Dissecting changes and differences of selective pressures in the evolution of human pigmentation

Xin Huang[1,*], Sijia Wang[1,2], Yungang He[3,*], Li Jin[1,2,*]

[1] Chinese Academy of Sciences Key Laboratory of Computational Biology, Chinese Academy of Sciences-Max Planck Society Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, 200031, China

[2] State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, 200433, China

[3] Institute of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai, 200032, China

[*] **Corresponding authors:** xin.huang07@gmail.com (X.H.), heyungang@fudan.edu.cn (Y.H.), lijin@fudan.edu.cn (L.J.)

Keywords: population genetics, natural selection, human evolution, human pigmentation, complex traits

Abstract

Human pigmentation is a highly diverse and complex trait among populations, and has drawn particular attention from both academic and non-academic investigators for thousands of years. This diversity may be adapted for ultraviolet radiation and shaped by natural selection. Previous studies detected selection signals in several human pigmentation genes, but few studies have examined how multiple genes contributed to the evolution of human pigmentation. Moreover, none has quantified selective pressures on human pigmentation over epochs and between populations. Here, we dissect changes and differences of selective pressures during different periods and between distinct populations with new approaches. We propose a new model with multiple populations to estimate historical changes of selective pressures by summarizing selective pressures on multiple genes. We collected genotype data of 19 genes associated with human pigmentation from 15 datasets, and obtained data for 2498 individuals of five representative populations from worldwide. Our results reveal not only a recent incremental change of selective pressure (13.32×10^{-3}/generation) in modern Europeans, but also a significant historical increase of selective pressure (14.60×10^{-3}/generation) on light pigmentation shared by all Eurasians during the Out-of-Africa event. Our results also suggest that several derived alleles recently associated with human dark pigmentation may be under directional selection. Thus, directional selection not only affects light pigmentation in Eurasians, but also influences dark pigmentation in Africans. Our study provides quantitative information on the evolution of human pigmentation, and may facilitate studies of the evolution of other complex traits.

Introduction

Human pigmentation—the color of human skin, hair, and eye—is one of the most diverse traits among populations. Its obvious diversity has attracted particular attention from both academic and non-academic investigators for thousands of years, as noted by Charles Darwin one century ago (Darwin 1889) and as noticed by ancient Egyptians more than 4000 years ago (Norton 2005). Why human pigmentation diverges, however, remains a central puzzle in human biology (Rees and Harding 2012). Some researchers have proposed that the diversity of human pigmentation is adapted for ultraviolet radiation (UVR) and driven by natural selection (Parra 2007). Dark skin may prevent sunburn amongst individuals in low latitude areas with high UVR, while light skin may protect infants against rickets in high latitude areas with low UVR (Parra 2007). A better understanding of how natural selection shapes the diversity of human pigmentation could provide relevant and beneficial information for public health (Parra 2007).

During the last decade, many studies have applied statistical tests to detect signals of natural selection in several human pigmentation genes (Izagirre et al. 2006; Voight et al. 2006; Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Pickrell et al. 2009; Beleza et al. 2013; Hider et al. 2013; Wilde et al. 2014). These genes encode different proteins, including: signal regulators—ASIP, KITLG, MC1R—stimulating the melanogenic pathway; possible enhancers—BNC2, HERC2—regulating pigmentation gene expression; important enzymes—TYR, TYRP1, DCT—converting tyrosine into melanin; putative exchangers—OCA2, SLC24A4, SLC24A5, SLC45A2, TPCN2—controlling the environment within melanosomes; and an exocyst complex unit and molecular motor—EXOC2, MYO5A—conveying vesicles and organelles within the cytoplasm (Abdel-Malek et al. 2001; Rebbeck et al. 2002; Duffy et al. 2004, 2007; Graf et al. 2005; Sulem et al. 2007, 2008; Anno et al. 2008; Han et al. 2008; Ito and Wakamatsu 2008; Kayser et al. 2008; Sturm and Duffy 2012; Visser et al. 2012, 2014; Guenther et al. 2014). These proteins work at different stages of the melanogenic pathway, illustrating that human pigmentation is a complex trait affected by multiple genes with different roles.

Previous studies applied two groups of methods to detect natural selection. One group of methods detects unusually long extended haplotype homozygosity (Izagirre et al. 2006;

Voight et al. 2006; Lao et al. 2007; Myles et al. 2007; McEvoy et al. 2009; Pickrell et al. 2009; Donnelly et al. 2012; Beleza et al. 2013). The other group of methods identifies extremely local population differentiation (Izagirre et al. 2006; Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Pickrell et al. 2009; Hider et al. 2013). By applying both groups of methods, previous studies have aimed to understand the evolution of individual pigmentation genes; however, few studies have examined how multiple genes contributed to the evolution of human pigmentation. Moreover, none of these studies have quantitatively investigated the historical selective pressures of pigmentation genes during different epochs, and compared the differences of selective pressures between different populations. Therefore, it is necessary to perform an extensive quantification of selective pressures on human pigmentation using a creative approach.

New Approaches

On the basis of a previous study (He et al. 2015), we measured selective pressures by (genic) selection coefficients. For any single nucleotide polymorphism (SNP) *L*, we can estimate the expectation of the selection (coefficient) difference per generation between populations *i* and *j* by

$$d_{ij}(L)=\left[\ln\frac{p_i(L)}{q_i(L)}-\ln\frac{p_j(L)}{q_j(L)}\right]/t_{ij}\,, \tag{1}$$

where *p* and *q* are the frequencies of derived and ancestral alleles in a population, respectively; and t_{ij} is the divergence time of the populations *i* and *j*. Details of the calculations are described elsewhere (He et al. 2015).

We then extended Eq. 1 to summarize selection differences of multiple SNPs, because a complex trait is usually affected by multiple loci. For a trait with *n* SNPs in linkage equilibrium, the expectation and variance of the total selection difference in this trait is

$$\begin{aligned} d_{ij}(L_1L_2\cdots L_n)&=\sum_{i=1}^{n}d_{ij}(L_k)\\ \mathrm{var}\left[d_{ij}(L_1L_2\cdots L_n)\right]&=\sum_{i=1}^{n}\mathrm{var}\left[d_{ij}(L_k)\right] \end{aligned}\,. \tag{2}$$

Based on Eq. 2, we developed a new approach for dissecting historical changes of selective pressures over epochs of human evolutionary history. We simplified the evolutionary

history of five representative human populations as a binary tree (Fig. 1). Let s_0 denote the selection coefficient of the population in the root. We can express the selection coefficients of populations in the different branches using s_0 plus some selection (coefficient) changes (c_i in Fig. 1). We further assumed $s_8 \approx s_0$, because modern humans may originated in East Africa. Thus, populations *r* and *x* may have had a similar environment, and the selective pressures between them did not change much. We can represent the selection differences of all the paired populations in Fig. 1 as

$$\mathbf{d} = \mathbf{T}\mathbf{c}, \tag{3}$$

where

$$\mathbf{d} = \begin{pmatrix} d_{21}t_1 & d_{31}t_1 & d_{41}t_1 & d_{51}t_1 & d_{32}t_2 & d_{42}t_2 & d_{52}t_2 & d_{43}t_3 & d_{53}t_3 & d_{54}t_4 \end{pmatrix}^{tr}$$

$$\mathbf{c} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 \end{pmatrix}^{tr},$$

and

$$\mathbf{T} = \begin{bmatrix} -t_1 & t_2 & 0 & 0 & 0 & 0 & 0 \\ -t_1 & 0 & t_3 & 0 & 0 & 0 & t_2 \\ -t_1 & 0 & 0 & t_4 & 0 & t_3 & t_2 \\ -t_1 & 0 & 0 & 0 & t_4 & t_3 & t_2 \\ 0 & -t_2 & t_3 & 0 & 0 & 0 & t_2 \\ 0 & -t_2 & 0 & t_4 & 0 & t_3 & t_2 \\ 0 & -t_2 & 0 & 0 & t_4 & t_3 & t_2 \\ 0 & 0 & -t_3 & t_4 & 0 & t_3 & 0 \\ 0 & 0 & -t_3 & 0 & t_4 & t_3 & 0 \\ 0 & 0 & 0 & -t_4 & t_4 & 0 & 0 \end{bmatrix}.$$

Here, the superscript *tr* denotes the transpose of a vector. To obtain the optimal solution, we propose a probabilistic approach. Under neutral evolution (NE), we consider each estimated *c* as an independent random variable following a normal distribution with a mean of zero and a variance of σ^2. For each solution with seven variables, the summation below follows a chi-square distribution with seven degrees of freedom:

$$\sum_{i=1}^{7} \frac{c_i^2}{\sigma_i^2} \sim \chi^2(7).$$

Therefore, we have $\Pr(|\mathbf{c}|^2 > |\mathbf{c}_a|^2 \mid \text{NE}) \geq \Pr(|\mathbf{c}|^2 > |\mathbf{c}_b|^2 \mid \text{NE})$, if $|\mathbf{c}_a|^2 \leq |\mathbf{c}_b|^2$ for solutions *a* and *b*. Here, $|\mathbf{c}|^2$ is the norm of a vector **c**. In other words, we can choose the most conservative solution with the least deviation from neutral evolution using a probabilistic

approach. Thus, the optimal solution $\mathbf{c}^*$ is

$$\mathbf{c}^* = \mathbf{T}^+\mathbf{d}, \tag{4}$$

where $\mathbf{T}^+$ is the pseudo-inverse of $\mathbf{T}$. We provide more details of our model in Supplementary Text S1.

Using matrix algebra, the variance of the ith element in the optimal solution $\mathbf{c}^*$ is

$$\mathrm{var}\left(c_i^*\right) = \mathbf{T}_i^+\mathbf{V}\left(\mathbf{T}_i^+\right)^{tr}, \tag{5}$$

where $\mathbf{V}$ is the variance and covariance matrix of selection differences between populations, and $\mathbf{T}_i^+$ is the ith row of the matrix $\mathbf{T}^+$. We estimated the variance of $d_{ij}t_{ij}$ as the previous study (He et al. 2015). To estimate the covariance of $d_{ij}t_{ij}$ and $d_{kl}t_{kl}$, we used the variances of the shared branches between $d_{ij}t_{ij}$ and $d_{kl}t_{kl}$, as suggested by a previous study (Nei and Jin 1989). Take Fig. 1 as an example. The covariance of $d_{31}t_1$ and $d_{21}t_1$ is the sum of the variances in the branches b_{rx} and b_{r1}, because b_{rx} and b_{r1} are the shared branches between $d_{31}t_1$ and $d_{21}t_1$. If $d_{ij}t_{ij}$ and $d_{kl}t_{kl}$ have no shared branches, then their covariance is zero. To obtain the variance of each branch, we assumed they contribute independently to the variances of selection differences, and estimated them with a least-squares approach. Thus, the variance of the ith branch is

$$\mathrm{var}\left(b_i\right) = \mathbf{A}_i^+\mathbf{v},$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathrm{var}\left(d_{21}t_1\right) \\ \mathrm{var}\left(d_{31}t_1\right) \\ \mathrm{var}\left(d_{41}t_1\right) \\ \mathrm{var}\left(d_{51}t_1\right) \\ \mathrm{var}\left(d_{32}t_2\right) \\ \mathrm{var}\left(d_{42}t_2\right) \\ \mathrm{var}\left(d_{52}t_2\right) \\ \mathrm{var}\left(d_{43}t_3\right) \\ \mathrm{var}\left(d_{53}t_3\right) \\ \mathrm{var}\left(d_{54}t_4\right) \end{bmatrix},$$

and $\mathbf{A}_i^+$ is the ith row of the pseudo-inverse of matrix $\mathbf{A}$. It should be noted that $\mathrm{var}\left(b_1\right)$ is the sum of the variances in the branches b_{rx} and b_{r1}. Because we assumed that c_i follows a

normal distribution, the 95% confidence interval (CI) of c_i is given by $c_i^* + 1.96 \cdot \sqrt{\mathrm{var}\left(c_i^*\right)}$.

Results

Selection changes over epochs

To apply our new approach, we first summarized selection differences in 25 individual SNPs with weak linkage disequilibrium ($r^2 < 0.2$), representing the total selective pressures on human pigmentation. Fig. 2 plots the total selection differences, and Table 1 provides more details. The maximum differences (30.538 and 30.795 $\times 10^{-3}$/generation) were observed between Europeans and the two African populations, while the minimum difference (-8.406 $\times 10^{-3}$/generation) was observed between West and East Africans. The estimated 95% CIs indicate that these selection differences significantly deviate from zero (Table 2). Our results of selection differences are consistent with those of skin reflectance differences between populations (Spearman correlation coefficient $\rho = 0.98$, $p < 10^{-6}$), suggesting that human pigmentation diversity may be explained by different strengths of selective pressures among populations.

We then solved the linear system (Eq. 4) with the estimated selection differences on human pigmentation (Table 1), and estimated the variances of the optimal solution (Eq. 5 and Supplementary Table S3). The optimal solution (Fig. 3 and Table 3) matches well with the changes of human skin color in native populations (Barsh 2003), and indicates that the selective pressure on human pigmentation changed markedly during the Out-of-Africa event (c_7 = 14.60 $\times 10^{-3}$/generation), probably because of a dramatic environmental change during this period. Our estimate shows that the modern European lineage had an incremental change of selective pressure (c_3 = 13.32 $\times 10^{-3}$/generation), suggesting that natural selection has continuously favored light pigmentation in Europeans since the Out-of-Africa event. Recent studies using ancient DNA could support our observation of recent directional selection in Europeans (Wilde et al. 2014; Mathieson et al. 2015), whereas the modern East Asian lineage had a smaller change (c_5 = -4.88 $\times 10^{-3}$/generation), possibly experiencing relaxation of selective pressures on light pigmentation during this period. These observations are consistent with the lower skin reflectance in East Asians compared with Europeans (Fig. 2). A microbial study also observed the fitness change was larger at the early stage than at the late stage

(Holder and Bull 2001). These diminishing returns suggest a general pattern of adaptation may exist in both microbe and human evolution (Orr 2005).

We applied SLiM 2 (Haller and Messer 2017) to examine whether the optimal solution could reproduce the selection differences (Fig. 2) we observed from the data. We set up a human demographic model according to previous studies (Schaffner et al. 2005; McEvoy et al. 2011), and used the optimal solution to calculate selective pressures as parameters for each period (Materials and Methods). The simulated selection differences are close to the data, and are little affected by the initial frequency of the beneficial allele (Supplementary Fig. S6).

Selection differences between populations

We also separately quantified selection differences of the 42 selected candidate SNPs associated with human pigmentation (Supplementary Table S1) to explore selection patterns of individual loci. Statistical tests suggest selective pressures in many loci differed significantly between populations. Based on the selection patterns of these SNPs, we categorized them into five groups (Fig. 4).

In the first group, derived alleles may be affected by Eurasian-shared selection (Fig. 4A). This result is consistent with the observation that reduced pigmentation occurred in Eurasians. Here, we observed significant selection differences in *ASIP*, *MC1R*, and *DDB1* between Eurasians and Africans (Table 4 and Supplementary Table S2). Among these SNPs, rs6119471 (*ASIP*) had the largest selection difference between Eurasians and Africans (Table 4). The derived allele of rs6119471 (*ASIP*) is almost fixed across Eurasians, but maintains a low frequency in Africans (Supplementary Fig. S2). Recent studies applied this SNP to predict dark/non-dark pigmentation phenotype in humans (Spichenok et al. 2011). This may be due to different selective pressures on this SNP among populations. Other SNPs, such as rs642742 (*KITLG*) and rs10756819 (*BNC2*), exhibited non-significant differences, suggesting that selection may contribute less to the diversity of these SNPs.

In the second group, derived alleles may be affected by African-specific selection (Fig. 4B). All these derived alleles are in/near two genes (*DDB1* and *MFSD12*) and were very recently associated with human dark pigmentation (Crawfold et al. 2017). Our results show that these SNPs differ significantly—possible signals of directional selection. From the first and second groups, we can observe that directional selection not only affects derived alleles

associated with light pigmentation in Eurasians, but also influences derived alleles associated with dark pigmentation in Africans. This observation suggests that human pigmentation is under directional selection with diversifying orientations. Thus, the previous view that the dark pigmentation in Africans is the result of purifying selection on ancestral alleles is incomplete. Moreover, these new loci illustrate that we still have little understanding of the genetic diversity within African populations, and more research in this area is needed (Lasisi and Shriver 2018).

The third and fourth groups display European- and Asian-specific selection, respectively (Fig. 4C and 4D). One notable SNP is rs1426654 (*SLC24A5*), which has the largest selection difference between Europeans and East Asians in our study (4.795×10^{-3}/generation). Previous studies reported that this SNP is under strong directional selection in Europeans (Izagirre et al. 2006; Voight et al. 2006; Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Pickrell et al. 2009). Another notable SNP is rs1800414 (*OCA2*), which had a large selection difference between East Asians and other populations. This reveals a potential role of rs1800414 (*OCA2*) on light pigmentation in East Asians. Several studies have suggested directional selection on this SNP in East Asians (Edwards et al. 2010; Yang et al. 2016). These large selection differences indicate the significant contributions of these SNPs to light pigmentation in Europeans and East Asians, respectively. Other SNPs in these groups also support the hypothesis that recent natural selection for light pigmentation independently occurred in Europeans and Asians since they diverged (Norton et al. 2007; Edwards et al. 2010; Yang et al. 2016).

The last group includes the five remaining SNPs (Fig. 4E), which exhibit specific selection differences between limited population pairs. Among them, the derived allele of rs1800401 (*OCA2*) and the ancestral allele of rs12896399 (*SLC24A4*) are both associated with dark pigmentation (Supplementary Table S2). Only rs12896399 (*SLC24A4*) differs significantly between West Africans and Eurasians (Supplementary Table S2). This may be a selection signal associated with dark pigmentation in West Africans, again indicating possible genetic diversity within African populations. We note that rs35264875 (*TPCN2*) and rs12821256 (*KITLG*) might be affected by selection in both East Africans and Europeans. Recent study showed that rs12821256 might have large effect on the skin pigmentation in

South Africans (Martin et al. 2017). The other two SNPs, rs3829241 (*TPCN2*) and rs642742 (*KITLG*), also differ between Eurasians and Africans (Fig. 4A). These similar patterns of *TPCN2* and *KITLG* suggest some connection between them.

Discussion

Our study provides the first attempt to resolve time-varied selective pressures in the evolution of human pigmentation. The fluctuation of selective pressures is an important factor in evolution (Crow and Kimura 2009), but few studies have revealed time-varied selective pressures in human populations. In the case of human pigmentation, as human populations migrated out from Africa, the UVR in the environment they encountered also changed. Thus, it is possible that the selective pressures on human pigmentation varied throughout human evolutionary history. Our results not only demonstrate independent selective pressures on light pigmentation in modern Europeans and East Asians as in previous studies (Myles et al. 2007; Norton et al. 2007; Edwards et al. 2010; Yang et al. 2016), but also indicate a large increase of selective pressure that favored light pigmentation in proto-Eurasian populations. This shared selection is consistent with other studies, revealing that *ASIP*, *BNC2*, and *KITLG* were under directional selection before the divergence of ancestral Europeans and Asians (McEovy et al. 2009; Beleza et al. 2013).

Our results also indicate that two SNPs in *MC1R* (rs2228479 and rs885479) largely differ between Eurasians and Africans (Table 4). Previous studies used variants in *MC1R* to solve a long-standing puzzle, regarding whether light pigmentation in low UVR areas is caused by directional selection or the relaxation of selective pressures (Rana et al. 1999; Harding et al. 2000; Wilde et al. 2014). The relaxation of selective pressures would suggest that the diversity of *MC1R* variants increased in Eurasians due to the lack of selective constraints. In this scenario, the genetic diversity of *MC1R* variants could be largely attributed to genetic drift. In contrast, directional selection would suggest that *MC1R* variants were under positive selection in Eurasians. In this scenario, genetic drift cannot explain the genetic divergence of *MC1R* between Africans and Eurasians. Our statistical results show that the divergences of rs2228479 and rs885479 between Eurasians and Africans are highly significant departure from neutral evolution (Supplementary Table S2), suggesting that directional selection is the more likely explanation. Experimental evidence suggests that the derived allele of rs2228479

could cause lower affinity for alpha-melanocyte stimulating hormone than the ancestral allele (Xu et al. 1996). Another study showed that the derived allele of rs885479 carries a lower risk of developing freckles and severe solar lentigines than the ancestral allele in East Asians (Motokawa et al. 2007). These studies revealed the potential roles of these *MC1R* variants in pigmentation phenotypes.

Compared with previous studies (Izagirre et al. 2006; Voight et al. 2006; Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Pickrell et al. 2009; Beleza et al. 2013; Hilder et al. 2013; Wilde et al. 2014), our study has three advantages. First, our approach considers the fluctuation of selective pressures that were ignored by previous studies. Our results provide more information about the changes of selective pressures during human evolution. Second, we summarize selective pressures based on multiple human pigmentation genes with a larger sample size and more representative populations. Therefore, our results could be more relevant to the evolution of human pigmentation. Third, we do not need to assume population continuity as in those ancient DNA studies (Wilde et al. 2014; Mathieson et al. 2015), because our study is based on genetic data from only present-day populations. Thus, our results could provide more solid evidence of the recent directional selection in Europeans.

We note that our investigation has several limitations. First, our model is based on the infinite population size model. The limited sample size would affect our results. Analysis of a larger sample size could improve our estimate, as more and more genomic datasets become available. Second, although we chose the solution that deviates least from neutral evolution as the optimal solution, we cannot exclude the possibility of other solutions. This reflects the difficulty of analyzing historical selective pressures, which is a well-recognized challenge in population genetics (Crow and Kimura 2009). Our solution provides a first step toward resolving the historical changes of selection in the evolution of human pigmentation. This solution may be improved by combining both ancient and modern human genetic data, as well as by using a Bayesian framework for inference. Third, our results may be affected by a severe bottleneck. A recent study (Terhorst et al. 2017) suggests a more severe Out-of-Africa bottleneck in human evolutionary history than in the model used in our simulation. This would probably reduce the selection differences between Eurasians and Africans, leading to an underestimation of the selection changes. Finally, our results may also be affected by

population migration and sub-structure. We applied two approaches to reduce these effects. One is that we used knowledge from previous studies, principle component analysis and F_3 test to rigorously prune potential admixed populations (Materials and Methods). The other is that we demonstrated that our estimate provides lower bounds of selection differences on human pigmentation when migration or sub-structure exists (Supplementary Text S2).

To summarize, we extended an established method (He et al. 2015) to dissect historical changes of selective pressures over epochs. Using genetic variants from worldwide populations, we quantified the changes of selective pressures on human pigmentation during different periods of human evolutionary history. Our results well explain the current features of human pigmentation among representative populations. Using individual variants of pigmentation genes, we thoroughly compared the differences of selective pressures between populations. Our results suggest that directional selection not only affects light pigmentation in Eurasians, but also influences dark pigmentation in Africans. Our study provides information about changes and differences of selective pressures during different epochs and between distinct populations. Further studies are in progress to verify our present views on the evolution of human pigmentation.

Materials and Methods

Data collection

Fifteen datasets (Li et al. 2008; Teo et al. 2009; The 1000 Genomes Project Consortium 2010; The International HapMap 3 Consortium 2010; Behar et al. 2010; Rasmussen et al. 2010; Metspalu et al. 2011; Pagani et al. 2012; Yunusbayev et al. 2012; Cristofaro et al. 2013; Fedorova et al. 2013; Xing et al. 2013; Kovacevic et al. 2014; Raghavan et al. 2014; Yunusbayev et al. 2015) containing genotype data from worldwide human populations were obtained from the listed resources (Supplementary Table S4 and S5).

Data preparation

All the genotype data were transferred to genomic coordinates using NCBI dbSNP (build 144) with the Human Reference Genome hg19. A merged dataset containing 6119 individuals and 81,364 SNPs was obtained after removing duplicate and related individuals (Supplementary Table S6 and S7). The total genotyping rate was larger than 0.99. SNPs in strong linkage disequilibrium were removed by applying a window of 200 SNPs advanced by 25 SNPs and

an r^2 threshold of 0.8 (--indep-pairwise 200 25 0.8) in PLINK 1.7 (Purcell et al. 2007). The remaining 80,377 SNPs were used for further analysis. One hundred and twenty-eight individuals (Supplementary Table S8) were identified as outliers by principal component analysis (PCA) using SMARTPCA (version: 13050) from EIGENSOFT 6.0.1 (Patterson et al. 2006; Price et al. 2006). Potential admixed populations were excluded according to previous studies (Li et al. 2008; Teo et al. 2009; The 1000 Genomes Project Consortium 2010; The International HapMap 3 Consortium 2010; Behar et al. 2010; Rasmussen et al. 2010; Metspalu et al. 2011; Pagani et al. 2012; Yunusbayev et al. 2012; Cristofaro et al. 2013; Fedorova et al. 2013; Xing et al. 2013; Kovacevic et al. 2014; Raghavan et al. 2014; Yunusbayev et al. 2015), PCA, and F_3 test (Supplementary Table S9) in ADMIXTOOLS 3.0 (Patterson et al. 2012). After removing individuals from potential admixed populations (Supplementary Table S10) and outliers identified by PCA, 2498 individuals were obtained and divided into five groups according to their geographic regions for further analysis (Supplementary Table S11). PCA plots (Supplementary Fig. S1) showed that these 2498 individuals were properly separated into different population groups.

Data imputation

Genotypes of 19 human pigmentation genes with 500-kb flanking sequences on both sides were obtained from the genotype datasets. Haplotype inference and genotype imputation were performed on the selected genotypes using BEAGLE 4.1 (Browning SR and Browning BL 2007; Browning BL and Browing SR 2016) with 1000 Genomes phase 3 haplotypes as the reference panel (Supplementary Table S12). During phasing and imputation, the effective population size was assumed to be 10,000 (N_e = 10000), and the other parameters were set to the default values. Fifty-two SNPs were selected for analysis due to their association with human pigmentation in published genome-wide association studies or phenotype prediction models (Supplementary Table S1). Ten SNPs (rs1110400, rs11547464, rs12203592, rs1800407, rs1805005, rs1805006, rs1805007, rs1805008, rs1805009, rs74653330) were removed because of their low frequencies in our datasets after imputation (Supplementary Fig. S2). Because rs12203592 (*IRF4*) was removed, 18 genes with the remaining 42 SNPs were used for further analysis.

Estimation of selection differences and changes

We used Eq. 1 to estimate the selection differences of the remaining 42 SNPs. We then used Eq. 2 and selected 25 SNPs with weak linkage disequilibrium ($r^2 < 0.2$) as well as known phenotypes to estimate the total selection differences on human pigmentation between populations (Supplementary Fig. S3 and Supplementary Table S2). These SNPs were: rs16891982, rs26722, rs4959270, rs1408799, rs12350739, rs11230664, rs35264875, rs3829241, rs1042602, rs1393350, rs642742, rs12821256, rs1407995, rs12896399, rs2402130, rs1800414, rs1800401, rs12913832, rs1426654, rs2228479, rs885479, rs56203814, rs6510760, rs6119471, rs2378249. To dissect selection changes over epochs, we applied Eq. 4 with the total selection differences from the selected 25 SNPs and the divergence times shown in Fig. 1.

Reproducing the observed selection differences from the optimal solution

We used SLiM 2 (Haller and Messer 2017) to simulate a demographic model of human evolution (Supplementary Fig. S5) to examine whether the optimal solution could reproduce the observed selection differences. We assumed $s_0 = 0$, and varied the initial frequency of the beneficial allele with 0.001 and 0.01. We divided the optimal solution (Table 4) by 25 to obtain the average selection changes for each SNP, because we used 25 SNPs to estimate the total selection differences on human pigmentation. We determined the selection coefficients during different periods in SLiM 2 based on these average selection changes (Supplementary Text S3). We used the effective population size of each population estimated by a previous study (McEvoy et al. 2011). We set both the mutation rate and the recombination rate to 1×10^{-8}/generation/site. In each run, we simulated a fragment with 10^6 base pairs, and set the 50,000th site under selection. We repeated each set of parameters more than 10,000 times, and analyzed those results in which beneficial alleles were not fixed or lost in all the populations. We compared the average selection differences from simulation with the observed selection differences (Supplementary Fig. S6). All the simulations were performed in Digital Ocean (https://cloud.digitalocean.com/) Optimized Droplets. The information of these droplets is as follows: CPU, Intel® Xeon® Platinum 8168 Processor; Random-access memory, 64 GB; Operating system, Ubuntu 16.04.4 x64.

We noticed that the selection coefficient in SLiM 2 measures differences in fitness between genotypes instead of alleles. We can transform the selection coefficient of genotypes

into that of alleles as follows. Let the fitness of the ancestral allele *A* be 1, and the relative fitness of the derived allele *a* is e^s. When *s* is close to 0, we can approximate e^s as 1 + *s* using the Taylor series. The fitness of genotype *aa* becomes $(1+s)^2 = 1+2s+s^2 \approx 1+2s$, and the fitness of genotype *Aa* is $1+s = 1+0.5s'$. If *s'* is the selection coefficient in SLiM 2, then *s'* = 2*s*; and the dominance coefficient becomes 0.5. Therefore, the selection difference is $d_{ij} = s_i - s_j = 0.5(s_i' - s_j')$.

Skin reflectance data

We selected skin reflectance data at 685 nm based on a previous study (Jablonski and Chaplin 2000) and calculated their pairwise differences. We used five populations: Yoruba (27.4, *n* = 3), Ethiopian (32.6, *n* = 4), German (66.9, *n* = 1), Ainu (59.1, *n* = 2), and Chinese (58.8, *n* = 2) to represent West Africans, East Africans, Europeans, North Asians and East Asians in our study, respectively.

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (91331109 to Y.H.; 31271338 and 31330038 to L.J.; 31322030 and 91331108 to S.W.). Y.H. is also grateful for support from the SA-SIBS scholarship program and the Youth Innovation Promotion Association of Chinese Academy of Science (2012216). L.J. was also supported by the Shanghai Leading Academic Discipline Project (B111) and the Center for Evolutionary Biology at Fudan University. S.W. was also awarded by the National Thousand Young Talents Award, the Max Planck-CAS Paul Gerson Unna Independent Research Group Leadership Award, and open projects from the State Key Laboratory of Genetic Engineering at Fudan University. X.H. is grateful to Dr. Minxian Wang, Dr. Yuchen Wang, Dr. Haiyi Lou and Dr. Lin Tang for comments on the manuscript.

References

Abdel-Malek ZA, Scott MC, Furumura M, Lamoreux ML, Ollmann M, Barsh GS, Hearing VJ. 2010. The melanocortin 1 receptor is the principal mediator of the effects of agouti signaling protein on mammalian melanocytes. *J Cell Biol*. 114:1019–1024.

Anno S, Abe T, Yamamoto T. 2008. Interactions between SNP alleles at multiple loci contribute to skin color differences between Caucasoid and Mongoloid subjects. *Int J Biol Sci.* 4:81–86.

Barsh GS. 2003. What controls variation in human skin color? *PLoS Biol.* 1:19–22.

Behar DM, Yunusbayev B, Metsplau M, Metsplau E, Rosset S, Parik J, Rootsi S, Chaubey G, Kutuev I, Yudkovsky G, et al. 2010. The genome-wide structure of the Jewish people. *Nature* 466:238–242.

Beleza S, Santos AM, McEvoy B, Alves I, Martinho C, Cameron E, Shriver MD, Parra EJ, Rocha J. 2013. The timing of pigmentation lightening in Europeans. *Mol Biol Evol.* 30:24–35.

Browning BL, Browning SR. 2016. Genotype imputation with millions of reference samples. *Am J Hum Genet.* 98:116–126.

Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies by use of localized haplotype clustering. *Am J Hum Genet.* 81:1084–1097.

Crawfold NG, Kelly DE, Hansen MEB, Beltrame MH, Fan S, Bowman BL, Jewett E, Ranciaro A, Thompson S, Lo Y, et al. 2017. Loci associated with skin pigmentation identified in African populations. *Science* 358:eaan8433.

Cristofaro JD, Pennarun E, Mazières S, Myres NM, Lin AA, Temori SA, Metspalu M, Metspalu E, Witzel M, King RJ, et al. 2013. Afghan Hindu Kush: where Eurasian sub-continent gene flows converge. *PLoS One* 8:e76748.

Crow JF, Kimura M. 2009. *An introduction to population genetics theory.* Caldwell: The Blackburn Press.

Darwin C. 1889. *The descent of man, and selection in relation to sex.* New York: D. Appleton and Company.

Donnelly MP, Paschou P, Grigorenko E, Gurwitz D, Barta C, Lu RB, Zhukova OV, Kim JJ, Siniscalco M, New M, et al. 2012. A global view of the *OCA2-HERC2* region and pigmentation. *Hum Genet.* 131:683–696.

Duffy DL, Box NF, Chen W, Palmer JS, Montgomery GW, James MR, Hayward NK, Martin NG, Sturm RA. 2004. Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. *Hum Mol Genet.* 13:447–461.

Duffy DL, Montgomery GW, Chen W, Zhao ZZ, Le L, James MR, Hayward NK, Martin NG, Sturm RA. 2007. A three-single-nucleotide polymorphism haplotype in intron 1 of *OCA2* explains most human eye-color variation. *Am J Hum Genet.* 80:241–252.

Edwards M, Bigham A, Tan J, Li S, Gozdzik A, Ross K, Jin L, Parra EJ. 2010. Association of the *OCA2* polymorphism His615Arg with melanin content in East Asian Populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genet.* 6:e1000867.

Fedorova SA, Reidla M, Metspalu E, Metspalu M, Rootsi S, Tambets K, Trofimova N, Zhadanov SI, Hooshiar Kashani B, Olivieri A, et al. 2013. Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): implications for the peopling of Northeast Eurasia. *BMC Evol Biol.* 13:127.

Graf J, Hodgson R, van Daal A. 2005. Single nucleotide polymorphisms in the *MATP* gene are associated with normal human pigmentation variation. *Hum Mutat.* 25:278–284.

Guenther CA, Tasic B, Luo L, Bedell MA, Kingsley DM. 2014. A molecular basis for classic blond hair color in Europeans. *Nat Genet.* 46:748–754.

Haller BC, Messer PW. 2017. SLiM 2: Flexible, interactive forward genetic simulations. *Mol Biol Evol.* 34:230–240.

Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, Hankinson SE, Hu FB, Duffy DL, Zhao ZZ, et al. 2008. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet.* 4:e1000074.

Harding RM, Healy E, Ray AJ, Ellis NS, Flanagan N, Todd C, Dixon C, Sajantila A, Jackson IJ, Birch-Machin MA, et al. 2000. Evidence for variable selective pressures at MC1R. *Am J Hum Genet.* 66:1351–1361.

He Y, Wang M, Huang X, Li R, Xu H, Xu S, Jin L. 2015. A probabilistic method for testing and estimating selection differences between populations. *Genome Res.* 25:1903–1909.

Hider JL, Gittelman RM, Shah T, Edwards M, Rosenbloom A, Akey JM, Parra EJ. 2013. Exploring signatures of positive selection in pigmentation candidate genes in populations of East Asian ancestry. *BMC Evol Biol.* 13:150.

Holder KK, Bull JJ. 2001. Profiles of adaptation in two similar viruses. *Genetics* 159:1393–1404.

Ito S, Wakamatsu K. 2008. Chemistry of mixed melanogenesis - pivotal roles of dopaquinone. *Photochem Photobiol*. 84:582–592.

Izagirre N, García I, Junquera C, de la Rúa C, Alonso S. 2006. A scan for signatures of positive selection in candidate loci for skin pigmentation in humans. *Mol Biol Evol.* 23:1697–1706.

Jablonski NG, Chaplin G. 2000. The evolution of human skin coloration. *J Hum Evol.* 39:57–106.

Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, Vermeulen M, Arp P, Jhamai MM, van ljcken WF, et al. 2008. Three genome-wide association studies and a linkage analysis identify *HERC2* as a human iris color gene. *Am J Hum Genet.* 82:411–423.

Kovacevic L, Tambets K, Ilumäe AM, Kushniarevich A, Yunusbayev B, Solnik A, Bego T, Primorac D, Skaro V, Keskovac A, et al. 2014. Standing at the gateway to Europe - the genetic structure of Western Balkan populations based on autosomal and haploid markers. *PLoS One* 9:e105090.

Lao O, de Gruijter JM, van Duijin K, Navarro A, Kayser M. 2007. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann Hum Genet.* 71:354–369.

Lasisi T, Shriver MD. 2018. Focus on African diversity confirms complexity of skin pigmentation genetics. *Genome Biol.* 19:13.

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, et al. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319:1100–1104.

Martin AR, Lin M, Granka JM, Myyrick JW, Liu X, Sockell A, Atkinson EG, Werely CJ, Möller M, Sandhu MS, et al. 2017. An unexpectedly complex architecture for skin pigmentation in Africans. *Cell* 171:1340–1353.

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528:499–503.

McEvoy B, Beleza S, Shriver MD. 2006. The genetic architecture of normal variation in human pigmentation: an evolutionary perspective and model. *Hum Mol Genet.* 15:R176–R181.

McEvoy BP, Powell JE, Goddard ME, Visscher PM. 2011. Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. *Genome Res.* 21:821–829.

Metspalu M, Romero IG, Yunusbayev B, Chaubey G, Mallick CB, Hudjashov G, Nelis M, Mägi R, Metspalu E, Remm M, et al. 2011. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am J Hum Genet.* 89:731–744.

Motokawa T, Kato T, Hashimoto Y, Katagiri T. 2007. Effect of Val92Met and Arg163Gln variants of the *MC1R* gene on freckles and solar lentigines in Japanese. *Pigment Cell Res.* 20:140–143.

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, Jin L. 1989. Variances of the average numbers of nucleotide substitutions within and between populations. *Mol Biol Evol.* 6:290–300.

Norton HL, Kittles RA, Parra E, McKeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, McEvoy B, Shriver MD. 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol.* 24:710–722.

Norton HL. 2005. *Human skin pigmentation variation: a phenotypic, genotypic and evolutionary perspective*. Ph.D. Dissertation, Pennsylvania State University. Available from https://etda.libraries.psu.edu/catalog/6658

Oppenheimer S. 2012. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. *Phil Trans R Soc B.* 367:770–784.

Orr HA. 2005. The genetic theory of adaptation: A brief history. *Nat Rev Genet.* 6:119–127.

Pagani L, Kivisild T, Tarekegn A, Ekong R, Plaster C, Gallego Romero I, Ayub Q, Mehdi SQ, Thomas MG, Luiselli D, et al. 2012. Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. *Am J Hum Genet.* 91:83–96.

Parra EJ. 2007. Human pigmentation variation: evolution, genetic basis, and implications for public health. *Am J Phys Anthropol.* Suppl 45:85–105.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhang Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192:1065–1093.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet.* 2:e190.

Pickrell J, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, et al. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* 19:826–837.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 38:904–909.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet.* 81:559–575.

Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford TW, Orlando L, Metspalu E, et al. 2014. Upper Paleolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505:87–91.

Rana BK, Hewett-Emmett D, Jin L, Chang BH, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde LB, Ramsay M, et al. 1999. High polymorphism at the human melanocortin 1 receptor locus. *Genetics* 151:1547–1557.

Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463:757–762.

Rebbeck TR, Kanetsky PA, Walker AH, Holmes R, Halpern AC, Schuchter LM, Elder DE, Guerry D. 2002. *P* gene as an inherited biomarker of human eye color. *Cancer Epidem Biomar.* 11:782–784.

Rees JL, Harding RM. 2012. Understanding the evolution of human pigmentation: Recent contributions from population genetics. *J Invest Dermatol.* 132:846–853.

Schaffner SF, Foo C, Gabriel S, Reich D, Daly MJ, Altshuler D. 2005. Calibrating a coalescent simulation of human genome sequence variation. *Genome Res.* 15:1576–1583.

Schiffel S, Durbin R. 2014. Inferring human population size and separation history from multiple genome sequences. *Nat Genet.* 46:919–925.

Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, Marjanovic D, Caragine T, Prinz M, Wurmbach E. 2011. Prediction of eye and skin color in diverse populations using seven SNPs. *Forensic Sci Int Gen*. 5:472–478.

Sturm RA, Duffy DL. 2012. Human pigmentation genes under environmental selection. *Genome Biol.* 13:248.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, et al. 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet.* 39:1443–1452.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, Steinberg S, Gudjonsson SA, Palsson A, Thorleifsson G, et al. 2008. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet.* 40:835–837.

Teo YY, Sim X, Ong RTH, Tan AKS, Chen J, Tantoso E, Small KS, Ku CS, Lee EJ, Seielstad M, et al. 2009. Singapore genome variation project: a haplotype map of three Southeast Asian populations. *Genome Res.* 19:2154–2162.

Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet.* 49:303–309.

The 1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. *Nature* 467:1061–1073.

The International HapMap 3 Consortium. 2010. Integrating common and rare genetic variation in diverse human populations. *Nature* 467:52–58.

Visser M, Kayser M, Palstra R. 2012. *HERC2* rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhance and the *OCA2* promoter. *Genome Res.* 22:446–455.

Visser M, Palstra R, Kayser M. 2014. Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearly *BNC2* pigmentation gene. *Hum Mol Genet.* 23:5750–5762.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. *PLoS Biol.* 5:e147.

Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, Unterländer M, Hollfelder N, Potekhina ID, Schier W, Thomas MG, et al. 2014. Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. *Proc Natl Acad Sci USA* 111:4832–4837.

Xing J, Wuren T, Simonson TS, Watkins WS, Witherspoon D, Wu W, Qin G, Huff CD, Jorde LB, Ge RL. 2013. Genomic analysis of natural selection and phenotypic variation in high-altitude Mongolians. *PLoS Genet.* 9:e1003634.

Xu X, Thörnwall M, Lundin L, Chhajlani V. 1996. Val92Met variant of the melanocyte stimulating hormone receptor gene. *Nat Genet.* 14:384.

Yang Z, Zhong H, Chen J, Zhang X, Zhang H, Luo X, Xu S, Chen H, Lu D, Han Y, et al. 2016. A genetic mechanism for convergent skin lightening during recent human evolution. *Mol Biol Evol.* 33:1177–1187.

Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, et al. 2012. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol.* 29:359–365.

Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, Akhmetova V, Balanovska E, Balanovsky O, Turdikulova S, et al. 2015. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genet.* 11:e1005068.

Tables

Table 1. Selection differences associated with human pigmentation between populations ($\times 10^{-3}$/generation).

	WAF	EAF	EUR	NAS	EAS
WAF	0	-6.276	-30.538	-22.588	-19.278
EAF	6.276	0	-30.795	-20.708	-16.504
EUR	30.538	30.795	0	13.11	18.576
NAS	22.588	20.708	-13.11	0	8.406
EAS	19.278	16.504	-18.576	-8.406	0

Populations are abbreviated as follows: WAF, West Africans; EAF, East Africans; EUR, Europeans; NAS, North Asians; EAS, East Asians.

Table 2. 95% confidence intervals of selection differences associated with human pigmentation between populations ($\times 10^{-3}$/generation).

	WAF	EAF	EUR	NAS	EAS
WAF	NA	[-9.831, -2.721]	[-35.550, -25.526]	[-28.639, -16.537]	[-25.162, -13.394]
EAF	[2.721, 9.831]	NA	[-37.116, -24.474]	[-28.274, -13.142]	[-23.942, -9.066]
EUR	[25.526, 35.550]	[24.474, 37.116]	NA	[6.070, 20.150]	[11.730, 25.422]
NAS	[16.537, 28.639]	[13.142, 28.274]	[-20.150, -6.070]	NA	[0.207, 16.605]
EAS	[13.394, 25.162]	[9.066, 23.942]	[-25.422, -11.730]	[-16.605, -0.207]	NA

Populations are abbreviated as follows: WAF, West Africans; EAF, East Africans; EUR, Europeans; NAS, North Asians; EAS, East Asians.

Table 3. The optimal solution ($\times 10^{-3}$/generation).

Epoch	Selection change	Standard deviation	95% CI
c_1	-10.97	1.20	[-13.32, -8.62]

c_2	-5.96	1.05	[-8.02, -3.89]
c_3	13.32	1.22	[10.92, 15.71]
c_4	3.52	1.46	[0.97, 6.07]
c_5	-4.88	1.78	[-7.92, -1.84]
c_6	-2.09	1.54	[-5.11, 0.94]
c_7	14.60	1.49	[11.67, 17.52]

Table 4. Mean selection differences between Eurasians and Africans in the first and second groups in Fig 4 ($\times 10^{-3}$/generation).

Group	Gene	SNP	Mean selection difference
1	*ASIP*	rs6119471	2.407
1	*ASIP*	rs6058071	1.259
1	*MC1R*	rs2228479	1.713
1	*MC1R*	rs885479	2.039
1	*DDB1*	rs11230664	2.229
1	*DDB1*	rs7120594	1.986
1	*DDB1*	rs1377457	2.469
2	*DDB1*	rs7948623	-1.985
2	*MFSD12*	rs56203814	-1.671
2	*MFSD12*	rs10424065	-2.044
2	*MFSD12*	rs6510760	-1.844
2	*MFSD12*	rs112332856	-2.482

Here, we averaged the selection differences of six population pairs. These pairs are: Europeans/West Africans, Europeans/East Africans, North Asians/West Africans, North Asians/East Africans, East Asians/West Africans, East Asians/East Africans.

Figure Captions

Fig. 1. Time-varied selective pressures on an evolutionary tree. We modeled the evolutionary history of five representative human populations as a binary tree. Here, s_i (i = 0, 1, ..., 8) denotes the selection coefficient of the i-th epoch. c_i (i = 1, 2, ..., 8) denotes the selection change of the i-th epoch, and can be obtained by estimating selection differences between paired populations. The numbers on the branches indicate different epochs. Divergence times between populations were based on previous studies (Schaffner et al. 2005;Oppenheimer 2012; Schiffel and Durbin 2014). We assumed one generation time is ~30 years.

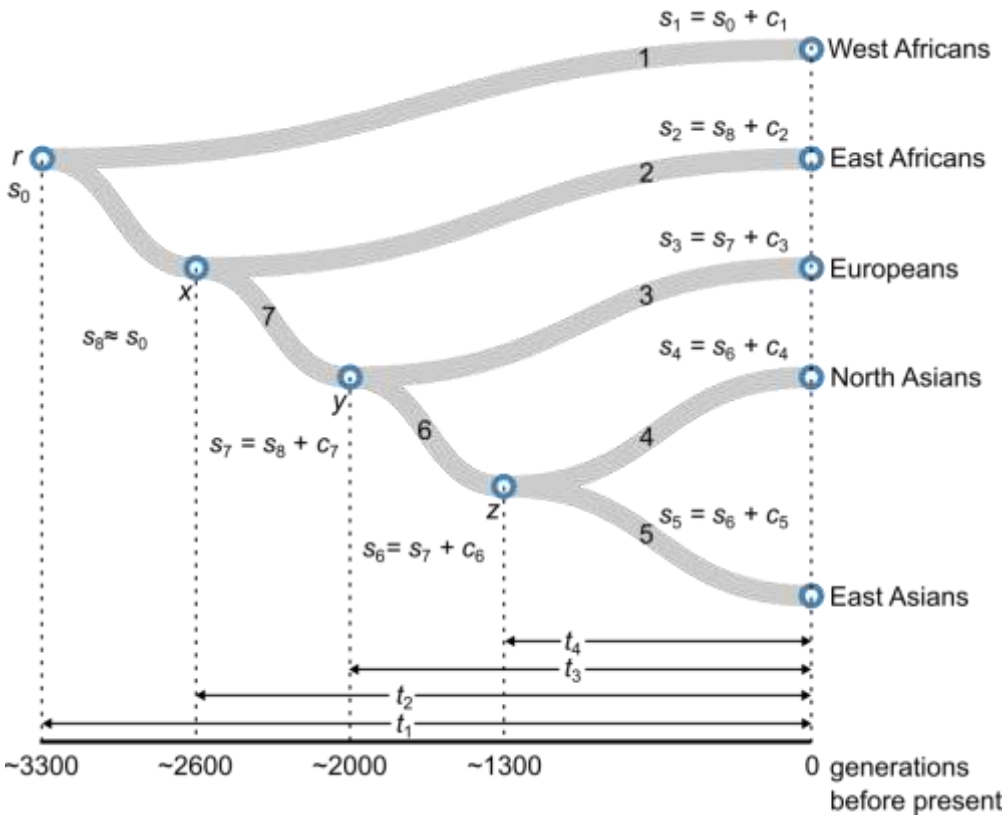
Fig. 2. The selection differences and skin reflectance differences on human pigmentation between populations. (A) We estimated the total selection differences on 25 selected SNPs in weak linkage disequilibrium. Red color (positive numbers) indicates that selective pressures of populations in rows are larger than those in columns; blue color (negative numbers) indicates that selective pressures of populations in rows are smaller than those in columns. (B) We calculated skin reflectance differences at 685 nm between populations. White color (positive numbers) indicates that skin reflectance of populations in rows are larger than those in columns; black color (negative numbers) indicates that skin reflectance of populations in rows are smaller than those in columns. We observed similar patterns of total selection differences on human pigmentation and skin reflectance differences, suggesting that our statistic (Eq. 2) may well represent human pigmentation diversity. Populations are abbreviated as follows: WAF, West Africans; EAF, East Africans; EUR, Europeans; NAS, North Asians; EAS, East Asians.

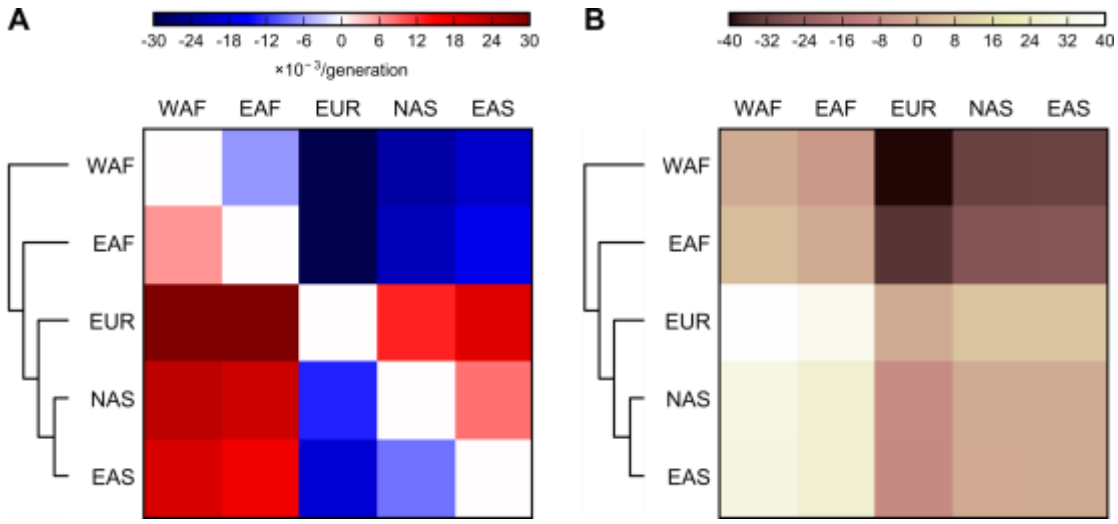
Fig. 3. Historical changes of selective pressures on human pigmentation. We determined historical changes of selective pressures on human pigmentation over epochs (Eq. 4). We chose the optimal solution to represent the possible historical selection on the world map. We observed not only a recent incremental change of selective pressure in modern Europeans, but also a significant historical increase of selective pressure on light pigmentation in proto-Eurasians during the out-of-Africa event. Here, s_0 is the selection coefficient of the ancestral population for all modern human populations. The numbers are the selection changes ($\times 10^{-3}$/generation) over epochs. The arrows indicate the direction of human

migration, and their color gradient indicates the trends of human skin color. Populations are abbreviated as follows: WAF, West Africans; EAF, East Africans; EUR, Europeans; NAS, North Asians; EAS, East Asians.

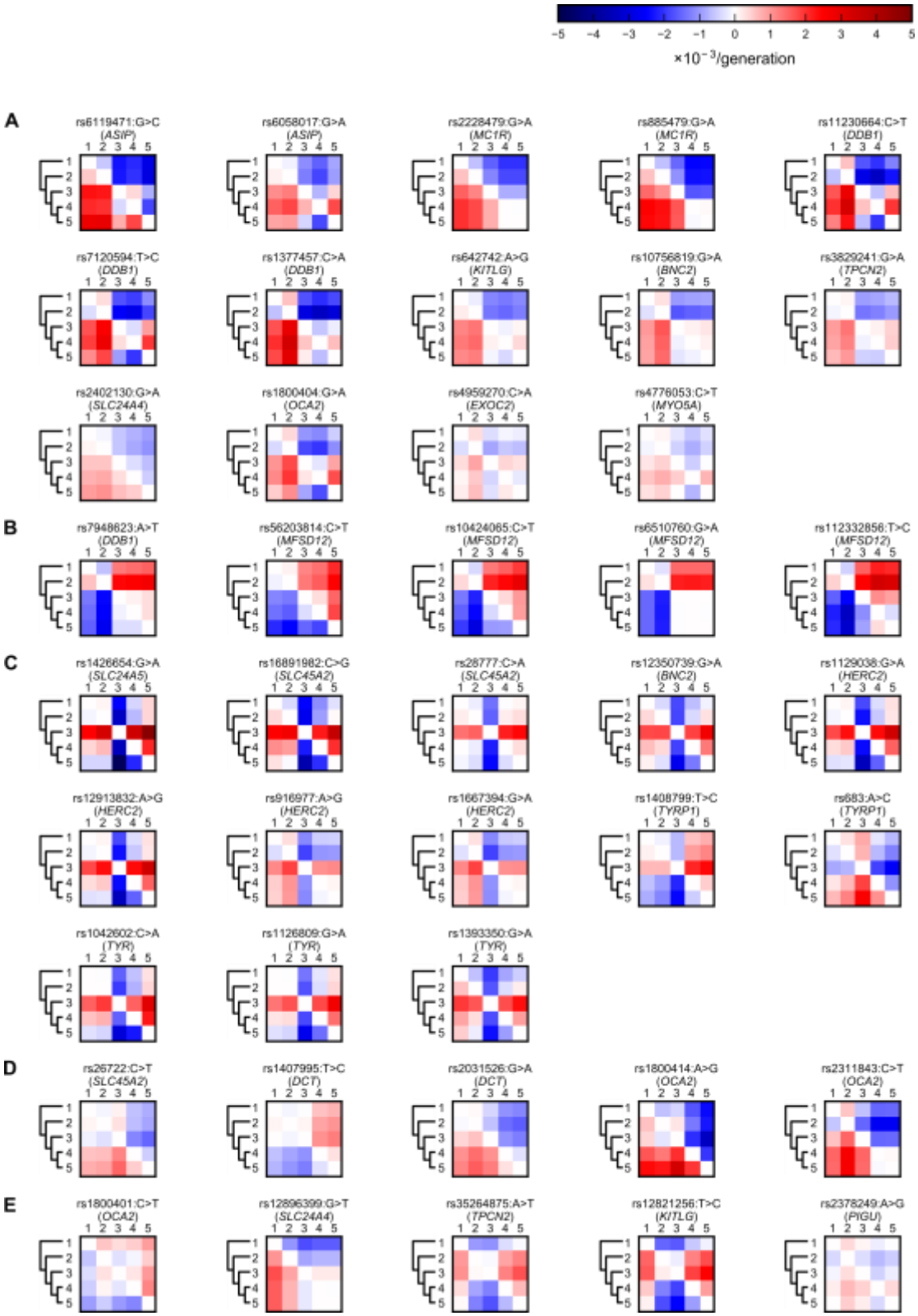
Fig. 4. Selection differences in individual loci between populations. We used Eq. 1 to quantify the selection differences of 42 SNPs associated with human pigmentation, and categorized them into five kinds of patterns: (A) Derived alleles affected by Eurasian-shared selection; (B) Derived alleles affected by African-specific selection; (C) Alleles affected by European-specific selection; (D) Alleles affected by Asian-specific selection; and (E) Others. Red color (positive numbers) indicates selective pressures of populations in rows are larger than those in columns; blue color (negative numbers) indicates selective pressures of populations in rows are smaller than those in columns. Populations are abbreviated as follows: 1, West Africans; 2, East Africans; 3, Europeans; 4, North Asians; 5, East Asians.

Figures


$s_1 = s_0 + c_1$
West Africans
$s_2 = s_8 + c_2$
East Africans
$s_3 = s_7 + c_3$
Europeans
$s_4 = s_6 + c_4$
North Asians
$s_5 = s_6 + c_5$
East Asians
r
s_0
x
y
z
$s_8 \approx s_0$
$s_7 = s_8 + c_7$
$s_6 = s_7 + c_6$
1
2
3
4
5
6
7
t_4
t_3
t_2
t_1
~3300
~2600
~2000
~1300
0 generations before present


A
-30 -24 -18 -12 -6 0 6 12 18 24 30
×10⁻³/generation
WAF EAF EUR NAS EAS
WAF
EAF
EUR
NAS
EAS
B
-40 -32 -24 -16 -8 0 8 16 24 32 40
WAF EAF EUR NAS EAS
WAF
EAF
EUR
NAS
EAS

NAS
c_4 = 3.52
EUR
c_3 = 13.32
c_6 = -2.09
EAS
c_5 = -4.88
c_2 = -5.96
EAF
c_7 = 14.60
WAF
c_1 = -10.97
s_0


−5 −4 −3 −2 −1 0 1 2 3 4 5
×10⁻³/generation
A
rs6119471:G>C (ASIP)
rs6058017:G>A (ASIP)
rs2228479:G>A (MC1R)
rs885479:G>A (MC1R)
rs11230664:C>T (DDB1)
rs7120594:T>C (DDB1)
rs1377457:C>A (DDB1)
rs642742:A>G (KITLG)
rs10756819:G>A (BNC2)
rs3829241:G>A (TPCN2)
rs2402130:G>A (SLC24A4)
rs1800404:G>A (OCA2)
rs4959270:C>A (EXOC2)
rs4776053:C>T (MYO5A)
B
rs7948623:A>T (DDB1)
rs56203814:C>T (MFSD12)
rs10424065:C>T (MFSD12)
rs6510760:G>A (MFSD12)
rs112332856:T>C (MFSD12)
C
rs1426654:G>A (SLC24A5)
rs16891982:C>G (SLC45A2)
rs28777:C>A (SLC45A2)
rs12350739:G>A (BNC2)
rs1129038:G>A (HERC2)
rs12913832:A>G (HERC2)
rs916977:A>G (HERC2)
rs1667394:G>A (HERC2)
rs1408799:T>C (TYRP1)
rs683:A>C (TYRP1)
rs1042602:C>A (TYR)
rs1126809:G>A (TYR)
rs1393350:G>A (TYR)
D
rs26722:C>T (SLC45A2)
rs1407995:T>C (DCT)
rs2031526:G>A (DCT)
rs1800414:A>G (OCA2)
rs2311843:C>T (OCA2)
E
rs1800401:C>T (OCA2)
rs12896399:G>T (SLC24A4)
rs35264875:A>T (TPCN2)
rs12821256:T>C (KITLG)
rs2378249:A>G (PIGU)
1 2 3 4 5

670