Journal of Human Evolution 75 (2014) 143–152

Contents lists available at ScienceDirect

Journal of Human Evolution

journal homepage: www.elsevier.com/locate/jhevol

Variation and signatures of selection on the human face

Jing Guo [a, 1], Jingze Tan [b, c, 1], Yajun Yang [b, c], Hang Zhou [a], Sile Hu [a], Agu Hashan [d], Nurmamat Bahaxar [d], Shuhua Xu [a], Timothy D. Weaver [e, f], Li Jin [a, b, c], Mark Stoneking [g], Kun Tang [a, *]

[a] CAS-MPG Partner Institute and Key Laboratory for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Yueyang Road 320, Shanghai 200031, China
[b] State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai 200433, China
[c] Fudan-Taizhou Institute of Health Sciences, 1 Yaocheng Road, Taizhou, Jiangsu 225300, China
[d] Department of Human Anatomy, Preclinical Medical College, Xinjiang Medical University, Urumqi, Xinjiang 830011, China
[e] Department of Anthropology, University of California, Davis, CA 95616, USA
[f] Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany
[g] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany

ARTICLE INFO

Article history:
Received 14 October 2013
Accepted 4 August 2014
Available online 1 September 2014

Keywords:
Natural selection
Sexual selection
Human facial morphology
3D imaging
Morphological divergence

ABSTRACT

There has been much debate about why humans throughout the world differ in facial form. Previous studies of human skull morphology found levels of among-population differentiation that were comparable to those of neutral genetic markers, suggesting that genetic drift (neutral processes) played an important role in influencing facial differentiation. However, variation in soft-tissue morphology has not been studied in detail. In this study, we analyzed high-resolution 3D images of soft-tissue facial form in four Eurasian populations: Han Chinese, Tibetans, Uyghur and Europeans. A novel method was used to establish a high-density alignment across all of the faces, allowing facial diversity to be examined at an unprecedented resolution. These data exhibit signatures of population structure and history. However, among-population differentiation was higher for soft-tissue facial form than for genome-wide genetic loci, and high-resolution analyses reveal that the nose, brow area and cheekbones exhibit particularly strong signals of differentiation (Q_{st} estimates: 0.3–0.8) between Europeans and Han Chinese. Our results suggest that local adaptation and/or sexual selection have been important in shaping human soft-tissue facial morphology.

Introduction

Humans are characterized by variation in many external features, including the shape of the brow area and nose, skin and eye color, various properties of body hair, and body size and proportions. The extent to which such variation can be explained by genetic drift (neutral processes), local adaptation, or sexual selection has been much debated, going back to Darwin (1871). One approach to unraveling the relative influence of neutral versus selective processes on the differentiation of human traits is to compare the relative amounts of within-population and between-population variance. It is well known that in humans the majority of the genetic variance (~90%) is found within continental regions, whereas only a minor portion (~10%) is accounted for by differences between regions (Weir et al., 2005; Barreiro et al., 2008). Such apportionment of diversity is generally accepted as the amount of differentiation expected under neutral evolution (Relethford, 2002). In genetic data, differentiation is mainly measured using Wright's fixation index (F_{st}) (Wright, 1950), while an analogous statistic has also been defined for phenotypic variation contributed to by genetic factors, usually called Q_{st} (Spitze, 1993). Consequently, the direct comparison of F_{st} and Q_{st} constitutes a useful neutrality test for phenotypic traits, where strong deviations of Q_{st} from neutral F_{st} levels are suggestive of non-neutral evolution (Miller et al., 2008).

In humans, studies of phenotypic diversity and apportionment have mainly focused on features that can be measured on skeletal

* Corresponding author.
E-mail addresses: tangkun@picb.ac.cn, kuncarl@gmail.com (K. Tang).
[1] These authors contributed to the study equally.

http://dx.doi.org/10.1016/j.jhevol.2014.08.001

remains (in particular, skulls). Based on 57 inter-landmark distances, Relethford (1994) first reported that the variation among continental regions accounted for ~10% of overall craniometric variation (Relethford, 1994), in good agreement with the apportionment based on neutral genetic loci. Later studies that utilized various sets of measurements (Harvati and Weaver, 2006; von Cramon-Taubadel, 2009a), different methods for partitioning variation (e.g., principal components analysis (Roseman and Weaver, 2004)), 3D landmark data (Harvati and Weaver, 2006; von Cramon-Taubadel, 2009b) or different samples (Harvati and Weaver, 2006; Hubbe et al., 2009; von Cramon-Taubadel, 2009b), have repeatedly come to the conclusion that the morphological variation in the human skull has been largely shaped by neutral evolution (Roseman and Weaver, 2007; von Cramon-Taubadel and Weaver, 2009; Relethford, 2010). The finding of a close correspondence between phenotypic distance and geographic distance is also consistent with the idea that human skull variation has been shaped by neutral evolutionary processes (Relethford, 2004a,b; 2009). In addition, as with genetic diversity, human craniometric variation can be used to infer population structure and history (Harvati and Weaver, 2006; Gunz et al., 2009; von Cramon-Taubadel, 2009b).

This notwithstanding, the relatively low levels of differentiation of craniometric features strongly contrasts with the situation for skin pigmentation, which exhibits the most variation (~80%) among populations (Relethford, 2002). It also contradicts the intuitive notion that there exists extensive population variation in facial features across the world (Nei and Roychoudhury, 1982; Wright, 1992; Howells, 1995; Gill, 1998; Hennessy and Stringer, 2001). Indeed, adaptive hypotheses have been proposed for a number of craniofacial features. For example, the shape of the nose has long been hypothesized to play an important role in climatic adaptation (Thomson and Buxton, 1923; Coon et al., 1950, 1955). Consistently strong correlations have repeatedly been found between the nasal index (ratio of nose breadth/height) and temperature and humidity (Thomson and Buxton, 1923; Davies, 1932; Weiner, 1954; Wolpoff, 1968; Hiernaux and Froment, 1976; Crognier, 1981; Franciscus and Long, 1991). Recent studies have also reported higher among-population differentiation values (maximum Q_{st} ~ 0.4) than expected under neutrality for several nasal measurements (Roseman, 2004; Roseman and Weaver, 2004; Hubbe et al., 2009). Nonetheless, these studies were all based on the skeletal elements of the nose, leaving the soft-tissue external nose poorly studied.

To date there has been no systematic study of the variation in soft-tissue facial form even though soft-tissue facial form may have experienced greater selection pressures than the underlying skull due to the direct exposure to the environment. Selection might shape the skin, cartilage or adipose tissue distribution, rather than the skull bones. We therefore applied a new approach to analyze variation in soft-tissue facial form. In brief, high-resolution 3D facial images were taken from individuals from four Eurasian populations: Han Chinese from East China (HAN), Tibetans (TIB), Uygur (UYG) (an admixed population with European and Chinese ancestry) and Europeans (EUR). A novel 3D facial surface alignment approach was applied to automatically annotate 15 facial landmarks, and to subsequently establish a dense point-to-point correspondence for ~30,000 3D point markers, with a resolution of one point per 1 mm × 1 mm surface. The high-density data were then aligned to the same Cartesian coordinate system using generalized partial Procrustes analysis (pGPA) (Dryden and Mardia, 1998). Analyses of population structure and variance apportionment were carried out on both the whole face and specific facial features. We find that variation in the soft tissue morphology of the human face has been influenced by both population history and selection.

Materials and methods

Ethics statement

Sample collection for this study was carried out with the approval of the ethics committee of the Shanghai Institutes for Biological Science and in accordance with the standards of the Declaration of Helsinki. Written informed consent was obtained from every participant.

Data and sample collection

The 3dMDface® system (www.3dmd.com/3dMDface) was used to collect high-resolution 3D facial images from volunteers who took part in this study. Four hundred Han Chinese (200 females and 200 males) who were 17–25 years old were sampled in Taizhou, Jiangsu Province. Three hundred and three Uyghur (200 females and 103 males) who were 17–25 years old were sampled in Kashi, Xinjiang. One hundred sixty-nine Tibetans (100 females and 69 males) who were 15–22 years old were sampled in Shigatse. All participants were required to have the same ancestry over three generations. Finally, 89 individuals of self-reported European ancestry (32 females and 57 males) between 16 and 57 years old were collected in Shanghai. They were required to have complete European ancestry over the last three generations. The country of origin of all three generations is shown in Appendix A, Supplementary Online Material (SOM), Fig. S1. Eighty-one percent of the individuals studied have the same place of origin as their parents, and 79% have the same place of origin as their parents and grandparents. The age distributions of all four samples are shown in SOM, Fig. S2. Individuals with obvious health problems or any history of facial surgery were excluded from the study.

High-density 3D facial image alignment

We developed a novel approach for aligning a dense set of quasi-landmarks (Rohr, 2001), evenly distributed on the facial surface, to enable facial comparisons (Guo et al., 2013). First, 15 salient facial landmarks (SOM, Fig. S3) are automatically recognized. In brief, the automatic landmark recognition starts with identifying the location of the pronasale by searching a semi-sphere centered on the nose, followed by pose normalization, which is to align all sample faces to a uniform frontal view. Shape depth (z axis) values and surface texture are then projected to the x–y 2D plane (the frontal portrait plane), where highly specific texture/shape signatures of endo/ecto-canthions and cheilions are identified by a principal components analysis (PCA) based approach. The remaining 10 landmarks are recognized by heuristic methods using geometric relations and texture constraints (Guo et al., 2013). Next, a reference face is chosen and the surface mesh is re-sampled to achieve an even density of one point per vertex of a 1 mm × 1 mm grid. In total, ~30,000 points are used to construct the reference mesh. Third, this reference facial mesh is warped to each face in the sample to ensure the proper matching of all of the 15 landmarks, via a thin-plate spline (TPS) transformation (see SOM, Fig. S4). Fourth, the grid points of the reference face are projected onto each face in the sample. The resulting points of projection, which have a one to one correspondence with the grid points of the reference, are used to define a mesh that describes the surface of each of the faces in the sample (SOM, Fig. S4). Finally, these sample grids are aligned to achieve a common coordinate system by pGPA, in which scaling is not used and size information is preserved. We did not remove size information so that potential differentiation involving size changes could be examined. Details of this alignment method are described elsewhere (Guo et al., 2013) and the corresponding software is

available upon request. Since the same reference face is mapped to all of the sample faces, this process results in a set of facial images aligned by the same dense set of point markers, which then can serve as landmarks. A random Han Chinese female face was chosen as the reference, and 32,251 quasi-landmarks were mapped in three-dimensional space for each face in this study. Therefore, the geometric surface of each sample face is represented by a vector of length $32{,}251 \times 3 = 96{,}753$.

Within population variance PCA analysis

We used PCA to reduce the high-dimensional phenotypic data to a smaller number of dimensions. For a dataset with known subgroup structure, the standard PCA decomposition of the total variance could produce principal components (PCs) with a high proportion of between-group variance in the top PCs, which would make the population differentiation analysis biased (Roseman and Weaver, 2004). To avoid such potential bias, the total 3D face dataset was first subjected to a PCA using the within-population variance/covariance matrix (hereafter referred to as PCA_{wg}, with each individual PC designated as PC_{wg}). To assess the deviations of individuals relative to the mean of their population, rather than relative to the overall mean, the within-population variance/covariance matrix was calculated as follows:

$$W = \frac{1}{n-g}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)(x_{ij}-\bar{x}_i)^T \quad (1)$$

where n is the total number of individuals, g is the number of populations, n_i is the number of individuals in the ith population, x_{ij} is a row vector containing the values for each of the variables for the jth individual of the ith population, and $\bar{x}_i$ is a row vector containing the mean values for each of the variables for the ith population. The T indicates a vector (matrix) transpose. Eigenvalue decomposition was then applied to W.

Total variance PCA analysis

A standard PCA using the total variance/covariance matrix was also carried out for comparison with PCA_{wg} (hereafter referred to as PCA_{tot}, with each individual PC designated as PC_{tot}). The prcomp function in the R 'stats' package (R Development Core Team, 2010) was used for this analysis. The raw data were not scaled, i.e., an equivalent scale was assumed for all of the points.

PLS regression analysis

The PCA methods are non-supervised, i.e., they do not make use of group information. To evaluate to what extent populations can be distinguished with the high-density 3D face data, we carried out partial least squares (PLS) regression analysis, using the reported group identities as one block and the matrix of 3D quasi-landmark coordinates as the other block. We used the plsr function in the R 'pls' package (Mevik et al., 2013) for this analysis, and three components were retained.

F_{st} and Q_{st} calculation

To evaluate the degree of differentiation of soft-tissue facial form, we used several Q_{st} estimators and compared them to levels of genetic differentiation. Genetic differentiation was estimated with F_{st} from genome-wide single nucleotide polymorphism (SNP) data. We retrieved the whole genome SNP data of 45 Han Chinese from Beijing and 10 Uyghur from the CEPH-HGDP panel (Li et al., 2008). We randomly sampled five individuals from each of the eight European populations from the CEPH-HGDP dataset to represent a pooled European population, and we obtained data for 46 Tibetans from a previous study (Xu et al., 2011). Each of these four genetic datasets was matched with one of the four facial datasets. In total, 187,290 SNPs were found to overlap among the four SNP datasets. Individual F_{st} values (Weir and Cockerham, 1984) and genome average F_{st} values (Miller et al., 2008) were calculated based on these overlapping SNPs.

To estimate the degree of differentiation from phenotypic data, in general, the total additive genetic variance σ_G^2 of a trait can be partitioned into between and within population variances as σ_{GB}^2 and σ_{GW}^2, respectively. Phenotypic variation contributed to by genetic factors (Q_{st}) can then be defined as follows (Spitze, 1993):

$$Q_{st} = \frac{\sigma_{GB}^2}{\sigma_{GB}^2 + 2\sigma_{GW}^2} \quad (2)$$

If differentiation is calculated from multiple traits, the Q_{st} estimator proposed by Relethford and Blangero (hereafter referred to as Q_{st}^{R-B}) is the most commonly used statistic (Relethford and Blangero, 1990). The dataset can be represented as a matrix $\boldsymbol{M}$ of s rows and t columns, where s is the number of individuals and t the number of traits. Each individual belongs to one of the g populations. Following equation (2), the Q_{st}^{R-B} estimator for multiple traits is then:

$$Q_{st}^{R-B} = \frac{\sum_{i=1}^{g}\omega_i C_{ii}}{2t + \sum_{i=1}^{g}\omega_i C_{ii}} \quad (3)$$

where C_{ii} are the diagonal elements of a codivergence matrix $\boldsymbol{C}$ calculated from $\boldsymbol{M}$ and ω_i is the weighting factor for the relative census population size, which is fixed as $1/g$ in this study under the assumption of equal population sizes (Relethford and Blangero, 1990). We calculated Q_{st} for each quasi-landmark based on Q_{st}^{R-B}, using the corresponding three coordinate values x, y, z as different traits. This Q_{st} estimator for each quasi-landmark is denoted as Q_m.

However, the calculation of Q_{st}^{R-B} is problematic for very high dimensional phenotypic data, such as our whole face data, because it involves inverting the within-population phenotypic variance/covariance matrix. In order to calculate differentiation levels based on the high-density 3D whole face data, we first carried out the PCA_{wg} decomposition. The $\boldsymbol{M}$ matrix was decomposed into the PCA matrix $\boldsymbol{N}_{s\times k}$, where $k = min(s, t)$. Multiple PCs can then be combined to calculate an overall differentiation estimator (hereafter referred to as Q_{comp}) as previously described (Roseman and Weaver, 2004). The Q_{comp} estimation for multiple PCs is as follows:

$$Q_{comp} = \frac{\sum_{i=1}^{s}\overline{\sigma}_{GB,PC(i)}^2}{\sum_{i=1}^{s}\overline{\sigma}_{GB,PC(i)}^2 + 2\sum_{i=1}^{s}\overline{\sigma}_{GW,PC(i)}^2} \quad (4)$$

where $\overline{\sigma}_{GB,PC(i)}^2$ is the average additive genetic variance between populations for the ith PC, and $\overline{\sigma}_{GW,PC(i)}^2$ is the average additive genetic variance within a population for the ith PC. For each PC, $\overline{\sigma}_{GB,PC(i)}^2$ is calculated as:

$$\overline{\sigma}_{GB,PC(i)}^2 = \frac{\sum_{j=1}^{g}\left(\mu_{j,PC(i)} - \overline{\mu}_{PC(i)}\right)^2}{g} \quad (5)$$

where $\mu_{j,PC(i)}$ is the mean value in population j for PC i, and $\overline{\mu}_{PC(i)}$ is the mean over all $\mu_{j,PC(i)}$.

A differentiation estimator can be also obtained from each individual PC, assuming that each PC represents an abstract facial

feature. In analyses of genetic variation data, PCA has been shown to accurately capture sub-structure and divergence patterns when genetic structure does exist (Patterson et al., 2006). Furthermore, individual PCs only account for fractions of the total variance, each of which likely enriches variation related to certain specific features, or complex shape changes along specific axes (Roseman and Weaver, 2004). Separate analyses of the individual PCs therefore allow closer examination of phenotypic differentiation. We calculated Q_{st} for the major PCs following equation (2), hereafter referred to as Q_p.

The consistency between the Q_{st}^{R-B} and Q_{comp} estimators

Since Q_{st}^{R-B} and Q_{comp} are both designed for multiple traits, we also checked the consistency between Q_{st}^{R-B} and Q_{comp}. Briefly, five quasi-landmarks were randomly sampled from the whole face and Q_{st}^{R-B} and Q_{comp} were calculated for this point set. This random sampling was repeated 1000 times and the Q_{st}^{R-B} and Q_{comp} values were compared for consistency.

Differentiation estimation based on inter-landmark distance data

Inter-landmark distances have been widely used as metric traits in previous studies, so we also estimated differentiation levels using inter-landmark distances. Furthermore, since inter-landmark distances are free from potential GPA alignment errors, they also provide an internal control for the high-density data analyses. While there are issues with the statistical analyses of these distances, the conclusions of our study do not depend only on the inter-landmark distances. In this study, 33 inter-landmark distances based on 15 landmarks were computed as an alternative phenotypic dataset (see SOM, Fig. S3), and their individual and overall Q_{st} was calculated without scaling, as previously described (Relethford, 1994).

Correction for incomplete heritability

In theory, Q_{st} should be estimated from the additive genetic variances (Relethford, 1994), but in empirical population studies the additive genetic variance cannot be directly measured. Instead, differentiation estimation can be first carried out based on the within-population phenotypic variance σ_P^2, which is composed of the additive genetic variance (σ_G^2) and the environmental variance (σ_E^2). The additive genetic variance can be obtained from the phenotypic variance as $\sigma_G^2 = h^2\sigma_P^2$, where h^2 is the heritability. When σ_P^2 equals σ_G^2, $h^2 = 1$ and Q_{st} acquires the minimum possible value, which is the differentiation calculated based on the within-population phenotypic variance (Relethford, 1994). If $h^2 < 1$, the corrected Q_{st}^c can be obtained as follows (Relethford, 1994):

$$Q_{st}^c = \frac{minQ_{st}}{minQ_{st} + h^2(1 - minQ_{st})} \quad (6)$$

In many human craniofacial morphometric studies, a generally accepted value for h^2 is 0.55 (Devor, 1986; Relethford and Harpending, 1994; Sparks and Jantz, 2002; Roseman and Weaver, 2004; Hubbe et al., 2009). We therefore corrected the Q_{st} estimates assuming $h^2 = 0.55$ and labeled the corrected estimates with a superscript letter "c", e.g., Q_{comp}^c.

Construction of neighbor joining trees

In order to clearly view the evolutionary relationships among the four populations, we built un-rooted neighbor-joining (NJ) trees based on pair-wise Q_{comp}^c or F_{st} values with Molecular Evolutionary Genetics Analysis (MEGA) (Tamura et al., 2011), with 1000 bootstrap replicates carried out by sampling the individuals with replacement.

Test of neutrality

The basis of the neutrality test is to compare the Q_{st} estimates (Q_{comp}, Q_p, Q_m, Q_{comp}^c, Q_p^c, Q_m^c) to the empirical F_{st} distribution obtained from genetic markers, as described previously (Roseman and Weaver, 2004). Assuming that the genome is largely neutral, the F_{st} distribution reflects the range of population differentiation values produced by neutral processes (genetic drift, mutation, gene flow). Under neutrality, Q_{st} should have the same distribution asF_{st}. Extreme Q_{st} values that are outliers from the F_{st} distribution are indicative of non-neutral evolution. Therefore, the Q_{st} estimates are ranked against the F_{st} distribution to find the upper percentiles as the empirical P values.

Results

Population structure revealed by the human soft-tissue facial form

It has been demonstrated previously that human skull variation can be used to infer population structure as well the evolutionary relationships among human populations (von Cramon-Taubadel, 2009b). We examined whether the soft-tissue facial form also provides such information. Throughout this paper, the male and female datasets were analyzed separately, and we mainly used the male dataset to demonstrate the results unless otherwise specified; the results for females were quite similar and are included in the SOM results. The PCA_{wg} analysis indicates that the first few components account for the majority of the variance, whereas the fractions explained by smaller PC_{wg}s decline to a plateau around $PC_{wg}12$ (SOM, Fig. S5). We therefore carried out individual PC analysis only on the top 12 PC_{wg}s (SOM, Table S1). As can be seen in Fig. 1A, the first two PC_{wg}s indeed mainly reflect the within-population variance, as the data clouds of the four populations largely overlap. The comparisons of Tibetans with Han Chinese and with Uyghur are highly significant on $PC_{wg}1$ ($P = 3.62 \times 10^{-10}$ and $P = 4.89 \times 10^{-8}$ respectively, SOM, Table S1). In order to visually access the phenotypic variation along synthetic dimensions such as PC_{wg}s, we calculated the average form for all images, and subtracted/added three standard deviations (SD) along the PC_{wg} eigenvector to derive the 'extreme' faces. These are denoted as $PC_{wg}i \pm$ where i indicates the index of the PC_{wg} mode (Fig. 1A). The comparison between $PC_{wg}1+$ and $PC_{wg}1-$ reveals that $PC_{wg}1$ mainly accounts for variation in overall size (Fig. 1A). $PC_{wg}2$ seems to explain variation in the horizontal/vertical ratio of the face as the $PC_{wg}2-$ face is more elongated and narrower compared with the shorter and broader $PC_{wg}2+$ face (Fig. 1A). Some other PC_{wg}s show relatively larger differences among the populations. $PC_{wg}4$ and $PC_{wg}5$ are associated with large between-group variances (only smaller than $PC_{wg}1$) among all four populations (SOM, Table S1). For $PC_{wg}4$, the differences mainly occur in the pairwise comparisons between Europeans and other three groups (EUR-HAN t-test $P = 3.92 \times 10^{-20}$; EUR-UYG t-test $P = 7.35 \times 10^{-15}$; EUR-TIB t-test $P = 4.9 \times 10^{-18}$; SOM, Table S1). For $PC_{wg}5$, all pairwise comparisons revealed highly significant differences except between Han Chinese and Tibetan (SOM, Table S1). When standard PCA_{tot} on the total variance was performed, even stronger population structure was observed, especially between Europeans and Han Chinese ($PC_{tot}3$, SOM, Fig. S6). Finally, we carried out partial least squares (PLS) regression analysis (see Methods). As can be seen in Fig. 1C, Europeans largely separate from Han Chinese along the Han-European PLS axis ($P = 4.93 \times 10^{-18}$). The mean value of the two groups on

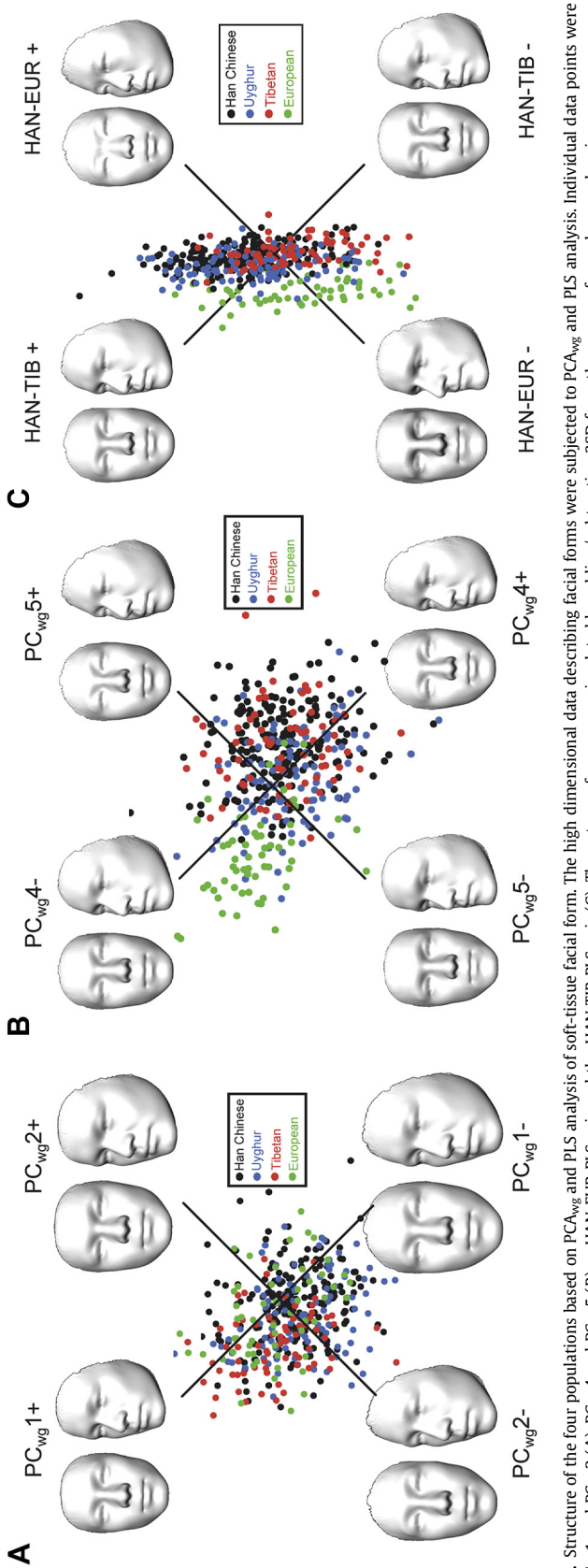

Figure 1. Structure of the four populations based on PCA_{wg} and PLS analysis of soft-tissue facial form. The high dimensional data describing facial forms were subjected to PCA_{wg} and PLS analysis. Individual data points were plotted along $PC_{wg}1$ and $PC_{wg}2$ (A), $PC_{wg}4$ and $PC_{wg}5$ (B) or HAN-EUR PLS axis and the HAN-TIB PLS axis (C). The extreme forms were simulated by adding/subtracting 3SD from the average form along each axis.

this axis, when used as a discriminant score, can assign 81% of the individuals correctly into the Han Chinese or European population. The results of this discriminant analysis should be taken as indicative of the extent to which the Chinese and European faces are differentiated along the PLS axis rather than of classification accuracy, because the same sample was used to calculate and evaluate the discriminant function. The extreme European face is narrower, longer, with a more pointed nose and recessed eye sockets, while the extreme Han Chinese face is wider and flatter, with more prominent cheekbones. Compared with this, the separation on the Han-Tibetan PLS axis was weaker (Fig. 1C), although a *t*-test was also highly significant ($P = 1.31 \times 10^{-11}$).

Overall differentiation

We report Q_{st}^c estimates and the raw Q_{st} estimates side by side (Fig. 2), as the latter indicate the minimum possible differentiation values (Relethford, 1994), serving as lower bounds for the amount of differentiation. In general, the three differentiation measurements (F_{st}, Q_{comp} and Q_{comp}^c) were found to be very consistent in their overall trends (Fig. 2). The Q_{comp}^c values were roughly twice the corresponding F_{st} values, except for the comparison between Uyghur and Han Chinese. The uncorrected Q_{comp} values were also generally higher than the F_{st} values (Fig. 2). Specifically, the highest differentiation was found between Europeans and the two East Asian populations, Han Chinese and Tibetans, for all three estimators (HAN-EUR: $F_{st} = 0.062$, $Q_{comp} = 0.083$, $Q_{comp}^c = 0.141$; TIB-EUR: $F_{st} = 0.059$, $Q_{comp} = 0.083$, $Q_{comp}^c = 0.141$). Uyghurs (UYG) had similar genetic distances to Europeans and Han Chinese (UYG-EUR: $F_{st} = 0.032$; HAN-UYG $F_{st} = 0.031$), but phenotypically, our Uyghur sample was more similar to Han Chinese than to Europeans (UYG-EUR: $Q_{comp} = 0.051$, $Q_{comp}^c = 0.089$; HAN-UYG: $Q_{comp} = 0.013$, $Q_{comp}^c = 0.024$). Among all four populations, the whole face Q_{comp}^c (0.137) was also more than twice as high as the genetic F_{st} (0.056). The neighbor-joining trees based on Q_{comp}^c resemble those based on F_{st} (SOM, Fig. S7), suggesting that human soft-tissue facial form does carry some information about population history, but the generally greater Q_{comp}^c and Q_{comp} values compared to the corresponding genetic F_{st} values indicates that the populations are more differentiated in soft-tissue facial form than the genome-wide average differentiation. Interestingly, the overall differentiation estimates calculated using the inter-landmark distances were strongly consistent with, but slightly greater than, the corresponding Q_{comp}^c and Q_{comp} values for the high-density 3D point data (SOM, Table S2), confirming that the higher phenotypic differentiation found for soft-tissue facial form is not an artifact of the new approach.

Feature-specific differentiation and signals of non-neutral evolution

The higher levels of differentiation exhibited by facial form suggest that it may have been subject to local adaptation and/or sexual selection. To evaluate these hypotheses, we examined the distribution of Q_m^c values for all pairs of populations across the face. The Q_m^c for the Han-European comparison clearly varies greatly across different facial regions (Fig. 3A–C) and is highly consistent between males and females. While Q_m^c is generally below 0.3, there is a marked increase around the nose, central brow area and cheeks (Fig. 3A,B).

When the density distributions of Q_m^c and F_{st} are compared side by side, the Q_m^c distributions show longer tails (Fig. 3C). Under neutrality, the Q_{st} distribution should closely match the F_{st} distribution, therefore a Q_{st} estimate that strongly deviates from the genome F_{st} distribution could signal a departure from neutrality (Whitlock and Guillaume, 2009). To evaluate the statistical

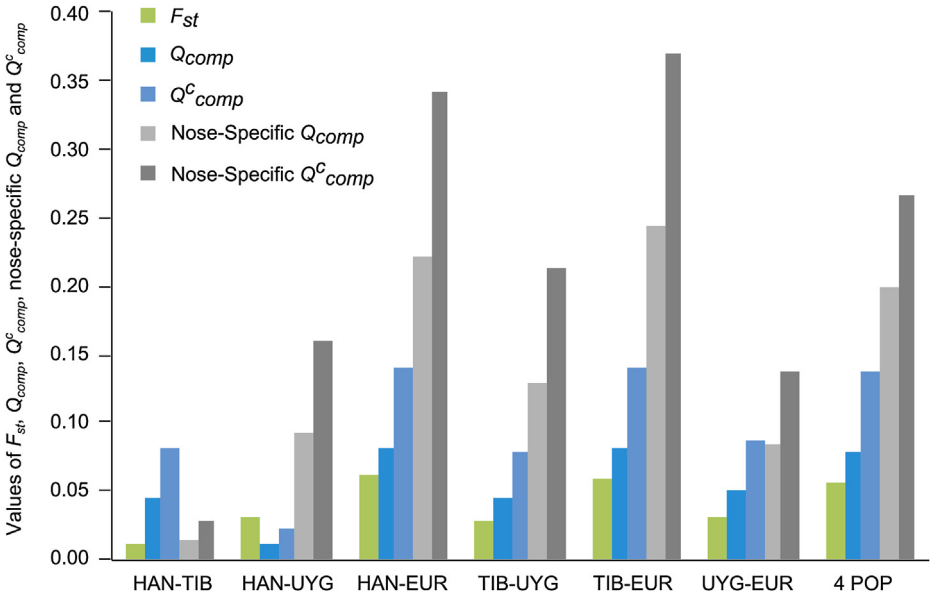

Figure 2. Estimated genetic and phenotypic differentiation for the whole face and the nose. The F_{st}, Q_{comp} and Q^c_{comp} values were calculated for the whole face and the nose for either pair-wise comparisons among the four populations or for the four populations together. HAN: Han Chinese; TIB: Tibetans; UYG: Uyghur; EUR: European. 4POP: all four populations calculated together.

significance of the Q_{st} estimates, we derived empirical P values by ranking the specific Q_{st} values against the genetic F_{st} distribution (see Methods). The highest Q^c_m values for the brow area, nose and cheeks are 0.470, 0.484 and 0.588, respectively, corresponding to the empirical P values of 0.0014, 0.0012 and 0.00024 (Table 1). The maximum Q^c_m point on the cheeks roughly marks the most prominent area. On the nose, the maximum point lies close to the nasion, and that of the brow area is the middle point bordering the

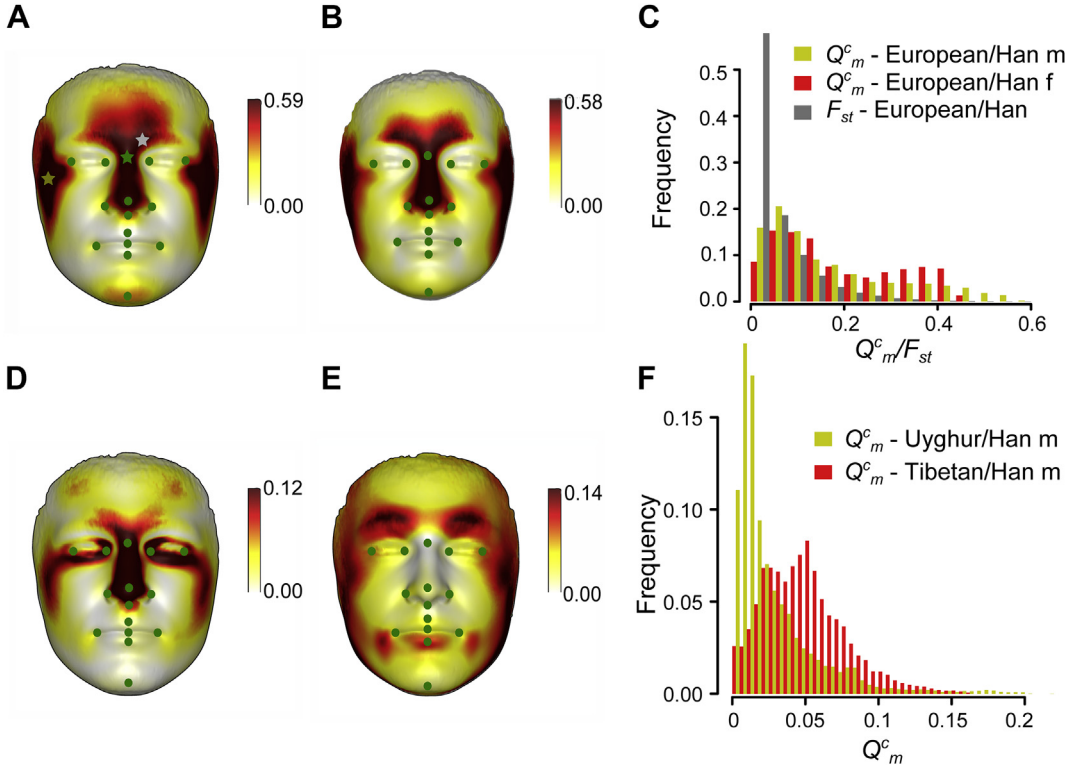

Figure 3. Patterns of differentiation in facial features between population pairs. Q_m for every quasi-landmark for each of the populations was calculated, and visualized with color gradients. (A) Differentiation pattern for the HAN-EUR comparison in males. Stars indicate the top three values, which are located in brow area (gray), nose (green) and cheekbone (olive). (B) Differentiation pattern for the HAN-EUR comparison in females. (C) Distribution of Q^c_m for males (yellow) and females (red) compared against the F_{st} distribution (gray). (D) Patterns for the HAN-UYG comparison in males. (E) Patterns for the HAN-TIB comparison in males. (F) Distribution of Q^c_m in HAN-UYG (yellow) and HAN-TIB (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Differentiation and significance for the HAN-EUR comparison in males: Q_{comp} for each specific feature, Q_p for the most significant PC_{wg} distinguishing Han and Europeans, the top Q_m for each facial feature, the corresponding corrected values (Qc_{comp}, Qc_m, Qc_p) and the empirical P values.

	Whole feature		Quasi-landmark		PC_{wg}	
	Q_{comp}/Qc_{comp}	P values	Q_m/Qc_m	P values	Q_p/Qc_p	P values
Nose	0.222	0.04	0.340	8.17×10^{-3}	0.673	5.67×10^{-5}
	0.342	8.02×10^{-3}	0.484	1.20×10^{-3}	0.790	1.13×10^{-5}
WG-cheeks	0.175	0.07	0.440	2.13×10^{-3}	0.233	0.032
	0.278	0.02	0.588	2.44×10^{-4}	0.356	6.68×10^{-3}
WOG-cheeks	0.248	0.03	0.440	2.13×10^{-3}	0.435	2.26×10^{-3}
	0.375	5.15×10^{-3}	0.588	2.44×10^{-4}	0.583	2.72×10^{-4}
Brow area	0.178	0.06	0.328	9.47×10^{-3}	0.479	1.32×10^{-3}
	0.282	0.02	0.470	1.43×10^{-3}	0.625	1.19×10^{-4}

extension of the nasal bridge (Fig. 3A). Similar maximum Q_m/Q_m^c values for the three features were also observed in females (SOM, Table S3). Interestingly, the strongest differentiation signals for the inter-landmark distances were for nasion-endocanthion ($Q_{st}^c = 0.56$) and pronasale-alare ($Q_{st}^c = 0.52$), which supports the results from the high-density data (see SOM, Fig. S3).

The HAN-UYG (Fig. 3D) and HAN-TIB (Fig. 3E) comparisons exhibit different but also highly specific patterns. The Q_m^c pattern of HAN-UYG is similar to the HAN-EUR comparison in some aspects; in particular, the nose and cheeks also show high differentiation. However, the differentiation for the brow area is moderate. Instead, the eye sockets, especially the inferior orbital areas that extend to the zygomatic bones, exhibit relatively high differentiation. The mixed pattern of differentiation indicates that some features of the Uygur face might have arisen independently from the ancestral European and Asian populations. On the other hand, the comparison between Tibetans and Han Chinese revealed less defined differentiation areas, other than the brow ridge and the lower mandible areas. It should be noted that the Q_m^c values of HAN-TIB and HAN-UYG span a much lower range (0–0.15, Fig. 3F) than that of HAN-EUR (0–0.4). Therefore, it is less clear whether the differentiation of these facial regions reflects genetic drift versus local adaptation or sexual selection.

Given the high differentiation of three facial features in the HAN-EUR comparison, we extracted their point subsets and analyzed them in more detail (SOM, Fig. S8). For the cheeks, we considered both sides together as a single feature. A pGPA was re-applied for each feature, and the PCA_{wg} decomposition was then carried out. The cheek data were also analyzed without redoing the pGPA to investigate if cheek differentiation mainly derives from the position and orientation relative to the whole face or from shape changes. For each facial feature we first calculated the whole feature Q_{comp} andQ_{comp}^c (Table 1). As expected, for the Han-European comparison, the feature specific Q_{comp} values are much higher than for the whole face. The nose and brow area have Q_{comp}^c values of 0.342 ($P = 8.02 \times 10^{-3}$) and 0.282 ($P = 0.02$), respectively. The differentiation is even larger, with a Q_{comp}^c value of 0.38 ($P = 5.15 \times 10^{-3}$) for the cheeks without pGPA re-alignment (hereafter referred to as WOG-cheeks). The relatively low Q_{comp}^c value of 0.278 ($P = 0.02$) for cheeks with pGPA (hereafter referred to as WG-cheeks) indicates that much of the inter-population variance in cheeks is explained by their relative position and orientation relative to the whole face (SOM, Fig. S9). It should be noted that Q_{comp}^c is in fact an average differentiation measurement. It is therefore conservative to compare the Q_{comp}^c values directly to single trait Q_{st} or to single locus F_{st} values. On the other hand, the Q_p^c/Q_p statistics measure differentiation along an individual PC_{wg}, and can be considered to be equivalent to a single trait Q_{st}. Since the data variance is accounted for mainly by the top PC_{wg}s, in each feature we restricted the individual PC analysis to the first 10 PC_{wg}s, where the fraction of the variance explained by the individual PC_{wg}s roughly plateaus (SOM, Fig. S5B–E). The top 10 PC_{wg}s accumulatively account for 87% of the variance for the nose, 96% for WG-cheeks, 94% for WOG-cheeks and 87% for the brow area (SOM, Fig. S5). We found that the Q_p^c values revealed even greater differentiation between Europeans and Han Chinese. The maximum Q_p^c for the nose is 0.79 ($P = 1.13 \times 10^{-5}$) for the $PC_{wg}3$, while the $PC_{wg}7$ of the brow area has a maximum Q_p^c of 0.625 ($P = 1.19 \times 10^{-4}$), and the $PC_{wg}2$ of the WOG-cheeks has a maximum Q_p^c of 0.583 ($P = 2.72 \times 10^{-4}$) (Table 1). Signals based on Q_p are similarly strong (Table 1). The top Q_p^c/Q_p values of nose, cheeks and brow area are similar in females (SOM, Table S3).

Our analyses show that the nose carries some of the strongest differentiation signals, so we examined variation in nasal form in further detail. The nose NJ trees show that the differentiation of EUR or UYG to the two East Asian populations is substantially increased compared to the whole face and the genetic trees, while the differentiation of HAN-TIB remains small (SOM, Fig. S7). Under the PCA_{wg} decomposition, the first two components seem to mainly correspond to nasal height ($PC_{wg}1$) and nasal breadth/protrusion ($PC_{wg}2$, Fig. 4A). This variation is not population specific (Fig. 4A). On the other hand, $PC_{wg}3$ and $PC_{wg}6$ account for substantial between-group variance; $PC_{wg}3$ shows the highest between-group variance, and it seems to strongly segregate the individuals into three clusters: Europeans, Uyghur and East Asians (HAN-EUR $P = 2.28 \times 10^{-43}$; EUR-TIB $P = 2.49 \times 10^{-45}$; HAN-UYG $P = 5.12 \times 10^{-32}$; UYG-EUR $P = 1.88 \times 10^{-26}$; UYG-TIB $P = 3.00 \times 10^{-23}$) (SOM, Table S4). Interestingly, Uyghur individuals lie almost exactly in the middle between Europeans and Han Chinese. The extreme forms indicate that the main changes along $PC_{wg}3$ involve the prominent nasal ridge in Europeans ($PC_{wg}3-$) compared with the recessive nose dorsum in the East Asians ($PC_{wg}3+$, Fig. 4B). Furthermore, the East Asian nose ($PC_{wg}3+$) has a much broader nasal base compared with that of Europeans ($PC_{wg}3-$). $PC_{wg}6$ also displays a large amount of inter-group variance (following $PC_{wg}3$, $PC_{wg}2$ and $PC_{wg}1$; SOM, Table S4), and it seems to distinguish all four groups from one another (Fig. 4B). The main form of variation seems to involve a recess of nose root from glabella in the $PC_{wg}6+$ type, compared with the less or non-recessed nose root in the $PC_{wg}6-$ type. The PLS analyses of the nose (Fig. 4C) produce results that are highly consistent with the PCA_{wg} plot. In particular, the HAN-EUR axis revealed similar differences as $PC_{wg}3$ among the European, Uyghur and East Asian populations (HAN-EUR, $P = 1.21 \times 10^{-43}$; HAN-UYG, $P = 7.23 \times 10^{-50}$; HAN-TIB, $P = 0.027$); and the form changes (HAN-EUR±) also resemble those found in $PC_{wg}3$ ($PC_{wg}3\pm$, Fig. 4B,C).

Effects of age and sample size

Age is known to affect facial morphology, so differences in the age distributions among the population samples could confound the analyses of morphological divergence. In order to evaluate the

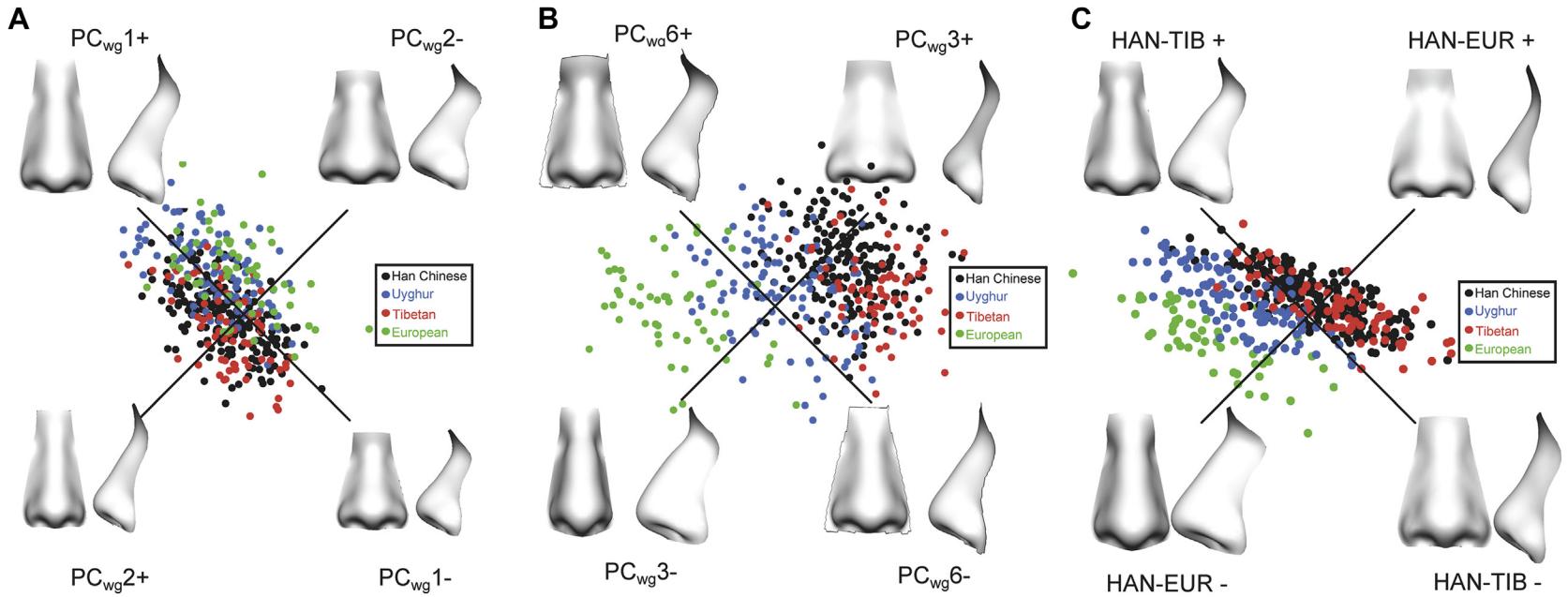

Figure 4. Structure of the four populations based on PCA_{wg} and PLS analysis of the nose. The high dimensional data of the external nose were subjected to PCA_{wg} and PLS analysis. Individual data points were plotted along $PC_{wg}1$ and $PC_{wg}2$ (A), $PC_{wg}3$ and $PC_{wg}6$ (B), or the HAN-EUR PLS axis and the HAN-TIB PLS axis (C). The extreme forms were simulated by adding/subtracting 3SD from the average form along each axis.

potential confounding effects of age, we calculated the within-group correlations between the self-reported age and face data for all of the PC_{wg} dimensions for which strong inter-group differentiation was observed. These include $PC_{wg}4$ and $PC_{wg}5$ in the whole face PCA_{wg}, $PC_{wg}3$ for the nose, $PC_{wg}2$ for WOG-cheeks, and $PC_{wg}7$ for the brow area. Other than the sporadic observations of marginally significant associations ($0.01 <$ nominal $P < 0.05$, before multiple-testing correction) in $PC_{wg}4$ and $PC_{wg}5$ of the whole face, and $PC_{wg}3$ of the nose (SOM, Table S5), there were no strong signals of association between age and facial forms. In females, the association signals between age and top PC_{wg}s were consistently low (SOM, Table S5). In addition, to control for the effects of varying ages and sample sizes, we extracted a subset of the male sample with matched age ranges (age 18–30) and sample size (38) among all four groups. Major analyses were repeated in this subset and revealed highly consistent patterns of morphological differentiation, which were comparable to the results for the complete dataset in all aspects (SOM, Fig. S10, Table S6). These analyses indicate that our results are not influenced by the varying sample ages and sizes of the groups.

Discussion

This study is, to our knowledge, the first comprehensive population differentiation analysis of the soft-tissue structures of the human face. The PCA_{wg} and PLS analyses demonstrate that soft-tissue facial form does vary among the populations, and moreover provides information about population structure (Fig. 1 and SOM, Fig. S6). Furthermore, the NJ trees reconstructed from the pairwiseQ^c_{comp} are in general consistent with the F_{st} tree based on genetic markers, and hence reflect the evolutionary relationships among the four populations. The whole face differentiation levels, estimated as Q^c_{comp} values, are approximately double the corresponding genomic F_{st} values (Fig. 2), which is a substantial deviation from neutral expectations. The higher differentiation of the soft-tissue facial form is further supported by the analyses of inter-landmark distances (SOM, Table S2).

One possible explanation for the higher differentiation values for the face versus genetic markers is that the actual heritability (h^2) may be higher than the assumed value of 0.55. However, the uncorrected whole face differentiation values, which assume the maximum possible h^2, calculated both on the high-density 3D data and the inter-landmark distances, are still higher than the genomic F_{st} values (Fig. 2). Such a pattern is different from the overall neutrality reported for human skull measurements (Relethford, 2002), suggesting that evolutionary processes other than genetic drift have substantially shaped human soft-tissue facial form.

Some previous studies of human skull variation found that certain features, such as nasion prosthion height, maximum cranial breadth and biauricular breadth, exhibited significant (although moderate) departures from neutrality (Roseman, 2004; Roseman and Weaver, 2004; Harvati and Weaver, 2006; Hubbe et al., 2009). We detected several facial regions with much higher differentiation than expected under neutrality. In particular, for the comparison of HAN-EUR, the whole feature Q_{comp} values for the nose, central brow area and cheeks (0.175–0.248) were two to three times higher than for the whole face (0.08). The point-wise differentiation estimates were even higher (maximum Q_m 0.33–0.44, Q^c_m 0.47–0.59; see Table 1). Interestingly, some inter-landmark distances involving the nose also had values higher than 0.55, consistent with the high point-wise differentiation found for the nose. On the other hand, the brow area and cheeks did not show high inter-landmark (Q^c_{st}) values, mainly because appropriate landmarks for these two facial features are lacking. This difference between the feature-based and the inter-landmark-based analyses demonstrates that high-density image analysis is crucial for capturing levels of phenotypic differentiation. This interpretation is supported by the fact that the highest differentiation signals appear in the composite complex shape changes, defined by individual principal components. Specifically, the maximum feature Q^c_p values of 0.625–0.79 (maxQ_p: 0.48–0.67; see Table 1) are much higher than the corresponding point-wise Q^c_m or the feature average Q^c_{comp} values (Table 1). This indicates that the most divergent traits may not correspond to single points or simple distances; instead, they may align with complex shape transformations of multiple anatomical structures in various aspects. This interpretation also holds for the nose. The extreme noses (Fig. 4B, $PC_{wg}3+$ and $PC_{wg}3-$) of $PC_{wg}3$ clearly show that the morphological differences involve changes in volume, height, breadth, and protrusion and dorsum shape. Similarly, the highest differentiation signals in the brow area correspond to the larger, more prominent brow ridges in Europeans (+3SD) compared with their relatively flatter shape in Han Chinese (−3SD), which also contribute to differences in eye socket depth (SOM, Fig. S11). For cheeks, the differentiation seems to mainly reflect the wider distance and more upwards/outwards prominence of the cheeks in Europeans (−3SD) compared with the flatter cheek shapes in Han Chinese (+3SD) (SOM, Fig. S9). These results support the previous observation that strong differentiation tends to involve multiple complex facial morphological variations, as captured by PCA (Roseman and Weaver, 2004).

The Q^c_p values found in the nose and brow area between Europeans and Han Chinese approach the high differentiation reported for skin pigmentation (Relethford, 2004b). This suggests that strong local adaptation may have shaped these facial features (Myles et al., 2007). For the nose, strong correlations have been found between the nasal index and temperature/humidity, supporting climate adaptation as the major selective force (Thomson and Buxton, 1923; Davies, 1932; Weiner, 1954; Wolpoff, 1968; Hiernaux and Froment, 1976; Crognier, 1981; Franciscus and Long, 1991). Models simulating airflow dynamics demonstrated that bigger nasal volumes, narrower shapes and downwardly pointing nares might enhance the airflow exposure of the mucosa and thereby facilitate the heating and humidification of the air (Churchill et al., 2004). The European nose shape thus may have resulted from adaptation to a colder climate. The relatively enlarged brow area in Europeans has also been noted previously (Russell et al., 1985). It has been argued that brow area size is positively correlated with the magnitude of the mechanical stresses resulting from mastication, so brow area shape differentiation (SOM, Fig. S11) could be the result of dietary differences (Russell et al., 1985). Adaptation to specific diets (Hubbe et al., 2009) and climate adaptation (Coon et al., 1950) have been hypothesized to explain the expanded zygomatics in Asians.

In addition to natural selection, sexual selection may also have played a major role in shaping inter-population variation in the human face. Selective mate choice based on facial appearance in humans is well documented as a universal condition in global populations (Wells et al., 2009). However, whether and to what extent sexual selection shaped human facial morphology has rarely been investigated. Fisher's runaway sexual selection model suggests that a positive feedback loop composed of an arbitrary trait involving appearance, and the accidental preference of this trait in the opposite sex, could initiate a powerful sexual selection process (Fisher, 1958). It is therefore possible that some of the strong differentiation signals involving the soft-tissue facial form may have resulted from sexual selection. Further studies that combine high-resolution 3D face analysis, studies of human behavior, and genetic analyses are necessary to delineate the possible roles of local adaptation versus sexual selection in explaining the relatively large between-population differentiation that we find for soft-tissues of the human face.

Acknowledgments

We would like to thank all of the study participants in this study. This work was supported by the Outstanding Junior Researchers (Grant No. 2011KIP201) from the CAS-SIBS (http://www.sibs.cas.cn/), the Key Research Direction Grant (No. KSCX2-EW-Q-1-12) from the CAS Knowledge Innovation Project (http://www.cas.cn/), the Max-Planck-Gesellschaft Partner Group Grant (http://www.mpg.de/), SA-SIBS 2011 Young Faculty Grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jhevol.2014.08.001.

References

Barreiro, L.B., Laval, G., Quach, H., Patin, E., Quintana-Murci, L., 2008. Natural selection has driven population differentiation in modern humans. Nat. Genet. 40, 340–345.
Churchill, S.E., Shackelford, L.L., Georgi, J.N., Black, M.T., 2004. Morphological variation and airflow dynamics in the human nose. Am. J. Hum. Biol. 16, 625–638.
Coon, C.S., 1955. Some problems of human variability and natural selection in climate and culture. Am. Nat. 89, 257–279.
Coon, C.S., Garn, S.M., Birdsell, J.B., 1950. Races: A Study of the Problems of Race Formation in Man. C.C. Thomas, Springfield.
Crognier, E., 1981. Climate and anthropometric variations in Europe and the Mediterranean area. Annls. Hum. Biol. 8, 99–107.
Darwin, C., 1871. The Descent of Man, and Selection in Relation to Sex. John Murray, London.
Davies, A., 1932. A re-survey of the morphology of the nose in relation to climate. J. R. Anthropol. Inst. 62, 337–359.
Devor, E.J., 1986. Transmission of human craniofacial dimensions. J. Craniofac. Genet. Dev. Biol. 7, 95–106.
Dryden, I.L., Mardia, K.V., 1998. Statistical Shape Analysis. John Wiley & Sons, New York.
Fisher, R.A., 1958. The Genetics Theory of Natural Selection, Second edition. Dover Publications, New York.
Franciscus, R.G., Long, J.C., 1991. Variation in human nasal height and breadth. Am. J. Phys. Anthropol. 85, 419–427.
Gill, G.W., 1998. Craniofacial Criteria in the Skeletal Attribution of Race. Advances in the Identification of Human Remains. Charles C. Thomas, Springfield.
Gunz, P., Bookstein, F.L., Mitteroecker, P., Stadlmayr, A., Seidler, H., Weber, G.W., 2009. Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario. Proc. Natl. Acad. Sci. 106, 6094–6098.
Guo, J., Mei, X., Tang, K., 2013. Automatic landmark annotation and dense correspondence registration for 3D human facial images. BMC Bioinformatics 14, 232.
Harvati, K., Weaver, T.D., 2006. Human cranial anatomy and the differential preservation of population history and climate signatures. Anat. Rec. A 288, 1225–1233.
Hennessy, R.J., Stringer, C.B., 2001. Geometric morphometric study of the regional variation of modern human craniofacial form. Am. J. Phys. Anthropol. 117, 37–48.
Hiernaux, J., Froment, A., 1976. The correlations between anthropobiological and climatic variables in sub-Saharan Africa: revised estimates. Hum. Biol. 48, 757–767.
Howells, W.W., 1995. Who's who in skulls: Ethnic identification of crania from measurements. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 82, pp. 1–108.
Hubbe, M., Hanihara, T., Harvati, K., 2009. Climate signatures in the morphological differentiation of worldwide modern human populations. Anat. Rec. 292, 1720–1733.
Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. Science 319, 1100–1104.
Mevik, B.H., Wehrens, R., Liland, K.H., 2013. pls: Partial Least Squares and Principal Component regression. R package version 2.4-3.
Miller, J.R., Wood, B.P., Hamilton, M.B., 2008. Fst and Qst under neutrality. Genetics 180, 1023–1037.
Myles, S., Somel, M., Tang, K., Kelso, J., Stoneking, M., 2007. Identifying genes underlying skin pigmentation differences among human populations. Hum. Genet. 120, 613–621.
Nei, M., Roychoudhury, A.K., 1982. Genetic relationship and evolution of human races. Evol. Biol. 14, 1–59.
Patterson, N., Price, A.L., Reich, D., 2006. Population structure and eigenanalysis. PLoS Genet. 2, e190.
Relethford, J.H., 1994. Craniometric variation among modern human populations. Am. J. Phys. Anthropol. 95, 53–62.
Relethford, J.H., 2002. Apportionment of global human genetic diversity based on craniometrics and skin color. Am. J. Phys. Anthropol. 118, 393–398.
Relethford, J.H., 2004a. Boas and beyond: migration and craniometric variation. Am. J. Hum. Biol. 16, 379–386.
Relethford, J.H., 2004b. Global patterns of isolation by distance based on genetic and morphological data. Hum. Biol. 76, 499–513.
Relethford, J.H., 2009. Race and global patterns of phenotypic variation. Am. J. Phys. Anthropol. 139, 16–22.
Relethford, J.H., 2010. Population-specific deviations of global human craniometric variation from a neutral model. Am. J. Phys. Anthropol. 142, 105–111.
Relethford, J.H., Blangero, J., 1990. Detection of differential gene flow from patterns of quantitative variation. Hum. Biol. 62, 5–25.
Relethford, J.H., Harpending, H.C., 1994. Craniometric variation, genetic theory, and modern human origins. Am. J. Phys. Anthropol. 95, 249–270.
Rohr, K., 2001. Landmark-Based Image Analysis: Using Geometric and Intensity Models. Kluwer Academic Publishers, Norwell.
Roseman, C.C., 2004. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. Proc. Natl. Acad. Sci. 101, 12824–12829.
Roseman, C.C., Weaver, T.D., 2004. Multivariate apportionment of global human craniometric diversity. Am. J. Phys. Anthropol. 125, 257–263.
Roseman, C.C., Weaver, T.D., 2007. Molecules versus morphology? Not for the human cranium. BioEssays 29, 1185–1188.
Russell, M.D., Brown, T., Garn, S.M., Giris, F., Turkel, S., İşcan, M.Y., Oyen, O.J., Jacobshagen, B., Pietrusewsky, M., Rightmire, G.P., 1985. The supraorbital torus:" A most remarkable peculiarity". Curr. Anthropol. 26, 337–360.
Sparks, C.S., Jantz, R.L., 2002. A reassessment of human cranial plasticity: Boas revisited. Proc. Natl. Acad. Sci. 99, 14636–14639.
Spitze, K., 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135, 367–374.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Thomson, A., Buxton, L.H.D., 1923. Man's nasal index in relation to certain climatic conditions. J. R. Anthropol. Inst. 53, 92–122.
Team R., 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria.
von Cramon-Taubadel, N., 2009a. Revisiting the homoiology hypothesis: the impact of phenotypic plasticity on the reconstruction of human population history from craniometric data. J. Hum. Evol. 57, 179–190.
von Cramon-Taubadel, N., 2009b. Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. Am. J. Phys. Anthropol. 140, 205–215.
von Cramon-Taubadel, N., Weaver, T.D., 2009. Insights from a quantitative genetic approach to human morphological evolution. Evol. Anthropol. 18, 237–240.
Weiner, J.S., 1954. Nose shape and climate. Am. J. Phys. Anthropol. 12, 615–618.
Weir, B., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.
Weir, B.S., Cardon, L.R., Anderson, A.D., Nielsen, D.M., Hill, W.G., 2005. Measures of human population structure show heterogeneity among genomic regions. Genome Res. 15, 1468–1476.
Wells, T., Dunn, A., Sergeant, M.J.T., Davies, M.N.O., 2009. Multiple signals in human mate selection: A review and framework for integrating facial and vocal signals. J. Evol. Psychol 7, 111–139.
Whitlock, M.C., Guillaume, F., 2009. Testing for spatially divergent selection: comparing QST to FST. Genetics 183, 1055–1063.
Wolpoff, M.H., 1968. Climatic influence on the skeletal nasal aperture. Am. J. Phys. Anthropol. 29, 405–423.
Wright, R.V.S., 1992. Correlation between cranial form and geography in *Homo sapiens*: CRANID-a computer program for forensic and other applications. Archaeol. Ocean. 27 (3), 128–134.
Wright, S., 1950. Genetical structure of populations. Nature 166, 247–249.
Xu, S., Li, S., Yang, Y., Tan, J., Lou, H., Jin, W., Yang, L., Pan, X., Wang, J., Shen, Y., Wu, B., Wang, H., Jin, L., 2011. A genome-wide search for signals of high-altitude adaptation in Tibetans. Mol. Biol. Evol. 28, 1003–1011.