Heredity (2010) 105, 274–281

www.nature.com/hdy

ORIGINAL ARTICLE

Artificial selection of the melanocortin receptor 1 gene in Chinese domestic pigs during domestication

J Li[1,2], H Yang[2], J-r Li[3], H-p Li[3], T Ning[1,2], X-R Pan[1], P Shi[2] and Y-P Zhang[1,2]
[1]*Laboratory for Conservation and Utilization of Bio-resources & Key Laboratory for Microbial Resources of the Ministry of Education, Yunnan University, Kunming, China;* [2]*State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China and* [3]*Department of Computational Genomics, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China*

Black coat colour is common in Chinese indigenous domestic pigs, but not among their wild ancestors, and it is thus presumed to be a 'domestication trait.' To determine whether artificial interference contributes to morphological diversification, we examined nucleotide variation from 157 Chinese domestic pigs and 40 wild boars in the melanocortin receptor 1 (*MC1R*) gene, which has a key role in the coat pigmentation of *Sus scrofa*. Compared with a pseudogene *GPIP*, our results showed that the joint effects of demography and selection have resulted in markedly low genetic diversity of *MC1R* in Chinese domestic pigs. Coalescent simulation and selection tests further suggest that the fixation of two non-synonymous substitutions associated with black colour is the result of artificial selection. In contrast, a much higher genetic diversity and only a single non-synonymous substitution were found among the wild boars, suggesting a strong functional constraint. Moreover, our conclusion is consistent with the preference for black colour in the ancient Chinese sacrificial culture. This case provides an interesting example of a molecular evaluation of artificial livestock selection and its associated cultural impact in ancient China.
Heredity (2010) **105,** 274–281; doi:10.1038/hdy.2009.191; published online 24 February 2010

Keywords: coat colour; domestic animal; population genetics; evolution

Introduction

Domestication promotes abundant morphological polymorphisms in domesticated animals relative to their wild ancestors (Andersson, 2001; Diamond, 2002). An obvious example is coat colour variation. For wild animals, colouration has a very important role in their ecological and behavioural adaptation and offers at least three broad categories of function: concealment, communication and regulation of physiological processes (Stoner *et al*., 2003). In contrast, coat colours of domestic animals mainly represent human needs or cultural preferences and result from strong artificial selection and domestication bottlenecks (Innan and Kim, 2004). Domestication is reflected at the molecular level as a 'footprint of artificial selection and demography' at domestication target loci. Therefore, analyses and comparisons of DNA sequence diversity between wild ancestors and their domestic counterparts provide insight into the genetic basis of morphological variation.

The Chinese domestic pig provides an ideal model for studying the molecular mechanism of phenotypic variation because of several special features. Unlike most other domestic animals, the wild ancestors of domesticated pigs and a number of outgroup species are still present in the world, which is convenient for examining ancestral and derived mutations as well as inferring the processes of artificial selection (Chen *et al*., 2007). Moreover, the origin and history of Chinese domestic pigs have been deciphered. As one of the first domesticated animals in China, the presence of domesticated pigs can be traced back to ~8000 years (Yuan and Rowan, 2002). Our previous research on mitochondrial DNA showed that Chinese domestic pigs were separately domesticated from wild boars in the Mekong region and in the middle and downstream regions of the Yangtze River (Wu *et al*., 2007). Since the initial domestication, more than 48 indigenous breeds have been domesticated so far (Zhang, 1986). It is interesting that the black colour is found to be predominant in most of the Chinese indigenous pig breeds and can be used to clearly distinguish it from its wild progenitors (Geng and Liu, 2003; Shi *et al*., 2004, 2006; Fang *et al*., 2009).

The coat colour differences between Chinese domestic pigs and their wild progenitors led to the hypothesis that the allele dominating the black coat colour was possibly under artificial selection during the domestication of Chinese domestic pigs. This hypothesis can be tested by examining the selective pressure on coat colour gene(s) in Chinese domestic pigs and wild boars.

Correspondence: Professor Y-P Zhang or Professor P Shi, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, 32 Jiaochangdong ST, Yunnan, Kunming 650223, China.
E-mails: zhangyp@mail.kiz.ac.cn or ship@mail.kiz.ac.cn
Received 5 July 2009; revised 1 November 2009; accepted 21 December 2009; published online 24 February 2010

The melanocortin receptor 1 (MC1R) has already been identified as a major determinant of pigment phenotype (Lin and Fisher, 2007). MC1R has a vital role in melanogenesis, as melanocytes produce black/brown eumelanin with active MC1R and red/yellow pheomelanin without MC1R signalling (Robbins *et al*., 1993). Polymorphisms in MC1R have also been documented and associated with coat colour variance in many domestic species, such as dogs, chickens, cats and cattle (Klungland *et al*., 1995; Eizirik *et al*., 2003; Kerje *et al*., 2003; Anderson *et al*., 2009). In pigs, six *MC1R* alleles associated with different colour phenotypes have been reported (Giuffra *et al*., 2000; Kijas *et al*., 1998, 2001). This evidence strongly suggested that *MC1R* might be an ideal marker to test our hypothesis.

In recent research, Fang *et al*. (2009) compared the non-synonymous/synonymous substitution rate ratio (dN/dS) between Asian and European domestic pigs and suggested that positive selection acted on the *MC1R* locus in domestic pigs. However, evidence of artificial selection on Chinese domestic pigs remains incomplete. On the one hand, the alleles were described from a limited sample size of Chinese domestic pigs ($n=23$) and wild boars ($n=3$); therefore, the results should be interpreted with caution. A larger population genetic data set is still to be developed to detect the effects of demographic events during pig domestication, because domestication bottlenecks can lead to patterns similar to those of selective sweep. On the other hand, it is widely known that pig domestication occurred independently in Europe and China (Giuffra *et al*., 2000; Larson *et al*., 2005). The large dN/dS ratio (23.5) between European and Chinese domestic pigs could be due to selection operating on European domestic pigs, in which a total of six non-synonymous sites were found, twice that of Chinese domestic pigs. Indeed, the dN/dS ratio (0.25) of Chinese domestic pigs was significantly less than one, which cannot be taken as an evidence for positive selection. Therefore, to provide direct evidence of selection during domestication, it is necessary to focus on the comparison between Chinese domestic pigs and their wild ancestors in addition to a comparison between the two lineages that underwent independent domestication (that is, European domestic pigs and Chinese domestic pigs).

To get a comprehensive understanding of the selection pressure acting on coat colour during the domestication of pigs in China, we extended previous research to detect the genetic variations and degree of artificial selection in newly acquired 1552 bp sequences of *MC1R*. Our research included the entire *MC1R*-coding region and partial 5′ and 3′ flanking regions from 157 samples of Chinese domestic pigs and 40 samples of Chinese wild boars. Moreover, to detect a potential bottleneck effect, we sequenced a 707-bp glucose phosphate isomerase pseudogene (*GPIP*) from a subset of the *MC1R* samples. Our analysis of amino-acid change patterns, haplotype structure and nucleotide diversity, together with our bottleneck and hitchhiking model simulations and selection tests, provide clear evidence that strong artificial selection forces operated on the *MC1R* locus in Chinese domestic pigs. This artificial selection correlates with the black colour preference in ancient China during animal sacrifice and reflects the impact of culture on domestication.

Materials and methods

Sample collection

A list of samples is given in Supplementary Table 1. A total of 157 blood/tissue samples from 34 Chinese domestic breeds were collected from 16 different provinces of China. To avoid collection of related samples, we only considered pigs with clear pedigrees. In addition, a total of 40 Chinese wild boar tissue samples from 10 different provinces of China were collected.

PCR amplification, cloning and sequencing

Total genomic DNA was extracted from blood/tissue according to a standard phenol–chloroform extraction protocol. The primers MF1 (5′-GTGCGGCGGCTCTGCG CTCCAA-3′) and MR1 (5′-CCCCCACTCCCCATGCCTC CTG-3′) were used to amplify a 1552-bp sequence of the *MC1R* gene, including 425 bp of the 5′ untranslated region (5′ UTR), 963 bp of the coding region and 164 bp of the 3′ UTR. Each PCR contained a total volume of 50 μl, consisting of 50 ng of template DNA, 3 μl of 10 × PCR Buffer (Mg^{2+} Plus), 20 pM of each primer, 200 μM of each dNTP and 1 U of rTaq polymerase (Takara, Dalian, China). Reaction profiles included a 3-min initial denaturation step at 95 °C, followed by 35 cycles of the following: 1 min denaturation at 94 °C, 1 min annealing at 64 °C, 1 min of extension at 72 °C and a final extension of 10 min at 72 °C.

The primers GF (5′-TGCAGTTGAGAAGGACTTTAC TT-3′) and GR (5′-GTATCCCAGATGATGTCATGAAT-3′) were used to amplify a 707-bp fragment of *GPIP* (Giuffra *et al*., 2000). PCR was performed with the same component concentrations as described above, but amplification was performed in different PCR conditions, which included a 3-min denaturation step at 95 °C followed by 35 cycles of the following: 1 min denaturation at 94 °C, 1 min annealing at 58 °C, 1 min of extension at 72 °C and a final 10 min extension at 72 °C.

The amplified products were purified with a gel extraction kit (Watson Biomedical Inc., Shanghai, China) and sequenced with an Applied Biosystems (Foster City, CA, USA) ABI 3730 sequencer using the primers listed qin Supplementary Table 2. To have a quality check for possible sequencing errors, ambiguous sequences and rare variants were run for confirmation using a second set of independent PCR products. To obtain two different alleles in the samples with more than one heterogeneous site, independent PCR products were cloned into vector pMD18-T using a Takara ligation kit (Takara). For each sample, at least five independent clones were sequenced. All sequences in this study have been evaluated by bi-directional sequencing and deposited as haplotypes in GenBank (Accession nos. FJ665467–FJ665499, GQ900667–GQ900673).

Data analysis

The sequences of *MC1R* and *GPIP* were aligned and edited by DNAstar Software (DNAstar Inc. Madison, WI, USA). Aligned sequence data were imported into MEGA3 (Kumar *et al*., 2004) for analysing nucleotide composition and variable sites. A median-joining network (Bandelt *et al*., 1999) was also constructed using program Network4.2.01 (http://www.fluxus-technology.com) and modified manually to reveal relationships among haplotypes.

Watterson's theta estimator (θ_W), nucleotide diversity (π), haplotype diversity and population genetic analyses were performed using DNASP4.0 software (Rozas *et al*., 2003). Tajima's *D* test was performed using coalescent simulation under the assumption of no recombination across genes, which makes the test more conservative (Tajima, 1989a, b). In the Hudson–Kreitman–Aguade (HKA) test (Hudson *et al*., 1987), a Red River Hog (*Potamochoerus porcus*) from the same family Suidae was used as an outgroup to calculate divergence.

To further analyse data under different scenarios, we assume that there is no recombination within loci. Simulations under neutrality and the hitchhiking model of selective sweeps were done according to the procedures previously described (Hudson, 1990; Kim and Stephan, 2002; Li and Stephan, 2006). Positive selection was assumed to be directional with co-dominant alleles.

Results and discussion

Sequence variation

We first surveyed sequence variation in a 963-bp region that covers the entire *MC1R*-coding region in 157 Chinese domestic pigs (representing 34 Chinese domestic breeds) and 40 Chinese wild boars. In the 80 chromosomes of our Chinese wild boar samples, 17 variable nucleotide positions were identified, which represent 16 synonymous and one non-synonymous substitution (G364A). However, to our surprise, only three variant sites were found in the 314 Chinese domestic pig chromosomes, indicating far fewer polymorphisms in domestic pigs compared with wild boars. Interestingly, two amino-acid replacements (G283A, T305C) that cause changes at V95M and L102P were fixed in all domestic pig samples but absent from all wild boars, making it easy to distinguish between Chinese domestic pigs and wild boars. To further trace the evolutionary history of these amino-acid changes (V95M and L102P), we sequenced the *MC1R*-coding region from a Red River Hog and found that the amino-acid sequence of the outgroup was similar to that of the wild boars despite the presence of many synonymous substitutions. Thus, this raised the possibility that the two derived amino acids in Chinese domestic pigs may have been under selection during domestication.

Haplotype structure

To elucidate the generation of variations in *MC1R*, we constructed haplotypes on the basis of the above genotypes using PHASE software (Stephens *et al*., 2001; Stephens and Scheet, 2005). We then confirmed all these haplotypes by cloning and sequencing heterozygous samples (Figure 1). Although four haplotypes, *MC1R**2 (Giuffra *et al*., 2000; Kijas *et al*., 1998, 2001) and *MC1R**21–*23, were observed in the Chinese domestic pig samples, *MC1R**2 accounted for 88.85% of the sampled chromosomes. The other haplotypes had very low frequency, ranging from 0.96 to 8.9%. In contrast, 16 unique haplotypes were observed in the Chinese wild boar samples, including two previously described haplotypes, *MC1R**1 and *MC1R**5 (Giuffra *et al*., 2000; Kijas *et al*., 1998, 2001). The frequency of haplotypes in Chinese wild boars ranged from 1.25 to 13.75%, indicating that many rare haplotypes were present in Chinese wild boars and

Haplotypes	Nucleotide positions																					Amino acid mutatio			Haplotype distribution	
	12	51	132	168	240	283	305	351	363	364	372	441	504	585	597	621	630	729	745	855	903	95	102	122	CWB	CDP
MC1R*5	C	G	C	C	C	G	T	G	C	G	C	G	C	C	C	G	C	G	C	C	C	V	L	V	11	
MC1R*1[a]	T	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5	
MC1R*7	.	.	.	.	T	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	2	
MC1R*8	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	10	
MC1R*9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	2	
MC1R*10	.	.	.	T	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	
MC1R*11[b]	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	8	
MC1R*12[c]	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	4	
MC1R*13	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	4	
MC1R*14	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	3	
MC1R*15	.	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	2	
MC1R*16	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	6	
MC1R*17[d]	T	A	.	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	.	.	.	.	8	
MC1R*18	T	A	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	4	
MC1R*19	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	I	8	
MC1R*20	T	A	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	I	2	
MC1R*2[e]	T	A	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	M	P	.		279
MC1R*21	T	A	.	.	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	M	P	.		28
MC1R*22	T	A	.	.	.	A	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	M	P	.		4
MC1R*23	T	A	.	.	.	A	C	.	.	A	.	.	.	.	.	.	.	A	.	.	.	M	P	I		3

Figure 1 Nucleotide mutations in the melanocortin receptor 1 (*MC1R*)-coding region defining 20 haplotypes and the haplotype distribution frequency. The codon and nucleotide positions are numbered from the first nucleotide of the start codon. Dots indicate that the identity of *MC1R**5. *MC1R**1, *2, *5 correspond to those reported previously (Kijas *et al*., 1998; Giuffra *et al*., 2000). *MC1R**3, *4, *6 were not found in this study. [a]*MC1R**25, [b]26–28, [c]29, [d]30 and [e]31 share the same haplotype in the coding region with *MC1R**1, 11, 12, 17 and 2, respectively.

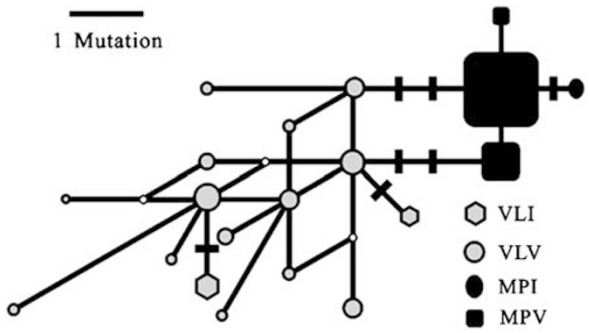

Figure 2 Median-joining networks of melanocortin receptor 1 (MC1R) haplotypes in Chinese wild boars and domestic pigs. The size of each node was proportional to the haplotype frequency in our data. Black nodes represent haplotypes from Chinese domestic pigs, whereas grey ones indicate haplotypes from Chinese wild boars. The lengths of lines show steps of synonymous mutation. Non-synonymous replacement mutations are demonstrated with thick short black lines. The VLI, VLV, MPV and MPI amino-acid haplotypes are indicated by pentagons, circles, ellipses and squares, respectively. The open circles represent predicted haplotypes that have not yet been found.

possibly reflecting recent population growth as described by Wu *et al*. (2007). In the Chinese wild boar samples, ~87.5% of the haplotypes differed only at synonymous sites, which possibly suggests that a strong functional constraint was acting on wild boars to maintain a camouflaging coat colour for survival enhancement in the wild. Moreover, the median-joining networks analysis also revealed a strikingly different genealogy pattern between Chinese wild boars and domestic pigs. Figure 2 depicts star-like topologies of haplotypes in Chinese domestic pigs that represent one common haplotype and several rare haplotypes, whereas the Chinese wild boar haplotypes exhibited several divergent clusters. Because different topologies are often associated with different population histories, different selective pressures or both, the star-like topology of Chinese domestic pigs might be due to a recent population expansion or positive selection that should strongly reduce genetic diversity at domestication target genes.

Genetic diversity

To investigate whether low genetic diversity existed in Chinese domestic pigs, we calculated the nucleotide diversity and found that it was 0.00022 in Chinese domestic pigs. This value is lower than other gene loci described to date in domestic pigs (Figure 3). For instance, the nucleotide diversity value π of the *IGF2* gene in Chinese domestic pigs was about 0.0029 (Ojeda *et al*., 2008b), which is 13-fold larger than that of the *MC1R* gene described here and approximately 11-fold lower than that of Chinese wild boars (0.00276). A similar finding in the *FABP5* gene has been previously described in Japanese and European wild boars (0.0023) (Ojeda *et al*., 2008a). In addition to the π value, Watterson's estimator (θ_w) was calculated from the observed number of polymorphic sites in samples, and this also suggested that the level of diversity in the *MC1R*-coding region was low in Chinese domestic pigs ($\theta_w = 0.00049$), comprising 14% of the diversity found in Chinese wild boars ($\theta_w = 0.00356$). Together, our findings show low levels of *MC1R*-coding region diversity in Chinese domestic pigs, which is in contrast to previous observations of higher levels of genetic diversity in domesticated pigs (Wu *et al*., 2007; Ojeda *et al*., 2006, 2008a, b). Our findings are consistent, however, with the logical theory of domestication events in that genetic diversity is generally thought to become lower after domestication. Lower genetic diversity could be caused by population dynamics or artificial selection, which evidently happens in several crops (for example, sorghum, rice and maize) (Wang *et al*., 1999; Hamblin *et al*., 2006; Olsen *et al*., 2006). By contrast, previous studies reported unexpectedly high sequence variation in domesticated pigs even when considering the influences of target genes on economically important traits and the presence of strong artificial selection (Ojeda *et al*., 2006, 2008a, b). More recently, Ramírez *et al*. (2009) conducted a joint analysis of mitochondrial microsatellite and Y-chromosome polymorphisms, and showed that levels of genetic variation are similar in pigs and wild boars. Therefore, on the basis of currently published reports and our current findings, *MC1R* is a rare case in which the nucleotide diversity is much lower in Chinese domestic pigs than it is in their wild ancestors.

Bottleneck and hitchhiking model simulation

The observation of the low *MC1R*-coding region diversity in Chinese domestic pigs could be explained by the bottleneck effect. To verify whether the bottleneck effect had any role in the low nucleotide diversity of our samples, we further sequenced a 707-bp fragment of *GPIP*, which is known as a pseudogene (Harbitz *et al*., 1993) and is considered a neutral marker (Giuffra *et al*., 2000) in the same wild boars ($n = 36$) and domestic pig ($n = 132$) samples selected for *MC1R* typing. We detected 12 and 10 variant sites in Chinese domestic pigs and wild boars, respectively (Supplementary Table 3). The θ_w estimation showed no statistically significant difference between Chinese domestic pigs (0.00277) and wild boars (0.00292) for this locus, although the value in Chinese domestic pigs was slightly lower than that in wild boars. In addition, the ratio of π between Chinese domestic pigs (0.00135) and wild boars (0.00237) for *GPIP* was 0.57, which is 7.15-fold higher than that of the *MC1R* locus. These results suggest that the low diversity of *MC1R* found in Chinese domestic pigs is not a result of sampling bias but may be the consequence of the bottleneck effect. Although selection would not affect the neutral pseudogene, the bottleneck effect has been shown to decrease the nucleotide diversity of the entire genome in various domestic species (Wright *et al*., 2005).

To further test the possibility of artificial selection on *MC1R*, we first estimated the severity of the bottleneck due to domestication using the DNA polymorphism data of *GPIP*, which is assumed to evolve neutrally. We then examined whether the bottleneck scenario can explain the reduced genetic diversity of *MC1R* in the domesticated pig population. Finally, we estimated the minimum strength of artificial selection on *MC1R*.

For *GPIP* in the wild boar population, $\pi = 0.00237$ per site. Under the assumption that the *Sus scrofa–P. porcus* divergence time is 10 million years (Randi *et al*., 1996), nine synonymous mutations contribute to a total sequence divergence of 0.0127 (9/705), we estimate that

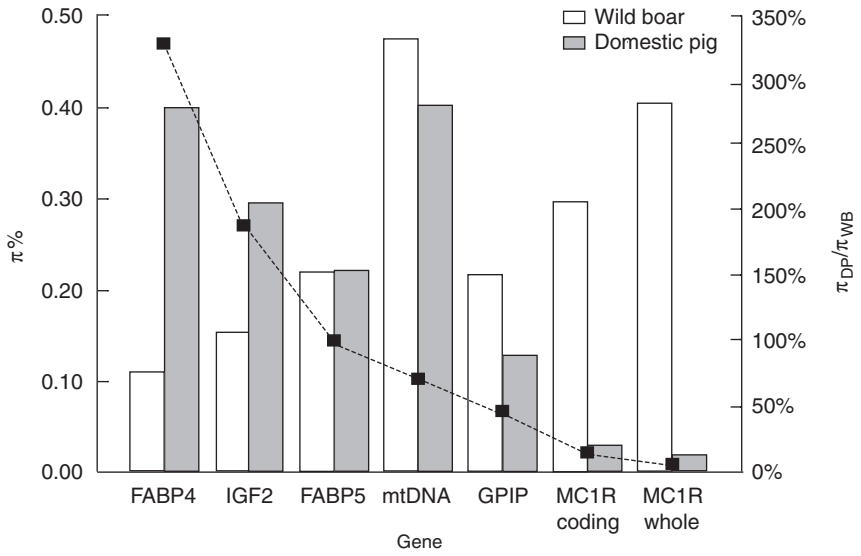

Figure 3 Nucleotide diversity (π%) in wild boars and domestic pigs. Each pair of bars represents a different locus/region. Black rectangles connected by broken lines show the ratio of π between domestic pigs and wild boars. 'MC1R coding' denotes the entire 963 bp coding region. 'MC1R whole' denotes the 1552 bp region, including the 5′ untranslated region (5′ UTR), the coding region and 3′ UTR. The glucose phosphate isomerase pseudogene (GPIP) and melanocortin receptor 1 (MC1R) data are from this study, and the other data are from previous reports: FABP4 (Ojeda *et al*., 2006), mitochondrial DNA (Wu *et al*., 2007), FABP5 (Ojeda *et al*., 2008a) and IGF2 (Ojeda *et al*., 2008b).

GPIP mutation rate is 1.27×10^{-9} per site per year. Assuming 3 years per generation for wild pigs, we estimate that the effective population size of the wild pig population is 1.56×10^5. On the basis of the *MC1R* gene, the estimated effective population size is the same (1.57×10^5). We modelled domestication by using an instantaneous bottleneck. We assumed that the effective population size of the ancestral wild pig population is N_0 and that a small number of individuals sampled from the ancestral population founded the domesticated pig population, which has an effective size of N_1. We also assume that the population size of the domesticated pig population increased instantaneously to N_0 after a certain time period (t_1, in units of $2N_0$ generations). Domestication occurred about 8000 generations ago (we presume 1 year per generation for domesticated pigs); therefore, the unknown parameter in the bottleneck model is the severity of the bottleneck ($f = t_1 N_0 / N_1$).

As more genetic diversity will be reduced due to domestication when f increases, f can be estimated by minimizing the difference between the expected and observed genetic diversity (π) of *GPIP* in the domesticated pig population. Using coalescent simulations, the estimated severity of the bottleneck is 0.585 (the expected value of π is estimated from the 1.0×10^5 simulated data). Given the severity of the bottleneck, the population parameter of the *MC1R* gene of the ancestral wild pig population ($\theta_w = 6.0259$) and the sample size ($n = 314$), the expected π of the *MC1R* gene in the domesticated pig population is 3.440, which is much higher than the observed value. Moreover, the probability that the expected π is less than the observed π (0.2135) in the domesticated pig population is 0.044. Therefore, the bottleneck alone cannot explain the lack of genetic diversity in the *MC1R* gene of the domesticated pig population.

As the neutral bottleneck scenario is rejected for the *MC1R* gene, we considered the hitchhiking model to interpret the lack of *MC1R* genetic diversity in the domesticated pig population. We assumed a neutral locus (that is, *MC1R* gene) that is tightly linked to a selected site. According to the model, all observed segregating sites on the *MC1R* gene of the domesticated pig population should arise after domestication. Moreover, less genetic diversity is expected when selection strength is stronger. Therefore, the selection coefficient(s) for the domestication can be estimated by minimizing the difference of the expected and observed genetic diversity (π) of *MC1R* in the domesticated pig population. Using simulations, we found that the estimated selection coefficient(s) is 0.0012 (the expected π is 0.2146).

Test for selection

The possible influence of artificial selection on the *MC1R* locus in Chinese domestic pigs was further examined by the Tajima's *D* test, which compares the differences between π and θ_w. Under neutral mutation-drift equilibrium, these two values would be equal and Tajima's *D* would be close to zero (Tajima, 1989b). In our analysis, *D* values were negative in both the Chinese domestic pigs and wild boars, although all of them insignificantly deviated from zero (Table 1). The insignificant *D* value in Chinese domestic pigs is likely due to the small number of segregating sites, which would have weakened the Tajima's *D* test power (Simonsen *et al*., 1995). Under such circumstances, the selection of wider gene-flanking regions may provide useful information. We sequenced 425 bp in the 5′ UTR and 164 bp in the 3′ UTR of the *MC1R* gene in all the Chinese domestic and wild samples considered above. For comparison, we also sequenced the same region from the Red River Hog. Interestingly, the π value of this flanking region in Chinese wild boars was 0.00571, which is about twofold greater than that of the coding region. However, in Chinese domestic pigs, the π value was significantly smaller in the flanking

Table 1 Polymorphism statistics and neutral tests in Chinese wild boars and domestic pigs

Population	*Gene*	*Base pairs*	N	*S*	π%	θ_w%	*Tajima's* D	*HKA test*[a]
Chinese wild boars	*MC1R* whole	1552	80	32	0.389	0.417	−0.212	1.28
	MC1R coding	963	80	17	0.276	0.356	−0.661	0.019
	GPIP	707	72	10	0.237	0.292	−0.506	
Chinese domestic pigs	*MC1R* whole	1552	314	4	0.014	0.041	−1.109*	20.367***
	MC1R coding	963	314	3	0.022	0.049	−0.848	4.762*
	GPIP	707	264	12	0.135	0.277	−1.207	

Abbreviations: HKA test, Hudson–Kreitman–Aguade test; *MC1R*, melanocortin receptor 1; *N*, Number of sequence; S, number of polymorphic sites.
MC1R whole represents the 1552 bp region including the 5′ untranslated region (5′ UTR), the coding region and the 3′UTR.
[a]*Potamochoerus porcus* was used as the outgroup. *$0.01 < P < 0.05$; ***$P < 0.001$.

region (0.00001) than it was in the coding region (0.00022). This result is in agreement with the selective sweep hypothesis that directional selection may have acted on this region by selectively sweeping old variants out of Chinese domestic pig populations. In addition, Tajima's *D* test in Chinese domestic pigs was marginally significant ($P = 0.04$) when both flanking regions and coding regions were considered.

Alternatively, a negative *D* value may result from population growth. To evaluate these alternatives, we used the HKA neutrality test to compare the polymorphisms and divergence of the sequenced *MC1R* gene with that of the neutral *GPIP* pseudogene. We first estimated the interspecific sequence variation between the Red River Hog and Chinese wild boars; then, we compared the interspecific variation with Chinese wild boars' intraspecific polymorphisms. The result showed no significant deviation from neutrality when the *GPIP* pseudogene was tested (χ^2-test, $P = 0.258$). In contrast, the HKA test yielded a highly significant result when the *MC1R* region was compared with *GPIP* in Chinese domestic pigs ($P = 0.029$). Mutation rate variation among loci would not result in significant HKA test results (Tajima, 1989a). Population dynamics in Chinese domestic pigs during domestication also cannot explain these results. The influence of population growth would equally affect all loci, whereas selection would affect only the target regions (Hudson *et al*., 1987). Our results suggest that directional selection may act on this region by selectively sweeping old variants out of Chinese domestic pigs.

If the lower than expected nucleotide diversity of the *MC1R* region in Chinese domestic pigs (as suggested by the HKA test and *D* statistics) was indeed a result of selective sweep, then we would expect that at least one of the two amino-acid changes (V95M, L102P) found in Chinese domestic pigs would be advantageous. To test this hypothesis, we compared the ratio of two non-synonymous substitutions with zero synonymous substitutions that had fixed between Chinese domestic pigs and wild boars with a ratio of one non-synonymous to 16 silent polymorphisms. This comparison showed a significant excess of fixed non-synonymous substitutions (Fisher's exact test, $P = 0.0175$) during domestication and supports our hypothesis that the non-synonymous substitutions were fixed by positive selection. Indeed, evidence from other species strongly suggests that the change from a leucine residue at codon 102 to a proline (L102P) can only result in black pigment (Kijas *et al*., 1998). For example, the L102P mutation had been identified and associated with a dominant black colour trait in cattle (Klungland *et al*., 1995). In mice, a functional analysis showed that the *MC1R* receptor with the mutated amino acid (proline) can be constitutively activated and related to the sombre phenotype (Robbins *et al*., 1993). It is interesting to note that although the majority (if not all) of Chinese wild boars had wild-type coat colour and the leucine residue at this codon, all the Chinese domestic pigs with dominant black coat colour possessed this amino-acid change (Figure 1). Therefore, the key amino acid involving in coat colour changes was under selection during the domestication.

The above observation raised the question of when this amino-acid change occurred in evolution. However, it is difficult to ascertain the exact time that the mutation appeared without ancient DNA data. Nevertheless, we could still speculate on some possible scenarios based on our analyses and the knowledge of the pig domestication in China. One scenario is that it changed during domestication, as shown in European domestic pigs (Fang *et al*., 2009). Another scenario is that the amino-acid change occurred in the wild ancestors, as was recently shown in horses (Ludwig *et al*., 2009). As there were multiple independent domestication events in China (Wu *et al*., 2007), it is unlikely that the two amino-acid changes (V95M, L102P) occurred in each course of domestication. However, the absence of this amino acid in our sequence survey of Chinese wild boars suggests that the amino-acid frequency in Chinese wild boars would be very low if it indeed existed. Therefore, a composite scenario with an initial occurrence in a wild ancestor of domesticated pigs, followed by artificial selection in domesticated pigs, is more reasonable. This hypothesis awaits further verification from the analysis of ancient DNA data.

The reason why the coat colour change between Chinese domestic pigs and their wild ancestors is still unknown. One possible explanation would be that a small number of wild individuals with black coat colour sampled from the ancestral population founded the domesticated pig population during domestication. However, it is most unlikely because our simulation analysis shows demographic events alone have less possibility to result in the observed nucleotide diversity in Chinese domestic pigs. Moreover, our further selection tests also support that the amino acids related to the coat colour changes were under selection. Another possible explanation is that the artificial selection of Chinese black-coloured pigs is likely reflecting the Chinese sacrificial culture. Archaeological studies indicate that

pigs were first domesticated ~8000 years ago in China (Yuan and Rowan, 2002). In ancient China, domesticated pigs initially served as the major protein source and gradually became the most common animal sacrifice in the Neolithic era of China (Wang, 1981, 1996; Kim, 1994; Okamura, 2002; Yuan and Rowan, 2005). In rituals, the uniformly coloured black pigs were especially preferred because they were thought to be sacred, clean and a representation of faithful respect to the holy gods and/or ancestors (Qu, 1998). Sacrificial offerings in China are believed to have emerged ~7000 years ago (He and Cheng, 2007), and this was immediately followed by the domestication of pigs. It would be interesting to test this hypothesis when more background information of Chinese domestic pigs and the ancient DNA data are available.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank three anonymous referees, Shan-yuan Chen, Guo-Dong Wang and Monojit Debnath, for valuable comments on the paper, and Shi-Fang Wu, Hon-Jun Han, Hua-Wei Wang, Rui Du, Ruo-Xi Fan and Tao Sha for laboratory assistance. The work was supported by grants of the National Basic Research Program of China (973 Program, 2007CB815700, and 2006CB102100), Chinese Academy of Sciences (KSCX2-YW-N-018), Bureau of Science and Technology of Yunnan Province, National Natural Science Foundation of China (30621092, U0836602) and the National 863 Project of China (2008AA101001) to Ya-Ping Zhang, and by a start-up fund of 'Hundreds-Talent Program' from Chinese Academy of Sciences Genetically modified organisms breeding major projects (2009ZX08009-149B) to Peng Shi.

References

Andersson L (2001). Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet* **2**: 130–138.

Anderson TM, VonHoldt BM, Candille SI, Musiani M, Greco C, Stahler DR *et al.* (2009). Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323**: 1339–1343.

Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**: 37–48.

Chen KF, Baxter T, Muir WM, Groenen MA, Schook LB (2007). Genetic resources, genome mapping and evolutionary genomics of the pig (Sus Scrofa). *Int J Biol Sci* **3**: 153–165.

Diamond J (2002). Evolution, consequences and future of plant and animal domestication. *Nature* **418**: 700–707.

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ (2003). Molecular genetics and evolution of melanism in the cat family. *Curr Biol* **13**: 448–453.

Fang MY, Larson G, Ribeiro HS, Li N, Andersson L (2009). Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genet* **5**: e1000341.

Geng SM, Liu XL (2003). Compendium of animal breed resources in China. Chinese Agriculture Press: Beijing. (in Chinese).

Giuffra E, Kijas JM, Amarger V, Carlborg Ö, Jeon JT, Andersson L (2000). The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* **154**: 1785–1791.

Hamblin MT, Casa AM, Sun H, Murray SC, Paterson AH, Aquadro CF *et al.* (2006). Challenges of detecting directional selection after a bottleneck: lessons from Sorghum bicolor. *Genetics* **173**: 953–964.

Harbitz I, Chowdhary BP, Kran S, Davies W (1993). Characterization of a porcine glucosephosphate isomerase-processed pseudogene at chromosome 1q1.6-1.7. *Mamm Genome* **4**: 589–592.

He G, Cheng LW (2007). Propagation of Gaomiao culture and its influence. *Relics From South* **2**: 51–60. (in Chinese).

Hudson RR (1990). Gene genealogies and the coalescent process. In: Harvey, PH and Artridge, LP (eds). *Oxford Surveys in Evolutionary Biology* Vol 7. Oxford University Press: New York. pp 1–44.

Hudson RR, Kreitman M, Aguadé M (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.

Innan H, Kim Y (2004). Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci USA* **101**: 10667–10672.

Kerje S, Lind J, Schutz K, Jensen P, Andersson L (2003). Melanocortin 1-receptor (*MC1R*) mutations are associated with plumage colour in chicken. *Anim Genet* **34**: 241–248.

Kijas JM, Moller M, Plastow G, Andersson L (2001). A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. *Genetics* **158**: 779–785.

Kijas JM, Wales R, Törnsten A, Chardon P, Moller M, Andersson L (1998). Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics* **150**: 1177–1185.

Kim SO (1994). Burials, pigs, and political prestige in Neolithic China. *Curr Anthropol* **35**: 119–141.

Kim Y, Stephan W (2002). Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* **160**: 765–777.

Klungland H, Vage DI, Gomez-Raya L, Adalsteinsson S, Lien S (1995). The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm Genome* **6**: 636–639.

Kumar S, Tamura K, Nei M (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**: 150–163.

Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, Robins J *et al.* (2005). Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* **307**: 1618–1621.

Li H, Stephan W (2006). Inferring the demographic history and rate of adaptive substitution in Drosophila. *PLoS Genet* **2**: e166.

Lin JY, Fisher DE (2007). Melanocyte biology and skin pigmentation. *Nature* **445**: 843–850.

Ludwig A, Pruvost M, Reissmann M, Benecke N, Brockmann GA, Castanos P *et al.* (2009). Coat color variation at the beginning of horse domestication. *Science* **324**: 485.

Ojeda A, Estellé J, Folch JM, Pérez-Enciso M (2008a). Nucleotide variability and linkage disequilibrium patterns at the porcine FABP5 gene. *Anim Genet* **39**: 468–473.

Ojeda A, Huang LS, Ren J, Angiolillo A, Cho IC, Soto H *et al.* (2008b). Selection in the making: a worldwide survey of haplotypic diversity around a causative mutation in porcine IGF2. *Genetics* **178**: 1639–1652.

Ojeda A, Rozas J, Folch JM, Perez-Enciso M (2006). Unexpected high polymorphism at the FABP4 gene unveils a complex history for pig populations. *Genetics* **174**: 2119–2127.

Okamura H (2002). Animal sacrifice in burials of ancient China. *Tôhô gappô* **74**: 1–182. (in Japanese).

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD (2006). Selection under domestication: evidence for a sweep in the rice waxy genomic region. *Genetics* **173**: 975–983.

Qu MA (1998). Slightly theory about Chinese antiquity religion sacrifice and diet culture. *Chinese History Research* **3**: 152–162. (in Chinese).

Ramírez O, Ojeda A, Tomàs A, Gallardo D, Huang LS, Folch JM *et al*. (2009). Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig breeds. *Mol Biol Evol* **26**: 2061–2072.

Randi E, Lucchini V, Diong CH (1996). Evolutionary genetics of the suiformes as rereconstructed using mtDNA sequencing. *J Mamm Evol* **3**: 163–194.

Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-Rehfuss L, Baack E *et al*. (1993). Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* **72**: 827–834.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Shi KR, Wang AG, Li N, Deng XM (2004). Single nucleotide polymorphism analysis on melanocortin receptor 1 (*MC1R*) of Chinese native pig. *Sci China C Life Sci* **47**: 287–292.

Shi KR, Wang AG, Yuan XF, Deng XM, Li N (2006). Analysis of the MC1R, KIT and ASIP loci in Chinese and European pigs. *Anim Genet* **37**: 300–302.

Simonsen KL, Churchill GA, Aquadro CF (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**: 413–429.

Stephens M, Scheet P (2005). Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* **76**: 449–462.

Stephens M, Smith NJ, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**: 978–989.

Stoner CJ, Caro TM, Graham CM (2003). Ecological and behavioral correlates of coloration in artiodactyls: systematic analyses of conventional hypotheses. *Behav Ecol* **14**: 823–840.

Tajima F (1989a). The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597–601.

Tajima F (1989b). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Wang JH (1996). Preliminary analysis of the burial of domesticated animals in prehistoric contexts. *Huaxia Archaeol* **1**: 24–36. (in Chinese).

Wang RL, Stec A, Hey J, Lukens L, Doebley J (1999). The limits of selection during maize domestication. *Nature* **398**: 236–239.

Wang RX (1981). The religious significance of the Neolithic practice of interring pig carcasses. *Cultural Relics* **2**: 71–85. (in Chinese).

Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD *et al*. (2005). The effects of artificial selection on the maize genome. *Science* **308**: 1310–1314.

Wu GS, Yao YG, Qu KX, Ding ZL, Li H, Palanichamy MG *et al*. (2007). Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome Biol* **8**: R245.

Yuan J, Rowan F (2002). Pig domestication in ancient China. *Antiquity* **76**: 724–732.

Yuan J, Rowan F (2005). New zooarchaeological evidence for changes in Shang Dynasty animal sacrifice. *J Anthropol Archaeol* **24**: 252–270.

Zhang ZG (1986). Pig breeds in China. Shanghai Scientific & Technical Publishers: Shanghai. (in Chinese).

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)