Darwinian positive selection on the pleiotropic effects of *KITLG* explain skin pigmentation and winter temperature adaptation in Eurasians

Zhaohui Yang[1,2,4,*], Hong Shi[1,4,*], Pengcheng Ma[2,*], Shilei Zhao[3], Qinghong Kong[5], Tianhao Bian[1,4], Chao Gong[1,4], Qi Zhao[1,4], Yuan Liu[5], Xuebin Qi[2], Xiaoming Zhang[2], Yinglun Han[6], Jiewei Liu[2], Qingwei Li[6], Hua Chen[3,†], Bing Su[2,7,†]

[1] Yunnan Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, 650500, China;

[2] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, China;

[3] CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, 100101, China;

[4] Yunnan Provincial Academy of Science and Technology, Kunming, 650051, China;

[5] Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming, 650500, China;

[6] College of Life Science, Liaoning Normal University, Dalian, 116029, China

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

*These authors contributed equally to this work.

†Correspondence to B.S. (sub@mail.kiz.ac.cn) or H.C. (chenh@big.ac.cn)

Keywords: *KITLG*; Darwinian positive selection; skin pigmentation; winter temperature; pleiotropic effect; genetic adaptation; Eurasian

Abstract

Human skin color diversity is considered an adaptation to environmental conditions such as UV radiation. Investigations into the genetic bases of such adaptation have identified a group of pigmentation genes contributing to skin color diversity in African and non-African populations. Here we present a population analysis of the pigmentation gene *KITLG* with previously reported signal of Darwinian positive selection in both European and East Asian populations. We demonstrated that there had been recurrent selective events in the upstream and the downstream regions of *KITLG* in Eurasian populations. More importantly, besides the expected selection on the *KITLG* variants favoring light skin in coping with the weak UV radiation at high latitude, we observed a *KITLG* variant showing adaptation to winter temperature. In particular, compared to UV radiation, winter temperature showed a much stronger correlation with the prevalence of the presumably adaptive *KITLG* allele in Asian populations. This observation was further supported by the *in vitro* functional test at low temperature. Consequently, the pleiotropic effects of *KITLG*, *i.e.* pigmentation and thermogenesis were both targeted by natural selection that acted on different *KITLG* sequence variants, contributing to the adaptation of Eurasians to both UV radiation and winter temperature at high latitude areas.

Introduction

Modern humans originated in Africa around 200,000-300,000 years ago, and they spread to the other parts of the world less than 75,000 years ago (Hublin, et al. 2017; Pagani, et al. 2016). During the "Out of Africa" migration, modern humans moved from a near-equator tropical environment into high-latitude sub-tropical and frigid regions with entirely different environmental conditions, such as reduced UV radiation and much colder weather in winter.

Skin color is thought to be one of the key human traits responding to the environmental changes (Jablonski and Chaplin 2010, 2000). Recent studies on diverse African populations indicated a complex architecture of skin pigmentation due to different local evolutionary pressures, with derived alleles in several genes associated with both light and dark skin (*e.g. SLC24A5*, *MFSD12*, *DDB1*, *TMEM138*, *SMARCA2/VLDLR* and *SNX13*) (Crawford, et al. 2017; Martin, et al. 2017). By contrast, Europeans and East Asians live in high-latitude environments with weak light, and selection on vitamin D synthesis, a key nutrient for bone development would favor a lighter skin in these populations (Loomis 1967). Previous investigations into skin lightening in Eurasians have found a set of pigmentation genes showing molecular signatures of Darwinian positive selection. For example, *OCA2* serves as the key gene for skin lightening in East Asians (Donnelly, et al. 2012; Eaton, et al. 2015; Edwards, et al. 2010; Gittelman, et al. 2016; Murray, et al. 2015; Rawofi, et al. 2017; Yang, et al. 2016; Yuasa, et al. 2007). In Europeans, there are other responsible genes including *SLC24A5*, *SLC45A2*, *KITLG*, *TYR, OCA2*, *MC1R*, *ASIP* and *IRF4* (Basu Mallick, et al. 2013; Beleza, et al. 2013a; Graf, et al. 2005; Han, et al. 2008; Izagirre, et al. 2006; Lamason, et al. 2005; Lao, et al. 2007; McEvoy, et al. 2006; Miller, et al. 2007; Norton, et al. 2007; Sabeti, et al. 2007; Soejima and Koda 2007; Sulem, et al. 2007; Yuasa, et al. 2006). Interestingly, most of the identified pigmentation genes responsible for skin lightening are not shared between East Asians and Europeans, suggesting convergent

evolution of skin pigmentation in non-African populations (Edwards, et al. 2010; Norton, et al. 2007; Yang, et al. 2016).

In addition to UV radiation, temperature is also an important environmental factor that could affect human skin color. During the Korean War in 1950s, it was observed that soldiers of African-American descent were more susceptible to cold injury than those of European-American descent (Post, et al. 1975). Also, it was reported that melanocytes are more susceptible to cold injury than the hair follicle. Cooling or freezing the skin of black rats could kill the melanocytes, but not the hair follicles which subsequently produced hair of normal shape and texture, albeit lacking pigment (Post, et al. 1975; Taylor 1949). This finding was further supported by another hair neogenesis study (Mikhail 1963). Consequently, it was proposed that cold climate might be a contributing factor during the emergence of "white" skin (Post, et al. 1975). However, during the past half a century, this “cold hypothesis” has not been rigorously tested.

Recently, the pigmentation gene *KITLG* (also named the stem cell factor, *SCF*) was reported involved in mitochondrial function and energy expenditure in brown adipose tissue under cold condition (Huang, et al. 2014; Nishio, et al. 2012). Overnight exposure of mice to cold temperature increases the level and ratio of soluble *Kitl* protein (encoded by *KITLG*) in brown adipose tissue and increases thermogenesis and reduces weight gain (Huang, et al. 2014). *KITLG* also serves as an essential factor of brown adipocytes derived from human pluripotent stem cells (Nishio, et al. 2012). Additionally, it was shown that working in the cold can retain brown adipose tissue in "strategic" places in human adults (Huttunen, et al. 1981).

Besides playing a key role in mitochondrial function and energy expenditure, the *KITLG* gene has been well known as a pigmentation gene. It is associated with synthesis of skin and hair pigment, skin sensitivity to sun and freckles (Guenther, et al. 2014; Han, et al. 2008; Mengel-From, et al. 2009; Miller, et al. 2007; Sulem, et al. 2007), familial progressive hyperpigmentation and hypopigmentation (FPHH) (Amyere, et al. 2011; Cuell, et al. 2015; Wang, et al. 2009) as well as with tumor pathogenesis and

Warrdensburg Syndrome Type 2 (with pigmentary abnormalities) (Rapley, et al. 2009; Zazo Seco, et al. 2015). *KITLG* encodes the ligand for the receptor tyrosine kinase KIT that rivet on the melanocyte. *KITLG* and its protein–tyrosine kinase receptor KIT are implicated in the regulation of diverse biological processes, particularly in melanogenesis, including melanoblast and melanocyte migration, proliferation and differentiation during embryogenesis, and also for maintaining postnatal cutaneous melanogenesis (Bedell, et al. 1995; Grichnik, et al. 1998; Kasamatsu, et al. 2008; Wehrle-Haller 2003). For example, a regulatory sequence change in *KITLG* contributed to natural variation in vertebrate pigmentation, and that similar mechanism underlies pigmentary change in both fish and humans (Miller, et al. 2007). Additionally, a strong signal of positive selection on the regulatory regions surrounding the gene was also reported in Eurasians, and two variants (rs642742 and rs12821256) located 300-350 kb upstream of the *KITLG* transcription start site had a significant effect on skin color of African-Americans and hair color of Europeans (Guenther, et al. 2014; Miller, et al. 2007; Williamson, et al. 2007). Taken together, these data suggest that *KITLG* is pleiotropic and may contribute to human adaptation to both UV radiation and temperature change.

In the current study, we conducted a population analysis of *KITLG*, and we showed that besides UV radiation, winter temperature may have an even stronger correlation with the prevalence of the presumably adaptive allele of *KITLG* in Asian populations. Natural selection had acted on different *KITLG* sequence variants, leading to the adaptation of Eurasians to both the weak UV radiation and the cold winter temperature at high latitude areas.

Results

Detection of positive selection and the molecular history of *KITLG* in Eurasians

We obtained the selection test results of *KITLG* (549 kb, including upstream and downstream of *KITLG*; chr12_88,803,520-89,352,520, hg19) from the 1000 Genomes

Selection Browser (http://hsb.upf.edu). The haplotype-based EHH test detects extended homozygosity surrounding the variant under positive selection, which is sensitive to relatively recent selection (<100,000 years) (Sabeti, et al. 2006). In the EHH test, we saw strong selection signals in both Europeans and East Asians, especially in the upstream region (126-375 kb upstream of *KITLG*) (fig. 1*A*), and this is in agreement with the previous observation (Miller, et al. 2007). In the downstream region (57.5 kb downstream of *KITLG*), we observed a signal of selection in East Asians. In contrast, we did not observe obvious selection signal in Africans in the EHH test. The selection signals of the *KITLG* upstream and downstream regions in Eurasians were further confirmed by the XP-CLR test (a population-differentiation method, and Africans was used as outgroup), in which the selection on the upstream region is strong in Europeans and relatively weak in East Asians. Also, the downstream region selection in East Asians was confirmed (fig. 1*B*).

The selection signals in Eurasian populations were also validated by the results of other tests, including Fay and Wu's *H* test, Fu and Li's test, Tajima's D test and XP-EHH test (supplementary fig. 1). Interestingly, we detected selection signals in Africans from three allele-frequency-based tests (Fay and Wu's *H* test, Fu and Li's test and Tajima's D test). However, Kumar and Liu have shown that the signal seen in the allele-frequency-based tests was not due to positive selection, and the EHH profile of Africans suggested a trend of neutral evolution (Kumar and Liu 2014). Taken together, these tests suggest selection signals in both Europeans and East Asians in the upstream region of *KITLG*, and an additional signal of selection in East Asians in the downstream of *KITLG*.

To look into the molecular history of *KITLG* in human populations, we constructed a maximum-likelihood phylogenetic tree using the 549 kb *KITLG* sequence data (a total of 540 haplotypes were phased) of three world populations from the 1,000 Genomes Project, including Africans (YRI), Europeans (CEU) and East Asians (CHB). The homologous sequences of chimpanzee and archaic humans (Neanderthal and

Denisovan) were used as outgroup. As shown in fig. 2*A*, there are three major clades in the population tree. The allelic status (red for major allele and blue for minor allele) of 727 variants with minor allele frequency larger than 0.05 is illustrated in fig. 2*B*. As expected, Clade-1 is the basal clade and is mostly composed of Africans, and this clade displays a major-allele block in the *KITLG* gene region (fig. 2*B*). Clade-2 is similar with Clade-1 in view of major-allele blocks and it includes more Eurasians. Clade-3 is a Eurasian clade and a large major-allele block was observed in the *KITLG* upstream and downstream region, consistent with the detected strong signal of selection of this region in both Europeans and East Asians (fig. 1).

The *KITLG* variants under selection are associated with light skin in Eurasians

Given the strong signal of selection on *KITLG* in Eurasian populations (fig. 1), presumably, skin pigmentation is expected to be associated with the *KITLG* variants under selection. Indeed, a previous study observed a significant association of rs642742 (325.5 kb upstream of *KITLG*) with skin pigmentation in African-Americans, a population with a mixed ancestry from West Africa and Europe (Miller, et al. 2007). To further test the association, we measured skin color of 301 Han Chinese individuals, including 119 males (average age$\pm$*S.D.*: 19.05$\pm$0.87) and 182 females (average age $\pm$*S.D.*: 18.73$\pm$0.96). Both constitutive skin color from skin areas not exposed to sunlight (underarm and buttock) and facultative skin color from skin area exposed to sunlight (back of hand) were measured. We used the Haploview software (D' algorithm) (Barrett, et al. 2005) to construct the linkage pattern of the 549 kb region in Han Chinese (CHB) from the 1000 Genomes Project (supplementary fig. 2). Based on the LD (linkage disequilibrium) map, we first picked five tag SNPs that show high allelic divergences between Africans and Eurasians (measured by F_{ST}) (Weir and Hill 2002), including three upstream variants (rs642742, rs1703078 and rs4073022), one downstream variant (rs6538148) and one within-gene variant (rs4590952) (Table 1, supplementary fig. 2). These variants represent the major LD blocks with signals of

selection. We also chose another upstream variant (rs428316) located in the predicted enhancer region (with peaks for H3K4Me1, data from ENCODE, Table 1) though it is in complete linkage with rs642742 in the tested Han Chinese (table 1). The LD map of the 6 variants is shown in supplementary fig.S3A. All tested variants are in Hardy–Weinberg equilibrium except for rs1703078 (P=0.0017, chi-square test). Further manual check did not see genotyping errors for rs1703078.

We adopted the additive genetic model in the association analysis. The results indicated that three of the four upstream variants (rs4073022, rs428316 and rs642742) showed significant association with skin darkness (P <0.05, after permutation test correction) for the constitutive skin areas (underarm and buttock), but not the skin area exposed to sunlight (back of hand) (table 1). When only northern Han Chinese were included (273 individuals with 106 males and 167 females), we saw the same association pattern, ruling out the potential effect of population substructure (supplementary table 2). The observed association of rs642742 in Han Chinese is consistent with the reported association in African-Americans (Miller, et al. 2007). As expected, the derived alleles (the major alleles) of the three upstream variants were associated with light skin in the tested Han Chinese, suggesting that natural selection favors skin lightening. However, we did not find significant association for rs6538148 (a downstream variant) and rs4590952 (a within-gene variant) in any of the three skin areas (table 1). Previously, there were reported associations of hair color in Europeans for four *KITLG* variants (rs12821256, rs1492354, rs1907702 and rs10777129) located in the upstream and the within-gene regions (Guenther, et al. 2014; Mengel-From, et al. 2009), and these variants are not in linkage with the three upstream variants showing association in Han Chinese (supplementary fig. 3B). Taken together, these results suggest that the upstream *KITLG* variants under selection affect pigmentation of Eurasian populations, and they may function by regulating the expression of *KITLG*. Hence, the following functional tests were focused on these three upstream variants.

In vitro functional tests of the *KITLG* upstream variants

Previous studies suggested that the *KITLG* upstream region containing the tested variants may function as enhancers (Guenther, et al. 2014; Myles, et al. 2007). To test whether the three upstream variants (rs4073022, rs428316 and rs642742) can affect expression regulation of *KITLG*, we performed dual-luciferase reporter experiments (enhancer assays) using two different cell lines, A375 (human melanoma cells) and HEK293 cells. Since these three variants are 9.2-253 kb apart from each other, we tested them individually. Under the normal culture condition at 37℃, the major alleles (presumably the adaptive alleles) of rs428316 (T allele) and rs642742 (C allele) displayed significantly lower enhancer activities when compared with the minor alleles, and the results were consistent between the two cell lines ($P< 0.05$, *t*-test) (fig. 3). This pattern would predict less *KITLG* production for the major allele carriers, and consequently lighter skin, which fits the prediction of direction of selection and the association data (table 1). Surprisingly, the other variant (rs4073022) showed an opposite pattern though statistically not significant ($P = 0.07$ for A375 cells and $P = 0.557$ for 293 cells, *t*-test) (fig. 3). Hence, the *in vitro* data indicated that the *KITLG* upstream variants may regulate the expression of *KITLG*, and the adaptive alleles of rs428316 and rs642742 tend to weaken the enhancer activity and may contribute to skin lightening in Eurasians.

Both UV radiation and winter temperature serve as the selective forces for skin pigmentation

In order to find the selective forces acting on *KITLG*, we carried out correlation analysis between environmental factors and the *KITLG* variants. The environmental data (22-years UV radiation level and temperature) were collected from *NASA* Langley Research Center. The insolation on horizontal surface data covered 22 years (Jul 1983 - Jun 2005), and the average annual total of UV radiation ($KWh/m^2/year$) for each geographic site was calculated via the global insolation on horizontal surface (Zhu 2005), and the monthly average temperature (℃) (air temperature at 10 m) data covered

22 years (Jan 1983 - Dec 2004) was used in the correlation analysis. The average temperatures in July and January were taken as the highest and lowest temperature, respectively. The 19 Asian populations from China, Cambodia and Thailand were included (supplementary tables 1, 3). We also collected environmental and allele frequency data of 14 African and European populations from the 1000 Genomes Project (supplementary table 3). The major allele frequencies of the six variants were used in the correlation analysis. The result of rs642742 (data not shown) was the same as that of rs428316 since these two variants are in nearly complete linkage (supplementary fig. 3).

Firstly, we observed a clear correlation of the major allele frequencies of rs4073022 and rs428316 with latitude when all 33 world populations were included (correlation coefficient $r^2 > 0.5$, $P<0.0001$, blue dot, F test; fig. 4*A* and *B*; supplementary table 3). Since both UV radiation and temperature would be affected by latitude, we next performed separate correlation analyses. As predicted, the UV intensity is negatively correlated with the major allele frequencies for all three variants ($r^2> 0.4$, $P<0.0001$, blue dot, F test; fig. 4*E* and 4*F*; supplementary table 3). In other words, the lower the UV intensity is, the higher is the frequency of the adaptive allele that leads to lighter skin. Interestingly, the winter temperature also showed similar levels of correlation ($r^2> 0.4$, $P<0.0001$, blue dot, F test; fig. 4*E* and 4*F*; supplementary table 3), but not the summer temperature ($P> 0.05$, F test; supplementary fig. 4).

As UV intensity and winter temperature are correlated ($r^2 = 0.4033$, $P<0.0001$, F test; supplementary fig. 5A), we therefore performed mediation analysis so that the effects of UV and temperature could be evaluated independently (Preacher and Hayes 2004). When winter temperature was taken as the mediator variable, the mediation effect was significant and explained 38.8% and 39.6% of the observed correlation (for rs4073022 and rs428316, respectively) between UV and allele frequency (supplementary table 4). Similarly, when UV was taken as the mediator variable, we saw similar percentages of mediation effect (supplementary table 4). Consequently,

both UV intensity and winter temperature contribute to the cline of the *KITLG* variants in world populations. Of note, summer temperature is not correlated with UV intensity among the 33 world populations (r^2 = 0.0079, *P* = 0.6237, F test; supplementary fig. 5B). Similar correlation patterns were seen for the other three variants (rs6538148, rs4590952 and rs1703078) (supplementary fig. 6; supplementary table 3).

We also conducted a local correlation analysis using the 19 Asian populations, among which there is no correlation between UV and temperature (winter or summer) due to the relatively narrow latitude range (supplementary fig. 5C, D). For rs428316, we saw similar levels of correlation for UV (r^2 = 0.441, *P*=0.002, red dot) and winter temperature (r^2 = 0.463, *P*=0.002, red dot) (fig. 4*D* and *F*; supplementary table 3). So were the correlations for rs1703078 and rs4590952 (supplementary fig. 6). By contrast, for rs4073022, winter temperature (r^2 = 0.446, *P*=0.002) showed a much stronger correlation than UV intensity (r^2 = 0.218, *P*=0.044) (the red dots in fig. 4C, 4E), which strengthened the pattern observed in the 33 world populations. No correlations were detected for rs6538148 (supplementary fig. 6). Hence, winter temperature may act as an independent selective force, and rs4073022 is likely the target of selection by winter temperature. Of note, the mediation analysis using the 19 Asian populations was not significant (supplementary table 4), likely due to the limited sample size.

In vitro functional test at low temperature

To directly test if the *KITLG* variants would react to winter temparature, based on the published protocol, we performed dual-luciferase reporter experiments (enhancer assays) at a relatively low temperature (31℃) (Kim, et al. 2003). Again, both A375 and HEK293 cells were used. The result is shown in fig. 5. Two variants (rs428316 and rs642742) showed the same pattern as seen at 37℃. In contrast, the opposite pattern of rs4073022 seen at 37℃ became consolidated at 31℃. For both cell lines, the major (derived) allele of rs4073022 had significantly increased luciferase signals (presumably stronger enhancer activities) compared with the minor (ancestral) allele (fig. 5A, 5*D*),

suggesting that rs4073022 responds to low temperature and the presumably adaptive (major) allele produces more *KITLG* at low temperature. According to the previous report (Huang, et al. 2014), the upregulation of *KITLG* at low temperature may help promoting the production of brown fat for heat generation.

Age estimation of selection onset and haplotype network analysis

Given the potential functional effects of the *KITLG* upstream variants, we estimated the onset time of selection onset of the two upstream variants, rs428316 (in nearly complete linkage with rs642742) and rs4073022. We assumed a model of soft-sweep in the time inference (refer to the method section for details). Assuming a generation time of 29 years, the selection onset time for rs4073022 are 30,176 years ago in East Asians (95% CI: 25,865-36,211 years ago), and 58,243 years ago in Europeans (95% CI: 51,253-67,439 years ago). The inferred onset time for rs428316 are relatively younger, which are 21,320 years ago in East Asians (95% CI: 18,379-25,381 years ago) and 34,939 years ago in Europeans (95% CI: 30,678-40,574 years ago), indicating possibly independent selection acting on *KITLG*.

We next performed haplotype network analysis by selecting two upstream segments containing rs4073022 and rs428316, respectively (10 kb for each segment, chr12_89,030,872-89,040,872 and chr12_89,285,545-89,295,545). The results are consistent with the hypothesis of independent selective events since the network topologies of these two segments are different though an Eurasian-enriched dominant haplotype (the signal of positive selection) can be seen in both networks (supplementary fig. 7). Altogether, the inferred selection onset time and the pattern of local haplotype networks suggest that there might have been recurrent selections on different sequence variants of *KITLG* in Eurasians during the last glacial period.

Discussion

Recent studies on African and non-African populations indicated a complex architecture of skin pigmentation due to different local evolutionary pressures (Crawford, et al. 2017; Martin, et al. 2017). In the current study, we present genetic evidence of Darwinian positive selection in both Europeans and East Asians in the upstream region of *KITLG*. The selection on *KITLG* in Eurasians likely occurred after the "Out of Africa" migration. Given the estimated selection ages (30-58 kya for rs4073022 and 21-35 kya for rs428316) and the shared high-frequency adaptive alleles of *KITLG* between Europeans and East Asians, *KITLG* might represent the earliest episode of Darwinian positive selection on pigmentation genes in non-African populations, contrasting the much younger selective events of the other pigmentation genes, *e.g. OCA2*, 12–18 kya in East Asians, *SLC45A2,* 11-19 kya and *SLC24A5,* 22-28 kya in Europeans (Basu Mallick, et al. 2013; Beleza, et al. 2013b; Yang, et al. 2016;). Consequently, besides the well-known relatively recent and probably independent selections on different pigmentation genes in Europeans and East Asians, the data of *KITLG* suggests an old selection during the early period when modern humans started to spread to high latitude regions.

Notably, the observed selection on the two upstream variants (rs4073022 and rs428316) seems to represent independent events in view of time and phenotypic outcomes. These two variants are 253 kb apart in the upstream of *KITLG*, and they are not in linkage disequilibrium (*D*' = 0.4; supplementary fig. 2*A*). The selection on rs4073022 occurred earlier according to the estimated onset time of 30 kya -58 kya, and the mainly selected phenotype is likely body heat generation in order to cope with winter temperature, supported by the much higher correlation coefficient of winter temperature compared with UV radiation and the *in vitro* enhancer activity test at low temperature in cultured cells. The selection on rs428316 was a later event (21 kya -35 kya) and the selected phenotype was skin color when populations were facing high latitude environments with relatively less UV radiation necessary for vitamin D

synthesis. It should be noted that given the difficulty of obtaining the paleoclimatology data of the sample locations, we used contemporary climate data (22 years) in the correlation analysis which may not accurately reflect the environmental conditions during the relevant selective events. On the other hand, since population expansion of Han Chinese was a recent event (< 5 kya) (Wen, et al. 2004), the observed strong correlation in Asian populations (12 of the 19 samples are Han ancestry) implies that the latitude-dependent selection on the *KITLG* variants might have lasted till recent time in eastern Asia. This speculation is reflected by the still unfixed adaptive alleles (mostly around 0.70-0.85) in Han Chinese populations. By contrast, in Europeans, two of the adaptive alleles (rs1703078 and rs4073022) are nearly fixed (0.93-0.95) (supplementary table 3). Alternatively, we cannot rule out the possibility that there was only one selective event for the upstream *KITLG* variants in the ancestral Eurasian population because reduced UV radiation and colder weather are generally correlated with each other when people moved to high latitude areas. Also, the estimated time intervals of the two variants are slightly overlapped (30-58 kya vs. 21-35 kya).

Previous study showed that the *KITLG* gene contributed to pigmentary change in fish and humans. *KITLG* is a secreted signal that has a cell nonautonomous effect on melanocyte development (Wehrle-Haller, 2003; Miller, et al. 2007). Many *KITLG* variants were shown to be associated with pigmentation changes in Europeans, including skin color, hair color as well as skin sensitivity to sun exposure and skin freckles (Guenther, et al. 2014; Han, et al. 2008; Mengel-From, et al. 2009; Miller, et al. 2007; Sulem, et al. 2007). Our genetic association analysis in Han Chinese suggests that the selected adaptive alleles cause skin lightening in East Asians. There are under-selection variants located in the upstream of *KITLG*, a region containing enhancer elements that were proven to regulate the expression of *KITLG* and change the pigmentation of gill and ventrums of sticklebacks and human skin pigmentation (Miller, et al. 2007). Our *in vitro* functional tests revealed that the adaptive *KITLG* alleles caused a reduced enhancer activity and presumably reduced expression of *KITLG* that would

influence the melanocyte development and eventually skin lightening. We also detected a signal of selection in the downstream region of *KITLG* in East Asians (fig. 1) though we did not observe an association of the downstream variant (rs6538148) with skin pigmentation. It was reported that rs6538148 was associated with mean corpuscular hemoglobin concentration and red blood cell distribution in Europeans (Astle, et al. 2016). Further analysis is needed to test the functional effect of the *KITLG* downstream variants.

In addition to UV radiation, temperature could serve as another potential selective force as demonstrated in our previous study of the p53 Arg72 polymorphism across eastern Asia (Shi, et al. 2009). Besides pigmentation, *KITLG* is also involved in mitochondrial function and energy expenditure in brown adipose tissue under cold condition (Huang, et al. 2014; Nishio, et al. 2012). We demonstrated that winter temperature showed a much stronger correlation than UV for rs4073022 (fig. 4C, 4E). This variant rs4073022 showed an opposite pattern compared with the other two upstream variants in the *in vitro* functional test at 37℃, which seemed to contradict the predicted selection on light skin. The data at low temperature (31℃) consolidated the pattern seen at 37℃ and the derived allele at rs4073022 showed a significantly increased enhancer activity, an indication of a response to cold temperature. Hence, unlike the other upstream variants, rs4073022 likely represents an independent selection not on skin color, but on winter temperature. Phenotypically, the opposite direction of selection on enhancer activity of rs4073022 does not contradict the selection on skin lightening as the effect only becomes significant at low temperature. More functional data is needed to test the inferred selection pattern and the underlying molecular mechanism.

It should be noted that although we presented a case of the involvement of pigmentation gene in winter temperature adaptation, this does not provide direct genetic evidence to the "cold hypothesis" because the *KITLG* variants under selection work in different ways and we did not find a variant causing both skin lightening and cold

adaptation. Rather, it is the pleotropic effect of *KITLG*, *i.e.* pigmentation and thermogenesis links the two phenotypes. This may represent the first human case that two different functions of a gene were both picked up by natural selection in order to adapt to the different environmental stresses (UV and coldness).

Materials and Methods

Test of selection and evolutionary analysis of KITLG

We first extracted data from the 1000 Genomes Selection Browser (http://hsb.upf.edu), a database containing the results of selection tests across the human genome. We employed multiple selection tests to detect signatures of selection on *KITLG*. The extended haplotype homozogysity (EHH) test was used to detect the genetic imprint of recent positive selection by analyzing long-range haplotypes (Sabeti, et al. 2002), each spot in the graph indicates the average value of a 3-kb window. The XP-CLR test detects selective sweeps, which was used to identify highly differentiated genomic regions as targets of selective sweeps (Chen, et al. 2010), each spot in the graph indicates the average value of a 2-kb window. We calculated the *P* values of the tests based on the genome Rank Score (http://hsb.upf.edu). In addition, we also downloaded the results of multiple selection tests (Fay& Wu's *H* test, Fu and Li's test, Tajima's *D* test and XP-EHH test).

We extracted population data over a 549-kb genomic region (chr12_ 88,803,520-89,352,520, hg19), covering both the upstream and the downstream noncoding regions of *KITLG* from the 1000 Genomes Project (http://www.1000genomes.org), and including East Asian ancestry (CHB, 97 samples), western European ancestry (CEU, 85 samples), and Yoruba in Ibadan (YRI, 88 samples). The homologous sequences of chimpanzee (panTro3), Neanderthal and Denisovan were also included. A maximum-likelihood (ML) tree was constructed with RaxML and sequence variants with allele frequency (MAF) ≥0.05 across the haploblock region were included (Capellini, et al. 2017; Stamatakis 2014). The RAxML code is: raxmlHPC -PTHREADS -s region.fasta

-n block1 -m GTRGAMMAI -o Chimp -p 999 -T 8. The Haploview software (Barrett, et al. 2005) was used to analyze the linkage disequilibrium pattern of the *KITLG* variants in different populations. We used fastPHASE to reconstruct haplotypes (Scheet and Stephens 2006), and the code is: fastPHASE.exe -T10 -ogene gene.inp. The median-joining algorithm was used in haplotype network analysis (Bandelt, et al. 1999).

The hidden Markov model (HMM) approach by Chen et al (2015) was adopted to infer the selection onset time. The haplotype data of the 97 CHB and 85 CEU individuals from the 1000 Genomes Project were used in the analysis as input data, and the physical and genetic positions of the SNPs were assumed known. The HMM method explored the haplotype structure around the putatively selected SNP locus, and the extent of ancestral haplotypes carrying the selected SNP locus were recorded. A likelihood function was built to infer the selection intensity and duration (or the selection onset time) of the selective process based on the extents of ancestral haplotypes (please see the original paper for more details of the approach). Bootstrapping over haplotypes was performed to achieve the confidence intervals of the parameters. A generation time of 29 years was used to convert generations to years. The executable software and the command lines for estimating selection onset time can be found from the webpage (http://chenlab.big.ac.cn/software/software.html).

Sample collection, genomic DNA preparation and skin color measurement

We collected 546 blood samples from China (372 Han Chinese, 32 Blang, 31 Wa, 29 De'ang, 27 Manchu and 14 Zhuang), Cambodia (30 Lun) and Thailand (11 LaWa), representing 19 geographic populations living at different latitudes (supplementary table S2). Among the 546 individuals, skin color measurements were performed for 301 Han Chinese, including 119 males (average age$\pm$*S.D.*: 19.05$\pm$0.87) and 182 females (average age$\pm$*S.D.*: 18.73$\pm$0.96). We used a CR-400 tristimulus colorimeter (Konica Minolta, Tokyo, Japan) to measure skin color, and the CIE lab system for quantification. The colorimeter generates three numerical readings: L^*, a^* and b^*. L^* indicating the

level of darkness, and the lower L* value is the darker the measured skin area would be. Both constitutive skin color from skin areas not exposed to sunlight (underarm and buttock) and facultative skin color that is exposed to sunlight (back of hand) were measured for comparisons. To minimize technical variations, each skin area was measured three times, and the average value was used for analysis.

DNA samples were prepared using the standard phenol-chloroform method. Written informed consent was obtained from each individual prior to their inclusion in the study. All protocols of this study were approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences.

Genotyping of KITLG tag SNPs and genetic association analysis

Based on the result of selection test and evolutionary analysis, we picked 6 tag SNPs for genetic association analysis with skin color in 301 Han Chinese individuals. These six SNPs are located in the 549-kb genomic region containing the *KITLG* gene region, as well as the upstream and downstream regions of *KITLG* showing strong signals (F_{ST}) (Weir and Hill 2002) of Darwinian positive selection.

Genotyping was conducted by using an ABI 3730 sequencer (Applied Biosystems, Forster city, California, USA). Except for rs6538148 (three dropouts) and rs4590952 (four dropouts), there are no missing data in the genotyping of other variants. Genetic association of the tag SNPs with L^* values (back of hand, underarm and buttock) were conducted by utilizing PLINK v1.07 (Purcell, et al. 2007) with age and sex as covariates. Permutation test was used for *P* value correction. The PLINK code is: --plink --file kitlg --pheno kitlgpheno --all-pheno --linear --genotypic --covar kitlg.txt --reference-allele mylist.txt --perm --out kitlga.

Prediction of potential functional effects of candidate SNPs

We predicted functional effects of the candidate SNPs by searching the ENCODE database (http://genome.ucsc.edu/encode). Potential enhancer-associated histone

marks (H3K4me1) were collected from 7 cell lines. The H3K4me1 histone mark is the mono-methylation of lysine 4 of the H3 histone protein, and as regulatory elements associated with enhancers (Consortium 2012).

Functional test using dual-luciferase reporter assay

To construct *KITLG* enhancer reporters, we amplified fragments of 1.5 kb of *KITLG* by PCR from genomic DNA of two individuals homozygous with respect to the corresponding genotypes rs4073022 (GG and AA), rs428316 (AA and TT) and rs642742 (TT and CC), using primers tailed with Nhe I and Xho I restriction sites for rs4073022, Kpn I and Xho I restriction sites for rs428316 and rs642742 and directionally subcloned them into the pGL3-promter expression vector (Promega). We verified all recombinant clones by sequencing. A375 (human malignant melanoma) and HEK293 cells were cultured in high-glucose Dulbecco's modified Eagle's mediun (DMEM) (Corning) with 10% fetal bovine serum (FBS) (Gibco). For luciferase reporter assays, A375 and HEK293 cells were transfected with the indicated plasmids by using the lipofectamine 2000 transfection reagent (Invitrogen) and incubated overnight at 37℃ or 31℃ (based on the published protocol(Kim, et al. 2003)) in 5% CO_2 condition. According to the manufacture's instruction, luciferase activity was measured at 36 hours after transfection by using the Dual-Luciferase Reporter Assay System (Promega). All assays were performed in at least three replicates.

Environmental data collection and comparison

The geographic locations of the sampling points (latitude and longitude) were obtained from Google Earth (http://earth.google.com). The environmental data (insolation levels and temperature) were obtained from the *NASA* Langley Research Center Atmospheric Science Data Center Surface meteorological (https://eosweb.larc.nasa.gov) and Solar Energy (SSE) web portal supported by the NASA LaRC POWER Project. The insolation on horizontal surface data covered 22

years (Jul 1983 - Jun 2005), and the average annual total of UV radiation (KWh/m^2/year) for each geographic site was calculated via the global insolation on horizontal surface (Zhu 2005). The monthly average temperature (℃) (air temperature at 10 m) data covering 22 years (Jan 1983 - Dec 2004) was used in the regression analysis. The average temperatures in July and January were taken as the highest and lowest seasonal temperatures, respectively. The GraphPad Prism 5 software (https://www.graphpad.com/) was used in analyzing the correlation between environmental factors and candidate SNPs. In addition to the populations collected in this study, we also collected genetic and geographic data of 14 world populations (supplementary table S1) from the 1000 Genomes Project (http://www.1000genomes.org). These samples had information of clear geographic regions or states, and we either selected the collection points or the capitals of the country as the environmental data location.

For mediation analysis (Preacher and Hayes 2004), the SPSS 20.0 software was used. Either UV intensity or winter temperature is taken as independent variable (X), and SNP frequency is taken as dependent variable (Y). When UV intensity is taken as "X", winter temperature is then taken as mediator variable (M) so that its effect on the correlation between UV intensity and SNP frequency can be quantitatively evaluated, *i.e.* the percentage explained by the mediator effect out of the total effect. Similarly, when winter temperature is "X", then UV intensity becomes "M".

Acknowledgements

We are grateful to all voluntary donors in this study. We thank Dr. Dongsheng Lu for his help in *KITLG* haplotype tree construction. This study was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13010000), the National Natural Science Foundation of China (31601016, 31371268, 91631307, 31571370, and 91631106), Science and technology project of Yunnan Province (2017FD092), the scientific research project of Yunnan Provincial Department of Education (2017ZZX140), the Personnel Training Project of Yunnan Province (KKSY201526061) and State Key Laboratory of Genetic Resources and Evolution grant (GREKF15-06), and the One Hundred Talents Program of the Chinese Academy of Sciences.

References

Amyere M, Vogt T, Hoo J, Brandrup F, Bygum A, Boon L, Vikkula M 2011. KITLG mutations cause familial progressive hyper- and hypopigmentation. J Invest Dermatol 131: 1234-1239.

Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, et al. 2016. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 167: 1415-1429 e1419.

Bandelt HJ, Forster P, Rohl A 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37-48.

Barrett JC, Fry B, Maller J, Daly MJ 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263-265.

Basu Mallick C, Iliescu FM, Mols M, Hill S, Tamang R, Chaubey G, Goto R, Ho SY, Gallego Romero I, Crivellaro F, et al. 2013. The light skin allele of SLC24A5 in South Asians and Europeans shares identity by descent. PLoS Genet 9: e1003912.

Bedell MA, Brannan CI, Evans EP, Copeland NG, Jenkins NA, Donovan PJ 1995. DNA rearrangements located over 100 kb 5' of the Steel (Sl)-coding region in Steel-panda and Steel-contrasted mice deregulate Sl expression and cause female sterility by disrupting ovarian follicle development. Genes Dev 9: 455-470.

Beleza S, Johnson NA, Candille SI, Absher DM, Coram MA, Lopes J, Campos J, Araujo, II, Anderson TM, Vilhjalmsson BJ, et al. 2013a. Genetic architecture of skin and eye color in an African-European admixed population. PLoS Genet 9: e1003372.

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

Capellini TD, Chen H, Cao J, Doxey AC, Kiapour AM, Schoor M, Kingsley DM 2017. Ancient selection for derived alleles at a GDF5 enhancer influencing human

growth and osteoarthritis risk. Nat Genet 49: 1202-1210.

Chen H, Patterson N, Reich D 2010. Population differentiation as a test for selective sweeps. Genome Res 20: 393-402.

Chen H, Hey J and Slatkin M 2015. A hidden Markov model for investigating recent positive selection through haplotype structure. Theor. Popul. Biol. 99:18-30.

Consortium EP 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57-74.

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

Cuell A, Bansal N, Cole T, Kaur MR, Lee J, Loffeld A, Moss C, O'Donnell M, Takeichi T, Thind CK, et al. 2015. Familial progressive hyper- and hypopigmentation and malignancy in two families with new mutations in KITLG. Clin Exp Dermatol 40: 860-864.

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.

Eaton K, Edwards M, Krithika S, Cook G, Norton H, Parra EJ 2015. Association study confirms the role of two OCA2 polymorphisms in normal skin pigmentation variation in East Asian populations. Am J Hum Biol 27: 520-525.

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.

Gittelman RM, Schraiber JG, Vernot B, Mikacenic C, Wurfel MM, Akey JM 2016. Archaic Hominin Admixture Facilitated Adaptation to Out-of-Africa Environments. Curr Biol 26: 3375-3382.

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.

Grichnik JM, Burch JA, Burchette J, Shea CR 1998. The SCF/KIT pathway plays a critical role in the control of normal human melanocyte homeostasis. J Invest Dermatol 111: 233-238.

Guenther CA, Tasic B, Luo L, Bedell MA, Kingsley DM 2014. A molecular basis for classic blond hair color in Europeans. Nat Genet 46: 748-752.

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.

Huang Z, Ruan HB, Xian L, Chen W, Jiang S, Song A, Wang Q, Shi P, Gu X, Gao X 2014. The stem cell factor/Kit signalling pathway regulates mitochondrial function and energy expenditure. Nat Commun 5: 4282.

Hublin JJ, Ben-Ncer A, Bailey SE, Freidline SE, Neubauer S, Skinner MM, Bergmann I, Le Cabec A, Benazzi S, Harvati K, et al. 2017. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. Nature 546: 289-292.

Huttunen P, Hirvonen J, Kinnula V 1981. The occurrence of brown adipose tissue in outdoor workers. Eur J Appl Physiol Occup Physiol 46: 339-345.

Izagirre N, Garcia I, Junquera C, de la Rua 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 2010. Colloquium paper: human skin pigmentation as an adaptation to UV radiation. Proc Natl Acad Sci U S A 107 Suppl 2: 8962-8968.

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

Kasamatsu S, Hachiya A, Higuchi K, Ohuchi A, Kitahara T, Boissy RE 2008. Production of the soluble form of KIT, s-KIT, abolishes stem cell factor-induced melanogenesis in human melanocytes. J Invest Dermatol 128: 1763-1772.

Kim DS, Park SH, Kwon SB, Joo YH, Youn SW, Sohn UD, Park KC 2003. Temperature regulates melanin synthesis in melanocytes. Arch Pharm Res 26: 840-845.

Kumar S, Liu L 2014. No positive selection for G allele in a p53 response element in Europeans. Cell 157: 1497-1499.

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, et al. 2005. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310: 1782-1786.

Lao O, de Gruijter JM, van Duijn 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.

Loomis WF 1967. Skin-pigment regulation of vitamin-D biosynthesis in man. Science 157: 501-506.

Martin AR, Lin M, Granka JM, Myrick JW, Liu X, Sockell A, Atkinson EG, Werely CJ, Moller M, Sandhu MS, et al. 2017. An Unexpectedly Complex Architecture for Skin Pigmentation in Africans. Cell 171: 1340-1353 e1314.

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 Spec No 2: R176-181.

Mengel-From J, Wong TH, Morling N, Rees JL, Jackson IJ 2009. Genetic determinants of hair and eye colours in the Scottish and Danish populations. BMC Genet 10: 88.

Mikhail GR 1963. Hair Neogenesis in Rat Skin. Arch Dermatol 88: 713-728.

Miller CT, Beleza S, Pollen AA, Schluter D, Kittles RA, Shriver MD, Kingsley DM 2007. cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. Cell 131: 1179-1189.

Murray N, Norton HL, Parra EJ 2015. Distribution of two OCA2 polymorphisms associated with pigmentation in East-Asian populations. Hum Genome Var 2: 15058.

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.

Nishio M, Yoneshiro T, Nakahara M, Suzuki S, Saeki K, Hasegawa M, Kawai Y, Akutsu H, Umezawa A, Yasuda K, et al. 2012. Production of functional classical brown adipocytes from human pluripotent stem cells using specific hemopoietin cocktail without gene transfer. Cell Metab 16: 394-406.

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.

Pagani L, Lawson DJ, Jagoda E, Morseburg A, Eriksson A, Mitt M, Clemente F, Hudjashov G, DeGiorgio M, Saag L, et al. 2016. Genomic analyses inform on migration events during the peopling of Eurasia. Nature 538: 238-242.

Post PW, Daniels F, Jr., Binford RT, Jr. 1975. Cold injury and the evolution of "white" skin. Hum Biol 47: 65-80.

Preacher KJ, Hayes AF 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput 36: 717-731.

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 tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.

Rapley EA, Turnbull C, Al Olama AA, Dermitzakis ET, Linger R, Huddart RA, Renwick A, Hughes D, Hines S, Seal S, et al. 2009. A genome-wide association study of testicular germ cell tumor. Nat Genet 41: 807-810.

Rawofi L, Edwards M, Krithika S, Le P, Cha D, Yang Z, Ma Y, Wang J, Su B, Jin L, et al. 2017. Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry. PeerJ 5: e3951.

Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature 419: 832-837.

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES 2006. Positive natural selection in the human lineage. Science 312: 1614-1620.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449: 913-918.

Scheet P, Stephens M 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum Genet 78: 629-644.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, et al. 2009. Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. Am J Hum Genet 84: 534-541.

Soejima M, Koda Y 2007. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. Int J Legal Med 121: 36-39.

Stamatakis A 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312-1313.

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.

Taylor AC 1949. Survival of rat skin and changes in hair pigmentation following freezing. J Exp Zool 110: 77-111.

Wang ZQ, Si L, Tang Q, Lin D, Fu Z, Zhang J, Cui B, Zhu Y, Kong X, Deng M, et al. 2009. Gain-of-function mutation of KIT ligand on melanin synthesis causes familial progressive hyperpigmentation. Am J Hum Genet 84: 672-677.

Wehrle-Haller B 2003. The role of Kit-ligand in melanocyte development and epidermal homeostasis. Pigment Cell Res 16: 287-296.

Weir BS, Hill WG 2002. Estimating F-statistics. Annu Rev Genet 36: 721-750.

Wen B, Li H, Lu D, Song X, Zhang F, He Y, Li F, Gao Y, Mao X, Zhang L, et al. 2004. Genetic evidence supports demic diffusion of Han culture. Nature 431: 302-305.

Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD, Nielsen R 2007. Localizing recent adaptive evolution in the human genome. PLoS Genet 3: e90.

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.

Yuasa I, Umetsu K, Harihara S, Kido A, Miyoshi A, Saitou N, Dashnyam B, Jin F, Lucotte G, Chattopadhyay PK, et al. 2006. Distribution of the F374 allele of the SLC45A2 (MATP) gene and founder-haplotype analysis. Ann Hum Genet 70: 802-811.

Yuasa I, Umetsu K, Harihara S, Kido A, Miyoshi A, Saitou N, Dashnyam B, Jin F, Lucotte G, Chattopadhyay PK, et al. 2007. Distribution of two Asian-related coding SNPs in the MC1R and OCA2 genes. Biochem Genet 45: 535-542.

Zazo Seco C, Serrao de Castro L, van Nierop JW, Morin M, Jhangiani S, Verver EJ, Schraders M, Maiwald N, Wesdorp M, Venselaar H, et al. 2015. Allelic Mutations of KITLG, Encoding KIT Ligand, Cause Asymmetric and Unilateral Hearing Loss and Waardenburg Syndrome Type 2. Am J Hum Genet 97: 647-660.

Zhu QL, Yu, G.R., Cai, F., Liu, X.A., Li, Z.Q., Su,W., and Hui,L. 2005. patialization research on ultraviolet radiation in China. Resources Science 27.

Table 1. Association of the six *KITLG* variants with skin pigmentation in Han Chinese. The predicted enhancer-associated histone marks (H3K4me1) and the between-population divergence (measured by F_{ST}, CHB *vs.* YRI/CEU *vs.* YRI) are listed. The additive genetic model was used for the association analysis. Age and sex were considered as covariates. Permutation test was used for multiple test correction.

SNP (derived allele)	Position (hg 19)	H3K4Me1 peak	F_{ST}	*P*-value (BETA value)		
				Hand	Underarm	Buttock
rs6538148(G)	chr12:88818479	3.0	0.609/0.513	1.000(0.10)	0.091(0.93)	0.667(0.20)
rs4590952(A)	chr12:88953659	29.6	0.293/0.411	0.857(0.09)	0.104(-0.63)	0.234(-0.62)
rs1703078(C)	chr12:89018883	7.0	0.373/0.744	0.600(-0.15)	0.414(0.36)	0.152(0.64)
rs4073022(A)	chr12:89037372	--	0.475/0.779	0.857(0.26)	**0.047**(0.77)	**0.008**(1.32)
rs428316(T)	chr12:89290545	156.6	0.591/0.735	0.375(0.52)	**0.031**(0.99)	**0.043**(1.24)
rs642742(C)	chr12:89299746	5.0	0.591/0.735	0.304(0.53)	**0.039**(0.99)	**0.043**(1.24)

Fig. 1. Test of selection on *KITLG* in Europeans and East Asians. A) The EHH score distribution across the 549 kb *KITLG* region. Each spot in the graph indicates the average value of a 3-kb window. **B)** The XP-CLR score distribution across the 549 kb *KITLG* region. Each spot in the graph indicates the average value of a 2-kb window. The red, blue and black dotted lines indicate $P = 0.05$ in East Asians (CHB), Europeans (CEU) and Africans (YRI), respectively. The black arrow indicates the genomic position and transcriptional direction of the *KITLG* gene.

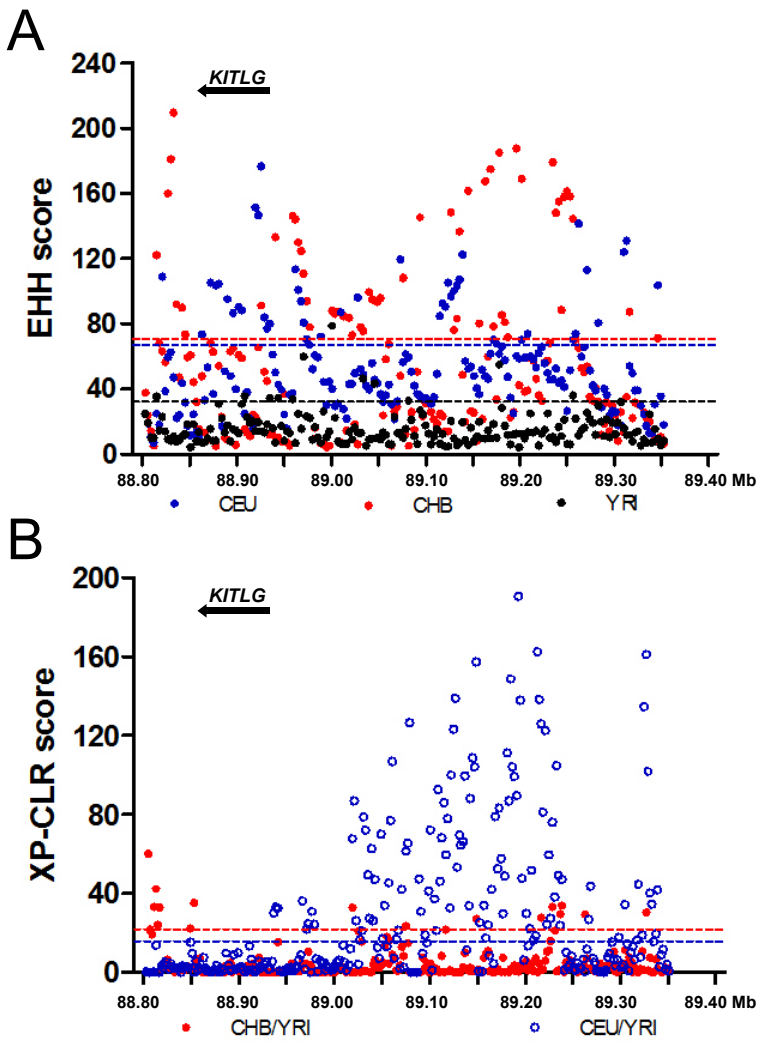

Fig. 2. The *KITLG* haplotype tree and the allelic status of each haplotype. A) The maximum-likelihood tree was constructed using the 549 kb *KITLG* sequence data (540 haplotypes) of three world populations (CHB, CEU and YRI) from the 1,000 Genomes Project. The homologous sequences of chimpanzee (circle), Neanderthal (diamond) and Denisovan (triangle) were used as outgroup; **B**) The allelic status (red for major allele and blue for minor allele) of the 727 polymorphic sites (minor allele frequency larger than 0.05) in the 549 kb *KITLG* region. The black arrow indicates the genomic position and transcriptional direction of the *KITLG* gene.

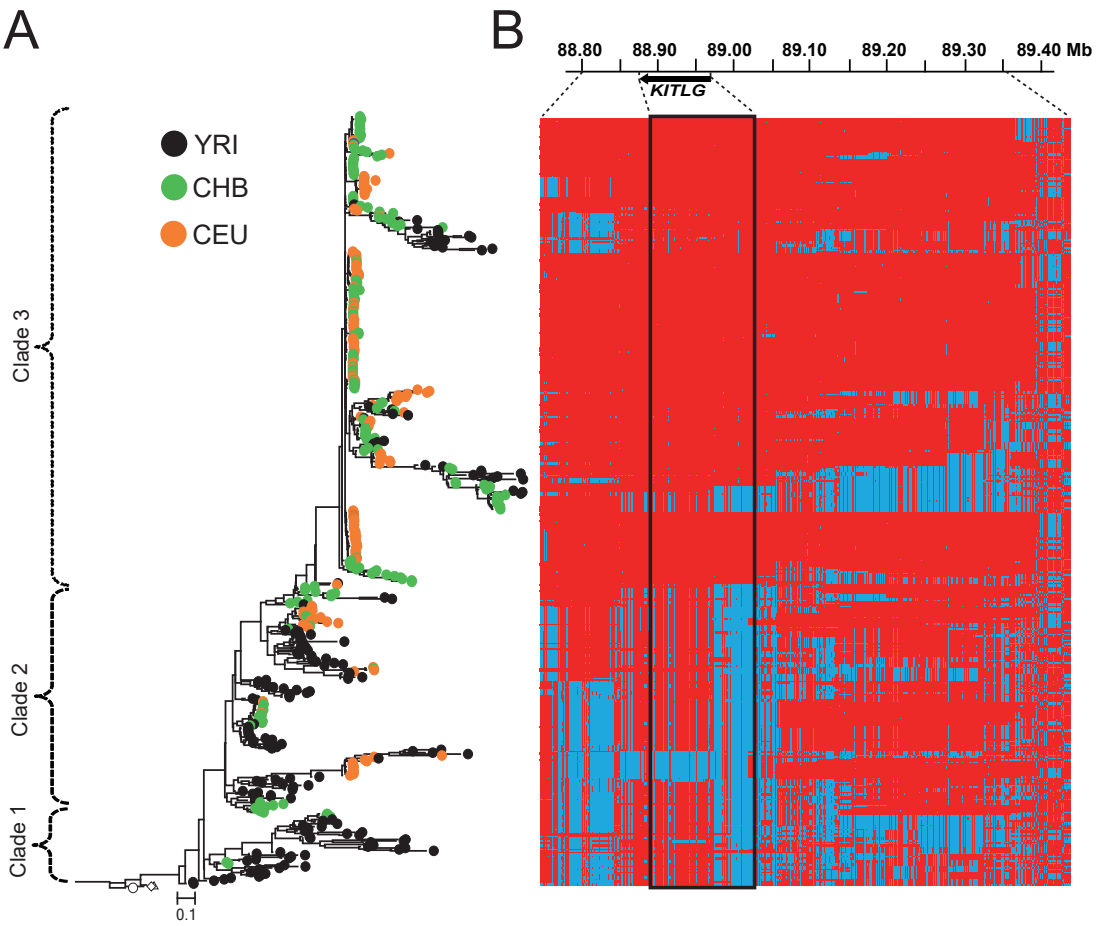

Fig. 3. Dual-luciferase enhancer reporter gene assay of three upstream *KITLG* variants at normal culture condition of 37℃. A375 (human malignant melanoma) and HEK293 cells were used. All assays were performed in at least three replicates, and an empty vector was used as control. The *P* values were calculated by *t*-test. The presumably adaptive (major) alleles are "A" for rs4073022, "T" for rs428316 and "C" for rs642742.

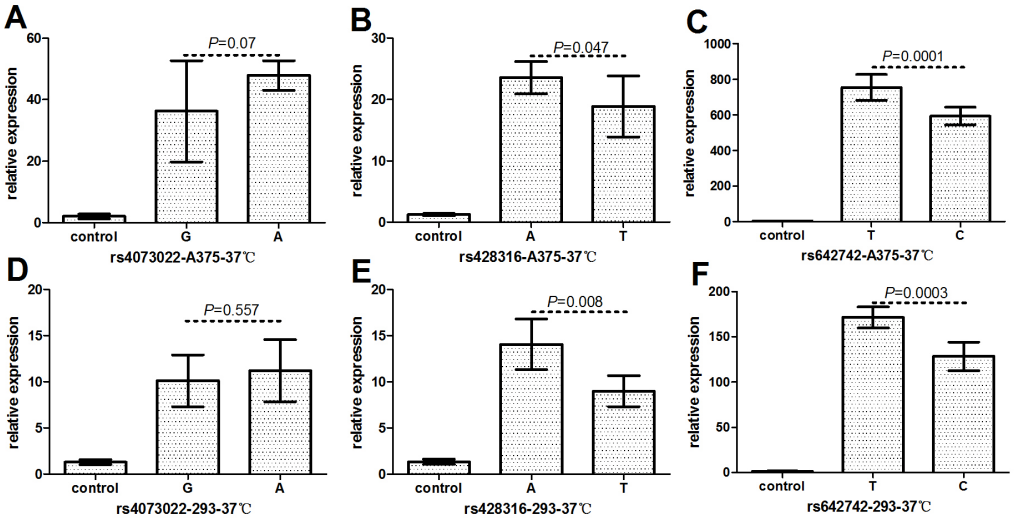

Fig. 4. Correlations of two *KITLG* variants with environment factors. A total of 33 world populations living in different latitudes were included (blue and red dots). The red dots indicate the 19 Asians populations in the 33 world populations. *P*-value was calculated using F test and the correlation coefficient (r^2) is indicated in every panel. The UV radiation levels and seasonal temperatures were collected from the *NASA* Langley Research Center. Detailed data are shown in supplementary table S3.

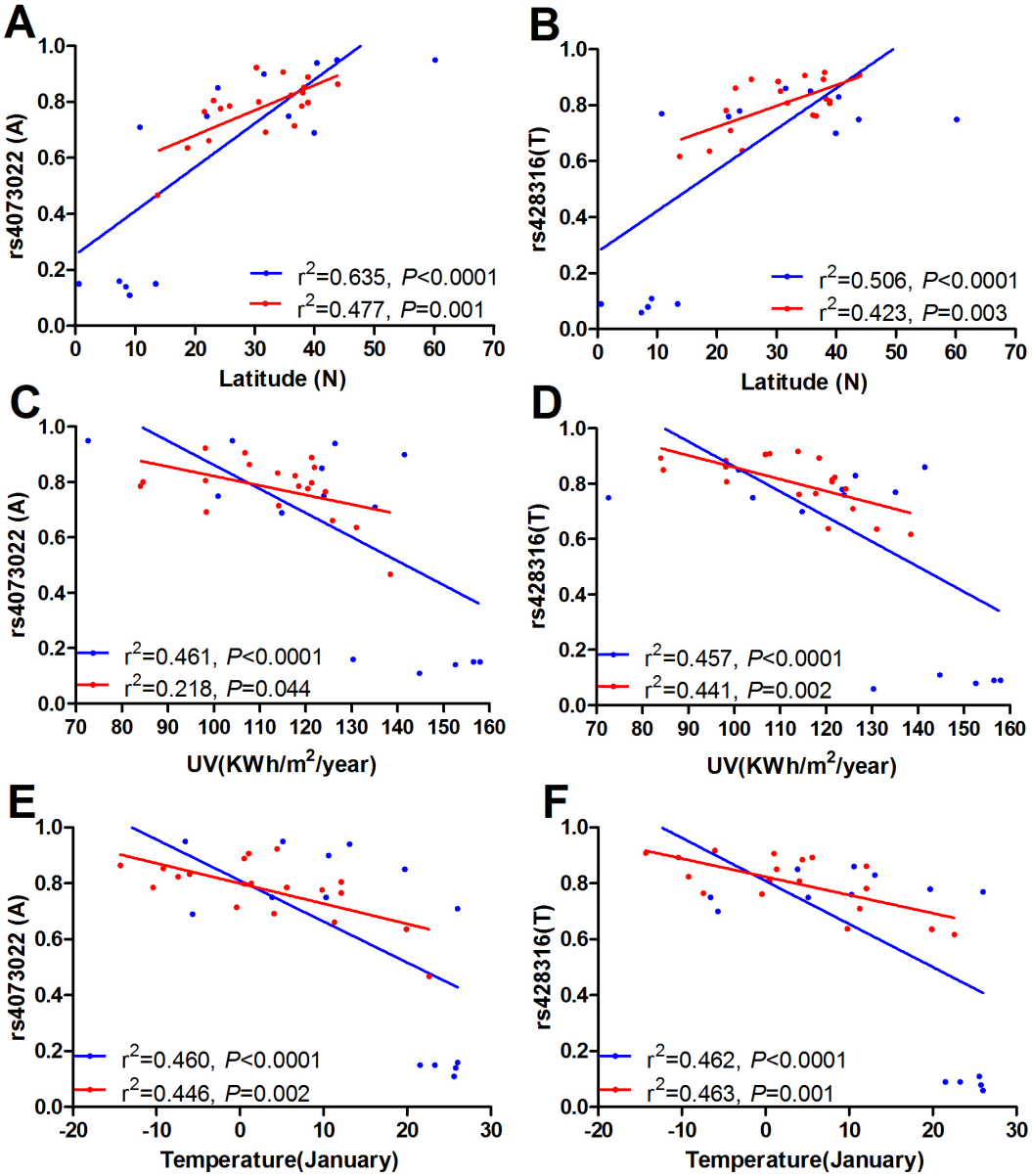

Fig. 5. Dual-luciferase enhancer reporter gene assay of three upstream *KITLG* variants at low temperature (31℃) culture condition. All assays were performed in at least three replicates, and an empty vector was used as control. The *P* values were calculated by *t*-test. The presumably adaptive (major) alleles are "A" for rs4073022, "T" for rs428316 and "C" for rs642742.