3.6.1	The R Foundation	https://www.r-project.org; RRID: N/A
PLS-PM	The R CRAN	https://github.com/gastonstat/plspm; RRID: N/A
PLINK v1.9	Purcell et al., 2007	https://www.cog-genomics.org/plink2; RRID: N/A
SHAPEIT	Delaneau et al., 2011	http://www.shapeitforum.com; RRID: N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
IMPUTE2	Howie et al., 2009	https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; RRID: SCR_013055
EIGENSOFT	Price et al., 2006	https://data.broadinstitute.org/alkesgroup/EIGENSOFT/; RRID: SCR_004965
qqman v0.1.4	Turner, 2014	https://cran.r-project.org/web/packages/qqman/index.html; RRID: N/A
BOLT-LMM v2.3.2	Loh et al., 2015	https://alkesgroup.broadinstitute.org/BOLT-LMM/BOLT-LMM_manual.html; RRID: N/A
MACH2DAT	Li et al., 2010	https://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output; RRID: SCR_009599
MERLIN-OFFLINE	Abecasis et al., 2002	https://sph.umich.edu/csg/abecasis/Merlin/; RRID: N/A
EMMAX	Kang et al., 2010	http://genetics.cs.ucla.edu/emmax_jemdoc/; RRID: N/A
METAL	Willer et al., 2010	http://csg.sph.umich.edu/abecasis/Metal/; RRID: SCR_002013
METASOFT	Han and Eskin, 2011	http://genetics.cs.ucla.edu/meta/; RRID: N/A
PAINTOR v2.1	Kichaev et al., 2014	https://github.com/gkichaev/PAINTOR_V3.0; RRID: N/A
GCTA-GREML	Lee et al., 2012; Yang et al., 2011	https://yanglab.westlake.edu.cn/software/gcta/; RRID: N/A
LocusZoom	Pruim et al., 2010	https://genome.sph.umich.edu/wiki/LocusZoom; RRID: SCR_009257
GREAT v4.0.4	McLean et al., 2010	http://great.stanford.edu/public/html/index.php; RRID: SCR_005807
Adobe Photoshop	Adobe	N/A
MxPro	Stratagene	N/A
Zen Imaging Software	Carl Zeiss	N/A
Other		
Greenbit DactyScan26	Neurotechnology	https://www.neurotechnology.com/fingerprint-scanner-green-bit-dactyscan-26.html
Epson Perfection V370 Photo (EPSONScanV370)	EPSON	https://www.epson.com.cn/products/scanners/206/V370/
Hanlintongxin PU-JY500U	Beijing Hanlin ICT Information Technology	N/A
Smiths Heimann Biometrics ACCO1394	Medland et al., 2007	https://www.smiths.com/news-and-media/2005/07/smiths-heimann-biometrics-and-cross-match-technologies-to-join-forces
Cytoscape	Kohl et al., 2011	https://apps.cytoscape.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Sijia Wang (wangsijia@picb.ac.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Full summary statistics for all SNPs in the eight cohorts (TZL, NSPT, JD, CKB, WeGene, ALSPAC, QIMR, and Pittsburgh) GWAS datasets can be found in the National Omics Data Encyclopedia (NODE) (http://www.biosino.org/node). The accession number for the data analyzed in this paper is NODE: OEP000198 (http://www.biosino.org/node/project/detail/OEP000198), and also listed in the Key Resources Table. Individual genotype and phenotype data cannot be shared due to privacy concerns. All other relevant data, such as mice dermatoglyph, are within the paper and its Supporting Information files.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the Lead Contact upon request.

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°C 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 centers. *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, of both sexes, were used (Table S4). Mice were killed by cervical dislocation and DNA from tail tips was 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.

METHOD 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 S6. 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 toward 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 (https://github.com/gastonstat/plspm) (Tenenhaus et al., 2005). 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 S3. 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 S6.

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.bristol.ac.uk/alspac/researchers/access/). 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 (The International HapMap Consortium, 2003) and 1000 Genomes Project (International HapMap et al., 2010) 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; $p < 1 \times 10^{-5}$) and differences in minor allele frequencies (MAF < 1% excluded) and heterozygosity. The chip genotype data were first 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 (Altshuler 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 ($p < 1 \times 10^{-5}$), 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 first 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 utilized 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 $p < 1 \times 10^{-6}$ 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 center, 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 ($p < 1 \times 10^{-5}$), 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 was 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 dataset of 7,124,171 SNPs and these dataset 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 $< 5 \times 10^{-7}$). 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 ($p < 1 \times 10^{-6}$) 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 (Altshuler 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 $\geq$ 0.05 and HWE $p > 10^{-6}$, in each of the populations YRI, CEU and CHB, (2) pairwise $r^2 \leq 0.1$ 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 was 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 show 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 (Loh et al., 2015). 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 χ2-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; Li et al., 2010). 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 versus 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 versus 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 $p < 5 \times 10^{-8}$ 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., $5 \times 10^{-8}/3$). 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., $5 \times 10^{-8}/14$).

Fine-mapping credible set analysis

First, we created a 3q26.2 signal (GWA SNPs) space including SNPs with $p < 5 \times 10^{-8}$ 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 datasets. 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 entering 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., 2011) and Ensemble genome browser 89 (Aken et al., 2016). We used HaploReg V4.1(Ward and Kellis, 2012), an integrative database browser combined of histone modification (ChIP-seq tracks) and ChIA-PET (Chromatin Interaction) from ENCODE (Davis et al., 2018) and Roadmap Epigenomics Project, eQTLs from GTEx (GTEx Consortium,

2017), and conserved regions from GERP and Phastcons to identify more regulatory annotations of genetic variants. RefSeq (Pruitt et al., 2014) and GENCODE v19 (Harrow et al., 2012) databases were used to obtain gene annotations. 3D genome browser (http://3dgenome.fsm.northwestern.edu) 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 S5) 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 S5) 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) (https://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) (https://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 BglII 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 BglII/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′-ATCGACGCGTACTGCC ATCTCAAGACTAAGC-3′) and rs7646897-cloning-R (5′-AGCTAGATCTCATCCTGCACATGTACCTCTG-3′), then the 1469 bp fragment was cloned into pGL3-318 vector with restriction enzymes MluI and BglII. The whole same approach was applied with SNP rs7623083. The genomic fragment containing the rs7623083 SNP was amplified using the primers rs7623083-cloning-F (5′-ATCG ACGCGTGAGATGACCCCAAAGGATGGG-3′) and rs7623083-cloning-R (5′-AGCTCTCGAGGTCACTGCCTTAATAGCTCCCC-3′), then the 1415 bp fragment was cloned into pGL3-promoter vector with restriction enzymes MluI and XhoI. 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. The amplified, cloned sequence carrying SNP rs7646897 was verified by Sanger sequencing and was consistent with the NCBI primary assembly sequence (NCBI: gi|224514994: 168695100-168696800 *Homo sapiens* chromosome 3, GRCh37.p13 Primary Assembly), though the target site (red box) was different due to site-directed mutagenesis. The actual sequence carrying rs7623083 (red box) was different from the NCBI primary assembly sequence (NCBI: NC_000003.11: 168709002-168710504 *Homo sapiens* chromosome 3, GRCh37.p13 Primary Assembly) at three sites (green box), which could be explained by the individual mutation of donor. The BLAST alignment for each fragment is presented at Data S1.

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°C 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 s, 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 (Figures 3B and 3C). 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 S4). 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 S3B).

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 was 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 homogenizer. 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 4°C 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 table in STAR Methods). 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* hybridization sections were blocked (5% goat serum/TBST) and incubated overnight at 4°C with primary antibodies (see table in STAR Methods) 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 primary antibody. Sections were washed with PBS, incubated with fluorescent secondary antibody at room temperature 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

Information on specific quantification methods such as collection and specification of fingerprint patterns, hand traits in human and quantification of dermatoglyph in mice are described in associated Method details, or main texts. Statistical tests were performed using R software, parameters such as number of sample size, the number of mice, the number of independent experiments, measures of center, dispersion, and precision (mean ± SEM and 95% confidence interval), statistical test and significance, are reported in Figures and Figure Legends or Method details.

Supplemental figures

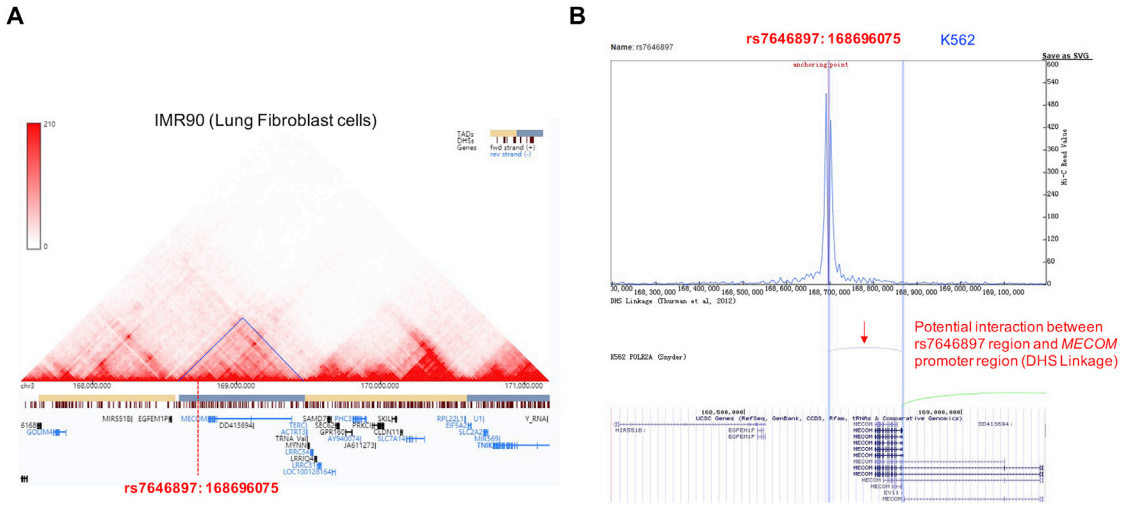


Figure S1. Visualizing chromatin interaction between the *EVI1* gene, enhancer, and the rs7646897 region, related to Figure 2

(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-versus-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 promoter region of the *EVI1* gene and enhancer harboring SNP rs7646897 (primarily supported by the DHS-linkage data).

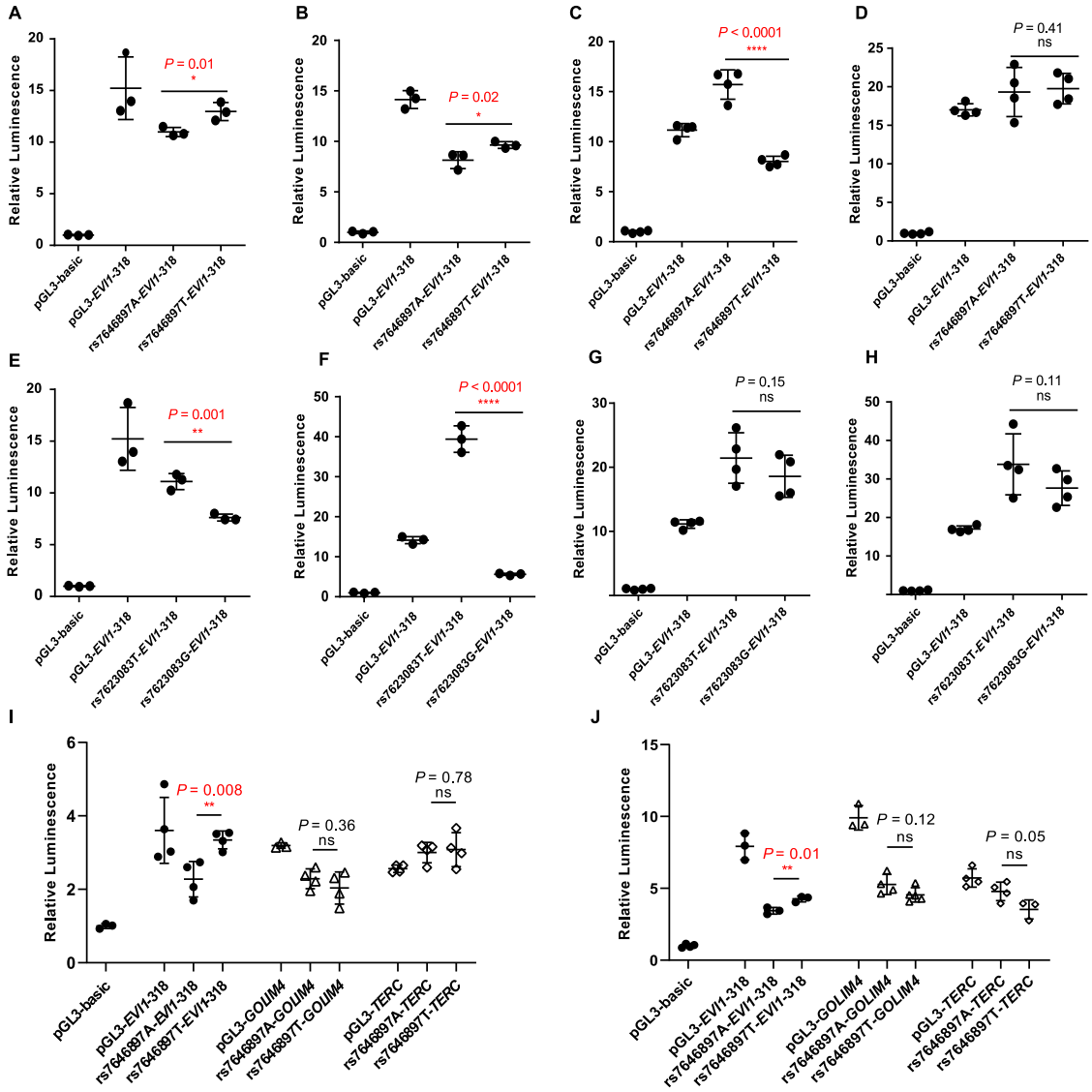


Figure S2. Independent luciferase reporter assays on candidate regulatory elements carrying alternate alleles at SNPs rs7646897 in HEK293T cells, related to Figure 2

(A-D) Four independent luciferase reporter assays to test the effect of substitutions at rs7646897 in modulating the expression of *EVI1*. Three out of four experiments confirmed regulatory activity. Two of them showed a consistent direction of change.

(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 of *EVI1* is inconsistent.

(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).

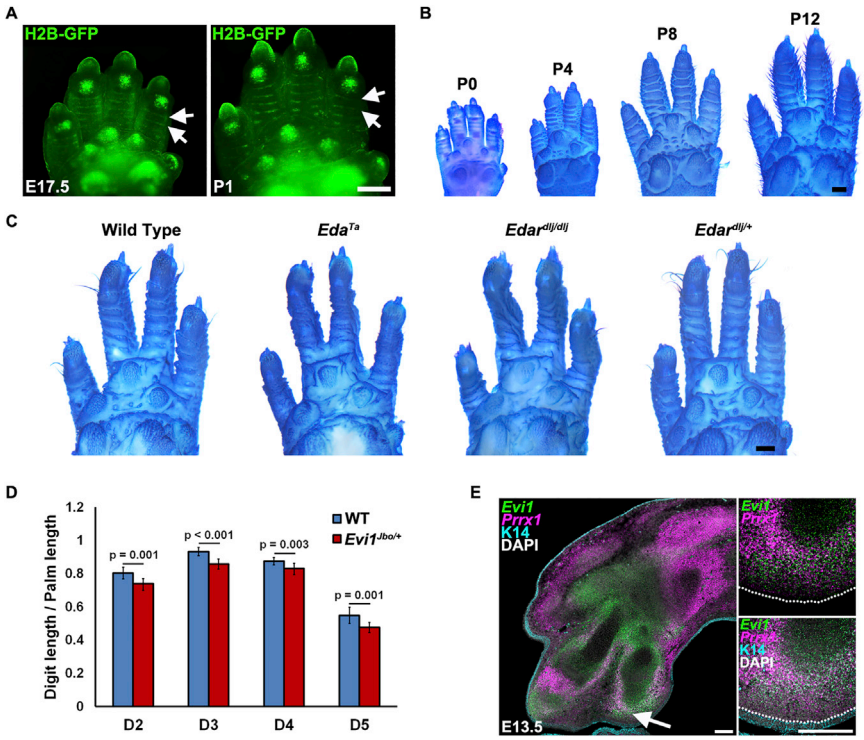


Figure S3. Development of digit ridge pattern in mouse and *Evi1*^{Jbo} effect on digit proportions and coexpression with *Prrx1*, related to Figure 3

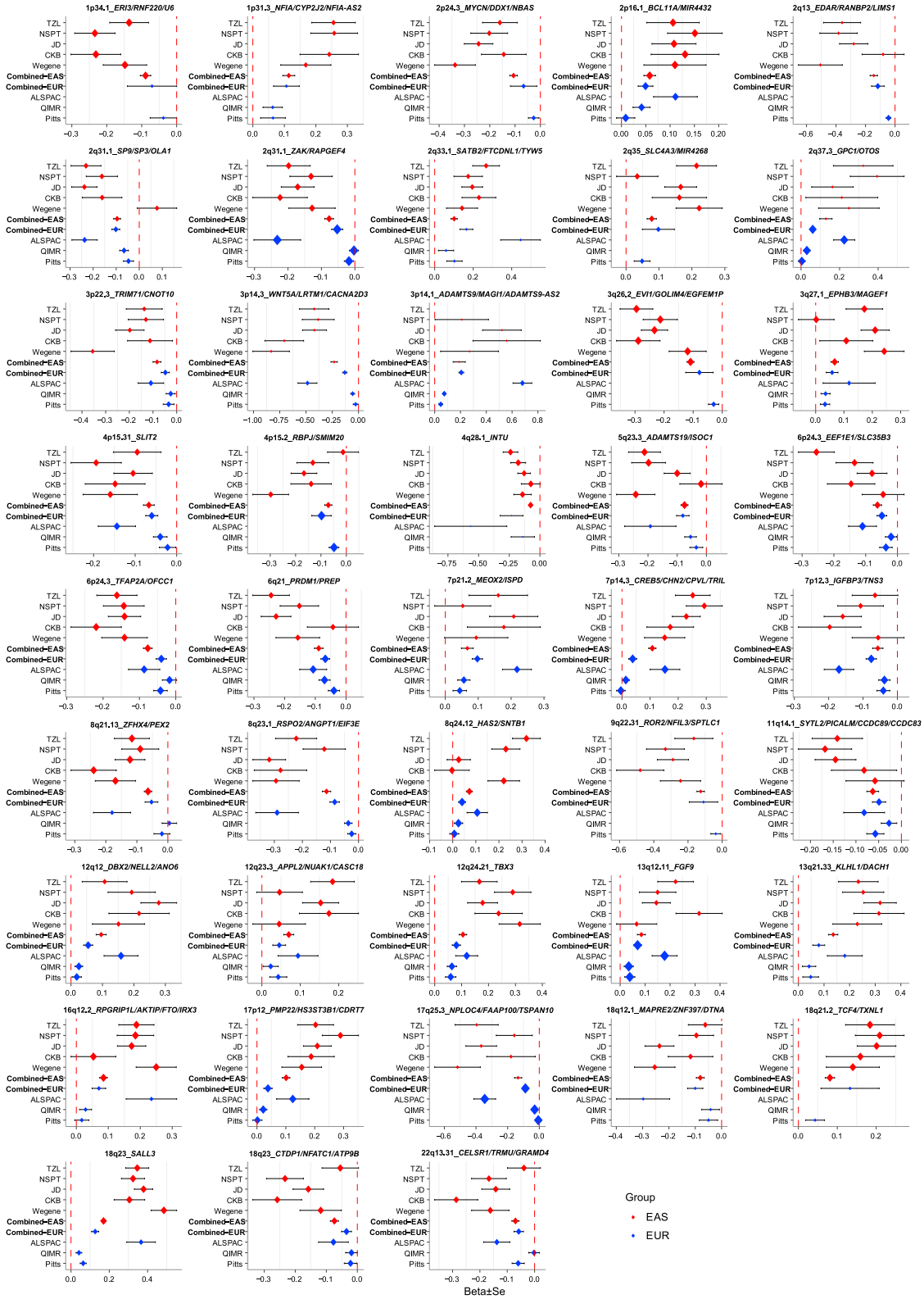
(A) Embryonic origin of digit ridge pattern in mouse. Ventral views of forelimbs of mice carrying a TCF/Lef::H2B-GFP WNT pathway reporter gene at embryonic day 17.5 (E17.5) and newborn (P1). Transverse digital ridges (examples indicated by arrows) are apparent from E17.5. Scale bar = 500 μm.

(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. Palmar dermal surface of toluidine blue stained right forepaws from wild-type mice. Scale bar, 500 μm.

(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 and the distorted ridge pattern in the mutants. Scale bar, 500 μm.

(D) Digit length, normalized to palm length, in wild-type (WT) and *Evi1*$^{Jbo/+}$ heterozygous adult mice. Digits are shorter in mutant animals.

(E) 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 mesenchymal cells, with *Evi1* also expressed 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.



(legend on next page)

Figure S4. Effect sizes (regression coefficients) for the derived allele at index 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 Allele1. 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.

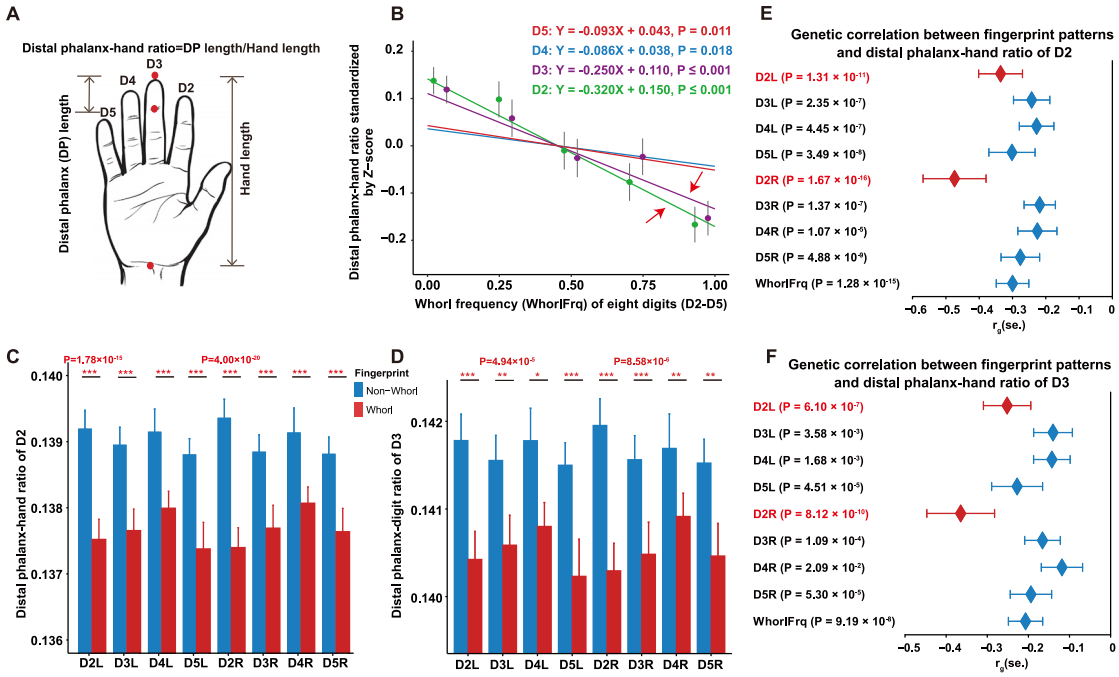


Figure S5. Association between fingerprint patterns and hand proportions, related to Figure 5

(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 short black lines were standard deviation for each group. The arrow indicates the linear regression passes the significance test.

(C-D) Bar plot of fingerprint pattern type (non-whorl versus whorl) of each digit (D2-D5) and the mean DPHR of D2 or D3. Error bars indicate SEM. *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 SEM.

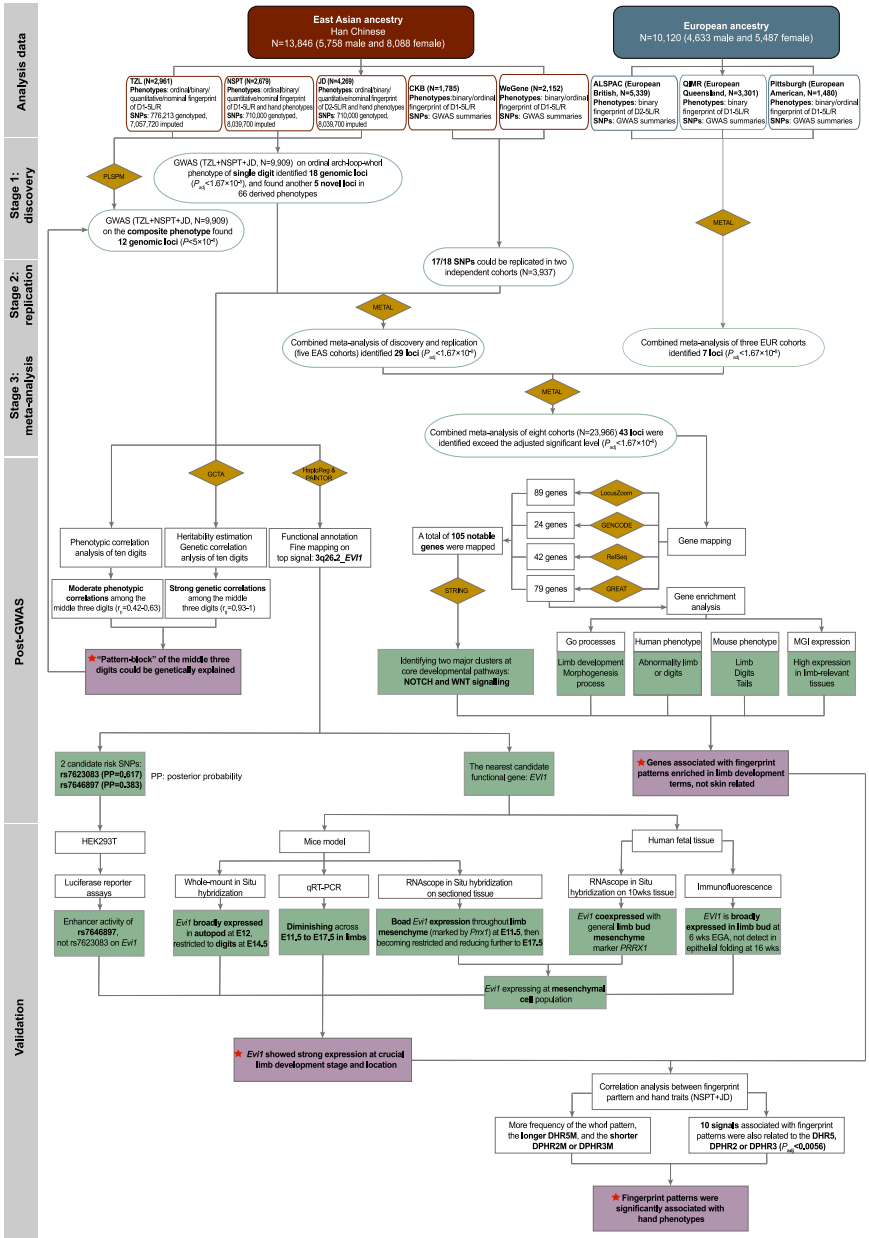


Figure S6. Flow chart of analyses conducted in the study of fingerprint patterns, related to STAR Methods

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.