Title

Genetic variants underlying differences in facial morphology in East Asian and European populations

Author List

Manfei Zhang[1,2,3,15], Sijie Wu[1,2,4,15], Siyuan Du[2,15], Wei Qian[1,2,3,15], Jieyi Chen[2,4], Lu Qiao[2], Yajun Yang[4], Jingze Tan[4], Ziyu Yuan[5], Qianqian Peng[2], Yu Liu[2], Nicolas Navarro[6,7], Kun Tang[2], Andrés Ruiz-Linares[1,4,8,9], Jiucun Wang[1,5], Peter Claes[10,11,12,13], Li Jin[1,2,5,*], Jiarui Li[2,10,11,*] & Sijia Wang[1,2,14,*]

Affiliations

[1] State Key Laboratory of Genetic Engineering, Human Phenome Institute, Zhangjiang Fudan International Innovation Center, Fudan University, Shanghai, China.

[2] CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China.

[3] School of Computer Science, Fudan University, 2005 Songhu Road, Shanghai, China.

[4] Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China.

[5] Fudan-Taizhou Institute of Health Sciences, Taizhou, Jiangsu 225326, China.

[6] Biogéosciences, UMR 6282 CNRS, EPHE, Université Bourgogne Franche-Comté, Dijon 21078, France.

[7] EPHE, PSL University, Paris 75014, France.

[8] Aix-Marseille Université, CNRS, EFS, ADES, Marseille, 13005, France.

[9] Department of Genetics, Evolution and Environment, and UCL Genetics Institute, University College London, London WC1E 6BT, UK.

[10] Department of Electrical Engineering, ESAT/PSI, KU Leuven, Leuven, Belgium.

[11] Medical Imaging Research Center, UZ Leuven, Leuven, Belgium.

[12] Department of Human Genetics, KU Leuven, Leuven, Belgium.

[13] Murdoch Children's Research Institute, Melbourne, Victoria, Australia.

[14] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223 China.

[15] These authors contributed equally: Manfei Zhang, Sijie Wu, Siyuan Du, Wei Qian.

[*] Corresponding authors.

Sijia Wang, wangsijia@picb.ac.cn

Jiarui Li, lijiarui@picb.ac.cn

Li Jin, lijin@fudan.edu.cn

Abstract

Facial morphology, a conspicuous feature of human appearance, is highly heritable. Previous studies on the genetic basis of facial morphology were mainly performed in European-ancestry cohorts (EUR). Applying a data-driven phenotyping and multivariate genome-wide scanning protocol to a large collection of 3D facial images of individuals with East Asian ancestry (EAS), we identified 244 variants in 166 loci (62 novel) associated with typical-range facial variation. A newly proposed polygenic shape analysis indicates that the effects of the variants on facial shape in EAS can be generalized to EUR. Based on this, we further identified 13 variants related to differences between facial shape in EUR and EAS populations. Evolutionary analyses suggest that the difference in nose shape in EUR and EAS populations is caused by a directional selection, mainly due to a local adaptation in Europeans. Our results illustrate the underlying genetic basis for facial differences across populations.

Main Text

Facial morphology has substantial variations at the individual and population level. Multiple studies show significant differences in craniofacial morphology across people from different geographic regions[1-18]. For example, individuals of European ancestries (EUR) have a more protruding nose and brow ridges than those of East Asian ancestries (EAS)[19]. Such differences must have a strong genetic basis, which remains unknown, largely due to the low number of studies in performed in EAS populations compared to EUR. Previous genetic studies collectively reported about 219 loci associated with facial morphology in EUR populations, but only 24 were reported in EAS or Eurasian-ancestry populations, about 20 in Latin American-ancestry populations, and four in African-ancestry populations (Supplementary Table 1)[3-5,9,12-14,16,17]. Large-scale studies in EAS populations and other non-European populations are much needed to provide a complete architecture of the genetic basis of facial morphology, particularly the observable differences across populations.

Here we performed a genome-wide association study (GWAS) based on a large collection of 3D facial images from the Han Chinese population. Using a data-driven phenotyping approach, we identified hundreds of associated variants[10,18]. In addition, we identified specific variants distinguishing facial appearance between EUR and EAS populations. We further provided evidence that those population-based facial differences, especially for nose shape, were under selection. A schematic overview of our study design can be found in Extended Data Figure 1.

GWAS on facial phenotypes discovered 244 leading variants.

To study facial variation from a global to local scale, we used 3D facial surface scans from a large-scale EAS population (Methods, Supplementary Table 2) in a discovery (n = 6,968) and replication cohort (n = 2,706), and subsequently combined them in a meta-analysis. A semi-supervised phenotyping procedure defined 63 hierarchically arranged facial segments using the discovery cohort (Methods). Next, we performed a canonical correlation analysis (CCA) based GWAS on each facial segment's group of principal components (Methods). Subsequently, we identified 50 independent tests using parallel analysis and permutation test (Methods, Supplementary Note)[18,20,21]. Thus, besides conventional genome-wide significant threshold ($P = 5\times10^{-8}$), we set a stricter study-wide significant threshold to $P = 9.8\times10^{-10}$ ($P = 5\times10^{-8}/51.41$) after Bonferroni correction for multiple testing. In the discovery datasets, we identified 153 genome-wide significant variants ($P < 5\times10^{-8}$, minor allele frequency (MAF)>0.05) in 124 loci after condition analysis and peak selection (Methods). In the replication dataset, 119 out of 153 (77.8%) variants were replicated at nominal significance (p<0.05), 118 (77.1%) were replicated at FDR<0.05, 75 (49.0%) were replicated at a Bonferroni-corrected significance ($P < 3.28\times10^{-4}$). For the 84 variants that passed the study-wide significance ($P < 9.8\times10^{-10}$), 80 out of 84 (95.2%) were replicated at nominal significance (p<0.05) and FDR<0.05, 65 (77.4%) were replicated at a Bonferroni-corrected significance ($P < 5.95\times10^{-4}$). To increase statistical power, we performed a meta-analysis using Stouffer's method to combine the P-values obtained from the discovery and replication cohort[18,22]. As a result, we identified 244 independent variants in 166 loci under the genome-wide threshold associated

with normal range facial variation (Fig. 1a; Supplementary Table 3) where 151 variants in 106 loci were study-wide significant. According to the anatomical structure, we classified the 244 genome-wide significant variants in ten facial regions (Supplementary Note), including forehead, glabella, eye, tempora, zygoma, nose, maxillary, upper mouth, lower mouth, and mandible (Fig. 1b; Supplementary Table 4). The nose was the feature associated with the most variants (107) out of the ten regions. The numbers of variants associated with the other nine regions were: glabella (35), upper mouth (32), eye (29), zygoma (28), maxillary (25), mandible (20), forehead (16), lower mouth (13) and tempora (12). In addition, we calculated the genome-wide heritability of each 3D facial segments at each level. The 3D facial segments' heritability ranged from 7.47% to 52.3%. As we expected, the nose segments were also among the most heritable area (42.17% to 46.86%, Supplementary Fig. 1).

We considered a variant as novel when it was not in linkage disequilibrium (LD, $r^2 < 0.1$) with previously reported variants in facial GWASs ($P < 5\times10^{-8}$, Supplementary Table 1) in any East Asian (EAS), European (EUR), and African (AFR) populations in the 1000 Genomes Project phase 3 (1000GP, Methods)[23]. As such, 130 of the 244 leading variants under the genome-wide threshold were novel, while 65 of the 151 under the study-wide threshold were novel (Supplementary Table 3). We considered a genetic locus as novel when it did not overlap with previously reported genomic loci associated with facial variation. As such, 62 out of 166 loci were novel.

We used FUMA and GREAT to annotate the leading variants[24,25]. As a result, we identified 206 candidate genes potentially associated with facial variation, among which 100 genes were not reported in previous facial GWASs

(Supplementary Table 3). We found that the genes associated with the leading variants were highly enriched in biological processes of skeletal system development and morphogenesis (Extended Data Fig. 2a). Moreover, the epigenome and transcriptome datasets showed that the leading variants were mainly enriched for enhancers in craniofacial tissues. Compared with the ectoderm at the later stage of fetal development, candidate genes were significantly highly expressed in the mesenchyme (Extended Data Fig. 2b-c), consistent with our expectation and previous studies[18,26-30].

We investigated associations ($P < 1\times10^{-5}$) with other complex traits for the 244 variants through PhenoScanner, a web-based GWAS repository (Methods)[31]. We found that the traits sharing high genetic components with facial shapes mainly involved physical measurements, body composition and hair morphology (Supplementary Table 5). The highest co-association was the risk of atrial fibrillation (*PITX2*, rs6843082, $P = 3.0\times10^{-155}$). This variant has been reported to be associated with cardioembolic stroke and ischemic stroke, suggesting that facial features might be a biomarker of cardiovascular disease.

Characteristics of specific variants in EAS and EUR

By comparing the 244 leading variants identified in our study with the 203 leading variants reported in a recently published EUR study using similar phenotyping and analysis framework, 89 variants were shared in both studies (Methods)[18]. The remaining 155 and 114 variants in the EAS and EUR studies respectively were different (Fig. 2a; Supplementary Table 6). Therefore, we defined three different groups: 89 shared variants, 155 EAS-specific variants, and 114 EUR-specific variants.

To understand these shared and population-specific variants' characteristics, we examined their allele frequencies in EAS and EUR populations based on the 1000GP[23]. Comparison of cross-population minor allele frequency (MAF) showed that the groups of population-specific variants had higher MAF in their respective population ($P_{EAS\text{-}specific} = 9.0 \times 10^{-12}$, $P_{EUR\text{-}specific} = 8.1 \times 10^{-9}$,). In contrast, the shared variants had no difference of MAF between the two populations ($P_{shared} = 0.1$) (Fig. 2b-e). These results suggest that higher MAFs may increase the statistical power to detect the variants associated with facial variation within the respective populations. Moreover, 77.5% (69 out of 89) of the shared variants passed the study-wide significance threshold while only 54.2% (84 out of 155) EAS-specific variants and 55.2% (63 out of 114) EUR-specific variants passed the study-wide significance threshold (chisq.test $P_{EAS} = 3.2\times10^{-4}$; $P_{EUR} = 4.8\times10^{-6}$).

To explore potential biological functional differences between the shared and population-specific variants, we used Metascape to compare the differences of enriched terms for their annotated genes[32]. We found that EAS-specific, EUR-specific, and shared genes were all associated with top terms that were previously reported in association with craniofacial variation (Fig. 2f). Moreover, we observed a considerable number of functional overlaps among the three groups (Fig. 2g). These results indicate that the associated genes across populations share substantially similar biological processes. We next compared epigenetic regulation patterns of the shared and population-specific variants in various cell-types or tissues. The shared and population-specific variants were all enriched for enhancers in craniofacial tissues (Fig. 2h), again indicating that the potential functions of facial variants across populations are analogous.

As expected, we found that the population-specific variants had a significantly higher Wright's fixation index (F_{ST}) than random variants from a genome-wide background both in EAS and EUR populations ($P_{EAS\text{-}specific}$= 2.42×10^{-10}; $P_{EUR\text{-}specific}$ = 0.0063, Extended Data Fig. 3), while the shared variants had no significant difference compared with random variants ($P_{EAS\text{-}shared}$ = 0.599)[33,34]. The same result applied to the Cross Population Extended Haplotype Homozygosity (XP-EHH) analysis using REHH2 in these two populations ($P_{EAS\text{-}specific}$ = 0.0078; $P_{EUR\text{-}specific}$ = 0.038; $P_{EAS\text{-}shared}$ = 0.449)[35,36]. These results suggest that facial variation across populations could be attributed to random drift and natural selection. The population-specific variants explained a larger proportion of natural selection while the shared variants may mainly explain random drift influencing facial variation.

In summary, the population-specific variants found in EAS and EUR were identified due to cross-population MAF differences and subtle effect sizes. Nonetheless, similar biological processes underlying facial variation were shared across populations.

Polygenic shape analysis generalizes results from EAS to EUR

To explore the genetic basis of EAS facial shape and the genetic factors contributing to the difference in facial shape between EAS and EUR, we first investigated whether the association results of the leading variants in our EAS study could be generalized to EUR.

We introduced a novel polygenic shape analysis to investigate whether the differential accumulated genetic effects between the two populations of the leading variants is in line with the actual population facial differences. Similar to

the classic polygenic score analysis, we defined the Polygenic Shape (PS) for an individual as the sum of the number of effect risk alleles weighted by risk allele effect size of all the leading variants[37]. Similarly, the Polygenic Population Shape (PPS) is the average polygenic shape for a given population (Methods). Using data of EUR (*n* = 404) and EAS (*n* = 208) individuals from 1000GP, we calculated the PPS of the two populations for the whole face and ten anatomical facial regions.

To visualize and compare this effect to the true population facial shape average for each facial region, we constructed EUR and EAS PPS-derived faces by adding and subtracting, respectively, $(\boldsymbol{PPS}_{EUR} - \boldsymbol{PPS}_{EAS})/2$ to a population neutral average face, which was constructed as the average of the EUR and EAS population average shapes (Methods). We used 3D facial scans of EAS and EUR individuals to calculate each population's average face and therefore generated EUR and EAS average faces. Compared with the average face of EUR, we found that EAS had more protrusion in the cheek; more concavity in the forehead, glabella, nose, and mandible (Fig. 3a-i), which were consistent with a previous study[19]. Interestingly, when we amplified the differentiated accumulated genetic effects five times, the PPS-derived faces looked very similar to EAS and EUR's actual average face (Fig. 3a-ii). The EUR and EAS PPS-derived faces showed similar facial variation to the ground truth, especially in the glabella and nose region (Fig. 3a-ii).

To test the generalization of the association results from our study to EUR, we compared the EUR and EAS PPS-derived faces of the whole face using all the 244 leading variants with the PPS of the whole face using 244 variants randomly chosen from the genome. The EAS and EUR PPS-derived faces

using the leading variants were significantly more similar to the true population average faces than that using random variants, either measured by Euclidean distance (P_{dis} = 0.007) or cosine similarity (P_{EUR} = 0.004; P_{EAS} = 0.005; Fig. 3b). Moreover, we calculated individual facial polygenic shape for EUR and EAS individuals from 1000GP and measured their East Asian ancestry facial appearance (EAS-FA), defined as the projected length of the individual's polygenic shape onto the explicit EAS-EUR shape difference (Methods, Supplementary Fig. 2). The EAS-FA of both EUR and EAS groups were significantly separated, with EUR individuals closer to the EUR face and EAS individuals closer to the EAS face (t.test $P < 2.2\times10^{-16}$, Fig. 3c).

We also performed the same analyses locally for ten anatomical facial regions. The EAS and EUR PPS-derived shapes using the leading variants were significantly more similar to the population average shapes than those using random variants, involving the upper mouth, nose, maxillary, glabella, eye, tempora and zygoma in all the three measurements of similarity including cosine similarity, Euclidean distance, and EAS-FA (Extended Data Fig. 4a). Notably, the nasal region performed the best among the ten anatomical regions ($P_{dis} = 4.8\times10^{-33}$; $P_{EUR} = 3.7\times10^{-37}$; $P_{EAS} = 3.4\times10^{-37}$; Fig. 3d-f). However, we obtained non-significant results for the mandible, forehead, and lower mouth (Extended Data Fig. 5). These results indicate that the PPS constructs morphological variations in most facial regions between EUR and EAS. Our results demonstrates that the PPS-derived faces using the leading variants are similar, both visually and statistically, to the true population average shapes at global and local scales, suggesting that the facial shape effects identified from the EAS study generalize well to EUR populations.

Variants contribute to EAS facial appearance (EAS-FA)

Among the leading variants, we aimed to find those variants that make EAS have more East Asian facial features, in other words, those variants that increase East Asian facial appearance (EAS-FA). We defined the contribution of a variant to EAS-FA as the effect allele frequency weighted projected length of its effect size vector onto the EAS-EUR shape difference (Methods, Supplementary Fig. 3, and Supplementary Table 7). A variant with a positive EAS-FA contribution may cause the EAS population's facial morphology to increase EAS-FA. In contrast, a variant with a negative EAS-FA contribution may cause the EAS population to increase EUR facial appearance. In brief, the variants with large positive EAS-FA are of our interest. In each facial region, we constructed a distribution of the variant's contribution to EAS-FA using 244 leading variants. We further calculated whether each leading variant has significantly higher contribution to EAS-FA than the distribution after Bonferroni correction ($P < 0.005$, Extended Data Fig. 6, and Supplementary Table 8). As a result, 13 variants that passed filtering were considered to increase EAS-FA (Table1; Supplementary Fig. 4).

The 13 variants had a higher F_{ST} than the other variants (t.test $P < 1.0\times10^{-16}$), indicating that they have significant allele frequency differences between EUR and EAS populations. The Population Branch Statistics (PBS) values of the variants were significantly higher in EAS ($P < 1.0\times10^{-16}$) relative to EUR and YRI (Yoruba), but not in EUR ($P = 0.188$) relative to EAS and YRI, which suggests that these variants may be under selection in EAS[38]. Thus, these variants potentially contribute to the morphological differences between EUR

and EAS. Most of the EAS-FA variants might be standing genetic variations, as the alternative allele frequency was relatively high, given the evolutionary time, as shown in Table 1. Notably, six variants had F_{ST} above 0.5 between EUR and EAS. Furthermore, three of these variants affected the glabella segments, and two affected the nasal region. This result suggests that local adaptation might play a role in forming facial variation between EAS and EUR. Among the 13 variants, six were reported to be associated with facial shape variation. Well-known facial genes such as *EDAR*, *TBX15* and *MRPS22* were associated with craniofacial morphology in many studies[3,4,6,10,13,15,16,18,39-43]. Our study showed a variant in an intron of *TBX15* (rs10923710, NC_000001.10:g.119502774G>T) contributing to maxillary and tempora shape in EAS, consistent with the observation that this locus had multiple spatial effects on the face[18]. The rs12632544 (NC_000003.11:g.138946868T>A) is an intergenic variant near *MRPS22* on chromosome 3q23. It is in LD (r^2 = 0.932) with rs12633011, which was reported associated with morphology of eyes in a previous EAS study[16]. *MRPS22* was reported to be associated with human earlobe size and a mouse skeleton phenotype[3,44,45]. A reanalysis of a GWAS study on cranioskeletal variation in outbred mice showed that variants in the region of chromosome 9, overlapping with *Mrps22*, were significantly associated with craniofacial variation (FDR < 5%, Supplementary Fig. 5). These variants were associated with the protrusion of the maxillary bone, the shrinkage of the eye and malar bone. In our study, rs12632544 contributes to EAS-FA in the glabella, eye and tempora (Extended Data Fig. 4b), in line with the result in the outbred mice. We also identified seven novel variants contributing to EAS-FA, five out of which are located closely to novel genes. Some of these have been reported in the

context of craniofacial dysmorphology. For instance, the rs6669519 (NC_000001.10:g.75584009T>A), which contributes to the shape of glabella, is an intergenic variant near *LHX8* on chromosome 1p31.1. *LHX8* (LIM Homeobox 8) was reported to be associated with cleft palate, forebrain neuron development and differentiation[46,47]. The rs12473319 (NC_000002.11:g.232880971G>C), an intronic variant of the *DIS3L2* gene, showed association with EAS-FA in glabella (Extended Data Fig. 4c). *DIS3L2* was found to affect a skeleton phenotype in a mouse genome study[45,48]. In addition, *DIS3L2* is a candidate gene for the Perlman syndrome, characterized by craniofacial abnormalities. Moreover, the frequency of the derived C allele is higher in EAS (47.8%) than EUR (2.2%). The estimated allele age of this variant is about 7,875 years old (Table 1), when EAS and EUR were already separated, suggesting that this variant arose in East Asian populations[49].

Interestingly, there are four independent variants associated with nose morphology in the *SOX9* locus in our EAS GWAS (Extended Data Fig. 7). Although *SOX9* is a well-known gene contributing to the variation in nose shape, rs8068343 (NC_000017.10:g.69447706T>C) is a novel variant that affects nose shape differences between populations[7,9,10]. In contrast, the other three variants may affect nose shape within populations. Compared with the variants previously identified near *SOX9*, the reference (T) allele of rs8068343 has a lower frequency in EUR (0.04) than in EAS (0.46). Moreover, this variant had a significantly higher F_{ST} (0.528) between EUR and EAS than any other variants, and the Integrated Haplotype Score (iHS) of this variant was also significantly higher in CHB (Han Chinese, 2.55)[50]. These results indicate that this variant contributes to EUR-EAS nose shape difference.

Nose-associated variants are under positive selection

To assess whether the variation of facial morphology in EAS and EUR populations is due to natural selection or random drift, we conducted several selection analyses of the leading variants discovered here. The F_{ST} enrichment test showed that regions of the whole face and nose had a significantly higher F_{ST} than random variants after Bonferroni correction ($P_{whole\ face}$ = 8.22×10^{-7}, P_{nose} = 1.00×10^{-4}, Supplementary Table 9; Fig. 4a), indicating that facial morphology has been under natural selection in EAS and EUR, especially in the nose region[51]. XP-EHH enrichment analysis showed a consistent result ($P_{whole\ face}$ = 3.78×10^{-3}, P_{nose} = 1.27×10^{-3}, Supplementary Table 9; Fig. 4b).

Next, the mean PBS values for the nose-associated loci were significantly higher than random variants in EUR (P = 6.90×10^{-4}) but not in EAS (P = 0.15) (Supplementary Table 9, Fig. 4c), indicating that nose shape may be under subtle local selection in Europeans rather than in East Asians. In addition, we found that mean PBS values for the nose-associated loci in a recent published EUR facial GWAS were significantly higher in EUR (P_{EUR} = 9.46×10^{-3}) but not in EAS (P_{EAS} = 0.464; Fig. 4d)[18]. The results support that nose shape may be under local selection in EUR rather than in EAS populations.

Based on a study of He et al., we estimated and tested differences using the selection coefficients for nose variants with positive contribution to East Asian nose appearance (EAS-FA in the nose segment)[52]. The EUR population showed higher selection coefficients for nose EAS-FA increasing variants than the EAS population (P = 4.88×10^{-2}, Supplementary Table 9, Fig. 4e). Moreover, by comparing the mean genetic prediction of EAS and EUR's facial variation to

the expected difference under random drift (Methods), the nose and glabella morphology in EUR was more protruding than EAS and the divergence of the nose was greater than expected under the neutral model (Fig. 4f)[34]. Furthermore, by comparing the PPS using the leading variants with the expected PPS under random drift in the EUR and EAS, we obtained the direction (and significance) of natural selection on facial morphology in each population[34]. Similar to the population differences, the nose, glabella, and zygoma were under significant natural selection in EUR (Fig. 4g). However, in EAS, the effects of natural selection were not significant (Fig. 4h). These results suggest that facial morphology in EUR may undergo local adaptation, producing a more protruded nose, glabella, and flatter zygoma.

Based on the above results, we speculate that facial features underlying EUR-EAS differences are potentially due to the adaptive selection that occurred in the EUR population, which makes European-ancestry populations have protruded and narrow noses, significantly different from those of East Asian-ancestry populations.

Discussion

As a large-scale East Asian population facial GWAS using a data-driven global-to-local phenotyping, our study broadens the knowledge of craniofacial genetics outside frequently investigated European-ancestry populations. Compared to previous facial GWASs, we identified 130 (out of 244) novel variants associated with typical range facial variation, which have similar biological functions as the variants identified previously[1-18]. A considerable number of shared variants were independently identified in a EUR study and

our EAS study, using the same facial phenotyping approach. Among the 114 known variants, 96 were associated with consistent facial regions reported in the previous facial GWASs. When we compared shared variants with EUR study of White et al., 82 out of 89 were associated with the same facial regions, which indicates that different segmentation patterns could obtain similar GWAS results (Supplementary Table 3, 6)[18]. These results suggest that the 244 variants identified in our study are reliable and the genetic factors underlying facial variation might be universal across populations.

We further extended the concepts of polygenic scores (PGS) to polygenic shapes (PS) in order to verify whether the association found in EAS could be generalized to EUR[37]. Both visual and statistical evidence supported this hypothesis on the whole face and major anatomical facial regions. However, the PPS of the mandible, forehead, and lower mouth showed some difference from the corresponding EUR and EAS average shapes, which is mainly due to the insufficient number of significant variants associated with these facial regions. Besides smaller phenotypic variation in these regions in EAS, environmental factors contributing to facial variation may also impact the results. In addition, the Q_{ST} analysis suggests that mandible, lower mouth, and forehead exhibit fewer signals of facial shape differences between EAS and EUR[19]. This could explain some reason for the inconsistency of the PPS for these facial regions. Of future interest is to calculate the PPS by combining all the variants identified in EAS and EUR studies. This might further improve the PPS in representing population facial shapes.

Our study also provides insights into the genetic basis of the facial shape differences between European-ancestry populations and East Asian-ancestry

populations. In addition to identifying 13 primary variants contributing to EUR-EAS facial differences, we provided a method to investigate the genetic factors associated with inter-population facial variation. These 13 variants all had positive and larger effects on EUR-EAS facial differences, shaping the faces of East Asian populations to be more EAS-FA. Again, corresponding with the PPS results, due to the innate limitation of GWAS, our study may overlook rare or fixed variants that lead to more EAS-FA in EAS population. By applying our method to EUR populations, additional variants affecting EAS-FA might be discovered. Moreover, for those rare or fixed variants with opposite alleles between EUR and EAS, a single population GWAS has limitations, and an admixture population instead is needed.

Due to the large number of significant variants identified in the nose region, our evolutionary analysis further supports the hypothesis proposed by Zaidi et al. that human nose shape has evolved in response to selection pressures[53]. Again, the PBS analyses showed that nose shape difference between EUR and EAS is mainly due to natural selection in European-ancestry populations rather than in East Asian-ancestry populations[18].

In conclusion, this study presents a large EAS population GWAS on 3D facial shapes. Our study identified a large number of novel variants associated with normal range facial shape variation. Using newly introduced polygenic shapes, we successfully depicted perceptually recognizable population average faces, making our results more tangible, comprehensive, and intuitive. We identified 13 variants contributing to more EAS-FA and revealed natural selection in shaping EUR-EAS nose shape difference. Our findings will greatly facilitate the understanding of human facial morphology across populations.

Acknowledgements

We would like to thank the participants of the NSPT, NHC, and TZL cohorts who consented to participate in research, and the related teams, which include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

This project was funded by the following grants and contracts: Shanghai Municipal Science and Technology Major Project (2017SHZDZX01 to L.J. and S.Wang, 2018SHZDZX01 to W.Q.); National Key Research and Development Project (2018YFC0910403 to S.Wang); CAS Interdisciplinary Innovation Team Project (to S.Wang); "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDB38020400 to S.Wang); Max Planck-CAS Paul Gerson Unna Independent Research Group Leadership Award (to S.Wang); National Natural Science Foundation of China (31521003 to L.J.,.31900408 to M.Z.); 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.); China Postdoctoral Science Foundation (2019M651352 to M.Z., 2020M670984 to W.Q.).

We thank for all suggestions collected during the poster exhibition of ASHG 2019 Annual Meeting. We thank all the participants in these studies.

Author Contributions Statement

S.Wang., L.J., J.L., and M.Z. conceptualized the study (ideas; formulation or evolution of overarching research goals and aims). M.Z, S.Wu., S.D., W.Q. and J.L. carried out the data curation (management activities to annotate, scrub

data and maintain research data for initial use and later re-use). M.Z, S.Wu., S.D., W.Q., J.L. and J.C. carried out the formal analysis (application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data). M.Z, S.Wu., S.D., W.Q. and J.L. did the visualization (preparation, creation and/or presentation of the published work, specifically visualization/data presentation). K.T. and L.Q., Y.Y. and J.T. collected the 3D facial scans of the TZL cohort. J.W., Z.Y., J.T, K.T. and L.Q. collected the 3D facial scans of the NSPT cohort. S.Wu, Y.L, and Q.P. contributed to generating the SNP array data. S.Wu., S.D., and J.L. registered the 3D facial scans of the Northern Han Chinese cohort and conducted the PCA of discovery cohort. N.N. and A.R.L. performed the analysis of *MPRS22*/Mprs22 and human/mouse craniofacial shape. M.Z, S.Wu., S.D., W.Q. and J.L. wrote the original draft. S.Wang., L.J., P.C., J.L., M.Z, S.Wu., S.D., and W.Q. reviewed and edited the final manuscript. All authors participated in preparing the manuscript by reading and commenting on drafts before submission.

Competing Interests Statement

The authors declare no competing interests.

Tables

Table 1: The 13 variants mainly associated with East Asian facial appearance (EAS-FA)

rsID	CytoBand	A1	A2	P-value	EUR	EAS	PBS_EAS	PBS_EUR	F_{ST}_EURvEAS	Seg	Candidate gene	Allele age
rs7516137	1p36.32	C	G	9.75×10^{-29}	0.318	0.553	0.077	0.036	0.107	18	*PRDM16*	537708
rs6669519*†	1p31.1	T	A	3.40×10^{-08}	0.173	0.781	0.831	0.000	0.547	24	*LHX8*	51373
rs10923710	1p12	G	T	1.20×10^{-44}	0.193	0.507	0.262	0.000	0.207	19, 26	*TBX15*	88185
rs3827760	2q12.3	A	G	2.17×10^{-13}	0.000	0.921	2.587	0.306	0.945	27	*EDAR*	36410
rs12473319*¶†	2q37.1	G	C	1.35×10^{-10}	0.022	0.478	0.554	0.161	0.511	24	*DIS3L2*	39385
rs12632544	3q23	T	A	1.87×10^{-65}	0.000	0.500	0.571	0.332	0.595	24, 25, 26	*MRPS22*	626678
rs147468294	6q14.3	A	AC	9.02×10^{-17}	1.000	0.690	0.299	0.207	0.397	7	*TBX18*	46640
rs111847181	8p23.1	G	GAC	5.28×10^{-09}	0.454	0.964	0.779	0.000	0.424	18	*PPP1R3B*	788343
rs4749259*¶†	10p12.1	T	C	3.88×10^{-29}	0.936	0.584	0.370	0.037	0.334	26	*MKX*	298528
rs12258832*¶†	10p12.1	A	G	1.61×10^{-24}	0.892	0.690	0.133	0.005	0.129	26	*MKX*	64715
rs3740550*	10q26.11	A	G	6.70×10^{-43}	0.994	0.875	0.127	0.029	0.145	19	*RAB11FIP2, EMX2*	64603
rs8068343*	17q24.3	C	T	3.32×10^{-51}	0.959	0.462	0.470	0.281	0.528	18	*SOX9*	1261665
rs9980535*¶†	21q21.3	A	G	3.99×10^{-11}	0.176	0.762	0.357	0.380	0.521	18	*LINC00161*	1030805

* Novel variants in our GWAS finding, which are not in LD ($r^2 < 0.1$) with variants reported in previous facial GWAS studies, see in ST1, ST3.

¶ Novel loci in our GWAS finding, which are not within the same loci (<500kb) in previous facial GWAS studies, see in ST1, ST3.

† Novel genes in our GWAS finding, which are not reported in previous facial GWAS studies.

Figure Legends

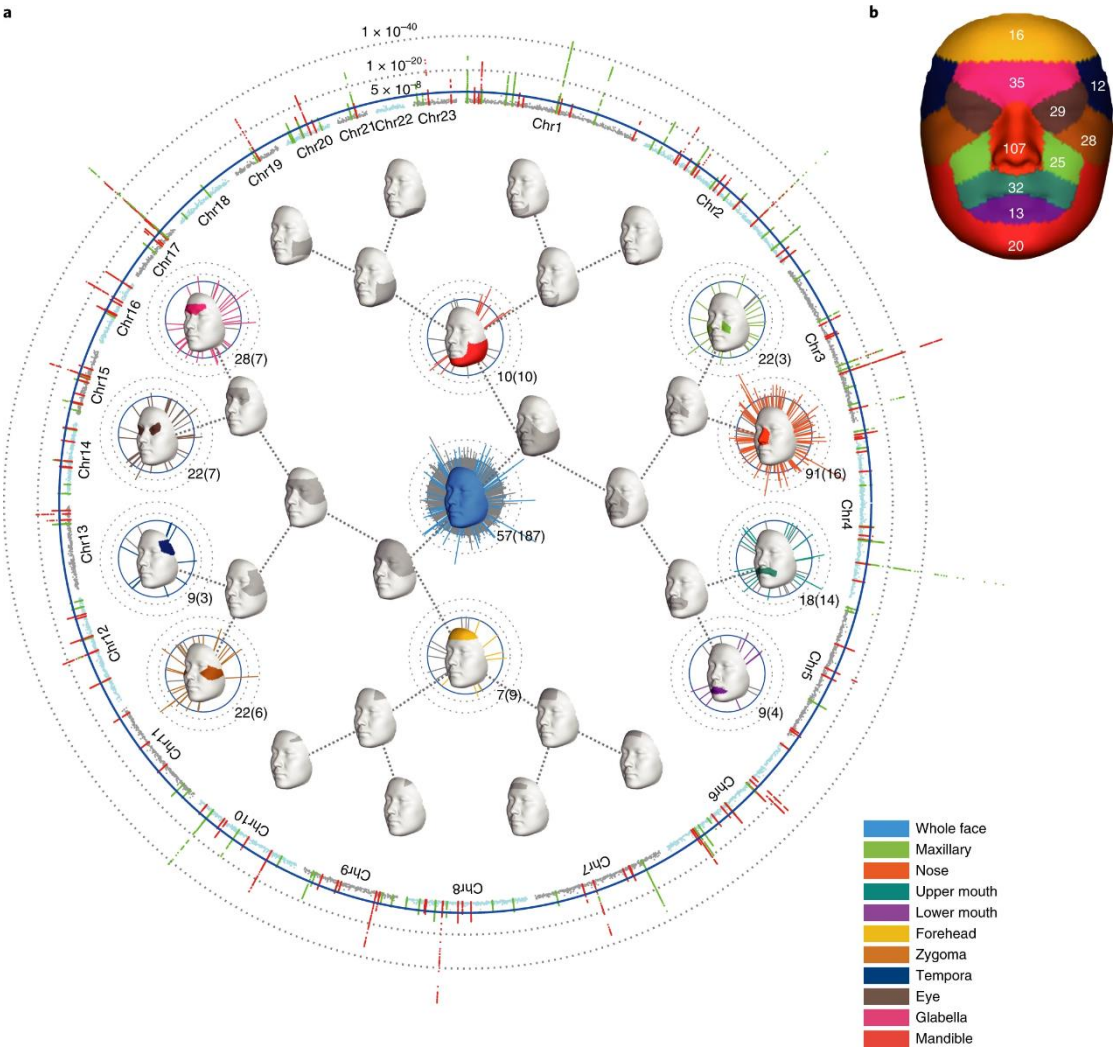


Fig. 1: Overall results of genome-wide association meta-analyses in Han Chinese cohort. a) The Manhattan plot on the ring represents the meta-analyses P-values, with chromosomes colored and labelled. P-values are -log10 scaled. Using 500kb windows, peak variants are colored in red (novel variants) and green (known variants). Inside the Manhattan plot, a binary tree of facial segments illustrates the hierarchical facial segmentation up to level four. On the whole face and ten selected facial segments (colored), variants in the segment itself (colored) and in its derivative segments (gray) are plotted, with the number of variants listed (#variants in derivative segments in parenthesis). b) Visualization of ten selected facial segments and number of significant variants.

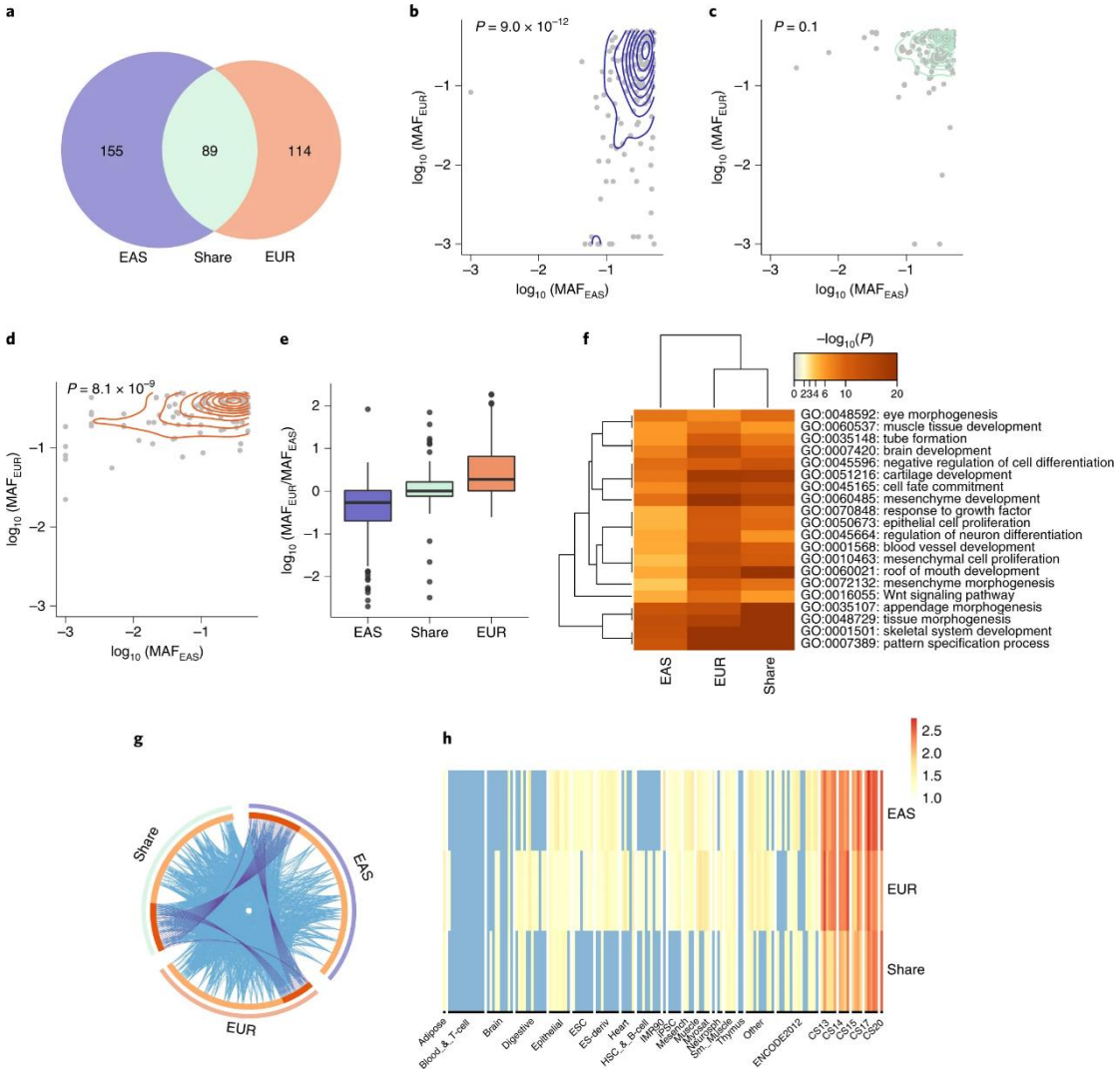


Fig. 2: Comparison of shared and population-specific variants. a) Number of specific and shared variants of two study cohorts (EAS and EUR). b-e) Cross-population MAF comparisons of b) specific variants of EAS (*n* = 155). c) Shared variants (*n* = 178 (89 in EAS study and 89 in EUR study), shared variants in both studies are used), and d) specific variants of EUR (*n* = 114). In b-d), P-values are provided using a two-sided Mann-Whitney U-test. When MAF < 0.001, MAF was truncated to 0.001 to fit the log scale (1000GP). e) MAF ratio comparison of three groups. Boxplots show medians (center lines), first and third quartiles (lower and upper box limits, respectively), 1.5-fold interquartile ranges (whisker extents) and outliers (black circles). Colors are corresponding with a). f) Metascape analysis shows the biological processes associated with genes in the three groups of genes. g) Each

outside arc represents a group, and each inside arc represent a gene list. On the inside, each arc represents a gene list, where each gene has a spot on the arc. The dark orange color represents overlapping genes among groups. Same genes (purple lines) and different genes fell into the same ontology term (blue lines) of three groups. h) Heatmap indicates the global enrichment of trait-associated variants of each group (y axis) in enhancer of different tissue (x axis).

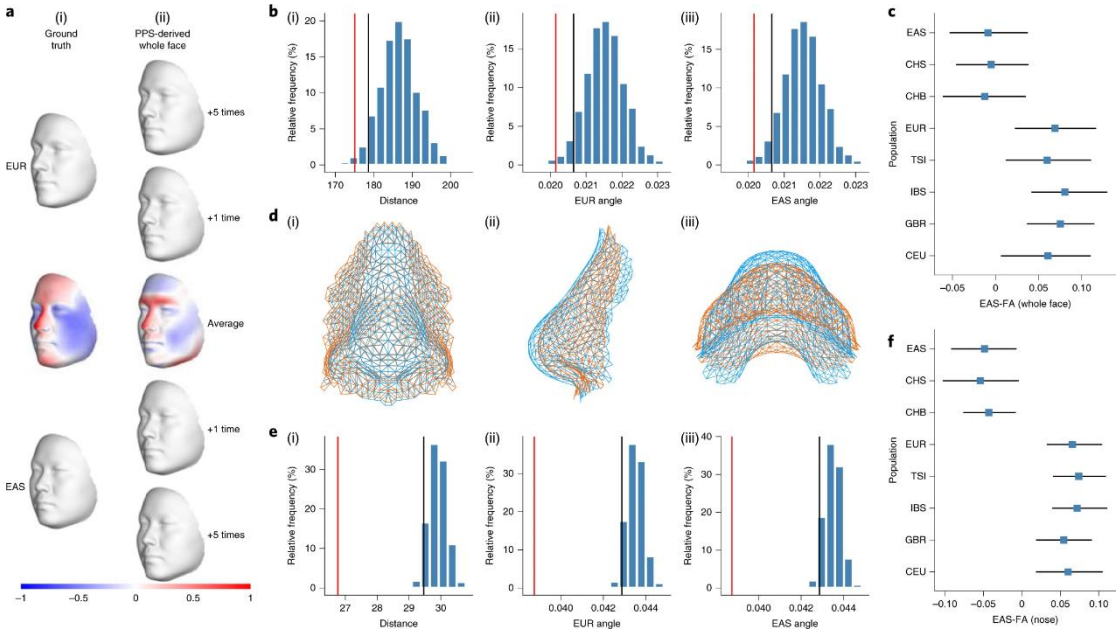


Fig. 3: Visualization of PPS derived whole faces and nose of EAS and EUR and the statistical validations for PPS approach. a) Visualization of facial morphology of EAS and EUR. i) from top to bottom, representing the ground truth of EUR average face, overall average face (EUR and EAS), and EAS average face respectively; ii) from top to bottom, the PPS derived whole faces of 244 leading variants by adding $\boldsymbol{PPS_{EUR} - PPS_{EAS}}$ +5, +1 to -1, -5 times on the overall average facial shape. Differences are visualized using the normal displacement (displacement in the direction locally perpendicular to the facial surface), blue and red refer to depression and protrusion in local shape respectively. b) The null distribution (blue) of i) Euclidean distance, ii) Cosine similarity with EUR average face, and iii) Cosine similarity with EAS average face using 1,000 simulations from 244 random variants on the whole face, red line infers the statistics of the 244 leading variants, black line infers 95% quantile of distribution from the random variants. c) The EAS-FA of polygenic shapes (whole face) for individuals in 1000GP. The squares represent the mean EAS-FA score and the horizontal lines represent 1st and 3rd quantile. d) The $\boldsymbol{PPS_{EUR} - PPS_{EAS}}$ difference of nose region in three views. i), ii), and iii) Front, side, and vertical views, respectively. The PPS derived nose of EUR and EAS are presented in the blue and orange color respectively. e) The distribution of i) Euclidean distance. ii) Cosine similarity with EUR

average nose and iii) Cosine similarity with EAS average nose using 1,000 simulations from random variants, red line infers the statistics of the 107 leading variants associated with nose, black line infers 95% quantile of distribution from the random variants. f) The EAS-FA of polygenic shapes (nose only) for individuals in 1000GP. The squares represent the mean EAS-FA score for the nose and the horizontal lines represent 1st and 3rd quantile.

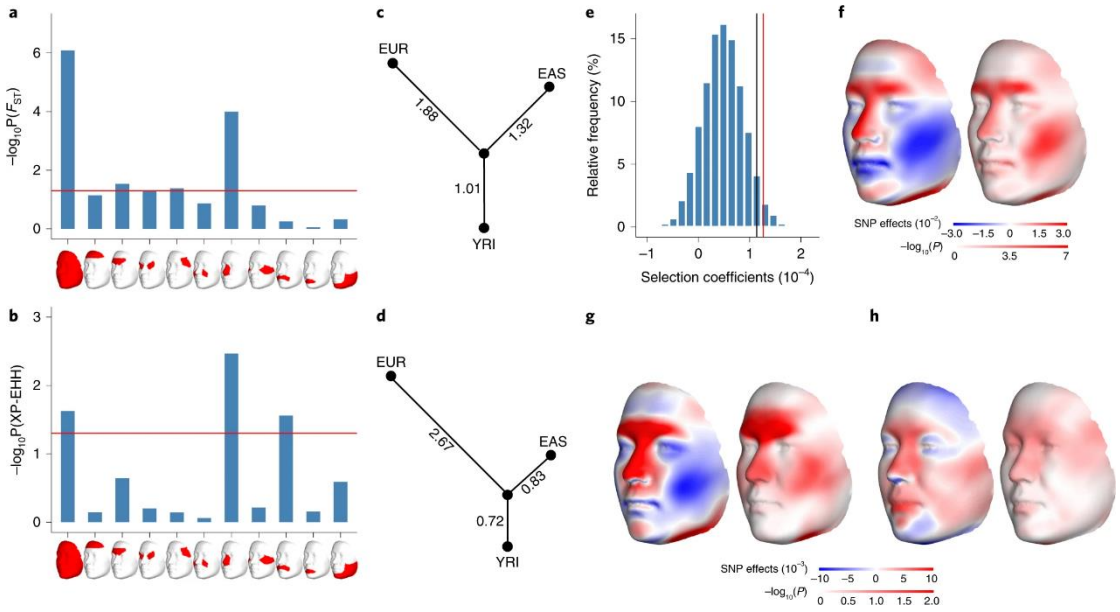
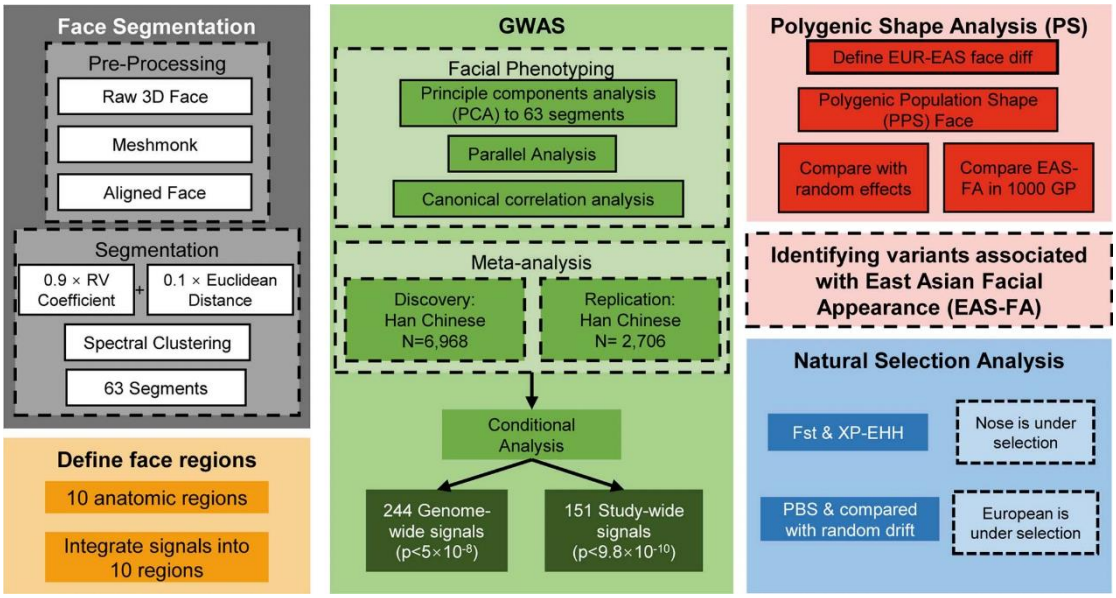
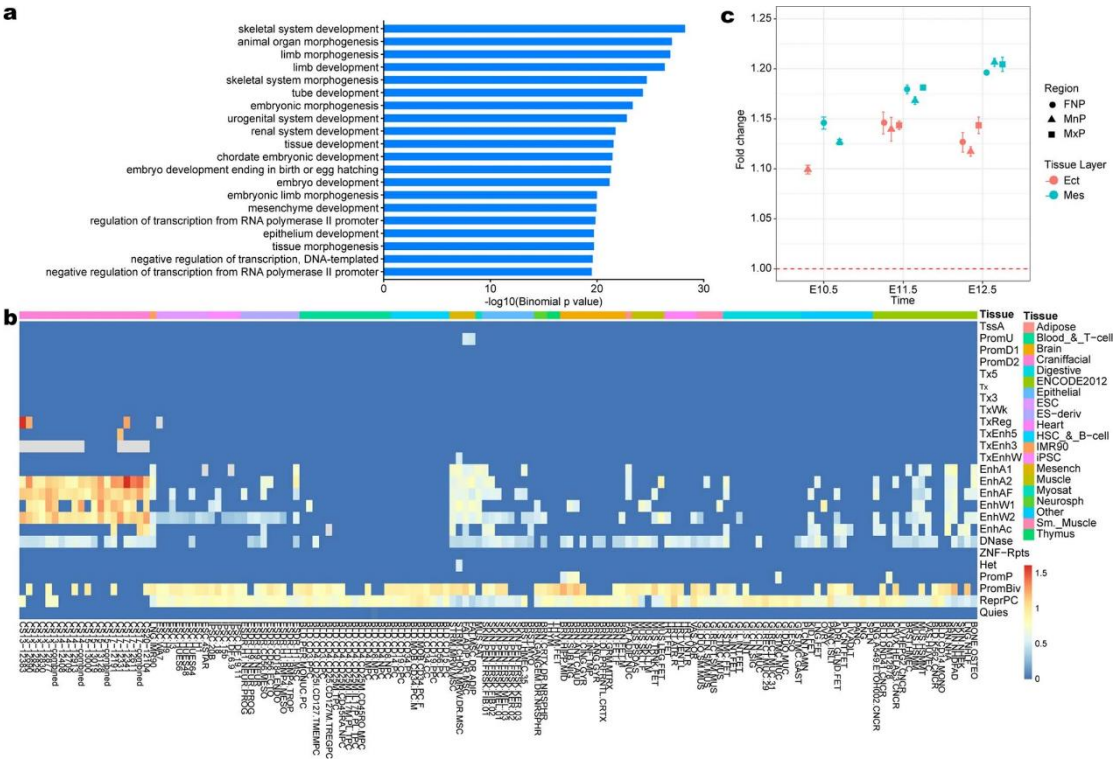


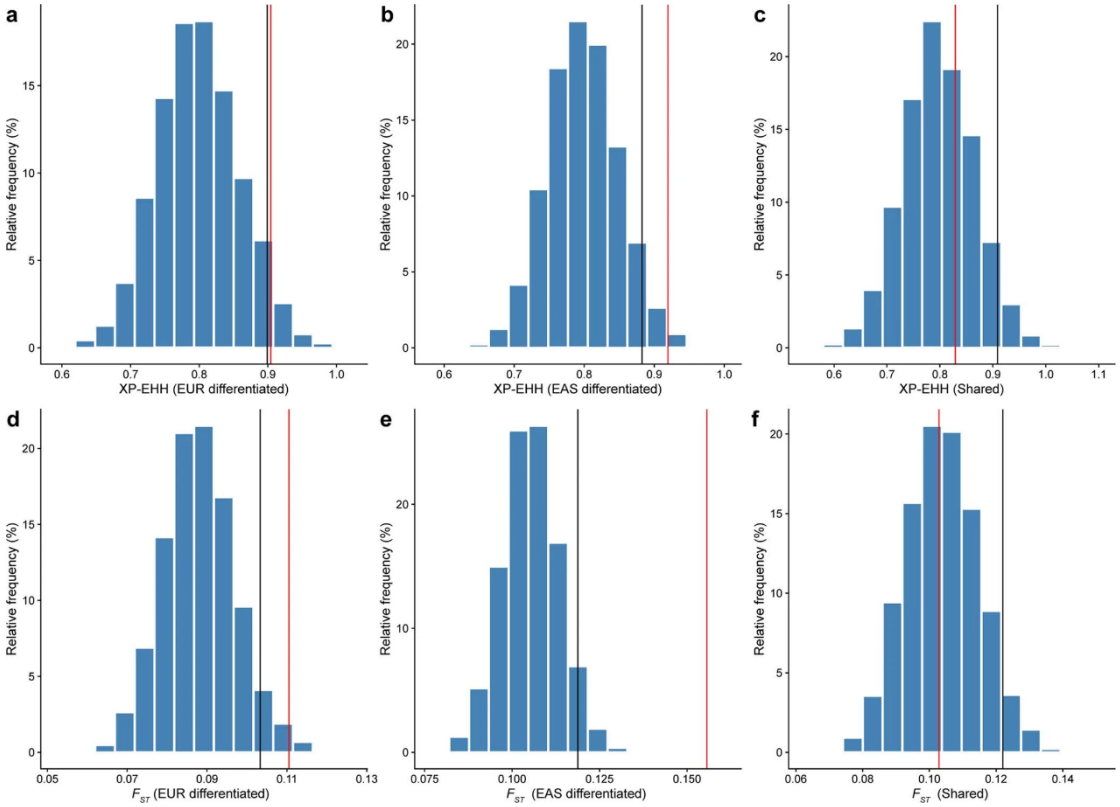
Fig. 4: Natural selection analyses and enrichment test of the differentiation of facial-associated variants among the EAS and EUR populations. a, b) P-values (-log10 scale) of a) F_{ST} and b) XP-EHH for the whole face and 10 anatomical regions. The red line is the P-value threshold of 0.05. c, d) Observed mean PBS value for the leading variants c) the 244 variants in this EAS study, and d) the 203 variants from study of White et al against the null distribution among EAS, EUR and YRI for the nose region. e) Selection coefficients for the nose region against the underlying null distribution (blue). The red line corresponds to the observed selection coefficients. The black line is 95% quantile of the null distribution. f) Differentiated accumulated genetic effects of the 244 leading variants (visualized using the local surface normal displacement) and P-values (-log10 scale) of each quasi-landmark. g) Effects and P-values (-log10 scale) of each quasi-landmark compared with random drift in the European population and in h) the East Asian population.



Extended Data Fig. 1: Study design. We first start with a face segmentation procedure to get 63 face segments from which we defined 10 anatomical face regions. Then by using a CCA based GWAS, we identified 244 variants with a P value lower than 5×10-8, in which 151 are also lower than 9.8×10-10. To investigate what affects the similarity of an EAS face, we used polygenic population shape (PPS) analyses to fit EUR and EAS faces and identified 13 variants mainly contributing to EUR-EAS facial differences. To investigate selection on facial variation, we used FST and XP-EHH to find which parts of the face are under selection. These results, we further compared with random drift and random PPS to find out, which from the two populations, EUR or EAS, experienced selection.

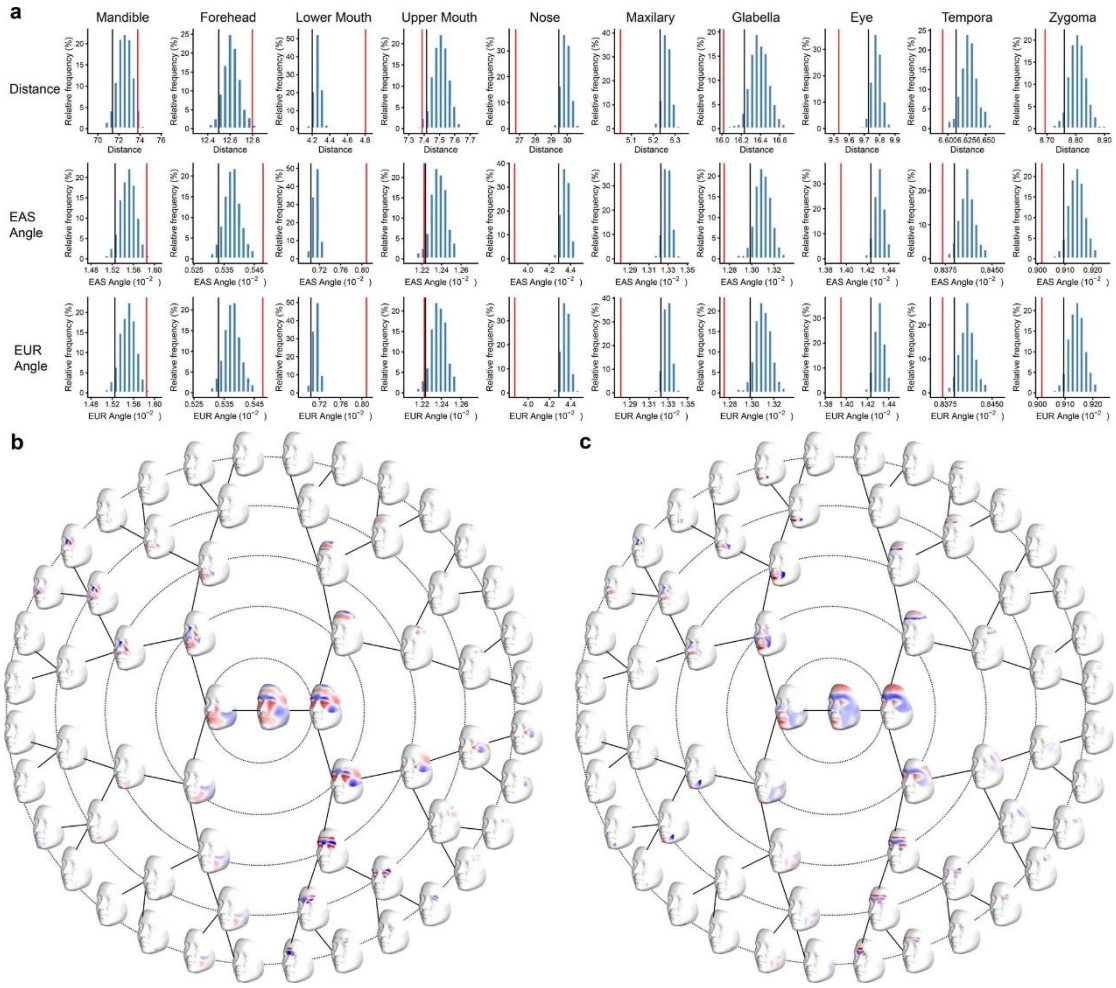


Extended Data Fig. 2: Enrichment analysis of leading variants. (a) Geno Ontology enrichment for genes annotated from leading variants by GREAT24. (b) Heatmap indicating the global enrichment of trait-associated variants in different chromatin state (y axis) and in different tissue (x axis). The fold change was calculated by GREGOR28. The embryonic craniofacial tissue was previously published by epigenomic atlas, while the other was previously published by Roadmap Epigenome27. The description of the 25-state chromatin model can be found at: https://egg2.wustl.edu/roadmap/web_portal/imputed.html#chr_imp. (c) Expression levels of the candidate genes in craniofacial tissues. Each point (n=3 biologically independent replicates for each condition) represents an estimated fold change compared to control genes at different times (E10.5, E11.5, E12.5), in different prominences (Frontonasal, FNP: circle; Maxillary, MxP: square; Mandibular, MnP: triangle), and tissue layer (Ectoderm, Ect: red; Mesenchyme, Mes: blue). Data are presented as mean values +/- 95% confidence intervals (1.96×SEM).

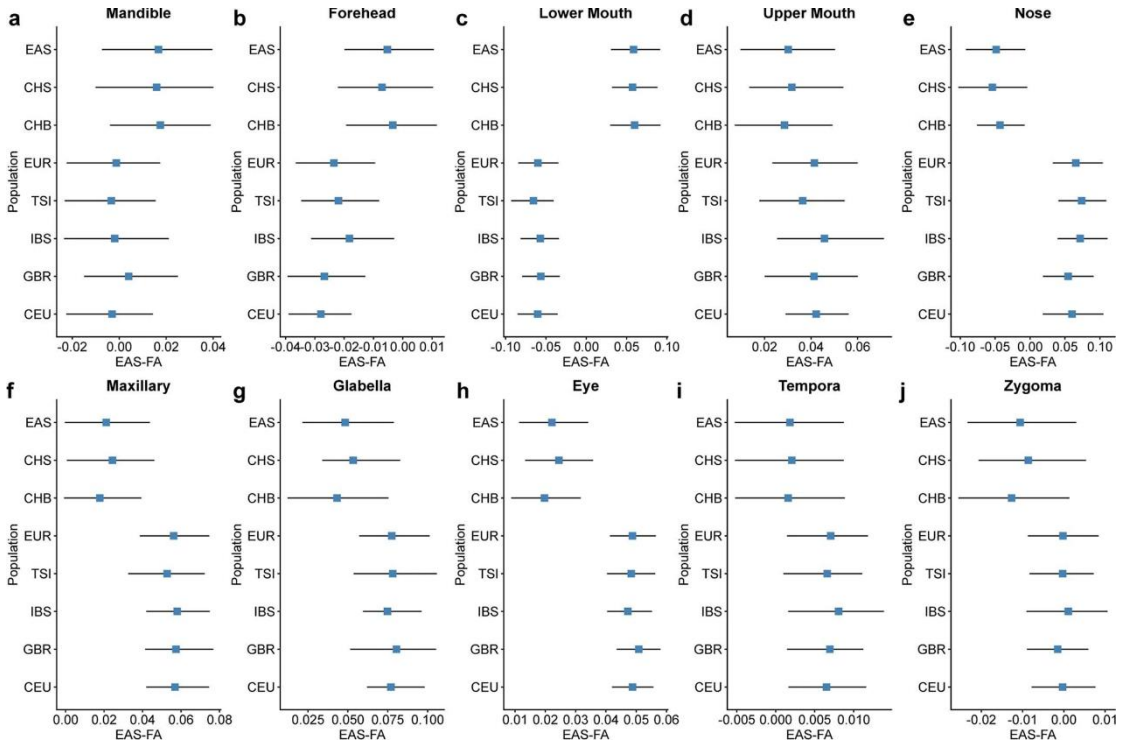

a
b
c
d
e
f
Relative frequency (%)
XP-EHH (EUR differentiated)
XP-EHH (EAS differentiated)
XP-EHH (Shared)
F_{ST} (EUR differentiated)
F_{ST} (EAS differentiated)
F_{ST} (Shared)

525

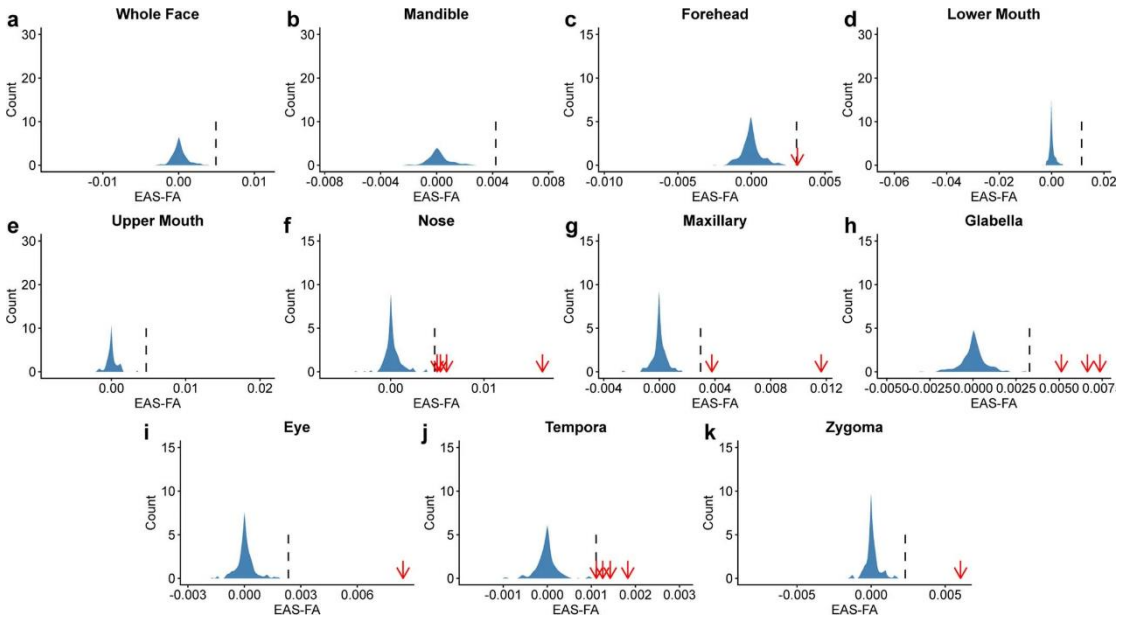
Extended Data Fig. 3 XP-EHH and FST enrichment analysis for shared and differentiated variants. XP-EHH and FST enrichment analysis for (a, d) EUR differentiated variants, (b, e) EAS differentiated variants, and (c, f) shared variants in EAS study. The blue color is the null distribution. The red line is the mean XP-EHH or FST score of shared or differentiated variants. The black line is the 95% quantile of the null distribution.



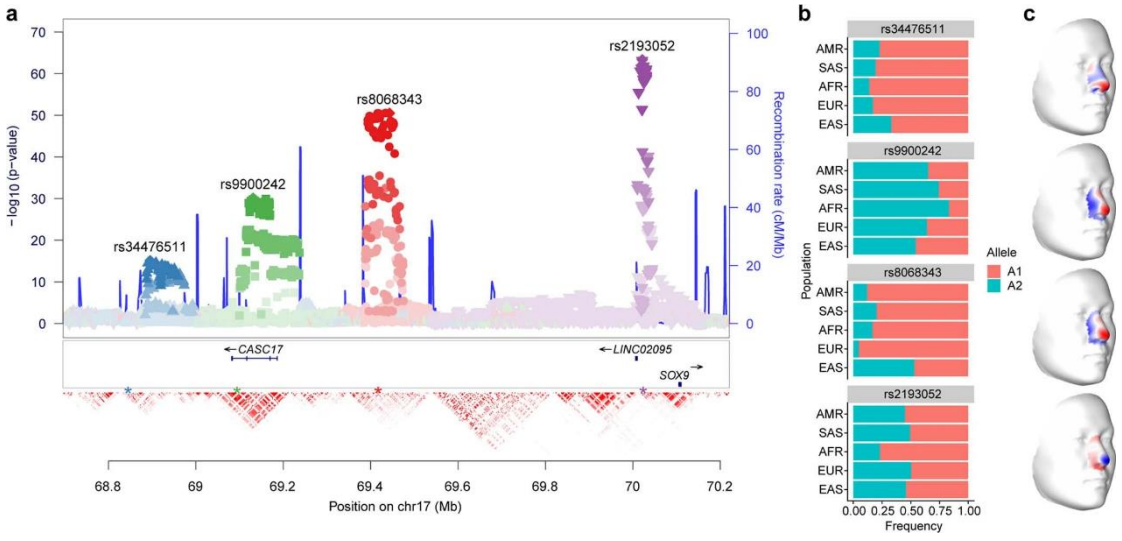
Extended Data Fig. 4 Validation of PPS in 10 anatomical segments. (a) The null distribution (blue) of Euclidean distance, cosine similarity with EUR average face and EAS average face using 1,000 simulations derived from random variants on the 10 anatomical regions, red line infers the statistics of the leading variants associated with corresponding regions; black line infers 95% quantile of distribution from the random variants with corresponding regions; (b) The genetic effects of rs12632544 and (c) rs12473319 weighted by their effect allele number difference of EUR and EAS (visualized using the local surface normal displacement).



Extended Data Fig. 5 The EAS-FA of polygenic shapes in 10 anatomical regions for EAS and EUR individuals in 1000GP. The EAS-FA of polygenic shapes in a) mandible, b) forehead, c) lower mouth, d) upper mouth, e) nose, f) maxillary, g) glabella, h) eye, i) tempora, and j) zygoma for EAS and EUR individuals in 1000GP. The squares represent the mean EAS-FA score in 10 anatomical regions and the horizontal lines represent 1st and 3rd quantile.



Extended Data Fig. 6 EAS-FA of the 244 leading variants on the EUR-EAS difference. The distributions (blue) of EAS-FA derived from 244 leading variants associated with a) whole face and b) - k) 10 anatomical segments. The black dotted line is the EAS-FA threshold of each region (mean + 3×SD). The red arrow is the variant over threshold.



Extended Data Fig. 7 Multi peak in 17q24.3 region. (a) Association variants in the SOX9 locus and genomic environment surrounding SOX9 across a 2-Mb window. Four independent variants, represented by (1) rs34476511 (blue), (2) rs9900242 (green), (3) rs8068343 (red), and (4) rs2193052 (purple) are observed; (b) Allele frequency in AMR, SAS, AFR, EUR and EAS population of the four variants from 1000GP; (c) The effects of the four variants in the nose region.

Reference

1. Liu, F. *et al.* A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genet* **8**, e1002932 (2012).
2. Paternoster, L. *et al.* Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. *Am J Hum Genet* **90**, 478-85 (2012).
3. Adhikari, K. *et al.* A genome-wide association study identifies multiple loci for variation in human ear morphology. *Nat Commun* **6**, 7500 (2015).
4. Adhikari, K. *et al.* A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nat Commun* **7**, 11616 (2016).
5. Cole, J.B. *et al.* Genomewide Association Study of African Children Identifies Association of SCHIP1 and PDE8A with Facial Size and Shape. *PLoS Genet* **12**, e1006174 (2016).
6. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-17 (2016).
7. Shaffer, J.R. *et al.* Genome-Wide Association Study Reveals Multiple Loci Influencing Normal Human Facial Morphology. *PLoS Genet* **12**, e1006149 (2016).
8. Lee, M.K. *et al.* Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. *PLoS One* **12**, e0176566 (2017).
9. Cha, S. *et al.* Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC Genomics* **19**, 481 (2018).
10. Claes, P. *et al.* Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nat Genet* **50**, 414-423 (2018).
11. Crouch, D.J.M. *et al.* Genetics of the human face: Identification of large-effect single gene variants. *Proc Natl Acad Sci U S A* **115**, E676-E685 (2018).
12. Qiao, L. *et al.* Genome-wide variants of Eurasian facial shape differentiation and a prospective model of DNA based face prediction. *J Genet Genomics* **45**, 419-432 (2018).
13. Li, Y. *et al.* EDAR, LYPLAL1, PRDM16, PAX3, DKK1, TNFSF12, CACNA2D3, and SUPT3H gene variants influence facial morphology in

a Eurasian population. *Hum Genet* **138**, 681-689 (2019).
14. Wu, W. *et al.* Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese. *Hum Genet* **138**, 601-611 (2019).
15. Xiong, Z. *et al.* Novel genetic loci affecting facial shape variation in humans. *Elife* **8**(2019).
16. Huang, Y. *et al.* A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. *J Genet Genomics* (2020).
17. Bonfante, B. *et al.* A GWAS in Latin Americans identifies novel face shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation. *Sci Adv* **7**(2021).
18. White, J.D. *et al.* Insights into the genetic architecture of the human face. *Nat Genet* **53**, 45-53 (2021).
19. Guo, J. *et al.* Variation and signatures of selection on the human face. *J Hum Evol* **75**, 143-52 (2014).
20. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221-7 (2005).
21. Kanai, M., Tanaka, T. & Okada, Y. Empirical estimation of genome-wide significance thresholds based on the 1000 Genomes Project data set. *J Hum Genet* **61**, 861-866 (2016).
22. Stouffer, S.A. Adjustment during army life. (1965).
23. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
24. McLean, C.Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501 (2010).
25. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
26. Bernstein, B.E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* **28**, 1045-8 (2010).
27. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
28. Schmidt, E.M. *et al.* GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. *Bioinformatics* **31**, 2601-6 (2015).
29. Wilderman, A., VanOudenhove, J., Kron, J., Noonan, J.P. & Cotney, J. High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. *Cell Rep* **23**, 1581-1597 (2018).
30. Som, P.M. & Naidich, T.P. Illustrated review of the embryology and development of the facial region, part 2: Late development of the fetal face and changes in the face from the newborn to adulthood. *AJNR Am J Neuroradiol* **35**, 10-8 (2014).
31. Staley, J.R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207-3209 (2016).
32. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
33. Weir, B.S. & Cockerham, C.C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358-1370 (1984).
34. Robinson, M.R. *et al.* Population genetic differentiation of height and body mass index across Europe. *Nat Genet* **47**, 1357-62 (2015).

35. Sabeti, P.C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-8 (2007).
36. Gautier, M., Klassmann, A. & Vitalis, R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol Ecol Resour* **17**, 78-90 (2017).
37. Choi, S.W., Mak, T.S. & O'Reilly, P.F. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* **15**, 2759-2772 (2020).
38. Yi, X. *et al.* Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75-8 (2010).
39. Singh, M.K. *et al.* The T-box transcription factor Tbx15 is required for skeletal development. *Mechanisms of Development* **122**, 131-144 (2005).
40. Lausch, E. *et al.* TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in Cousin syndrome. *Am J Hum Genet* **83**, 649-55 (2008).
41. Ding, H.L., Clouthier, D.E. & Artinger, K.B. Redundant roles of PRDM family members in zebrafish craniofacial development. *Dev Dyn* **242**, 67-79 (2013).
42. Kamberov, Y.G. *et al.* Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* **152**, 691-702 (2013).
43. Tan, J. *et al.* The adaptive variant EDARV370A is associated with straight hair in East Asians. *Hum Genet* **132**, 1187-91 (2013).
44. Shaffer, J.R. *et al.* Multiethnic GWAS Reveals Polygenic Architecture of Earlobe Attachment. *Am J Hum Genet* **101**, 913-924 (2017).
45. Gaudet, P., Livstone, M.S., Lewis, S.E. & Thomas, P.D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform* **12**, 449-62 (2011).
46. Zhao, Y. *et al.* Isolated cleft palate in mice with a targeted mutation of the LIM homeobox gene lhx8. *Proc Natl Acad Sci U S A* **96**, 15002-6 (1999).
47. Haenig, C. *et al.* Interactome Mapping Provides a Network of Neurodegenerative Disease Proteins and Uncovers Widespread Protein Aggregation in Affected Brains. *Cell Rep* **32**, 108050 (2020).
48. Astuti, D. *et al.* Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nat Genet* **44**, 277-84 (2012).
49. Albers, P.K. & McVean, G. Dating genomic variants and shared ancestry in population-scale sequencing data. *PLoS Biol* **18**, e3000586 (2020).
50. Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. A map of recent positive selection in the human genome. *PLoS Biol* **4**, e72 (2006).
51. Guo, J. *et al.* Global genetic differentiation of complex traits shaped by natural selection in humans. *Nat Commun* **9**, 1865 (2018).
52. He, Y. *et al.* A probabilistic method for testing and estimating selection differences between populations. *Genome Res* **25**, 1903-9 (2015).
53. Zaidi, A.A. *et al.* Investigating the case of human nose shape and climate adaptation. *PLoS Genet* **13**, e1006616 (2017).

Materials and methods

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 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 Northern Han Chinese (NHC) cohort 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 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). 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.

Sample and recruitment details

The samples in this study were collected from three independent cohorts, the National Survey of Physical Traits (NSPT) cohort (*n* = 3,322), the Northern Han Chinese (NHC) cohort (*n* = 4,767), and the Taizhou Longitudinal Study (TZL) cohort (*n* = 2,881). For the NSPT sample, individuals were recruited at three Chinese cities: Nanning, Guangxi province (*n* = 1,326); Taizhou, Jiangsu Province (*n* = 986); Zhengzhou, Henan province (*n* = 1,010). In the NHC cohort,

participants were recruited in Tangshan, Hebei province. These two cohorts constituted the discovery dataset. The TZL cohort, where individuals were recruited in Taizhou, Jiangsu province, were used as the replication dataset. The characteristics of the datasets is shown in Supplementary Table 2.

Genotyping and imputation

Since we used two different genotyping platforms in the discovery and replication datasets (details in the Supplementary Note), we chose to impute the two data sets separately, then combine the imputed results.

For each dataset, standard data cleaning and quality assurance practices were performed based on the GRCh37 genome assembly. Phasing was performed using SHAPEIT2 (v2.17) [54] and imputation to the 1000GP Phase 3 reference panel using IMPUTE2 (v2.3.2) [55]. After post-imputation quality control, 8,018,212 shared variants were obtained for analysis.

3D image acquisition, registration, and quality control

3D images of all individuals in the three cohorts were captured and acquired using the 3dMDface (3dMD) camera system. When capturing, participants were asked to close mouth, open eyes, and hold faces with a neutral expression.

The 3D surface images were registered using the MeshMonk (v0.0.6)[56] in MATLAB™ 2018a. This performed a homologous configuration of 7,906 spatially dense landmarks, allowing the 3D image data to be standardized. We performed Generalized Procrustes analysis (GPA) and symmetrization, then investigated every mapped image manually and identified outlier images. 3D

facial images with poor quality were removed or re-processed, with details available in the Supplementary Note.

As a result, 6,968 (*n* = 4,089 in NHC cohort, *n* = 2,879 in NSPT cohort) and 2,706 unrelated individuals with good quality 3D images in the discovery and replication dataset are used for further analysis.

Facial phenotyping

Like the approach of White et al., we performed a semi-supervised facial segmentation based on the phenotypic correlation between facial landmarks using the discovery dataset[18]. To calculate the phenotypic correlations, we first corrected the symmetrized facial shapes for the covariates of age, age squared, sex, body mass index (BMI), and four SUGIBS components using a partial least-squares regression (PLSR, function plsregress from MATLAB™ 2018a) in both the discovery and replication cohort[18]. SUGIBS is a method for genetically robust genome-wide ancestry inference based on the spectral (S) decomposition of an unnormalized genomic (UG) relationship matrix generalized by an Identity-by-State (IBS) similarity degree of individuals' matrix, which was also used by White et al[18,57].

To study global and local effects on facial variation, we refined the data-driven facial segmentation method to avoid isolated point and cluster specific facial regions. We performed a hierarchical spectral clustering on a combined matrix, as 0.9×RV similarity matrix + 0.1×distance matrix, up to level five, resulting in a total of 63 facial segments (Supplementary Fig. 6, Supplementary Note). In each segment, we performed principal component analysis (PCA) on the PLSR residuals of the discovery cohort and obtained the PC scores as the phenotypic

scores for the discovery cohort. In the replication cohort, we projected the PLSR residuals onto the PCA space build in the discovery cohort to obtain the PC scores as the phenotypic scores. We described the methods in detail in the Supplementary Note.

Multivariate genome-wide association meta-analyses

The association analysis is similar to that in White et al[18]. For all three phases (discovery, replication, and meta-analysis), the genotypes were coded as the number of major alleles present (0, 1 or 2). In the discovery phase, in each of the 63 facial segments, we used canonical correlation analysis (CCA) to define the linear combination of the facial segments PCs that are mostly correlated with each variant, which represent the phenotypic effect in shape space. When one of the two sets of variables has only one variable, CCA reduces to multiple regression[58]. The resulting vector $\boldsymbol{\omega}_i$ is the effect size vector of this variant in the shape PCA space (Supplementary Note). The correlation can be tested for significance based on Rao's exact F-test (one-sided, right tail)[59]. For each variant, we obtained a direction $\boldsymbol{\omega}_i$ in the shape PCA space most correlated with the genotype of that variant and a P-value representing the strength of correlation in the discovery phase. In the replication phase, we first projected the phenotypic scores onto the CCA direction and calculated the Pearson's correlation between the projected scores and the genotypes in the replication cohort. To test the correlation's significance, we used the Student's t-test where the t-statistics is defined as $t = \frac{\sqrt{\rho^2(1-\rho^2)}}{\sqrt{n-2}}$. We performed a one-sided right tail test for each variant to ensure that the effective direction of the variant within

the two datasets is the same. Next, the P-values obtained in the discovery and replication phase were combined in a meta-analysis using Stouffer's method weighted by the sample sizes[22]. We used the corresponding implementations of these methods in the SNPLIB package (available at https://github.com/jiarui-li/SNPLIB) to accelerate the analyses.

Conditional analysis and GWAS peak selection

For every variant, the meta-analysis described above yielded 63 P-values representing 63 facial segments. In the conditional analysis and peak selection, we selected the lowest P-value for each variant. For the initial selection, we selected the variants with P-value below the genome-wide threshold ($P = 5\times10^{-8}$) and calculated the pairwise r^2 between these variants. In each chromosome, we grouped the selected variants consecutively in a way that the r^2 between every two neighbor selected variants in the group is larger than 0.05, which resulted in 230 groups. Then, we selected the variant with the lowest P-value as the conditional variant for each group and performed association tests of the remaining variants on the condition of the conditional variant. The variant with the most significant P-value still lower than the genome-wide threshold was selected as conditional variants. We repeated these two steps until no variant remains significant. Finally, we obtained 244 leading SNPs from all groups. We considered ±500 kb genomic region of each leading variant as a genomic locus. If nearby genomic regions overlapped, we merged them into one genomic locus.

Permutation test of study-wide P-value threshold

To determine the study-wide Bonferroni P-value threshold, we calculated the number of independent tests by both the eigenvalues of the correlation matrix of the segments and the permutation analysis scheme used in the study of White et al[18,20,21]. The numbers of independent tests obtained from the eigenvalues of the correlation matrix and the permutation analysis are 50 and 51.41, respectively. Here, we used the more stringent threshold $5\times10^{-8}/51.41 = 9.8\times10^{-10}$. The details can be found in the Supplementary Note.

Heritability of Facial Segments

In each facial segment, we first calculated the genome-wide heritability of each retained shape PCs using the discovery cohort and calculated the facial segment's heritability as the mean of the PCs' heritability weighted by each PC's eigenvalue (variance). To calculate the genomic relationship matrix (GRM) for heritability estimation, we first removed SNPs based on high levels of pairwise LD by PLINK v1.9 with window size of 50, step size of 5 bases, and $r^2>0.1$, remaining 266, 241 SNPs. The heritability of all facial segment PCs was estimated by GCTA[60,61].

Gene mapping and functional annotation

Candidate genes of the genome-wide-significant leading SNPs were first identified using the NCBI, HaploReg V4.1, UCSC genome browser and Ensemble genome browser[62-65]. We also used three gene-mapping criteria implemented in Functional Mapping and Annotation (FUMA, v1.3.6) to identify the most likely candidate gene per leading variant[25]. First, we map variants to

genes based on physical distance (within a 10,000 base pair window) from the known protein-coding genes in the human reference assembly. Second, we included the genes which have a significant cis- expression quantitative trait locus (eQTL; 1 Mb distance to the leading variant) association with the leading variants, using 10 tissue types from the GTEx v8 database[66-68]. We used an FDR of 0.05 to define significant eQTL associations. Finally, we also identified candidate genes for each leading variant if there is chromatin interaction. To further prioritize candidate genes, we limited interaction-mapped genes to those who interact with a predicted enhancer region identified in any of the 111 tissues or cell types from the Roadmap Epigenomics Mapping Consortium (ROADMAP) and/or a gene promoter region (from 250 bp upstream to 500 bp downstream of the transcription start site (TSS) and predicted by the ROADMAP to be a promoter region)[27]. We expected that the resulting candidate genes are more likely to have a plausible biological function. We used an FDR of 1×10^{-6} to define significant interactions. To further narrow down the candidate genes, we investigated whether any gene in the window was previously associated with craniofacial development or morphology through normal-range facial association studies, genetic disorders with facial dysmorphology phenotypes, or animal models. To investigate the potential biological process of the candidate genes, FUMA (v1.3.6) and GREAT (v4.0.4) were performed using preset parameters[24,25].

Chromatin state association in embryonic craniofacial tissue

We used GREGOR (v1.4.0) to evaluate global enrichment of trait-associated variants in different chromatin states[28]. This method tests for an increase in the

number of facial-associated index variants, or their LD proxies ($r^2 > 0.8$), overlapping with the regulatory feature more often than expected by chance by comparing to permuted control sets (random control variants are selected across the genome that match the index variant for several variants in LD, minor allele frequency and distance to nearest intron). The reference epigenomes of 127 human tissues and cell types were obtained from the NIH Roadmap Epigenomics Mapping Consortium[26]. The human embryonic craniofacial chromHMM states were obtained from each Carnegie stage by Wilderman et al[29].

Gene expression enrichment analysis

We selected a set of transcriptome datasets from critical periods of mouse face formation that enable gene expression to be analyzed with respect to time, prominence, and tissue layer (GSE62214). We evaluate the expression level (fold change) of the candidate genes compare to a set of control groups where the genes were randomly selected from the genome. Then we regressed the fold change of gene expression on time, prominence, and tissue layer to test their associations.

The EAS and EUR average facial shapes

We recruited 89 individuals with self-reported European ancestry (32 females and 57 males) between 16 and 57 years old in Shanghai[25]. They were required to have complete European ancestry over the last three generations. Their 3D facial images were captured using the same protocol as used in the Chinese cohort. In each segment, we aligned these EUR facial shapes to the

corresponding sample full Procrustes mean shape. We then calculated the male and female average facial shapes separately and used the average facial shapes of the two average shapes as the EUR average facial shapes. To calculate the EAS average facial shapes, we selected five individuals in the Han Chinese cohort with matched age and gender to the individuals in the EUR cohort. We finally selected 445 (5×89) individuals to calculate the EAS average facial shapes in the same manner as in the EUR cohort.

Polygenic shape analysis (PSA)

One can calculate the effect size vector $\boldsymbol{\beta}_{\mathrm{i}}$ in the original shape space by:

$$\boldsymbol{\beta}_{\mathrm{i}} = \mathbf{V}_{\mathrm{k}}\boldsymbol{\Sigma}_{k}\boldsymbol{\omega}_{i} \quad (1)$$

where $\boldsymbol{\Sigma}_{k}$ is a diagonal matrix of the largest k singular values and the column vectors of $\mathbf{V}_{\mathrm{k}}$ are the corresponding k right singular vectors obtained from the PCA (Supplementary Note).

Thus, the polygenic shape (PS) of an individual could be calculated as:

$$\mathbf{PS} = \sum_{\mathrm{i}}^{\mathrm{n}} \boldsymbol{\beta}_{i}\boldsymbol{g}_{i} \quad (2)$$

where $\boldsymbol{g}_{i}$ is the genotype value of variant i[37]. Subsequently, we calculate the polygenic population shape (PPS) by:

$$\mathbf{PPS} = 2\sum_{\mathrm{i}}^{\mathrm{n}} \boldsymbol{\beta}_{i}\boldsymbol{a}_{i} \quad (3)$$

where $\boldsymbol{a}_{i}$ is the effect allele frequency of variant i and two times $\boldsymbol{a}_{i}$ is the average number of effect alleles in a given population.

Next, we used the PPS difference between EUR and EAS ($PPS_{EUR} - PPS_{EAS}$) calculated by leading variants compared with random variants to evaluate

whether leading variants could effectively fit the EUR-EAS shape difference. We calculated the PPS derived shapes as following:

$$F_{EAS}^{d} = F_{AVG} - \frac{PPS_{EUR} - PPS_{EAS}}{2} \quad (4)$$

$$F_{EUR}^{d} = F_{AVG} + \frac{PPS_{EUR} - PPS_{EAS}}{2} \quad (5)$$

where F_{EAS}^{d} and F_{EUR}^{d} are the corresponding PPS derived shapes, F_{AVG} is the average facial shape of the population average shapes of EUR and EAS (i.e., a population neutral average face).

We performed 1,000 simulations to calculate the random $PPS_{EUR} - PPS_{EAS}$. In each simulation, random variants with the same effect allele frequencies in EAS were chosen to calculate random $PPS_{EUR} - PPS_{EAS}$. Subsequently, the cosine similarity and the Euclidean distances between the PPS and the corresponding average face were used as measures of shape similarity. P-values of each approach were then calculated using the null distribution established by these 1,000 simulations.

Variant's contribution to EAS-FA

We used the projected (vector) length to quantify a variant's contribution to the EUR-EAS face difference:

$$l_i = \frac{\mathbf{2}\left(a_i^{EUR} - a_i^{EAS}\right) \times \boldsymbol{\beta}_i \times (\overline{\boldsymbol{F}}_{EUR} - \overline{\boldsymbol{F}}_{EAS})}{|\overline{\boldsymbol{F}}_{EUR} - \overline{\boldsymbol{F}}_{EAS}|} \quad (6)$$

If a variant has a positive sign of projected length, we regard this variant to be linked to EAS individuals having more EAS features. In contrast, a variant with a negative sign is linked with EAS individuals having more EUR characteristic features.

Genetic information used from 1000 Genome Project

We used the individuals' genetic information from 1000 GP Phase 3 for related analysis including calculating MAF, PPS analysis, F_{ST} and PBS calculation, etc. EAS including Han Chinese in Beijing (CHB, n = 103), and Southern Han Chinese (CHS, n = 105); EUR including Toscani (TSI, n = 107), British (GBR, n = 91), Iberian (IBS, n = 107), and Utah residents (CEU, n = 99), without Finnish (FIN); AFR including Yoruba (YRI, n = 103) were used for analysis.

Calculation of natural selection signatures

We calculated genome-wide natural selection signatures based on XP-EHH using REHH2 (v3.2.0)[36]. The genome-wide XP-EHH z-scores were standardized through normalization within each derived allele frequency bin (bin widths = 0.01). We estimated two-tailed P-values of the variant according to the normalized z-scores. We calculated the F_{ST} and PBS for different sets of a population, by using the observed allele frequencies estimated from the 1000GP Phase 3[33,38]. On the basis of a previous study, we measured selective pressures by (genic) selection coefficients, which the details of the calculations are described in He et al[52].

Phenome-wide selection signature analysis

Similar to the approach used in Guo et al., we compared the mean F_{ST}/PBS value of the leading variants with that of the control variants with MAF and LD score matched[51]. First, we divided all the variants (1000GP) into 20 MAF bins from 0 to 0.5 with an increment of 0.025 (excluding the SNPs with MAF < 0.01). Each of the MAF bins was further grouped into 20 bins according to the 20

quantiles of LD score distribution. The MAF and LD score values were computed from the EAS or EUR samples in the 1000GP described above. Second, we allocated the leading variants to the MAF and LD stratified bins, randomly sampled a matched number of "control" variants from each bin, computed a mean F_{ST}/PBS value for the control variants sampled from all bins, and repeated this process 10,000 times to generate a distribution of mean F_{ST}/PBS under drift. Third, a P-value was computed from a two-tailed test by comparing the observed mean F_{ST}/PBS value for the leading variants against the null distribution quantified by the control variants, assuming normality of the null distribution. Regarding enrichment analysis of the selection signatures by XP-EHH, we obtained the sum of the squared values of the normalized XP-EHH z-scores of the variants (or the proxy variants in LD when available; $r^2 > 0.6$ in the CHB or CEU data from 1000 Genome Project), which was compared with the X^2 distribution with the degree of freedom equal to the number of the variants.

Direction of genetic differentiation

The analysis below uses a similar method introduced in Robinson et al. to quantify the population genetic differentiation of a complex trait[34]. The leading variants' coefficients were randomized across variants 10,000 times, and 10,000 genetic predictors were created in the EAS or EUR samples from the 1000GP described above. By keeping the effect sizes consistent but attributing these effects across variants at random, the genetic predictors generated reflect the action of genetic drift.

Data Availability

The Meta-analysis GWAS summary statistics are available on the National Omics Data Encyclopedia.

NODE: OEP002283 (https://www.biosino.org/node/project/detail/OEP002283).

The participants making up the NSPT, NHC and TZL datasets were not collected with broad data sharing consent. Given the highly identifiable nature of both facial and genomic information and unresolved issues regarding risk to participants, we opted for a more conservative approach to participant recruitment. Broad data sharing of the raw data from these collections would thus be in legal and ethical violation of the informed consent obtained from the participants. This restriction is not because of any personal or commercial interests. Additional details can be requested from L.J. for the NSPT dataset, and S.Wang for the NHC and TZL datasets. Data usage shall be in full compliance with the Regulations on Management of Human Genetic Resources in China.

Publicly available data used were: the 1000 GP Phase 3 data: (https://www.internationalgenome.org/category/phase-3/)[23], The Roadmap Epigenomics Project: (http://www.roadmapepigenomics.org)[26], NCBI dbSNP database: (http://www.ncbi.nlm.nih.gov/SNP)[62], UCSC genome browser: (http://genome.ucsc.edu)[63], HaploReg v4.1: (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php)[64], Ensemble genome browser: (http://asia.ensembl.org/Homo_sapiens/Info/Index)[65], GTEx v8: (https://gtexportal.org/home/)[66-68], Human genome dating: (https://human.genome.dating/)[49], and the transcriptome resource from separated ectoderm and mesenchyme of the developing mouse face

(GSE62214):

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62214).

Code availability

The statistical analyses in this work were based on functions of the statistical toolbox in MeshMonk (https://github.com/TheWebMonks/meshmonk, v0.0.6)[56], MATLAB™ 2018a, R (v3.6.1), ggplot2 (v3.1.0), Python (v3.5.0), PLINK v1.9, SHAPEIT2 (v2.17), IMPUTE2 (v2.3.2), SNPLIB (https://github.com/jiarui-li/SNPLIB), GCTA-GREML, FUMA (v1.3.6), GREAT (v4.0.4), GREGOR (v1.4.0), Metascape (https://metascape.org), LocusZoom (https://genome.sph.umich.edu/wiki/LocusZoom), and REHH2 (v3.2.0) as mentioned throughout the Methods.

Reference for Methods

54. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-81 (2011).
55. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
56. White, J.D. *et al.* MeshMonk: Open-source large-scale intensive 3D phenotyping. *Sci Rep* **9**, 6085 (2019).
57. Li, J. *et al.* Robust genome-wide ancestry inference for heterogeneous datasets: illustrated using the 1,000 genome project with 3D facial images. *Sci Rep* **10**, 11850 (2020).
58. Puntanen, S. Methods of Multivariate Analysis, Third Edition. *J International statistical review* **81**, 328-329 (2013).
59. Olson & Chester, L. On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin* **83**, 579-586 (1976).
60. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-9 (2010).
61. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
62. Sherry, S.T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-11 (2001).
63. Navarro Gonzalez, J. *et al.* The UCSC Genome Browser database: 2021 update. *Nucleic Acids Res* **49**, D1046-D1057 (2021).
64. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-4 (2012).
65. Aken, B.L. *et al.* The Ensembl gene annotation system. *Database (Oxford)* **2016**(2016).
66. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
67. Carithers, L.J. & Moore, H.M. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank* **13**, 307-8 (2015).
68. Keen, J.C. & Moore, H.M. The Genotype-Tissue Expression (GTEx) Project: Linking Clinical Data with Molecular Analysis to Advance Personalized Medicine. *J Pers Med* **5**, 22-9 (2015).