A Genetic Mechanism for Convergent Skin Lightening during Recent Human Evolution

Zhaohui Yang,†,1,2,13 Hua Zhong,†,3 Jing Chen,†,4 Xiaoming Zhang,†,1 Hui Zhang,† Xin Luo,†,2 Shuhua Xu,5 Hua Chen,6 Dongsheng Lu,5 Yinglun Han,7 Jinkun Li,8 Lijie Fu,8 Xuebin Qi,† Yi Peng,1 Kun Xiang,1 Qiang Lin,1,2 Yan Guo,† Ming Li,† Xiangyu Cao,† Yanfeng Zhang,† Shiyu Liao,† Yingmei Peng,† Lin Zhang,4 Xiaosen Guo,9,10 Shanshan Dong,9 Fan Liang,9 Jun Wang,9,10 Andrew Willden,1 Hong Seang Aun,11 Bun Serey,11 Tuot Sovannary,11 Long Bunnath,11 Ham Samnom,12 Graeme Mardon,3 Qingwei Li,7 Anming Meng,*5 Hong Shi,*1,13 and Bing Su,*1

†These authors contributed equally to this work.

*Corresponding author: E-mail: sub@mail.kiz.ac.cn; mengam@mail.tsinghua.edu.cn; shihong@kbimed.com.

Associate editor: Anna Di Rienzo

Abstract

Skin lightening among Eurasians is thought to have been a convergence occurring independently in Europe and East Asia as an adaptation to high latitude environments. Among Europeans, several genes responsible for such lightening have been found, but the information available for East Asians is much more limited. Here, a genome-wide comparison between dark-skinned Africans and Austro-Asiatic speaking aborigines and light-skinned northern Han Chinese identified the pigmentation gene OCA2, showing unusually deep allelic divergence between these groups. An amino acid substitution (His615Arg) of OCA2 prevalent in most East Asian populations—but absent in Africans and Europeans—was significantly associated with skin lightening among northern Han Chinese. Further transgenic and targeted gene modification analyses of zebrafish and mouse both exhibited the phenotypic effect of the OCA2 variant manifesting decreased melanin production. These results indicate that OCA2 plays an important role in the convergent skin lightening of East Asians during recent human evolution.

Key words: OCA2, East Asians, skin lightening, pigmentation genes, adaptation, natural selection.

Introduction

Light skin is thought to be an adaptation gained by modern humans needed to cope with the new environmental and biological factors encountered after exiting Africa about 100,000 years ago (Loomis 1967; Jablonski and Chaplin 2000; Chaplin 2004). Among many environmental factors, ultraviolet (UV) is required for the synthesis of vitamin D, a key nutrient for human bone development (Loomis 1967), and selection on vitamin D synthesis would favor a lighter skin for people living in high latitude regions with relatively weak UV radiation.

Previous investigations into the bases of skin lightening among Europeans found a set of pigmentation genes (e.g.,
SLC24A5 and SLC45A2) responsible for their light skin (Graf et al. 2005; Lamason et al. 2005; Izagirre et al. 2006; Yuasa et al. 2006; Lao et al. 2007; Norton et al. 2007; Sabeti et al. 2007; Soejima and Koda 2007; Han et al. 2008; Beleza et al. 2013; Wilde et al. 2014). The Ala111Thr variant of SLC24A5 proved the most striking, being nearly fixed in Europeans (> 95%) due to local Darwinian positive selection, but extremely rare in Africans and East Asians (Lamason et al. 2005). Curiously, while East Asians share similarly light skin, the identified mutations of the pigmentation genes found in Europeans are rare in East Asians; instead, other pigmentation genes (e.g., OCA2, DCT, and ATRN) have been reported under selection in East Asians. For example, the variants of OCA2 (rs1800414 and rs74653330) were shown to be associated with skin melanin content (Edwards et al. 2010; Eaton et al. 2015). The DCT gene harbored a signature of positive selection in Chinese using multiple neutrality tests (Lao et al. 2007; Myles et al. 2007), and many loci in ATRN showed extended blocks in East Asians, an indication of selection (Norton et al. 2007). These data suggest convergent evolution of light skin in both East Asians and Europeans (Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Edwards et al. 2010). However, no functional analysis has been conducted to reveal the underlying genetic mechanism of how these pigmentation genes affect melanin production in the skin and cause skin lightening in East Asians.

Compared with Europeans north of the Mediterranean, East Asian populations living between the equator and North Pole manifest a wider spectrum of skin pigmentation (fig. 1A). Previous studies proposed that in the population tree, the dark-skinned Austro-Asiatic speaking populations (abbreviated as A-A) currently residing in southwestern China and Southeast Asia occupy a more basal clade than the light-skinned northern populations who migrated northward some 25,000–30,000 years ago (supplementary fig. S1, Supplementary Material online) (Su et al. 1999; Shi et al. 2005; Consortium et al. 2009). By extension, the A-A group may represent the skin type of early modern human settlers in eastern Asia. Accordingly, a comparative genomic comparison between light-skinned northern populations (e.g., northern Han Chinese, CHN) and dark-skinned A-A as well as African populations may help identify the pigmentation gene(s) responsible for skin lightening in East Asians. Meanwhile, the phenotypic effect of selected mutations in the pigmentation gene(s) can be modeled in laboratory animals to reveal the genetic mechanism.

**Results**

**Skin Pigmentation in CHN and A-A Populations**

We measured the level of darkness of the constitutive skin areas (underarm and buttock, exposed to sunlight) and the facultative skin area (hand back, exposed to sunlight) of 1,518 unrelated individuals from A-A (1,159 individuals from 15 geographic regions) and CHN (359 individuals from 8 northern provinces of China) populations (supplementary table S1, Supplementary Material online). The L* value (a measurement of skin pigmentation reflectance) was used to quantify the level of darkness, that is, lower L* values denote darker skin. A-A populations showed markedly darker skin than the CHN population for both the constitutive and facultative areas (fig. 1A and supplementary fig. S2, Supplementary Material online). After adding published data from African Americans and European Americans (underarm) (Shriver and Parra 2000), there was a clear similarity between A-A and African Americans and also between CHN and European Americans (fig. 1B).

**Genome-Wide Analysis of Sequence Variations**

Using an Affymetrix Genome-wide Human Single Nucleotide Polymorphism (SNP) Array 6.0 (> 900,000 SNPs), we genotyped 48 individuals randomly selected from A-A populations...
of southwestern China. Paired with published Affymetrix 6.0 data of other populations (International HapMap et al. 2010), including Han Chinese from Beijing (CHB, 89 individuals), Han Chinese from Denver (CHD, 90 individuals), Japanese (JPT, 90 individuals), Europeans (CEU, 118 individuals), and Africans (YRI, 120 individuals), principal component analysis (PCA) indicated a close relationship among Asian populations (A-A, CHB, CHD, and JPT) (supplementary fig. S3A, Supplementary Material online), with some local divergence (supplementary fig. S3B, Supplementary Material online).

To identify the genes responsible for skin pigmentation divergence, we looked for genes showing extraordinarily large divergence between CHN (with CHB used to represent CHN) and A-A and Africans (with YRI used to represent Africans) by calculating between-population allele frequency difference of each SNP (measured by $F_{ST}$) (Weir and Hill 2002). Compared with the randomly selected SNPs, SNPs located in pigmentation genes showed larger enrichment scores ($P < 0.001$, $\chi^2$ test) (supplementary fig. S4, Supplementary Material online), suggesting that pigmentation genes may be under selection in CHB.

In total, we identified 1,283 (CHB vs. A-A), 1,640 (CHB vs. YRI), and 1,706 gene regions (YRI vs. A-A) showing large divergences ($F_{ST}$, top 1% regions). Overlapping these gene regions with 171 known pigmentation genes from the human genome (Color Genes database, http://www.espcr.org/mice mut/, last accessed September 03, 2015) yielded three gene sets (15, 21 and 19 genes, respectively), among which three pigmentation genes (OCA2, EDAR, and PAH) showed consistent signals (supplementary table S2, Supplementary Material online). In the genome-wide $F_{ST}$ analysis, we pooled the samples of three A-A populations from China, and the potential population substructure may cause false negatives. Consequently, the listed gene sets are conservative callings of genes under selection (supplementary table S2, Supplementary Material online). The identified candidate pigmentation genes are mostly different from those in Europeans, supporting the hypothesized convergent evolution of skin pigmentation in East Asia and Europe.

Compared with EDAR and OCA2, the signal of PAH is relatively weak and its functional role in pigmentation remains elusive (Schallreuter et al. 1999). While EDAR has accumulated East Asian specific mutations due to local selection, its functional effect lies in an increased number of active eccrine glands (Kamberov et al. 2013). In contrast, OCA2 showed the largest between-population (light-skinned vs. dark-skinned) divergence and was previously reported to be associated with melanin content (Edwards et al. 2010; Eaton et al. 2015). For further analysis, we focused on OCA2.

OCA2 Accounts for Light Skin in CHN

To search for functional mutation(s) of OCA2, we conducted amplicon sequencing of 40 randomly selected A-A individuals (25 from China and 15 from Cambodia), and obtained sequences of the entire genomic region of OCA2 (344.5 kb) which we compared to data from other populations collected within the 1000 Human Genomes Project (Genomes Project et al. 2012). First, the maps of linkage disequilibrium (LD) (supplementary fig. S5, Supplementary Material online) showed that CHB exhibited the strongest LD, which could be explained either by Darwinian positive selection or population history (e.g., bottleneck effect).

Within the gene region of OCA2, we observed a peak of allele frequency divergence (measured by $F_{ST}$) (Weir and Hill 2002) between CHB and A-A, spanning about 44.2 kb, in which rs1800414 was among a group of SNPs showing large $F_{ST}$ values (fig. 2A). Further tests via extended haplotype homozygosity (EHH) (Sabeti et al. 2002) and Fay and Wu’s $H$ statistics (Fay and Wu 2000) (the 44.2 kb region was analyzed) supported the signal of selection in CHB (supplementary fig. S6 and table S3, Supplementary Material online), in line with the previous report using other tests of selection (LSBL statistics, InR$H$, and Tajima’s $D$ test) (Edwards et al. 2010). Conversely, no selective signal was observed in A-A and Europeans (supplementary table S3, Supplementary Material online). rs1800414 is a nonsynonymous SNP located in exon 17 of OCA2, and an A to G mutation causes an amino acid change from histidine to arginine (His615Arg). This missense mutation is located in the melanosomal cytoplasmic topological domain (http://www.uniprot.org/, last accessed September 15, 2015). The geographic distribution of this SNP across world populations (data from HGD database: http://hgdp.uchicago.edu, last accessed July 10, 2015) showed that the derived G allele of rs1800414 is highly prevalent in most East Asian populations (G allele > 50% in Han Chinese), as well as in American Indians, but is in low frequency in A-A populations (G allele = 16.0% [A-A from China]; 21.3% [A-A from Cambodia]), totally absent in Africans, and absent or extremely rare in western Europeans (G allele < 0.3%) (fig. 2B). This piece of evidence suggests that rs1800414 is likely an East Asian-specific mutation, consistent with the proposed local selection on this SNP and an independent evolution of skin pigmentation in the area (Edwards et al. 2010).

Assuming rs1800414 being the functional SNP under selection, we estimated the selective intensity and the time of selection using 1000 Human Genomes Project data (Genomes Project et al. 2012). In CHB, the selective coefficient was estimated to be 0.0157 (95% confidence interval [CI]: 0.0127–0.0186). Assuming a generation time of 25 years, the time of selection onset was estimated at 15,224 years ago (95% CI: 12,818–18,750 years ago), which is slightly older than a recently reported age (10,660 years ago with 95% CI of 8,070–15,780 years ago) using a Hidden Markov model (Chen and Slatkin 2015). The estimated age falls into the Upper Paleolithic, following the proposed northward migration of modern humans in East Asia some 25,000–30,000 years ago (Shi et al. 2005).

We conducted skin phenotyping of CHN (359 individuals) to determine whether the OCA2 SNPs affect skin pigmentation. We analyzed five SNPs located in the 44.2 kb region and three SNPs outside this region with strong selective signals. Under the additive genetic model, we detected a significant association of rs1800414 with skin pigmentation for both exposed ($P = 5.12 \times 10^{-4}$, hand back) and unexposed ($P = 5.34$...
in each cell line using the total amount of proteins as a control according to published protocols (Cook et al. 2009). As predicted, homozygotes of the derived G allele (Arg615) had the least amount of melanin (correspondent to a relatively lighter skin), whereas the homozygotes of the ancestral A allele (His615) had the highest amount (correspondent to a relatively darker skin), leaving the heterozygotes between the two types of homozygotes (fig. 3B). Surprisingly, the ancestral allele of rs1800414 (His615) is highly conserved among species from zebrafish to human (fig. 3C). The measurement of GERP’s (Genomic Evolutionary Rate Profiling) score suggested a strong functional constraint on this site during evolution (GERP’s score = 366, $P = 2.93 \times 10^{-72}$) (http://mendel.stanford.edu/SidowLab, last accessed October 20, 2015; Cooper et al. 2005), implying that the Arg615 substitution may change OCA2 function.

Equivalent A and G Alleles of rs1800414 in Zebrafish Have Different Functions

The deficiency of Oca2 leads to hypopigmentation in the skin melanophores and the retinal pigment epithelium, indicating an essential role in melanin synthesis (Beirl et al. 2014). The zebrafish Oca2 protein has His610 residue equivalent to His615 of human OCA2. The underpigmentation of melanophores depleted of oca2 allows us to test functional relevance of the His610Arg mutation genocopying the A to G substitute at rs1800414 in human. The expression of oca2 could be effectively blocked by injection of an oca2-specific antisense morpholino (oca2-MO) (supplementary fig. S7A, Supplementary Material online) and oca2 morphants exhibited progressive loss of melanophore pigmentation in a dose-dependent manner at 36 or 48 h postfertilization (supplementary fig. S7B, Supplementary Material online), which phenocopied oca2 mutants (Beirl et al. 2014). Next, we performed rescue experiments using wild-type zebrafish oca2 mRNA (wt(A) with His610) or mutant oca2 mRNA (mt(G) with His610Arg mutation). Results indicated that coinjection of 100 pg oca2 wt (A) mRNA largely recovered melanophore pigmentation in oca2 morphants whereas coinjection of 100 pg oca2 mt (G) mRNA failed to do so (fig. 4A and B). Hence, His610 of zebrafish Oca2 protein is important for melanin synthesis, supporting the idea that the rs1800414 variant of human OCA2 may be functionally associated with skin coloration.

Mice Homozygous for G Allele-Like Oca2 (His610Arg) Display Light Coloration

To further test the phenotypic effect of rs1800414 on melanin production and skin pigmentation, using the CRISPR/Cas9 approach (Li et al. 2013; Wang et al. 2013; Yang et al. 2013), we generated two gene-modified mouse strains using black hair C57BL/6 mice (WT), including the Oca2 knockout strain (KO) (a frame-shift mutation leading to a truncated Oca2) and the rs1800414 targeted modification strain with His610Arg mutation resembling human His615Arg mutation (#1 mutation, supplementary fig. S8A and B, Supplementary Material online). We also generated a control strain carrying

![Diagram](https://example.com/diagram.png)
two Oca2 synonymous mutations (WT) (2 and 3 mutations, supplementary fig. S8A and B, Supplementary Material online) in order to rule out any potential effect of these two synonymous mutations introduced when generating the targeted modification using CRISPR/Cas9. No off-target effect was detected when sequencing the top five predicted potential off-targeted sites in all mouse strains (supplementary fig. S8, Supplementary Material online). Sequencing of Oca2 RNA transcripts did not detect any change of gene alternative splicing among the mouse strains (supplementary fig. S8C, Supplementary Material online), suggesting that the phenotypic effect of rs1800414 is caused by protein functional alteration due to the observed amino acid change.

We compared the distributions of melanosomes in the tail skin among different mouse strains by transmission electron microscope. At cellular level, the TM mice showed a clear reduction of mature melanosomes compared with the WT and WT' mice and the KO mice appeared to lack mature melanosomes (fig. 4D). These results imply that the derived allele of rs1800414 alters the function of Oca2, ultimately affecting the maturation and/or transportation of melanosomes in the melanocytes.

**Discussion**

Here, we identified a set of pigmentation genes showing large genetic divergence between dark-skinned (A-A and YRI) and light-skinned populations (CHN) (supplementary fig. S4 and table S2, Supplementary Material online), which comprise potential candidates contributing to skin lightening among East Asians. The sustained dark skin of A-A people is likely due...
to their settling in low latitude tropical areas of Asia, which are similar to those inhabited by our African ancestors. When these groups migrated northward in East Asia about 25,000–30,000 years (Jin and Su 2000; Shi et al. 2005), natural selection would seemingly favor lighter skin, due to the reduced UV radiation at high latitudes. Though multiple pigmentation genes might be involved in this transition, our results strongly indicate that OCA2 served as a major-effect gene conferring skin lightening of East Asians.

OCA2 is one of the key players in the pigmentation pathway, and previous reports associated OCA2 with abnormal skin pigmentation, oculocutaneous albinism type 2 (OCA2)
(Gardner et al. 1992; Rinchik et al. 1993; Lee et al. 1994). The protein encoded by OCA2 (a 12-transmembrane protein) is localized on the melanosomal membrane (Rinchik et al. 1993; Rosenblat et al. 1994), and may be involved in the control of L-tyrosine transport (Toyofuku et al. 2002), pH adjustment in melanosome (Puri et al. 2000), glutathione metabolism (Staleza et al. 2002), and tyrosinase processing (Chen et al. 2002). We showed that the nonsynonymous SNP rs1800414 is the most diverged variant between CHN and A-A, and significantly associated with skin pigmentation for the tested skin areas in CHN (fig. 3A). Moreover, the derived G allele homozgyotes produced the least amount of melanin in the cultured melanocytes, and showed the lightest skin in northern Han Chinese (fig. 3B). This phenotypic effect was confirmed in the gene-modified zebrafishes and mice (fig. 4), suggesting that the His615Arg mutation at rs1800414 affects the maturation and/or transportation of melanosomes, eventually leading to skin lightening.

Given the demonstrated role of OCA2 for skin lightening of East Asians and previous reports of different pigmentation genes responsible for light skin in Europeans (SLC24A5 and OCA2) (Graf et al. 2005; Lamason et al. 2005; Izagirre et al. 2006; Yuasa et al. 2006; Lao et al. 2007; Norton et al. 2007; Sabeti et al. 2007; Soejima and Koda 2007; Han et al. 2008; Beleza et al. 2013; Wilde et al. 2014), our data support a convergent evolution of skin pigmentation in East Asians (Lao et al. 2007; Myles et al. 2007; Norton et al. 2007; Edwards et al. 2010). Although OCA2 also showed a potential signal of selection in Europeans (supplementary table S2, Supplementary Material online; Lao et al. 2007; Norton et al. 2007), the causal SNP rs1800414 enriched in East Asians is nearly absent in Europeans. Intriguingly, we found that the proteins of OCA2, SLC24A5, and OCA2 are all trans-membrane proteins located on the melanosomal membrane, implying that similar molecular checkpoints were acted on by natural selection, eventually leading to independent skin-lightening in both East Asians and Europeans, who live in high latitude areas with weak light conditions.

There are also pigmentation genes (DCT, EGFR, DRD2, and MC1R) reportedly under selection in dark-skinned people (e.g., Africans) (Harding et al. 2000; John et al. 2003; Lao et al. 2007), which have been explained by the protection conferred by melanin against the harmful effects of UV radiation, including protection against sunburn and folate destruction. For light-skinned people, because UV is required for the synthesis of vitamin D, a key nutrient for bone development (Loomis 1967), and selection on vitamin D synthesis would favor a lighter skin at these high latitude regions and thereby explain the selective signal of OCA2 in northern Han Chinese. However, we cannot rule out other environmental and biological factors (e.g., diet change and/or sexual selection) that may also create selective pressures on skin lightening. Despite differences in the underlying causes, our evidence for a convergent evolution of skin lightening in Europe and East Asia represents an excellent case of how natural selection has shaped our phenotypic diversity in recent human evolutionary history.

Materials and Methods

Sample Collection and Skin Pigmentation Measurement

We collected blood samples from 1,159 Austro-Asiatic speaking individuals (716 females and 443 males) living in southwestern China (Yunnan province) and northeastern Cambodia as well as 359 blood samples (210 females and 149 males) from northern Han Chinese of China (supplementary table S1, Supplementary Material online). Sample collection was conducted randomly and biologically related individuals were excluded from this study. DNA samples were prepared using the standard phenol–chloroform method.

We used a CR-400 tristimulus colorimeter (Konica Minolta, Tokyo, Japan) to measure skin pigmentation, which generates three numerical readings that can express the color of any object: L* (levels of darkness), a* (amount of green or red), and b* (amount of yellow or blue). Both constitutive skin pigmentation from skin not exposed to sunlight (underarm and buttock) and facultative skin pigmentation that is exposed to sunlight (back of hand) were measured for comparisons. To minimize technical variations, each area was measured three times, and the average used for statistical analysis. Due to the known seasonal variation in skin phenotype, skin pigmentation measurements were conducted in the same season (summer of 2011–2012) for all populations, the same seasons with the published data (Shriver and Parra 2000). GraphPad Prism 5 was used to plot L* distribution.

Written informed consent was obtained from each individual prior to their inclusion in the study. Additional written consent was obtained for use of portrait photos. All protocols of this study were approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences (approval No: KMDWS-SEYX-20120101010).

Genome-Wide SNP Analysis

We employed an Affymetrix Genome-wide Human SNP Array 6.0, which includes 906,600 SNPs with an average inter-SNP distance of 1.5 kb, to scan 48 randomly selected Austro-Asiatic individuals from southwestern China (20 Wa, 10 Bulang, and 18 De‘ang) (supplementary table S1, Supplementary Material online). The published Affymetrix data of Han Chinese from Beijing (CHB, 89 individuals) (International HapMap et al. 2010) was used to represent northern Han Chinese and compared with the A-A population.

PCA was used to show the relationship of the A-A samples with other populations. The SNPs were pruned based on LD for PCA, no particular genomic regions were removed prior to the PCA.

We also calculated the enrichment score for each FST bin when comparing A-A and CHB among genes involved in the pigmentation pathways and those not involved. We divided the FST values into five bins with the value interval of 0.05 in each bin. Simply, the FST values assigned to each bin were (0–0.05), (0.05–0.1), (0.1–0.15), (0.15–0.2), and (≥0.2), as
shown in supplementary fig. S4, Supplementary Material online. The genes spanning the SNP were assigned to the relevant $F_{ST}$ bin based on the $F_{ST}$ value of that SNP. The enrichment score in each bin was calculated based on 1,000 bootstrapping data sets from random SNPs or pigmentation-associated SNPs, the score was the fraction of SNPs in each bin, divided by the fraction of all SNPs from known genes except for the pigmentation associated genes.

To identify genomic regions showing deep divergence between dark-skinned (A-A and YRI) and light-skinned populations (CHN), we used the moments estimate of $F_{ST}$ described by Weir and Hill to measure between-population genetic divergence (Weir and Hill 2002). Overlapping with 171 known pigmentation genes from the human genome (Color Genes database, http://www.espcr.org/micemut/, last accessed September 20, 2015) yielded 15 (CHB vs. A-A), 21 (CHB vs. YRI), and 19 (YRI vs. A-A) pigmentation gene regions with large divergence ($F_{ST}$, top 1% gene regions).

Resequencing of OCA2 and Molecular Evolution Analysis

For sequencing, we randomly selected 40 A-A individuals (25 from China and 15 from Cambodia) and designed a series of primers to cover the entire 344.483 kb genomic region of OCA2. Using data from the 1000 Human Genomes Project (Genomes Project et al. 2012), we extracted all OCA2 SNPs of individuals from populations of Han Chinese (CHB 97), western Europeans (CEU 85), and Africans (YRI 88).

Haploview software (D’ algorithm) (Bandelt et al. 1999) was used to analyze LD pattern of OCA2 in different populations. Genetic divergence between A-A and CHN in the sequenced OCA2 genomic region (~344.5 kb) was measured by the averaged $F_{ST}$. The averaged $F_{ST}$ of each position is the simple moving average with a sliding window size of 10 kb. We also used the EHH test described by Sabetti et al. (2002) for detecting recent positive selection displaying incomplete selective sweep, that is, the selected allele has not reached fixation. We carried out tests of neutrality following the method developed by Fay and Wu (2000), using the program DnaSPV5 (Librado and Rozas 2009). A significant negative Fay and Wu’s $H$ indicates an excess of high-frequency derived alleles due to Darwinian positive selection.

Inferring Allele Age and Intensity of Selection of rs1800414

The SNP data covering a region of 942.7 kb for 97 CHB individuals was obtained from the 1000 Genomes Project (Genomes Project et al. 2012). We used fastPHASE (Scheet and Stephens 2006) to infer the haplotype phase from SNP genotypes. The inferred haplotypes were then used in further analysis. Starting from the putatively selected mutant position, we ran a hidden Markov model to identify ancestral haplotypes retained around the vicinity of the selected mutant. The extent of ancestral haplotypes was recorded and the likelihood function was built based on the haplotype extent. We adopted the importance sampling approach by Chen and Slatkin (2013) to infer the parameters, including selection intensity and allele age (see the original paper for the details of the Materials and Methods). A generation time of 25 years was used.

Genetic Association Analysis

According to the result of molecular evolution analysis of OCA2, we chose eight weak linkage SNPs as possible for association analysis, including five SNPs located in the 44.2 kb region with strong selective signal and other three SNPs outside this region. A total of 359 blood samples (210 females and 149 males) from Han Chinese individuals (CHN, light-skinned) from northern China (supplementary table S1, Supplementary Material online) were used for skin pigmentation association analysis. Genotyping was conducted by the SnapShots method on an ABI 3130 sequencer (Applied Biosystems, Forster city, CA, USA). Skin pigmentation association of single SNP with L* (hand back, underarm, and buttock) were tested by utilizing PLINK v1.07 (Purcell et al. 2007), using the linear regression option, with age and sex as covariates.

Melanocyte Culture and In Vitro Analysis of Melanin Synthesis

We collected foreskin tissue samples of male individuals (8–20 years old) who had circumcisions at the local hospitals. Written informed consent was obtained from each donor or donor’s parents prior to sample collection. The published method by Cook et al. (2003) was used to culture and purify melanocytes. We used the melanin analytical method by Cook et al. (2009) to calculate the total amount of melanin and protein of melanocytes in each cell culture dish. Total melanin and protein were isolated from cells with different rs1800414 genotypes (4 AA, 3 AG, and 4 GG). The protocol of this study was approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences (approval No: KMDWS-SEYX-2012010101).

Functional Study of oca2 in Zebrafish Embryos

Wild-type embryos of Tuebingen strain were used. Embryos were incubated in Holtfreter’s solution at 28.5°C and staged as previously described (Kimmel et al. 1995). Ethical approval was obtained from the Animal Care and Use Committee of Tsinghua University.

The coding sequence of oca2 was amplified and cloned into the pXT7 vector using enzymatic assembly method (Gibson et al. 2009), with Kozak sequence added to enhance the translation efficiency. The A > G nucleotide substitution at + 1829, which resulted in His610Arg mutation, was made with site direct mutation.

The oca2-MO, synthesized by Gene Tools, LLC, targets 5'-C TGCCAAATTAGTAAATGAAATGAT-3’ in the 5’-untranslated region (UTR) of oca2. The reporter construct poca2-EGFP was made by fusing 5’-UTR and adjacent coding sequence to EGFP cds in frame and used to test the effectiveness of oca2-MO. Capped mRNAs were synthesized using T7 mMessage mMACHINE Kit (Ambion) and purified using RNeasy Mini Kit (QIAGEN) according to the manufacturer’s instructions. MO was injected alone or in combination with
mRNA into the yolk of zebrafish embryos at the one-cell stage. Pigment phenotypes were observed and imaged with stereomicroscope (Nikon SMZ 1500). Because higher doses (>150 pg) of oca2 mRNA impaired early embryonic development, we chose to combine a lower dose (100 pg) of mRNA with a lower dose of oca2-MO.

Targeted Gene Modification Experiment in Mouse
Using a CRISPR design tool (http://crispr.mit.edu, last accessed July 10, 2015), we selected the sgRNA target sequence (TAAAGCAGACGATTTCAGAC) for murine Oca2 to direct Cas9 cleavage. The murine Oca2 sgRNA sequence was cloned into pDR274 (a sgRNA cloning vector from Addgene, Cambridge MA) to form a T7 promoter-mediated sgRNA expression vector (Hwang et al. 2013). With Dral digestion, the linearized expression vector was purified using a QIAquick Gel Purification Kit (QIAGEN) and used as a DNA template to produce sgRNA by a MAXiScript T7 kit (Life Technologies). For Cas9 mRNA production, T7 promoter-contained px330 vector (Cong et al. 2013) was digested with NotI and then purified using a QIAquick Gel Purification Kit (QIAGEN). Subsequently, the linearized and purified vector was used as a DNA template to synthesize Cas9 mRNA using a mMESSAGE mMACHINE T7 Ultra Kit (Life Technologies). Cas9 mRNA and sgRNA were purified using RNA Clean & Concentrator-25 (ZYMO Research) and dissolved in RNase-free water. RNA concentrations were measured using a NanoDrop ND1000. Finally, 2 μl each of sgRNA and Cas9 mRNA were mixed with an equal volume of formamide, re-sequenced to obtain thin slices and viewed under a JEOL JEM-1101 transmission electron microscope at room temperature. Refer to the published method of Varani et al. (2001) for details.

Supplementary Material
Supplementary material is available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments
The authors are grateful to all the voluntary donors in this study. Special thanks to persons who agreed to provide their portrait photos for publication in this paper. The authors thank Prof. Yongtang Zheng for his help in collecting blood samples. They also thank Xiaofeng Ma, Shuangjuan Yang, Wenting Wan and Xiaoli Li for their technical assistance. This study was supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13010000); the National Natural Science Foundation of China (91131001, 91231203, 31321002, 31123005, 31371268, and 31371269); the Natural Science Foundation of Yunnan Province (2010CI044); State Key Laboratory of Genetic Resources and Evolution grant (GREGF15-06); and the Personnel Training Project of Yunnan Province (KKSY201526061).

References
1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, et al. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65.
Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16:37–48.
Beirl AJ, Linbo TH, Cobb MJ, Cooper CD. 2014. oca2 regulation of chromatophore differentiation and number is cell type specific in zebrafish. Pigment Cell Res. 27:178–189.
Beleza S, Johnson NA, Candille SI, Absher DM, Coram MA, Lopes J, Campos J, Araujo IL, Anderson TM, Vilhjalmsson BJ, et al. 2013. Genetic architecture of skin and eye color in an African-European admixed population. PLoS Genet. 9:e1003372.
Chaplin G. 2004. Geographic distribution of environmental factors influencing human skin coloration. Am J Phys Anthropol. 125:292–302.
Chen H, Hey J, Slate M. 2015. A hidden Markov model for investigating recent positive selection through haplotype structure. Theor Popul Biol. 99:18–30.
Chen H, Slate M. 2013. Inferring selection intensity and allele age from multilocus haplotype structure. G3 3:1429–1442.
Chen K, Manga P, Orlow SJ. 2002. Pink-eyed dilution protein controls the processing of tyrosinase. Mol Biol Cell. 13:1953–1964.

Quantitative measurements of hair and skin pigmentation were conducted for the four mouse strains. Adult mice (~5 months) were measured (4 WT, 4 WT', 6 TM, and 6 KO) with a CR-400 tristimulus colorimeter. Hair pigmentation was measured in three different regions of the back (left, middle, and right), and we also obtained skin pigmentation measures of the tail (supplementary fig. S9, Supplementary Material online).

Transmission electron microscopic analysis was performed. Mouse skin tissues were prepared by conventional methods: fixations, staining, and dehydration, embedding in resin and sectioned to obtain thin slices and viewed under a JEOL JEM-1101 transmission electron microscope at room temperature. Refer to the published method of Varani et al. (2001) for details.
Singh P, Schimenti JC, Bolcun-Filas E. 2015. A mouse geneticist’s practical guide to CRISPR applications. Genetics 199:1–15.

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.

Staleva L, Manga P, Orlow SJ. 2002. Pink-eyed dilution protein modulates arsenic sensitivity and intracellular glutathione metabolism. Mol Biol Cell. 13:4206–4220.

Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J, et al. 1999. Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. Am J Hum Genet. 65:1718–1724.

Toyofuku K, Valencia JC, Kushimoto T, Costin GE, Virador VM, Vieira WD, Ferrans VJ, Hearing VJ. 2002. The etiology of oculocutaneous albinism (OCA) type II: the pink protein modulates the processing and transport of tyrosinase. Pigment Cell Res. 15:217–224.

Varani J, Spearman D, Perone P, Fligiel SE, Datta SC, Wang ZQ, Shao Y, Kang S, Fisher GJ, Voorhees JJ. 2001. Inhibition of type I procollagen synthesis by damaged collagen in photoaged skin and by collagenase-degraded collagen in vitro. Am J Pathol. 158:931–942.

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153:910–918.

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

Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, Unterlander 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 U S A. 111:4832–4837.

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 154:1370–1379.

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.

Zhong H, Chen Y, Li Y, Chen R, Mardon G. 2015. CRISPR-engineered mosaicism rapidly reveals that loss of Kcnj13 function in mice mimics human disease phenotypes. Sci Rep. 5:8366.