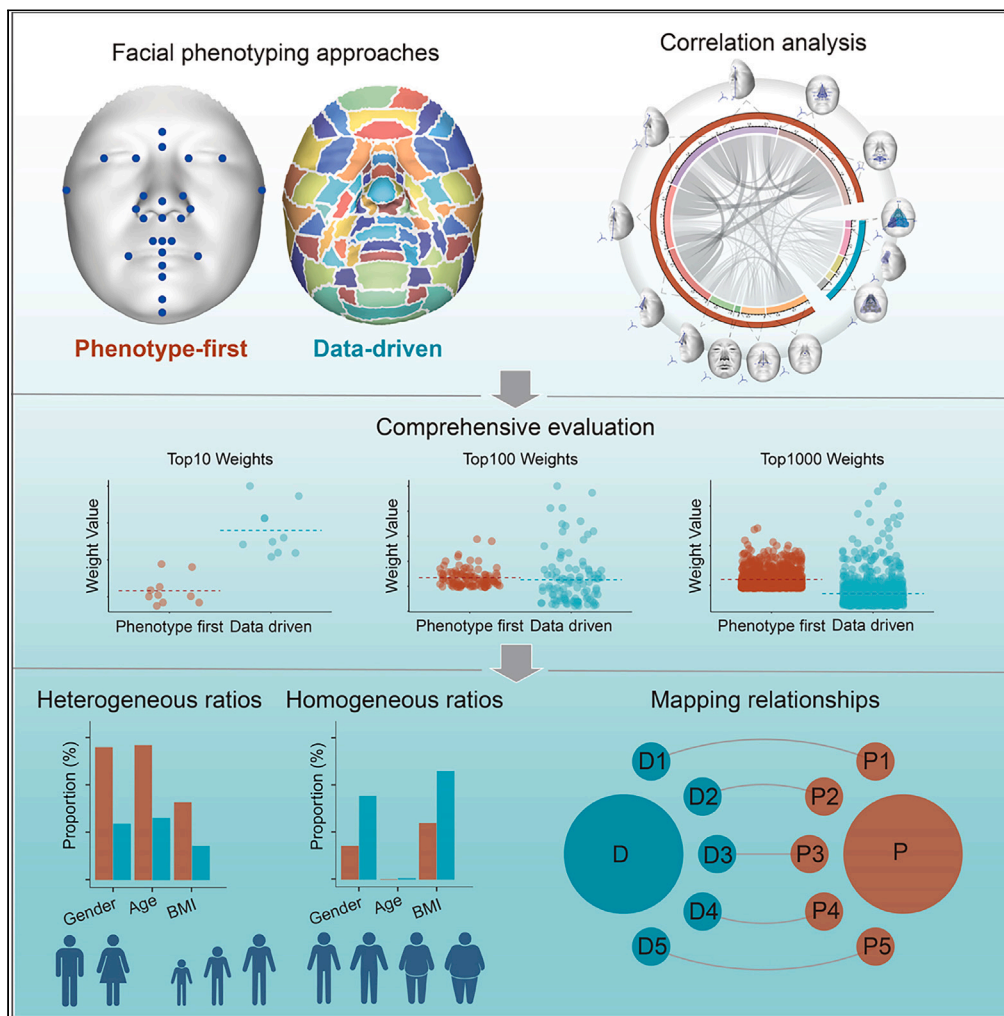


Article

# A comprehensive evaluation of the phenotype-first and data-driven approaches in analyzing facial morphological traits



Hui Qiao, Jingze Tan, Jun Yan, ..., Menghan Zhang, Shuhua Xu, Li Jin

mhzhang@fudan.edu.cn (M.Z.)  
xushua@fudan.edu.cn (S.X.)  
lijin@fudan.edu.cn (L.J.)

**Highlights**

Phenotype-first approach reveals facial phenotypic heterogeneity in Han Chinese

Data-driven approach (DDA) characterizes shared facial phenotypes in Han Chinese

A greater correlation of facial phenotypes with sex, age, and BMI captured by PFA

The deficiency of biological significance for DDA can be supplemented by PFA

Qiao et al., iScience 27, 109325  
March 15, 2024 © 2024 The Author(s).  
<https://doi.org/10.1016/j.isci.2024.109325>



## Article

# A comprehensive evaluation of the phenotype-first and data-driven approaches in analyzing facial morphological traits

Hui Qiao,<sup>1,2,7</sup> Jingze Tan,<sup>1,2,7</sup> Jun Yan,<sup>2,7</sup> Chang Sun,<sup>2</sup> Xing Yin,<sup>2</sup> Zijun Li,<sup>2</sup> Jiazi Wu,<sup>2</sup> Haijuan Guan,<sup>2</sup> Shaoqing Wen,<sup>3</sup> Menghan Zhang,<sup>2,4,5,8,\*</sup> Shuhua Xu,<sup>1,6,\*</sup> and Li Jin<sup>1,\*</sup>

**SUMMARY**

**The phenotype-first approach (PFA) and data-driven approach (DDA) have both greatly facilitated anthropological studies and the mapping of trait-associated genes. However, the pros and cons of the two approaches are poorly understood. Here, we systematically evaluated the two approaches and analyzed 14,838 facial traits in 2,379 Han Chinese individuals. Interestingly, the PFA explained more facial variation than the DDA in the top 100 and 1,000 except in the top 10 phenotypes. Accordingly, the ratio of heterogeneous traits extracted from the PFA was much greater, while more homogenous traits were found using the DDA for different sex, age, and BMI groups. Notably, our results demonstrated that the sex factor accounted for 30% of phenotypic variation in all traits extracted. Furthermore, we linked DDA phenotypes to PFA phenotypes with explicit biological explanations. These findings provide new insights into the analysis of multidimensional phenotypes and expand the understanding of phenotyping approaches.**

**INTRODUCTION**

The morphological diversity of the human face is the basis of many anthropological studies, and the human face is a highly complex and variable structure resulting from the intricate coordination of numerous factors.<sup>1</sup> Dissecting and quantifying facial morphological diversity plays an important role in understanding the homogeneity and heterogeneity both within and among human populations. Traditionally, facial anthropometry has been acquired through the direct measurement of subjects in a clinical setting, using calipers or metric tape to measure parameters between landmarks.<sup>2</sup> This method is referred to as the phenotype-first approach (PFA). A PFA based on distinct landmarks relies greatly on anthropologists' or experts' knowledge and experiences. This approach has contributed to the study of human evolution and comparisons among different human populations.<sup>3</sup> Genetic studies on facial anthropometric phenotypes have led to the identification of genes underlying measurements of nose width, the distance between two nostrils, and the nasolabial angle.<sup>4–6</sup> Instead of using only a limited number of distinct landmarks, Claes et al. developed a data-driven approach (DDA) to exploit both the partial and integrated information from three-dimensional (3D) facial images, facilitating the identification of genetic effects on the facial shape at multiple levels of organization from global to local levels.<sup>7</sup> Various aspects of facial analyses are now using this innovative approach increasingly.<sup>8–11</sup> In addition to facial phenotyping, DDAs such as unsupervised machine learning can be used to discover novel phenotypes of various diseases, including diabetes,<sup>12</sup> sepsis,<sup>13</sup> dilated cardiomyopathy,<sup>14</sup> pulmonary arterial hypertension,<sup>15</sup> heart failure,<sup>16</sup> gliomas,<sup>17</sup> and primary prostate cancer,<sup>18</sup> that may help in elucidating the mechanisms of diseases and treatment effects.

Previous studies on human facial phenotypes employed either a PFA or a DDA. However, the following unanswered questions remain: (1) What is the relationship between these two approaches? (2) What are the differences between the two approaches in capturing human facial characteristics? To address these questions, this study aimed to perform a comprehensive evaluation and comparison of the two approaches. We dissected 3D facial images of 2,379 individuals from a Han Chinese population into 14,838 facial morphological traits, including PFA phenotypes such as the coordinates, distances, and curvatures of landmarks, as well as DDA phenotypes such as the principal components and

<sup>1</sup>State Key Laboratory of Genetic Engineering, Center for Evolutionary Biology, Human Phenome Institute, Zhangjiang Fudan International Innovation Center, School of Life Sciences, Fudan University, Shanghai 200438, China

<sup>2</sup>Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai 201203, China

<sup>3</sup>Institute of Archaeological Science, Fudan University, Shanghai 200433, China

<sup>4</sup>Institute of Modern Languages and Linguistics, Fudan University, Shanghai 200433, China

<sup>5</sup>Research Institute of Intelligent Complex Systems, Fudan University, Shanghai 200433, China

<sup>6</sup>Department of Liver Surgery and Transplantation Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai 200032, China

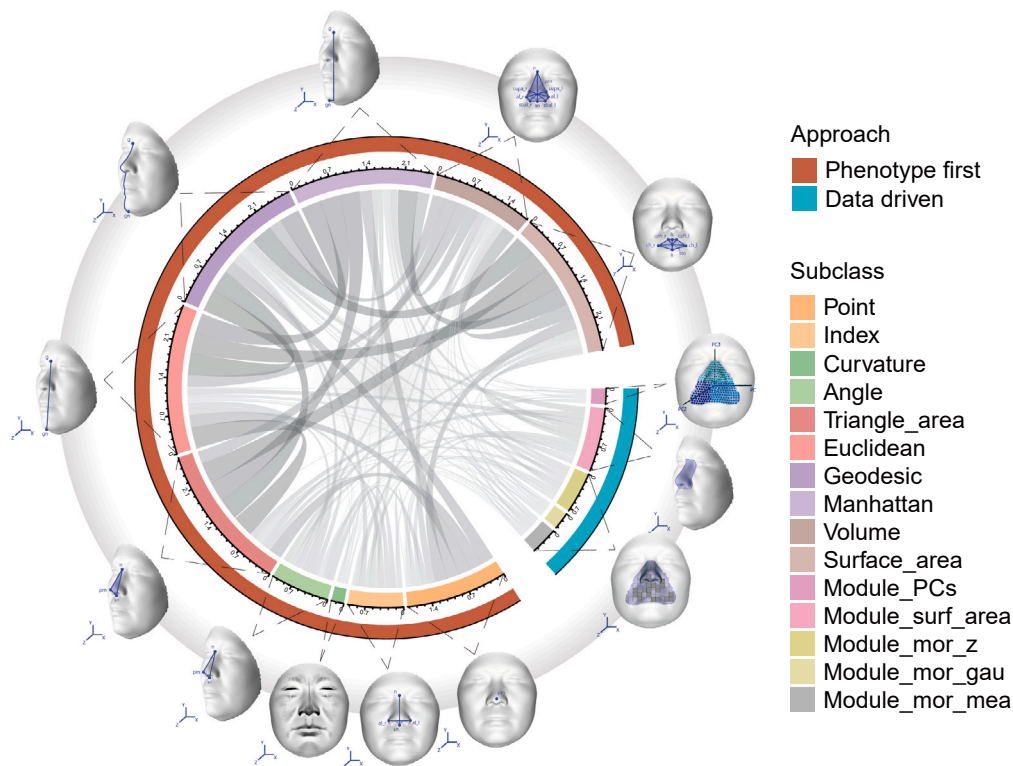
<sup>7</sup>These authors contributed equally

<sup>8</sup>Lead contact

\*Correspondence: mhzhang@fudan.edu.cn (M.Z.), xushua@fudan.edu.cn (S.X.), lijin@fudan.edu.cn (L.J.)

<https://doi.org/10.1016/j.isci.2024.109325>





**Figure 1. Association analysis among different phenotypes**

Circos plot illustrates correlations among phenotypes. Each color represents one type of phenotype. The wider is the band, the greater is the correlation. The subclass of phenotypes are abbreviated as point coordinates (Point), proportion indices (Index), curvatures (Curvature), angular measurements (Angle), triangle area measurements (Triangle\_area), Euclidean distances (Euclidean), geodesic distances (Geodesic), Manhattan distances (Manhattan), voluminal measurements (Volume), surface area measurements (Surface\_area), principal components of the module (Module\_PCs), surface area of the module (Module\_surf\_area), Moran's I of the module Z coordinate (Module\_mor\_z), Moran's I of the module Gaussian curvature (Module\_mor\_gau), and Moran's I of the module mean curvature (Module\_mor\_mea), respectively.

See also [Figure S1](#).

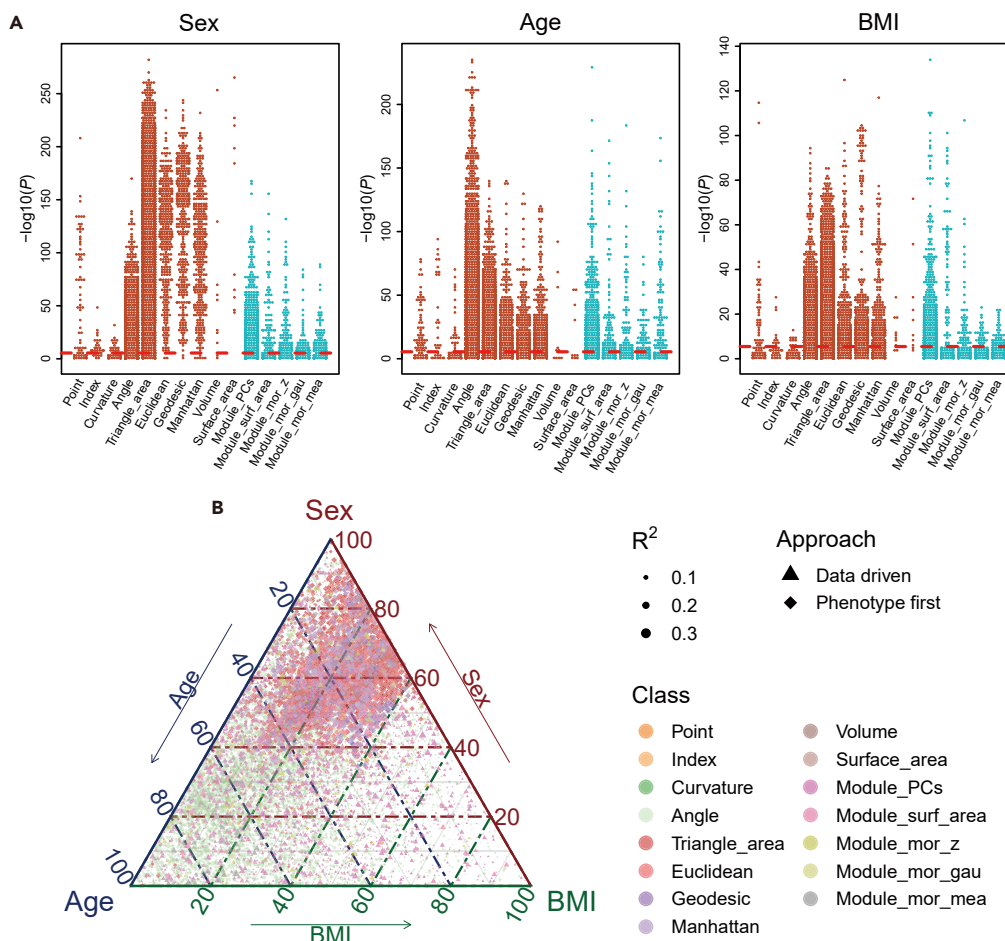
surface areas. Our findings are likely to bear some implications in fundamental and applied sciences, including human genetics, developmental biology, evolutionary biology, medical genetics, forensics, and the design of facial products.

## RESULTS

### The inter-association among facial phenotypes is widespread

To validate the associations among phenotypes, we performed a Spearman correlation analysis on the PFA and DDA. We observed that most traits were related to each other ([Figure 1](#)). The strongest mean correlation coefficient was observed within geodesic distance itself ( $r = 0.455$ ), followed by Euclidean distance and geodesic distance ( $r = 0.442$ ) in the PFA. The most relevant data-driven categories were the surface area of the module and the Moran's I of the module Z coordinate ( $r = 0.068$ ), indicating the weak correlation in the DDA. If considering the relationships between PFA and DDA, the surface area of the module and geodesic distance had the greatest mean correlation coefficient ( $r = 0.128$ ). These results were consistent with expectations and suggested that correlations within different types of traits may be explained in part by sets of shared common genetic components.

To gain further insight into the relationships among different types of phenotypes, we used multi-dimensional scaling (MDS) to identify clusters and patterns of high-dimensional phenotypic data in the lower dimensional space. According to the correlation analysis mentioned previously, we conducted MDS analysis to further derive these similarity or dissimilarity measurements. The results showed that point coordinates, triangle area measurements, surface area measurements, voluminal measurements, and three kinds of distance were clustered together ([Figures S1A–S1C](#)). In addition, the graphs based on the angular measurements, the principal components of the module, and the Moran's I of the module Z coordinate shared some common characteristics. The remaining phenotypes formed another cluster in Dim1 and Dim2. Proportion indices were an outlier for other types in Dim1 and Dim3. The point coordinates, triangle area measurements, surface area measurements, voluminal measurements, Euclidean distances, and geodesic distances formed the largest cluster, and the remaining phenotypes grouped together. From Dim2 and Dim3, the curvatures, the surface area of the module, and the Moran's I of the module mean or Gaussian curvature gathered together. The proportion indices clustered tightly and separated from other traits, while the voluminal



**Figure 2. Relationship between phenotypes and basic variables**

(A) The relationship between facial phenotypes and basic variables. Scatterplot for phenotypes against sex, age, and BMI.

(B) Ternary plot diagram showed the distribution of the load weight constrained by sex, age, and BMI from various phenotypes. Each dot represents one phenotype, and the same type of trait is indicated by the common color. The solid red line represents multiple correction significance thresholds ( $p = 3.37 \times 10^{-6}$ ), and phenotypes above the red line indicate significant differences or correlations.

See also [Tables S1](#), [S2](#), and [S3](#).

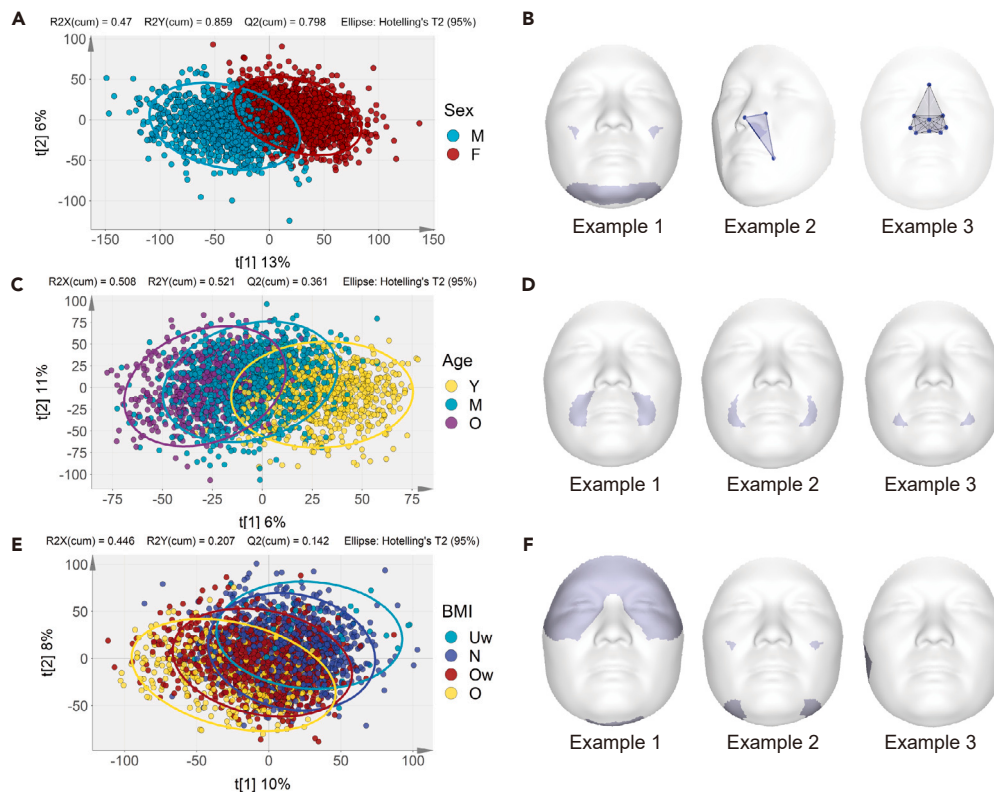
measurements, surface area measurements, triangle area measurements, Euclidean distances, and geodesic distances formed another cluster. To further verify the internal relationships among various phenotypes, t-distributed stochastic neighbor embedding was used to analyze the correlation matrix of each phenotype, and we found that the similarities among facial features were consistent with the MDS results ([Figure S1D](#)). These observations illustrated that facial traits using the same phenotyping approach were more similar, and the differences between PFA and DDA were greater than their internal differences.

### Facial phenotypes show significant associations with sex, age, or BMI

To probe the roles of these basic variables in different kinds of phenotypes, we performed a non-parametric test to examine the correlations between sex, age, or BMI, and facial traits. As shown in [Figure 2A](#), most of the traits showed significant associations with sex, age, or BMI after multiple adjustments ( $p < 3.37 \times 10^{-6}$ ). The triangle area of the pronasale, left cheilion, and left superior alar groove showed the greatest difference between males and females using Mann-Whitney U test ( $p = 1.6 \times 10^{-283}$ ) ([Table S1](#)). Age correlated most strongly with the angle of the left endocanthion, right exocanthion, and right cheilion ( $r = 0.604$ ,  $p = 1.3 \times 10^{-236}$ ), while the first principal component of the 38th module showed the greatest difference among BMI groups ( $r = -0.476$ ,  $p = 4.24 \times 10^{-135}$ ) ([Tables S2](#) and [S3](#)).

### Sex, age, and BMI play different roles in shaping the phenotypic diversity

Understanding the loadings of basic variables (sex, age, and BMI) on the various phenotypes can help to describe the associations between basic variables and phenotypes. Here, we used canonical correlation analysis (CCA) to estimate the possible associations among



**Figure 3. Scatterplot of sex, age, and BMI groups and representative facial features**

(A) Partial least squares-discriminant analysis (PLS-DA) scatterplot showed separation between male (green) and female (red) clustering with individual phenotypic data. M, male; F, female.

(B) Visualization example of the variable importance in projection (VIP) phenotypes. Example 1: The fifth principal component of the 11th module. Example 2: The triangular area of the pronasale, left cheilion, and left superior alar groove. Example 3: The nasal surface area.

(C) PLS-DA scatterplot showed separation among age clustering with individual phenotypic data. Y, young; M, middle-aged; O, old.

(D) Visualization example of the VIP phenotypes. Example 1: Mean curvature Moran index of the 15th module. Example 2: Mean curvature Moran index of the 62nd module.

(E) PLS-DA scatterplot showed separation among BMI clustering with individual phenotypic data. Uw, underweight; N, normal; Ow, overweight; O, obese.

(F) Visualization example of the VIP phenotypes. Example 1: The seventh principal component of the fourth module. Example 2: The second principal component of the 22nd module. Example 3: The first principal component of the 86th module. Colored circles represent 95% confidence intervals. Colored dots represent individual samples.

See also [Figures S2 and S3](#) and [Tables S4–S9](#).

different basic variables and different phenotypes. The results showed that most of the facial features significantly correlated with various basic variables ( $p < 3.37 \times 10^{-6}$ ). The ternary diagram illustrated that most of the phenotypes correlated with sex, followed by age ([Figure 2B](#)), but BMI had a relatively small loading values on phenotypes. In particular, we found sex affected the point coordinates, distances, proportion indices, triangle areas, volumes, and surface areas. Age affected curvatures, while both age and sex affected the angular measurements. Compared to the PFA, the three basic variables had little effect on the data-driven phenotypes with similar loading values. In summary, our findings showed that sex, age, and BMI played different roles in shaping the phenotypic diversity in Han Chinese populations.

### Facial features display significant differences between males and females

We performed the partial least squares-discriminant analysis (PLS-DA) and the Adonis to examine the facial differences across sex, age, or BMI. The PLS-DA scatterplot shown in [Figure 3A](#) illustrated the obvious facial differences between males and females with  $R2Y$  and  $Q2 > 0.5$  ( $R2X = 0.47$ ,  $R2Y = 0.859$ ,  $Q2 = 0.798$ ). We assessed the significance of our classification using permutation tests. The  $p$  values for permutation testing were  $pR2Y = 0.001$  and  $pQ2 = 0.001$ . The Adonis analysis also supported the result, and sex could explain about 30% of the variance in facial features ( $R^2 = 0.299$ ,  $p = 0.001$ ). Three examples of facial features contributing to facial discrimination across sexes are shown in [Figure 3B](#). This finding was consistent with the results of the previous association analysis and CCA analysis of facial phenotypes.

### Facial features display significant differences among age groups

As age is an important factor in human facial features, we examined the facial differences among three age groups of young, middle-aged, and older people using PLS-DA. As shown in Figure 3C, the three age groups formed distinct clusters in the PLS-DA scatterplots ( $R^2X = 0.508$ ,  $R^2Y = 0.521$ ,  $Q^2 = 0.361$ ). The  $p$  values for permutation testing were  $pR^2Y = 0.001$  and  $pQ^2 = 0.001$ . Adonis analysis showed a significant difference in the facial phenotypes among the three age groups, and age explained 3% facial variance ( $R^2 = 0.03$ ,  $p = 0.001$ ) (Table S4), although there were some overlaps among age groups which indicated some degree of potentially shared features. We provided three examples of facial features with variable importance in projection (VIP) > 1 that could well distinguish the age groups (Figure 3D). In addition, we also observed these patterns in either females or males (Figure S2, Tables S5 and S6).

### Facial features display significant differences among BMI groups

We further used PLS-DA and Adonis analysis to investigate the facial differences among four groups of underweight, normal, overweight, and obese people based on individual BMI. In particular, the PLS-DA scatterplots displayed distinct clusters of the four BMI groups ( $R^2X = 0.446$ ,  $R^2Y = 0.207$ ,  $Q^2 = 0.142$ ) (Figure 3E). The  $p$  values for permutation testing were  $pR^2Y = 0.001$  and  $pQ^2 = 0.001$ . And the Adonis analysis also showed an overall significant difference among the four BMI groups (Total:  $R^2 = 0.07$ ,  $p = 0.001$ ). However, no significant difference was observed between the underweight and normal groups, suggesting a relationship between certain BMI groups and facial morphology in the Han Chinese population (Table S7). We provided three examples of facial features with VIP > 1 that could well distinguish the BMI groups (Figure 3F). These results were further replicated in both male and female sub-datasets (Figure S3, Tables S8 and S9).

### PFA and DDA characterize different aspects of facial phenotypes

To further evaluate the performance of PFA and DDA, we adopted the criteria importance through inter-criteria correlation<sup>19</sup> to estimate the objective weights (OWs) for all facial phenotypes extracted using the two approaches. As shown in Figure 4A, the average OW obtained from PFA was significantly greater than that from DDA ( $p < 2.2 \times 10^{-16}$ ). By ranking the OWs for each facial phenotype, we next compared PFA and DDA in the top 10, 100, or 1,000 facial phenotypes (Figure 4B). The results showed that the average weight of the top 10 facial phenotypes was smaller in the PFA compared with the DDA ( $p = 1.6 \times 10^{-4}$ ). This indicated that DDA better explained the facial phenotypic variation than PFA. However, the situation was different for the top 100 and 1,000 phenotypes, for which the average OWs for PFA were larger than those for the DDA (top 100:  $p = 1.1 \times 10^{-3}$ ; top 1,000:  $p < 2.2 \times 10^{-16}$ ). Generally, we discovered an overall greater proportion of differentiated facial phenotypes when using PFA than DDA for different sex, age, and BMI groups (Figure 4C). On the other hand, we discovered an overall smaller proportion of the common facial phenotypes with PFA than DDA (Figure 4D). These were also observed using the Adonis analysis (Figure 4E).

### Unclear biological explanations of DDA phenotypes can be explained by PFA phenotypes

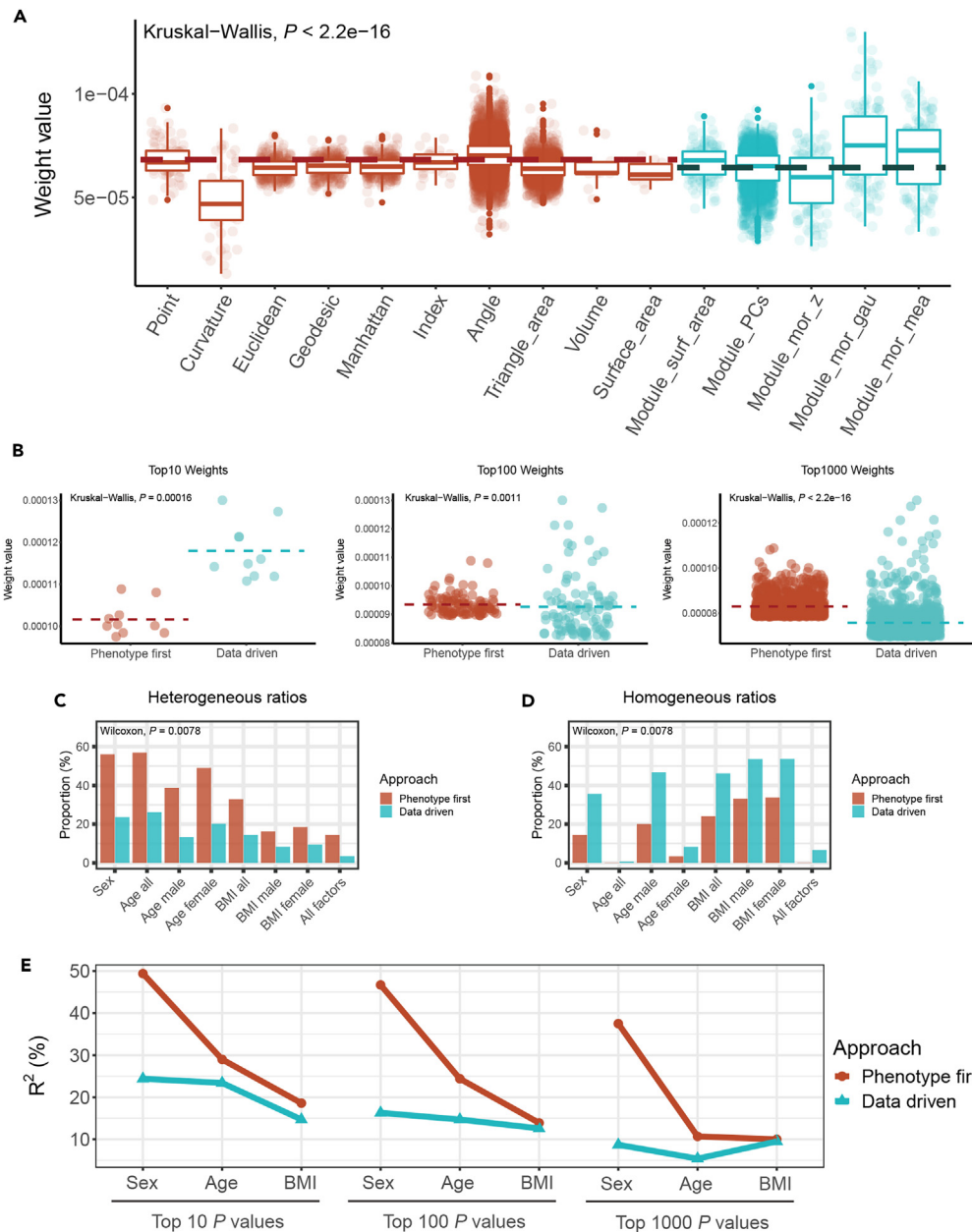
To provide a concrete meaning of the DDA phenotype in the context of the traditional understanding of facial morphology, we attempted to link DDA phenotypes to PFA phenotypes with explicit biological or physical explanations. We identified several such links using correlation analysis and local area mapping. For example, the DDA phenotype Gau\_Moran17 was associated with three PFA phenotypes for the morphological change of the forehead slope (Figure 5A). Similarly, Module6\_pc1 using principal component analysis could be explained by four PFA phenotypes related to the alae nasi projection (Figure 5B). The Module21\_pc8, with a low-rank OW, was strongly correlated with three angular phenotypes representing the chin projection in the morphological observation (Figure 5C). These cases demonstrated a prevalent corresponding relationship between PFA and DDA phenotypes.

## DISCUSSION

In this study, using both PFA and DDA, we analyzed 14,838 facial morphological phenotypes in the Han Chinese population. The findings indicated that there was a broad correlation among PFA phenotypes, but a lower correlation among DDA phenotypes (Figure 1). The nature of the DDA likely resulted in the lower correlation; its computational principle made it easier for each segment to exist independently.<sup>7</sup> Importantly, we proposed that PFA can better characterize the highly differentiated facial traits, while DDA can better characterize the shared facial traits among various groups defined based on sex, age, and BMI.

It has been well documented that facial features vary based on sex and age. The timing and duration of surges in facial growth tend to be different for males and females and between populations contributing to overall facial variation.<sup>20–22</sup> Based on 3D image analysis, a previous study reported that women's noses were shorter and that the nose tip was more pronounced than men's.<sup>23</sup> Interestingly, another physical anthropological study of the Han Chinese also showed that males had thinner eyes and higher nose wings, while females had more slanted eyes, more prominent cheekbones, a more prominent nose tip, and more Mongolian folds.<sup>24</sup> In this study, we observed significant sex differences in most phenotypes, in particular, the males showed larger mean values in 9,219 of 14,838 phenotypes than females.

With respect to age, a previous study in the Korean population reported that the height and thickness of the skull decreased with age, while the length of the skull did not change significantly.<sup>25</sup> Another study based on the Malaysian population samples reported significant differences in facial phenotypes among age groups.<sup>26</sup> Regarding the Han Chinese population, we found negative correlations with age in several facial phenotypes, including the head width, head circumference, lip height, face width, head length height index, and mouth width index, while the mouth width, morphological surface height, nose width, and upper lip skin height correlated with age positively.<sup>27</sup> Here, we



**Figure 4. Comprehensive evaluation of the phenotype-first approach (PFA) and the data-driven approach (DDA)**

(A) Objective weights of PFA and DDA. The dotted line represents the mean weight.

(B) Top 10, 100, and 1,000 objective weights of the PFA and DDA. The dotted line represents the mean weight.

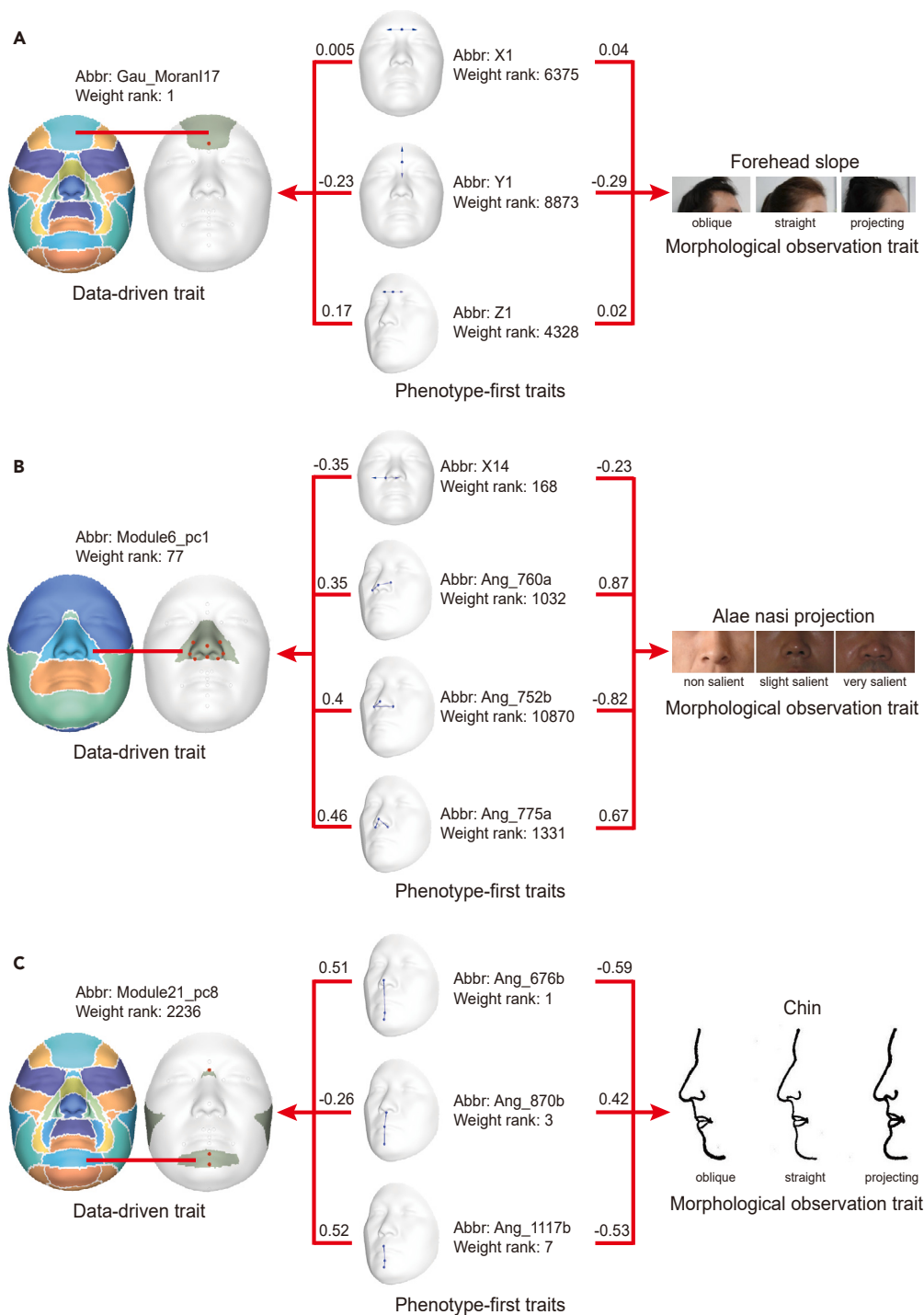
(C) Heterogeneous ratios in the PFA and DDA.

(D) Homogeneous ratios in the PFA and DDA. Age all and BMI all mean a combination of male and female data. All factors mean sex, age, and BMI were considered.

(E) The Adonis tests of the top 10, 100, and 1,000 smallest  $p$  values for sex, age, and BMI differences. Colored dots/bars represent each phenotype: phenotype-first (red) and data-driven (green).

showed that age played a critical role in the phenotypic diversity of the facial morphology for both the PFA and DDA. Overall, facial phenotypes largely correlated with sex, age, and BMI.

In the previous study, the goal was to investigate whether different types of phenotypes were correlated with basic variables. Utilizing several analytic strategies, our results demonstrated that the facial phenotypes of 15 types differed in basic variables. The next natural question is whether some phenotypes are not variations from all the basic variables. Therefore, this study screened for non-significant phenotypes



**Figure 5. Example of mapping relationships among morphological observations, phenotype-first approach (PFA), and data-driven approach (DDA) phenotypes**

(A) The biological explanation of Gau\_Moran17.

(B) The biological explanation of Module6\_pc1.

(C) The biological explanation of Module21\_pc8. The number on the horizontal red line represents the correlation between the two traits. The weight rank represents the weight size of traits obtained using CRITIC. The larger the weight value, the smaller the rank. The traits are abbreviated as the Gaussian curvature Moran index of the 17th module (Gau\_Moran17); the X coordinate direction of the glabella (X1); the Y coordinate direction of the glabella (Y1); the Z coordinate direction of the glabella (Z1); the first principal component of the sixth module (Module6\_pc1); the X coordinate direction of the right alare



**Figure 5. Continued**

(X14); the angle of the right alare, pronasale, and left superior alar groove (Ang\_760a); the angle of the pronasale, right alare, and left alare (Ang\_752b); the angle of the right subalare, pronasale, and left subalare (Ang\_775a); the eighth principal component of the 21st module (Module21\_pc8); the angle of the pronasale, sublabiale, and pogonion (Ang\_676b); the angle of the subnasale, stomion, and pogonion (Ang\_870b); and the angle of the labral superius, sublabiale, and pogonion (Ang\_1117b).

in the previous variation analysis. The results showed that 3,711 facial phenotypes had no significant differences in basic variables, among which 1,836 phenotypes were obtained through DDA ( $p \geq 3.37 \times 10^{-6}$ ) (Figure S4). In addition, we identified 259 facial phenotypes showing no considerable correlation with these three variables by an iterative Adonis test, of which 220 were DDA phenotypes. The Adonis test confirmed that regardless of sex, age, or BMI groups, the  $R^2$  interpretation was very tiny, and the  $p$  values also verified that there was no difference among subgroups (Tables S10 and S11). These results indicated the complexity of the mechanisms and driving forces of facial morphological diversity.

Several studies have already applied either PFA or DDA to the exploration of the relationship between genes and phenotypes.<sup>5,6,9,28–40</sup> However, to the best of our knowledge, the present study is the first research to evaluate the effects of characterizing facial features with PFA and DDA. In this study, we found that the PFA had a greater average weight than the DDA, indicating that the PFA could better characterize facial features, and also better explain the data conditioning on factors such as sex, age, and BMI. The PFA could extract more heterogeneous features, while the DDA could extract more homogenous features. Therefore, the PFA is a knowledge-based method that is valuable in physical anthropological studies, while DDA is a promising approach for knowledge-free phenotyping. We proposed that both PFA and DDA should be considered not only for facial phenotypes but also for the study of other physical anthropological traits, with a well-designed analysis considering the pros and cons of the two approaches.

**Limitations of the study**

In future studies, more efforts are needed to collect richer phenotype data to facilitate facial morphological studies, which will further improve the power of locating genes associated with facial phenotypes. It would be interesting to expand the study into non-Han-Chinese populations to explore whether these findings can be broadly applied.

**STAR★METHODS**

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

- KEY RESOURCES TABLES
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Participants
- METHOD DETAILS
  - 3D facial image acquisition and preprocessing
  - Dissecting facial phenotypes
- QUANTIFICATION AND STATISTICAL ANALYSIS

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.109325>.

**ACKNOWLEDGMENTS**

We are very grateful to the subjects who volunteered for the project. We thank the Fudan University Taizhou Institute of Health Sciences for its role in volunteer recruitment and data entry. We are indebted to the investigator's support of the Ministry of Education Key Laboratory of Contemporary Anthropology in Fudan. We also thank the local staff and institutions at recruitment sites for their assistance in sample collection. We are also very grateful to all of the cooperation agencies for generously donating their time to our project and to the present and former lab members who worked tirelessly to make these analyses possible. This study was supported by the Basic Science Center Program (32288101), the National Science and Technology Basic Research Project (2015FY111700 to L.J. and 2023YFC2605400 to S.X.), and the National Natural Science Foundation of China (NSFC), China (No. 32271186, 31771325, 32030020, T2122007, and 32070577). We thank LetPub ([www.letpub.com](http://www.letpub.com)) for its linguistic assistance during the preparation of this manuscript.

## AUTHOR CONTRIBUTIONS

L.J. conceived the project and provided the main resources. J.T., M.Z., and S.X. supervised the study. H.Q. and J.Y. conducted 3D facial image analysis, manual landmarks annotation, and data analysis and drafted the manuscript. C.S., X.Y., Z.L., J.W., H.G., and S.W. performed research. H.Q., L.J., M.Z., and S.X. revised the manuscript. All authors reviewed the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 8, 2023

Revised: January 17, 2024

Accepted: February 20, 2024

Published: February 24, 2024

## REFERENCES

- Liu, D., Ban, H.J., El Sergani, A.M., Lee, M.K., Hecht, J.T., Wehby, G.L., Moreno, L.M., Feingold, E., Marazita, M.L., Cha, S., et al. (2021). PRICKLE1 x FOCAD Interaction Revealed by Genome-Wide vQTL Analysis of Human Facial Traits. *Front. Genet.* 12, 674642.
- Farkas, L.G. (1994). *Anthropometry of the Head and Face in Clinical Practice* (Raven Press).
- Xi, H.J., and Chen, Z. (2010). *Anthropometric Methods*, 2nd (Science Press).
- Boehringer, S., van der Lijn, F., Liu, F., Günther, M., Sinigerova, S., Nowak, S., Ludwig, K.U., Herberz, R., Klein, S., Hofman, A., et al. (2011). Genetic determination of human facial morphology: links between cleft-lips and normal variation. *Eur. J. Hum. Genet.* 19, 1192–1197.
- Liu, F., van der Lijn, F., Schurmann, C., Zhu, G., Chakravarty, M.M., Hysi, P.G., Wollstein, A., Lao, O., de Bruijne, M., Ikram, M.A., et al. (2012). A Genome-Wide Association Study Identifies Five Loci Influencing Facial Morphology in Europeans. *PLoS Genet.* 8, e1002932.
- Cha, S., Lim, J.E., Park, A.Y., Do, J.H., Lee, S.W., Shin, C., Cho, N.H., Kang, J.O., Nam, J.M., Kim, J.S., et al. (2018). Identification of five novel genetic loci related to facial morphology by genome wide association studies. *BMC Genom.* 19, 481.
- 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.
- Sero, D., Zaidi, A., Li, J., White, J.D., Zarzar, T.B.G., Marazita, M.L., Weinberg, S.M., Suetens, P., Vandermeulen, D., Wagner, J.K., et al. (2019). Facial recognition from DNA using face-to-DNA classifiers. *Nat. Commun.* 10, 2557.
- White, J.D., Indencleef, K., Naqvi, S., Eller, R.J., Hoskens, H., Roosenboom, J., Lee, M.K., Li, J., Mohammed, J., Richmond, S., et al. (2021). Insight into the genetic architecture of the human face. *Nat. Genet.* 53, 45–53.
- Hoskens, H., Liu, D., Naqvi, S., Lee, M.K., Eller, R.J., Indencleef, K., White, J.D., Li, J., Larmuseau, M.H.D., Hens, G., et al. (2021). 3D facial phenotyping by biometric sibling matching used in contemporary genomic methodologies. *PLoS Genet.* 17, e1009528.
- Liu, D., Alhazmi, N., Matthews, H., Lee, M.K., Li, J., Hecht, J.T., Wehby, G.L., Moreno, L.M., Heike, C.L., Roosenboom, J., et al. (2021). Impact of low-frequency coding variants on human facial shape. *Sci. Rep.* 11, 748.
- Ahlqvist, E., Storm, P., Käräjämäki, A., Martinell, M., Dorkhan, M., Carlsson, A., Vikman, P., Prasad, R.B., Aly, D.M., Almgren, P., et al. (2018). Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol.* 6, 361–369.
- Seymour, C.W., Kennedy, J.N., Wang, S., Chang, C.C.H., Elliott, C.F., Xu, Z., Berry, S., Clermont, G., Cooper, G., Gomez, H., et al. (2019). Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. *JAMA* 321, 2003–2017.
- Verdonschot, J.A.J., Merlo, M., Dominguez, F., Wang, P., Henkens, M.T.H.M., Adriaens, M.E., Hazebroek, M.R., Masé, M., Escobar, L.E., Cobas-Paz, R., et al. (2021). Phenotypic clustering of dilated cardiomyopathy patients highlights important pathophysiological differences. *Eur. Heart J.* 42, 162–174.
- Sweatt, A.J., Hedlin, H.K., Balasubramanian, V., Hsi, A., Blum, L.K., Robinson, W.H., Haddad, F., Hickey, P.M., Condliffe, R., Lawrie, A., et al. (2019). Discovery of distinct immune phenotypes using machine learning in pulmonary arterial hypertension. *Circ. Res.* 124, 904–919.
- Segar, M.W., Patel, K.V., Ayers, C., Basit, M., Tang, W.H.W., Willett, D., Berry, J., Grodin, J.L., and Pandey, A. (2020). Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis. *Eur. J. Heart Fail.* 22, 148–158.
- Rui, W., Zhang, S., Shi, H., Sheng, Y., Zhu, F., Yao, Y., Chen, X., Cheng, H., Zhang, Y., Aili, A., et al. (2023). Deep Learning-Assisted Quantitative Susceptibility Mapping as a Tool for Grading and Molecular Subtyping of Gliomas. *Phenomics* 3, 243–254.
- Li, Y., Li, F., Han, S., Ning, J., Su, P., Liu, J., Qu, L., Huang, S., Wang, S., Li, X., and Li, X. (2023). Performance of 18F-DCFPyL PET/CT in Primary Prostate Cancer Diagnosis, Gleason Grading and D'Amico Classification: A Radiomics-Based Study. *Phenomics* 3, 576–585.
- Diakoulaki, D., Mavrotas, G., and Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Comput. Oper. Res.* 22, 763–770.
- Kau, C.H., Richmond, S., Zhurov, A., Ovsenik, M., Tawfik, W., Borbely, P., and JD, E. (2010). Use of 3-dimensional surface acquisition to study facial morphology in 5 population. *Am. J. Orthod. Dentofacial Orthop.* 137, S56.e51–S56.e59.
- Hopman, S.M.J., Merks, J.H.M., Suttie, M., Hennekam, R.C.M., and Hammond, P. (2014). Face shape differs in phylogenetically related populations. *Eur. J. Hum. Genet.* 22, 1268–1271.
- Richmond, R.C., Sharp, G.C., Herbert, G., Atkinson, C., Taylor, C., Bhattacharya, S., Campbell, D., Hall, M., Kazmi, N., Gaunt, T., et al. (2018). The long-term impact of folic acid in pregnancy on offspring DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). *Int. J. Epidemiol.* 47, 928–937.
- Hennessey, R.J., Mclearie, S., Kinsella, A., and Waddington, J.L. (2005). Facial surface analysis by 3D laser scanning and geometric morphometrics in relation to sexual dimorphism in cerebral-craniofacial morphogenesis and cognitive function. *J. Anat.* 207, 283–295.
- Zheng, L., Li, Y., and H, X. (2017). *Study on Physical Anthropology of Han Chinese* (Science Press).
- Kim, H.J., Kim, K.D., Choi, J.H., Hu, K.S., Oh, H.J., Kang, M.K., and Hwang, Y.I. (1998). Differences in the metric dimensions of craniofacial structures with aging in Korean males and females. *Korean J. Phys. Anthropol.* 11, 197–212.
- Al-Khatib, A.R., Rajion, Z.A., Masudi, S.M., Hassan, R., Anderson, P.J., and Townsend, G.C. (2012). Stereophotogrammetric analysis of nasolabial morphology among Asian Malays: influence of age and sex. *Cleft Palate. Craniofac. J.* 49, 463–471.
- Li, Y., Zheng, L., Yu, K., Lu, S., Zhang, X., Li, Y., Wang, Y., Xue, H., and Deng, W. (2013). Variation of head and facial morphological characteristics with increased age of Han in Southern China. *Chin. Sci. Bull.* 58, 517–524.
- Paternoster, L., Zhurov, A.I., Toma, A.M., Kemp, J.P., St Pourcain, B., Timpson, N.J., McMahon, G., McArdle, W., Ring, S.M., Smith, G.D., et al. (2012). Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. *Am. J. Hum. Genet.* 90, 478–485.
- Cole, J.B., Manyama, M., Kimwaga, E., Mathayo, J., Larson, J.R., Liberton, D.K., Lukowiak, K., Ferrara, T.M., Riccardi, S.L., Li,

- M., et al. (2016). Genome wide association study of african children identifies association of SHIP1 and PDE8A with facial size and shape. *PLoS Genet.* 12, e1006174.
30. Shaffer, J.R., Orlova, E., Lee, M.K., Leslie, E.J., Raffensperger, Z.D., Heike, C.L., Cunningham, M.L., Hecht, J.T., Kau, C.H., Nidey, N.L., et al. (2016). Genome-wide association study reveals multiple loci influencing normal human facial morphology. *PLoS Genet.* 12, e1006149.
  31. Pickrell, J.K., Berisa, T., Liu, J.Z., Séguérel, L., Tung, J.Y., and Hinds, D.A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* 48, 709–717.
  32. Lee, M.K., Shaffer, J.R., Leslie, E.J., Orlova, E., Carlson, J.C., Feingold, E., Marazita, M.L., and Weinberg, S.M. (2017). Genome-wide association study of facial morphology reveals novel associations with *FREM1* and *PARK2*. *PLoS One* 12, e0176566.
  33. Crouch, D.J.M., Winney, B., Koppen, W.P., Christmas, W.J., Hutnik, K., Day, T., Meena, D., Boumertit, A., Hysi, P., Nessa, A., et al. (2018). Genetics of the human face: Identification of large-effect single gene variants. *Proc. Natl. Acad. Sci. USA* 115, 676–685.
  34. Qiao, L., Yang, Y., Fu, P., Hu, S., Zhou, H., Peng, S., Tan, J., Lu, Y., Lou, H., Lu, D., et al. (2018). Genome-wide variants of Eurasian facial shape differentiation and a prospective model of DNA based face prediction. *J. Genet. Genom.* 45, 419–432.
  35. Wu, W., Zhai, G., Xu, Z., Hou, B., Liu, D., Liu, T., Liu, W., and Ren, F. (2019). Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese. *Hum. Genet.* 138, 601–611.
  36. Li, Y., Zhao, W., Li, D., Tao, X., Xiong, Z., Liu, J., Zhang, W., Ji, A., Tang, K., Liu, F., and Li, C. (2019). *EDAR*, *LYPLAL1*, *PRDM16*, *PAX3*, *DKK1*, *TNFSF12*, *CACNA2D3*, and *SUPT3H* gene variants influence facial morphology in a Eurasian population. *Hum. Genet.* 138, 681–689.
  37. Xiong, Z., Dankova, G., Howe, L.J., Lee, M.K., Hysi, P.G., de Jong, M.A., Zhu, G., Adhikari, K., Li, D., Li, Y., et al. (2019). Novel genetic loci affecting facial shape variation in humans. *Elife* 8, e49898.
  38. Huang, Y., Li, D., Qiao, L., Liu, Y., Peng, Q., Wu, S., Zhang, M., Yang, Y., Tan, J., Xu, S., et al. (2021). A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. *J. Genet. Genom.* 48, 198–207.
  39. Liu, C., Lee, M.K., Naqvi, S., Hoskens, H., Liu, D., White, J.D., Indencleef, K., Matthews, H., Eller, R.J., Li, J., et al. (2021). Genome scans of facial features in East Africans and cross-population comparisons reveal novel associations. *PLoS Genet.* 17, e1009695.
  40. Adhikari, K., Fuentes-Guajardo, M., Quinto-Sánchez, M., Mendoza-Revilla, J., Camilo Chacón-Duque, J., Acuña-Alonso, V., Jaramillo, C., Arias, W., Lozano, R.B., Pérez, G.M., et al. (2016). A genome-wide association scan implicates *DCHS2*, *RUNX2*, *GLI3*, *PAX1* and *EDAR* in human facial variation. *Nat. Commun.* 7, 11616.
  41. Aynechi, N., Larson, B.E., Leon-Salazar, V., and Beiraghi, S. (2011). Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition. *Angle Orthod.* 81, 245–252.
  42. Qiao, H., Tan, J., Wen, S., Zhang, M., Xu, S., and Jin, L. (2023). De novo dissecting the three-dimensional facial morphology of 2,379 Han Chinese individuals. *Phenomics*, 1–12. <https://doi.org/10.1007/s43657-023-00109-x>.
  43. Bookstein, F.J. (1992). *Morphometric Tools for Landmark Data: Geometry and Biology* (Cambridge University Press).
  44. Team, C.R. (2013). *R Development Core Team. R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing).
  45. Surazhsky, V., Surazhsky, T., Kirsanov, D., Gortler, S.J., and Hoppe, H. (2005). Fast exact and approximate geodesics on meshes. *ACM Trans. Graph.* 24, 553–560.
  46. Sture, H. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
  47. Buja, A., Swayne, D.F., Littman, M.L., Dean, N., Hofmann, H., and Chen, L. (2008). Data Visualization With Multidimensional Scaling. *J. Comput. Graph Stat.* 17, 444–472.
  48. Laurens, V.D.M., and Hinton, G. (2008). Visualizing Data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
  49. Lee, L.C., Liong, C.Y., and Jemain, A.A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* 143, 3526–3539.
  50. Wheelock, Á.M., and Wheelock, C.E. (2013). Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Mol. Biosyst.* 9, 2589–2596.
  51. Chong, I.-G., and Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* 78, 103–112.
  52. Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 32–46.
  53. Oksanen, J., Simpson, G.L., Guillaume Blanchet, F., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Peter, S., Stevens, M.H.H., Szoecs, E., et al. (2022). *vegan: Community Ecology Package*. R package version.
  54. Thompson, B. (2005). Canonical Correlation Analysis (Encyclopedia of Statistics in Behavioral Science).

## STAR★METHODS

## KEY RESOURCES TABLES

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
FaceXenius	GFM	NA
MATLAB R2016b	MathWorks	<a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a>
RStudio (R version 3.5.1)	The R Foundation	<a href="https://www.r-project.org">https://www.r-project.org</a>
SIMCA-P 14.0	Umetrics	NA
Adobe Illustrator CS6	Adobe	NA

## RESOURCE AVAILABILITY

## Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Menghan Zhang ([mhzhang@fudan.edu.cn](mailto:mhzhang@fudan.edu.cn)).

## Materials availability

This study did not generate reagents.

## Data and code availability

- All data reported in this paper will be shared by the [lead contact](#) upon request.
- The paper does not report the original code.
- Any additional information required to re-analyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

## Participants

A total of 2,379 Han Chinese individuals, consisting of 904 males and 1,475 females aged 17 to 83 (mean = 48.9, SD = 12.7), were recruited over five years in China for the collection of 3D data and other relevant information. In addition to 3D data, each participant's age, sex, height, and weight were also recorded. Informed consent was obtained from all individual participants included in the study. The sample collection process for this study was approved by the Ethics Committee of Human Genetic Resources of School of Life Sciences, Fudan University, Shanghai (14117) and has been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

## METHOD DETAILS

## 3D facial image acquisition and preprocessing

The GFM FaceScan 3D System, developed by GFMeasstechnik GmbH (Teltow, Germany), was used to obtain 3D images. The system was calibrated against a target image, containing a known pattern of dots on a white background, at the beginning and regular intervals during each acquisition session. This ensured that each 3D facial image had accurate dimensions. Individuals with a history of craniofacial trauma, congenital malformation, or surgery were excluded. Before scanning, an experienced operator marked critical bone structures on the participant's face with an erasable pen for further annotation.<sup>41</sup> During image capture, participants were instructed to tilt their heads slightly upward, maintain a neutral facial expression with their mouths closed and eyes open, and gaze directly forward. Images with incorrect characteristics were discarded, and new images were taken to meet the set criteria.<sup>41</sup> Each image was composed of a 3D triangular mesh, with about 50,000 vertices representing the frontal facial surface and the corresponding texture map. To account for variations in the original pose and position of the meshes, the acquired 3D images were transformed to a common orientation under the same coordinate space. The preprocessing pipeline was performed using the GFM FaceScan 3D System.

## Dissecting facial phenotypes

This paper implements two approaches for 3D facial morphological phenotyping to compare their differences in characterizing facial traits. The first approach, PFA, was described in detail by the authors in our previous study.<sup>42</sup> Briefly, one trained operator manually placed 26

landmarks across the facial surface twice. The coordinate values (x, y, and z) of these 26 facial landmarks were recorded for each 3D facial image, resulting in a total of 78 coordinate values. Generalized Procrustes analysis standardized the location and orientation of all point configurations.<sup>43</sup> The reliability of image capture and manual landmarking has been reported elsewhere in our previous study.<sup>42</sup> Using RStudio 3.5.1<sup>44</sup> and MATLAB R2016b, a total of 11,557 phenotype-first measurements were calculated including the point coordinates, mean curvatures, Gaussian curvatures, Euclidean distances, Manhattan distances, geodesic distances,<sup>45</sup> proportion indices, angular measurements, triangle area measurements, voluminal measurements, and surface area measurements. The second approach, DDA, was performed using the pipeline described by Claes et al.<sup>7</sup> The global-to-local facial phenotyping partitioned each facial surface into 127 segments, each of which consisted of several point clouds and was represented by multiple dimensions of variation (principal components, PCs). The extracted phenotypes included the principal components of the module, the surface area of the module, the Moran's I of the module Z coordinate, the Moran's I of the module mean curvature, and the Moran's I of the module Gaussian curvature. A total of 3,281 phenotypes were obtained in our previous study<sup>42</sup> (Figure S5, and Table S12).

## QUANTIFICATION AND STATISTICAL ANALYSIS

The Spearman rho correlation, a nonparametric correlation of the Pearson correlation, was used to analyze the facial phenotypes within the differences of sex, age, and BMI. *P* values were adjusted for multiple testing using the Holm-Bonferroni method;<sup>46</sup> for example, an  $\alpha$  threshold for statistical significance for 14,838 comparisons was determined to be  $3.37 \times 10^{-6}$  (i.e.,  $\alpha = 0.05/14,838$ ). Multidimensional scaling (MDS) and t-distributed stochastic neighbor embedding (t-SNE) were used to visualize the similarity and dissimilarity among different types of phenotypes.<sup>47,48</sup> We used SIMCA-P software, version 14.0 (Umetrics, Umea, Sweden), to identify key phenotypes that could discriminate between different subgroups. Partial least squares-discriminant analysis (PLS-DA)<sup>49</sup> was used to investigate the differences between subgroups and to identify the discriminative phenotypes responsible for the separation of the groups. The statistical models were validated using the default internal cross-validation method (10-fold CV) and evaluated by permutation test statistics. The quality of the models was described by the R2X, R2Y, and Q2 parameters. The R2 value is a cross-validation parameter that defines the proportion of variance in the data explained by the models, while the Q2 value is an internal cross-validation parameter that indicates the predictability of the model. Furthermore, the contributions of variates were evaluated in the classification models using the variable importance in projection (VIP) score.<sup>50</sup> VIP values > 1.0 indicate maximum discriminatory power, whereas those with values < 1.0 indicate minimal discriminatory power.<sup>51</sup> Permutational multivariate analysis of variance (PERMANOVA), also known as Adonis, was utilized to test the significant differences in phenotypic composition among various subgroups and to determine the extent to which different grouping factors explained sample differences.<sup>52</sup> The PERMANOVA procedures were implemented using the *adonis* function of *Vegan* package<sup>53</sup> in R. Canonical correlation analysis (CCA),<sup>54</sup> as implemented in the function *canoncorr* in MATLAB 2016b, was used as a straightforward multivariate testing framework. CCA was used to extract the linear combination of basic variables from each facial segment and calculate the weights of three components. Objective weights of the PFA and the DDA were examined based on the criteria importance through inter-criteria correlation (CRITIC) to detect which approach could better characterize facial features.<sup>19</sup>