



Y-Single Nucleotide Polymorphisms Diversity in Chinese Indigenous Horse

Haoyuan Han, Qin Zhang¹, Kexin Gao, Xiangpeng Yue, Tao Zhang², Ruihua Dang,
Xianyong Lan, Hong Chen, and Chuzhao Lei*

College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China

ABSTRACT: In contrast to high genetic diversity of mitochondrial DNA (mtDNA), equine Y chromosome shows extremely low variability, implying limited patriline in the domesticated horse. In this study, we applied direct sequencing and restriction fragment length polymorphism (RFLP) methods to investigate the polymorphisms of 33 Y chromosome specific loci in 304 Chinese indigenous horses from 13 breeds. Consequently, two Y-single nucleotide polymorphisms (SNPs) (Y-45701/997 and Y-50869) and one Y-indel (Y-45288) were identified. Of those, the Y-50869 (T>A) revealed the highest variation frequency (24.67%), whereas it was only 3.29% and 1.97% in Y-45288 (T/-) and Y-45701/997 (G>T) locus, respectively. These three mutations accounted for 27.96% of the total samples and identified five Y-SNP haplotypes, demonstrating genetic diversity of Y chromosome in Chinese horses. In addition, all the five Y-SNP haplotypes were shared by different breeds. Among 13 horse breeds analyzed, Balikun horse displayed the highest nucleotide diversity ($\pi = 5.6 \times 10^{-4}$) and haplotype diversity ($h = 0.527$), while Ningqiang horse showed the lowest nucleotide diversity ($\pi = 0.00000$) and haplotype diversity ($h = 0.000$). The results also revealed that Chinese horses had a different polymorphic pattern of Y chromosome from European and American horses. In conclusion, Chinese horses revealed genetic diversity of Y chromosome, however more efforts should be made to better understand the domestication and paternal origin of Chinese indigenous horses. (**Key Words:** Chinese Indigenous Horse, Y Chromosome, Single Nucleotide Polymorphisms, Diversity)

INTRODUCTION

As well known, horse is one of the earliest domesticated animals and has played a very significant role in the human civilization (Diamond, 2002; Gupta, 2004; Ludwig et al., 2009). However, when and where it has been domesticated still remains elusive. Archaeological data and coat color variations indicated that horse was probably first domesticated in Eurasian steppe around 5,000 years ago (Ludwig et al., 2009; Outram et al., 2009). In contrast, the mitochondrial DNA (mtDNA) analyses of ancient and modern domestic horses revealed extensive variations within and among breeds, with little congruence of

haplogroup distribution to breeds or geographic areas, suggesting that horse had multiple maternal origins (Vilá et al., 2001; Lei et al., 2009; Cieslak et al., 2010). Prezewalski's horse, the only alive wild horse, was ever deemed the progenitor of the current domestic horses during the process of investigating the origin of domestic horse (Deng et al., 2000; Zhang et al., 2004). Nevertheless, this standpoint was not supported by mtDNA evidence (Ishida et al., 1995; Oakenfull and Ryder, 1998; Oakenfull et al., 2000; Goto et al., 2011) and Y chromosome evidence because two fixed Y chromosome variations were detected in domesticated horses and Prezewalski's horses (Wallner et al., 2003). The maternal origins of domestic horse is so complicated that more researches were needed to investigate Y chromosome diversity and paternal origins of the domestic horse.

As for China, in which there are very rich genetic resources of domestic animals such as pig, yak, chicken and swamp buffalo (Guo et al., 2006; Liu et al., 2006; Lei et al., 2007; Larson et al., 2010). Furthermore, China is believed

* Corresponding Author: Chuzhao Lei. Tel: +86-29-87099384, Fax: +86-29-87092164, E-mail: leichuzhao1118@126.com

¹ College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, China.

² Shaanxi University of Technology, Hanzhong, Shaanxi 723001, China.

Submitted Oct. 8, 2014; Revised Dec. 16, 2014; Accepted Mar. 4, 2015

to be an important domestication center for domestic animals based on previous publications (Savolainen et al., 2002; Larson et al., 2005; Guo et al., 2006; Liu et al., 2006; Lei et al., 2007). There are 15 local domestic horse breeds and many populations throughout 14 provinces in Northwestern, Southwestern, Northeastern, and Central China (Xie, 1987). Previous studies have proven that China is also a potential domestication center for horse. The mtDNA sequencing analysis of Eastern and European horse populations revealed a higher incidence of haplogroup F and a relatively fewer occurrences of haplogroup D in Eastern populations, while European populations revealed an opposite pattern, suggesting that there has been a degree of genetic isolation and differentiation of Eastern and European horse populations (McGahern et al., 2006), therefore, haplogroup F may have been taken into the domestic gene pool in the Far East (Lei et al., 2009). This viewpoint was further supported by the mtDNA data of ancient horses between 2,000 and 4,000 years ago in China (Cai et al., 2009).

The Y chromosome, because of its male specific feature in the mammalian genome, unpairing with X chromosome during the meiosis, provides a powerful tool to trace the paternal origin of mammals (Skaletsky et al., 2003). Despite most of the studies on Y chromosome in domestic horse revealed no genetic variation (Lindgren et al., 2004; Wallner et al., 2004; Lau et al., 2009; Brandariz-Fontes et al., 2013), a recent study on equine Y chromosome short tandem repeats (Y-STR) detected two Y chromosome haplotypes in Chinese domestic horses (Ling et al., 2010), though, no Y-STR polymorphism was found in horses from Portugal, Spain and France (Brandariz-Fontes et al., 2013), indicating horse Y chromosome is not homogeneous and Chinese horse might display more abundant Y chromosome diversity. In addition, Chinese indigenous horse revealed a higher diversity level in comparison with Spanish horses,

German draught horses, Swiss horses, Norwegian horses, Portuguese Sorraia and Friesian horse and Japanese native horse breeds using 27 microsatellite markers (Ling et al., 2011). Recently, a few novel genetic variations on horse Y chromosome were identified by Wallner et al. (2013) that six haplotypes were first found in the European modern horses using high throughput sequencing technology. Of which, haplotype HT1 was the most ancestral with the highest frequency, and the remaining five haplotype were thought to arise on the background of HT1 by mutation or gene conversion after domestication. These polymorphisms have not been verified in Chinese horses and the genetic differences between horses from Europe and China still remain mysterious. Hence, to obtain deeper insights into the Y chromosome diversity, paternal origin and genetic relationship of Chinese domestic horse breeds, this study investigated the genetic diversity of 33 Y-specific loci and found two Y-SNPs and one Y-indel in Chinese domestic horses.

MATERIALS AND METHODS

Specimen collection and DNA extraction

Blood samples of 304 male horses were collected from 13 Chinese domestic breeds distributed in northwestern and southwestern China (Table 1). In addition, two female horses were also collected as negative controls to verify the male specificity of primers used in this study. These 13 breeds can be divided into 5 groups (Xie, 1987) based on their history, ecological environment and body size. The genomic DNA was extracted using a standard phenol-chloroform method (Sambrook and Russell, 2002).

DNA amplification and identification of Y chromosome diversity

A total of 33 pairs of primers were used to investigate

Table 1. Sample information of 13 horse breeds in China

Distribution	Breed	Abbreviation	Sample size	Source region
Northwestern China	Chaidamu	CDM	25	Chaidamu city, Qinghai
	Chakouyi	CKY	16	Tianzhu county, Gansu
	Datong	DT	10	Qilian county, Qinghai
	Yanji	YJ	50	Hejing county, Xinjiang
	Balikun	BLK	14	Balikun county, Xinjiang
	Kazakh	KZK	18	Changji city, Xinjiang
	Ningqiang	NQ	7	Ningqiang county, Shaanxi
	Yushu	YS	51	Yushu city, Qinghai
	Hequ	HQ	25	Maqu county, Gansu
	Guanzhong	GUZ	7	Fufeng county, Shaanxi
Southwestern China	Baise	BS	40	Baise county, Guangxi
	Debao pony	DB	19	Debao county, Guangxi
	Guizhou	GZ	22	Guiyang city, Guizhou
	Total		304	

the Y-SNPs in Chinese horse breeds (Table 2). Firstly, all the primers used were needed to verify their male specific, which was confirmed by a polymerase chain reaction (PCR) containing two female horse DNA paralleled with two male

Table 2. Detailed information of 33 Y chromosome loci in horse

No.	Locus	GenBank	Primer 5'-3'	Size (bp)	Tm (°C)	References
1	AMELY-1	AB091794	F: CTGGGTAGGTACAGGGAA R: ATACTTACGGGCATAGCA	365	56.4	This study
2	AMELY-6	AB091794	F: TTCAGTTCCCACTACCAG R: CTTGAGCCCAGATTACAG	390	56.4	This study
3	CLY032	BV140826	F: GGTCCAGAATGCCTGAGTAA R: AGAGACCTTTTGTGGGTGGA	396	64.2	Raudsepp et al. (2004)
4	CLY042	BV140783	F: CAGACCAGAAGCTGAAGAAGAG R: GGGCTGCATACAAGGAAAAGT	331	63.1	Raudsepp et al. (2004)
5	CLY048	BV140835	F: AACAGAACCCTGCCTAAACC R: CAGATTCCTTGGCTGACC	358	61	Raudsepp et al. (2004)
6	CLY050	BV140837	F: GCCTCAAGTAGAACACATCC R: GCATCCAGAACAGCAAACC	300	64.2	Raudsepp et al. (2004)
7	ETY2	EU687556	F: TAAGGCTTCCTCCTCCAAT R: CCAGTGACCCGACATACTGA	850	57.8	Paria et al. (2011)
12	SMCY	EU687564	F: AACAGCGAGCCAATGTTTTT R: GCAAAATTCTGGGAAATCCA	420	57.8	Agulnik et al. (1997)
13	SMCY1		F: TGACAATTTTCARRTTACTCCT R: CTGTAAAGGTCCAAGATCTT	1000	55	Hellborg and Ellegren (2003)
14	SMCY2	AY532886	F: ATTCCAATGTGGAGMGGAA R: CATAGCCACCTTCCTCMAT	300	61	Hellborg and Ellegren (2003)
15	SMCY3	AY532887	F: ATTTACCTTATGAAATRTTT R: TCAAATGGGTGWGTGYACAT	1000	64.2	Hellborg and Ellegren (2003)
16	SMCY5		F: GGTCCCAAGATGATRGGCT R: AGTTGGGGGGCARTSAC	200	61.8	Hellborg and Ellegren (2003)
17	SMCY6		F: CTACGGAAGAATCACAGCAG R: ATTTACGGAAGSGGTGGYAA	900	64.2	Hellborg and Ellegren (2003)
18	SMCY7	AY532888	F: TGGAGGTGCCCRAATGTGA R: AACTCTGCAAASTRITACTCCT	400	60.5	Hellborg and Ellegren (2003)
19	SMCY11		F: CTGCCCTGYRCCATGCAT R: TCCACCTGTSMAGRACAT	600	56.6	Hellborg and Ellegren (2003)
20	SMCY17	AY532889	F: AATGATCTGTGCAAGTGCTC R: GTCAAATGACTCAGCYCGAAT	200	55.5	Hellborg and Ellegren (2003)
21	SH3-B-7	BV005719	F: TGAAAGGGCTAATGAACGAT R: CCTCAGGTTTCGGGATT	436	64.2	Raudsepp et al. (2004)
22	USP9Y	EU687569	F: GGTATGAAATGGTCTCTGC R: CGAGTCTGTCCATCAGGATC	228	60.5	Paria et al. (2011)
23	Y3B1	G72336	F: TGGGTAAATGGGATTTGGTG R: CAAGCACAGCTCTGTATCAA	508	64.2	Raudsepp et al. (2004)
24	Y3B8	G72337	F: CCAAGTTCCCTTGCCATC R: AAATTGAAGAGGCCCAAAG	472	64.2	Raudsepp et al. (2004)
25	Y3B12	G72338	F: GGGAGGCACTGGAAAGTACA R: GGTGGAGGAATCAGCTGGAG	400	65	Raudsepp et al. (2004)
26	Y3B19	BV005720	F: AAGCCTTTCATGGAATTTGG R: TTACGCAGACATCCTGGACA	255	56.1	Raudsepp et al. (2004)
27	Y-25345	JX647030	F: CCTCCGGCCTTATGTCTTAG R: TTGGGCTGCAGTATACAACG	246	58	Wallner et al. (2013)
28	Y-45288	JX646942	F: CATTTAGTAACCTCCCTC R: CGTGAGACACTAGGTTTTA	232	55	This study
29	Y-50869	JX646950	F: GGCCTAAGTTGTTCCGAGAG R: TGAAGGTTGGTGTCCAGTGT	264	58	Wallner et al. (2013)
30	Y-45701/997	JX646942	F: CCAACACACGTAACAGCTC R: GGCTTAGGCCACTGATGGTA	444	58	Wallner et al. (2013)
31	ZFY	EU687571	F: TGAGCTATGCTGACAAAAGGTG R: TCTTCCCTTGTCTTGCTTGA	250	60.5	Paria et al. (2011)
32	ZFY last	DQ520652	F: GATGAAGCTAGTATGTATCGG R: CTAATGTGTGTGTTCTGAATG	666	60.6	Lau et al. (2009)
33	ZFY-ntl	DQ520642	F: TAGCAGTTCAAGGATTGCATA	450	50.2	Lau et al. (2009)

horse DNA. Subsequently, a total of 15 DNA pools (each pool is 200 μ L comprised of 20 individuals) were blended to scan Y-SNPs. The PCR was performed in a 25 μ L reaction containing 40 ng pooled genomic DNA, 10 pM each primer, 12.5 μ L 2 \times PCR Mix buffer (including 0.75 U Taq DNA polymerase, 2 \times PCR buffer, 37.5 μ M MgCl₂ and 5 μ M dNTPs) with the following conditions: 5 min at 95°C, followed by 35 cycles for 30 s at 95°C, 40 s at annealing temperature (Table 2), 30 s at 72°C, a final extension of 10 min at 72°C, at last storing at 4°C. The PCR products from the pooled DNA amplified by 33 pairs of primers were directly sequenced (Shanghai Sangon Biotech Company, Shanghai, China) and the sequences were aligned using DNASTAR 5.0 (DNASTAR, Madison, WI, USA). Those Y chromosome loci of 33 primer pairs (Table 2) suspected to be polymorphic were further genotyped by PCR-restriction fragment length polymorphism (RFLP) if had appropriate restriction enzymes site, otherwise were directly sequenced one by one.

Three Y chromosome loci (Y-45288, Y-50869, and Y-45701/997) were suspected to be polymorphic in above DNA pool sequencing analyses. Among which, an indel found in Y-45288 locus can be distinguished by a *TruII* restriction enzyme. In the digestion process, 2.5 μ L of PCR products were mixed with 5 μ L of restriction enzymes buffer (contains 2 U restriction enzymes) and then digested at 65°C for 3 h. Subsequently, digested products were visualized in native polyacrylamide gel electrophoresis. For the Y-50869 and Y-45701/997 locus, polymorphisms in 304 horses were genotyped by direct sequencing.

Data analysis

Haplotype diversity and nucleotide diversity for each breed were estimated using the DnaSPv5 (Librado and Rozas, 2009). A network was constructed to investigate the relationship among haplotypes by Network 4.6.1.1 (Fluxus Technology Ltd., 2012, Kiel, Germany). The polymorphisms in the analyzed segments were exported using MEGA 5.0 (Tamura et al., 2011). Neighbor-joining (NJ) tree with terminal node as breed based on D_A was constructed which was evaluated by 1,000 bootstrap (Nei et al., 1972). All positions containing gaps were instead of substitution in the analysis because of insufficient information to construct phylogenetic tree lack of haplotypes with gap.

RESULTS

In current study, we investigated polymorphisms of 33 Y-specific loci among 304 Chinese horses, and for the first time identified two Y-SNPs (in Y-50869 and Y-45701/997) and one Y-indel (in Y-45288). Within the locus Y-45288, only 10 horses harbored a T deletion (3.29%), which can be

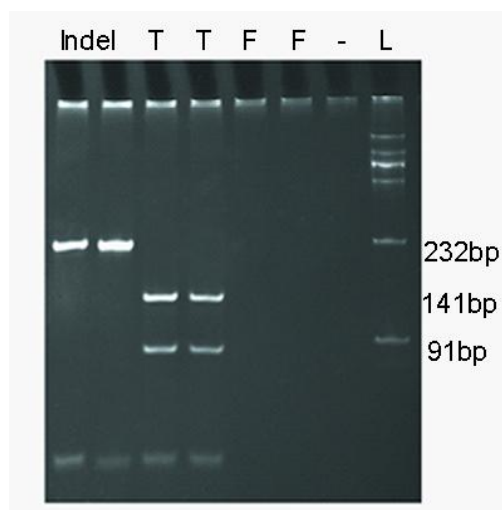


Figure 1. Polymerase chain reaction (PCR) products of Y-45288 locus analyzed by PCR-restriction fragment length polymorphism (RFLP). Individuals with T nucleotide yielded fragments of 91 and 141 bp, while individuals with T deletion only have 232 fragments (Indel, individuals with T deletion; T, individuals with T nucleotide; F, female; -, blank; L, ladder).

distinguished by a RFLP method (Figure 1). In locus Y-50869, 75 horses (24.67%) showed an A allele, whereas the remaining 229 horses displayed a T allele at the 264 bp. In the locus Y-45701/997, 298 horses were G allele at position 229 bp, whereas only six horses (1.97%) were T allele (two individuals in Hequ, one individual in Chakouyi, Datong, Yanji and Yushu horse, respectively) (Table 3). In all 13 breeds analyzed, Guanzhong horse revealed the highest mutation frequency in loci Y-45288 and Y-50869 (Table 3).

To further analyze the above three variations of Y chromosome loci, we made a haplotype analysis by combining them together. Hence, a total of five Y-SNP haplotypes (CHT1, CHT2, CHT3, CHT4, and CHT5) were identified (Table 4). And their frequencies among 13 Chinese domestic horse breeds were calculated and presented in Table 5. Among the 13 breeds studied, 12 horse breeds had at least two Y-SNP haplotypes, while Ningqiang breed only had CHT1 haplotype (Table 5). In addition to Guanzhong, Kazakh and Yanji breeds, the CHT1 is a predominant haplotype, consisting of 219 individuals, which accounted for 72.04% of 304 samples. In contrast, the remaining four haplotypes only contained 4, 69, 6 and 6 individuals, respectively (Table 5). The NJ tree based on D_A was showed in Figure 2. The dendrogram showed that all five haplotypes could be clustered into two branches. Furthermore, in the first branch haplotype CHT1 and CHT4 gathered together. In the second branch, haplotype CHT2 and haplotype CHT5 gathered firstly, then the haplotype CHT3 was gathered into the branch (Figure 2). All the Y-SNP haplotypes were not restricted to a specific region or breed (Figure 3). The Median-joining (MJ) network showed

Table 3. Allele frequencies of three polymorphic loci in 13 horse breeds

Breed	Samples	Y-45288		Y-50869		Y-45701/997	
		T	-	T	A	G	T
CDM	25	0.9200	0.0800	0.6000	0.4000	1.0000	0.0000
CKY	16	1.0000	0.0000	0.8750	0.1250	0.9375	0.0625
DT	10	0.9000	0.1000	0.9000	0.1000	0.9000	0.1000
YJ	50	1.0000	0.0000	0.3600	0.6400	0.9800	0.0200
BS	40	0.9750	0.0250	0.9250	0.0750	1.0000	0.0000
DB	19	0.9474	0.0526	1.0000	0.0000	1.0000	0.0000
GUZ	7	0.4286	0.5714	0.1429	0.8571	1.0000	0.0000
GZ	22	1.0000	0.0000	0.9545	0.0455	1.0000	0.0000
NQ	7	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
BLK	14	1.0000	0.0000	0.5714	0.4286	1.0000	0.0000
KZK	18	1.0000	0.0000	0.4444	0.5556	1.0000	0.0000
YS	51	1.0000	0.0000	0.9216	0.0784	0.9804	0.0196
HQ	25	0.9600	0.0400	1.0000	0.0000	0.9200	0.0800
Total	304	0.9671	0.0329	0.7533	0.2467	0.9803	0.0197

CDM, Chaidamu; CKY, Chakouyi; DT, Datong; YJ, Yanji; BLK, Balikpapan; KZK, Kazakh; NQ, Ningqiang; YS, Yushu; HQ, Hequ; GUZ, Guanzhong; BS, Baise; DB, Debao pony; GZ, Guizhou.

Table 4. The five haplotypes based on three Y chromosome loci

Haplotype	Y-45288	Y-50869	Y-45701/997	Percentage (%)
CHT1	T	T	G	72.04
CHT2	-	T	G	1.32
CHT3	T	A	G	22.70
CHT4	T	T	T	1.97
CHT5	-	A	G	1.97

the relationship among five haplotypes involving relevant breed information. It clearly revealed that CHT1 was in the center position and shared by different breeds, which indicated CHT1 maybe a common and ancestral haplotype (Figure 3). Furthermore, the regional haplotype data clearly demonstrated the increase of genetic variations in a

southwest–northwest direction.

The Y-SNP haplotype diversity and nucleotide diversity for each breeds were calculated using software DnaSPv5 (Librado and Rozas, 2009) without considering locus Y-45288, because the DnaSPv5 software cannot distinguish deletion. As shown in Table 6, the average haplotype diversity was 0.402, ranging from 0.000 to 0.527 and the average nucleotide diversity was 0.00044, ranging from 0.00000 to 0.00056. The Balikpapan and Kazakh horse breeds from Xinjiang autonomous region demonstrated the highest haplotype diversity ($h = 0.527$ for Balikpapan; $h = 0.523$ for Kazakh) and nucleotide diversity ($\pi = 5.6 \times 10^{-4}$). In contrast, the Ningqiang horse had the lowest haplotype diversity ($h = 0.000$) and nucleotide diversity ($\pi = 0.00000$). Furthermore, the horse breeds in this study were divided into 5 types,

Table 5. The frequencies of five Y chromosome haplotypes in 13 Chinese horse breeds

Breed	Samples	CHT1	CHT2	CHT3	CHT4	CHT5
CDM	25	0.5600 (14)	0.0400 (1)	0.3600 (9)	0.0000	0.0400 (1)
CKY	16	0.8125 (13)	0.0000	0.1250 (2)	0.0625 (1)	0.0000
DT	10	0.8000 (8)	0.0000	0.0000	0.1000 (1)	0.1000 (1)
YJ	50	0.3400 (17)	0.0000	0.6400 (32)	0.0200 (1)	0.0000
BS	40	0.9000 (36)	0.0250 (1)	0.0750 (3)	0.0000	0.0000
DB	19	0.9476 (18)	0.0526 (1)	0.0000	0.0000	0.0000
GUZ	7	0.1429 (1)	0.0000	0.2857 (2)	0.0000	0.5714 (4)
GZ	22	0.9545 (21)	0.0000	0.0455 (1)	0.0000	0.0000
NQ	7	1.0000 (7)	0.0000	0.0000	0.0000	0.0000
BLK	14	0.5714 (8)	0.0000	0.4286 (6)	0.0000	0.0000
KZK	18	0.4444 (8)	0.0000	0.5556 (10)	0.0000	0.0000
YS	51	0.9020 (46)	0.0000	0.0784 (4)	0.0196 (1)	0.0000
HQ	25	0.8800 (22)	0.0400 (1)	0.0000	0.0800 (2)	0.0000
Total	304	0.7204 (219)	0.0132 (4)	0.2270 (69)	0.0197 (6)	0.0197 (6)

CDM, Chaidamu; CKY, Chakouyi; DT, Datong; YJ, Yanji; BS, Baise; DB, Debao pony; GUZ, Guanzhong; GZ, Guizhou; NQ, Ningqiang; BLK, Balikpapan; KZK, Kazakh; YS, Yushu; HQ, Hequ.

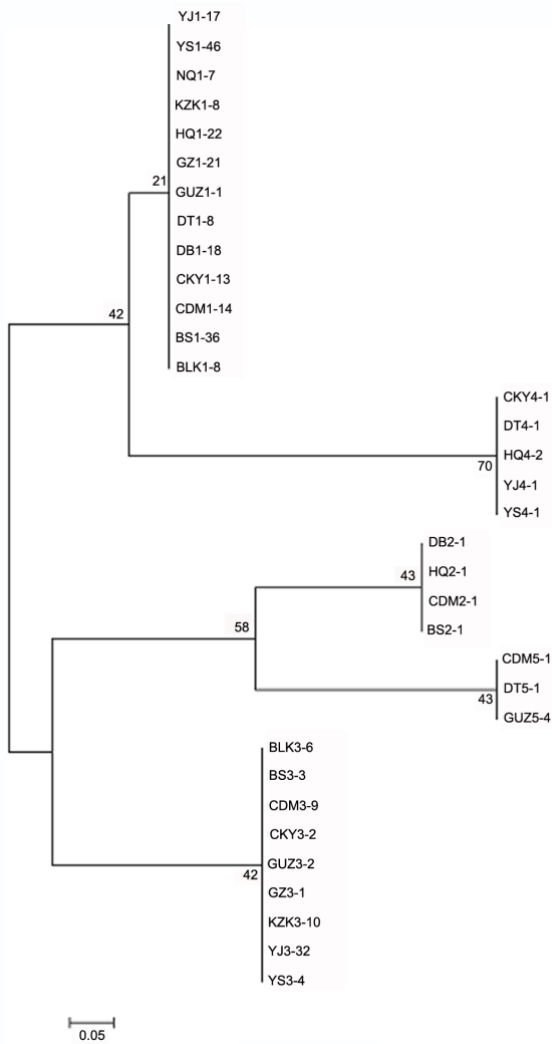


Figure 2. Neighbor-joining tree among 13 horse breeds based on the five haplotypes. The number before dash is the haplotype they belong to, the number after dash is the number of individuals.

including Mongolia, Southwest, Kazakh, Tibet, and Hequ (Xie, 1987). The results revealed that Mongolia type showed the highest diversity (Table 6). Meanwhile, horses belonging to Mongolia type and Kazakh type also revealed abundant Y chromosome diversity.

DISCUSSION

Since very limited polymorphism was found on the horse Y chromosome (Lindgren et al., 2004; Wallner et al., 2004; Lau et al., 2009), it has been proposed for a long time that horse was domesticated in a restricted geographical region, resulting in the incorporation of very limited Y chromosome haplotypes into the breeding stock. However, recent studies have observed polymorphisms of Y-STR in Chinese indigenous horses (Ling et al., 2010) and Y-SNPs in European and American horses (Wallner et al., 2013). In current study, we screened polymorphisms of 33 Y chromosome loci in 13 horse breeds, which further indicated that Y chromosome is polymorphic in Chinese horses.

Though the polymorphisms of four loci (Y-25345, Y-45288, Y-50869, and Y-45701/997) have been investigated in European and north American horses (Wallner et al., 2013), the current study found different polymorphic patterns in Chinese horses. First, no polymorphism was detected in locus Y-25345, which showed a substitution of G/A in nine Icelandic horses. Second, Y-indel at locus Y-45288 was 100% linked with Y-SNP T>A at locus Y-50869 in European and north American horses, whereas it is only 60% linked in Chinese horses. Furthermore, to our surprise, a total of 20 SNPs and indels which were identified at the locus Y-45701/997 in 15 Norwegian Fjord horses are absent in Chinese breeds, while only one substitution (G>T) was

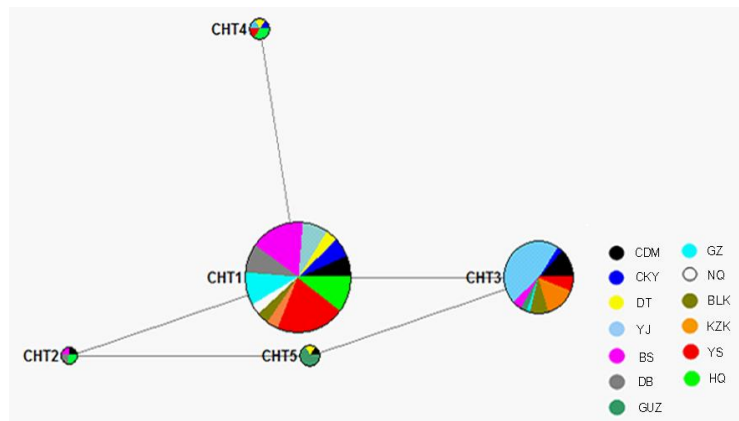


Figure 3. A Median-joining (MJ) network calculation of Chinese horse Y chromosome. Colored circles represent sequence haplotypes, the area being proportional to the frequency of the haplotype in total samples. Branch length is proportional to number of mutations. Different colors represent different horse breed. The network clearly revealed that CHT1 was in the center position, which indicated CHT1 maybe an ancestral haplotype. CDM, Chaidamu; CKY, Chakouyi; DT, Datong; YJ, Yanji; BS, Baise; DB, Debao pony; GUZ, Guanzhong; GZ, Guizhou; NQ, Ningqiang; BLK, Balikun; KZK, Kazakh; YS, Yushu; HQ, Hequ.

found (Table 4). The different polymorphic patterns may be caused by the genetic isolation and differentiation between Eastern and Western horse populations, which has been proved in mtDNA analysis (MoGahern et al., 2006; Lei et al., 2009).

The haplotype network (Figure 3) showed that all Y-SNP haplotypes were present in more than one breed, and no haplotype was restricted to a certain region. The constant procurement of horses from many regions by the imperial court, the high mobility of horses and the widely use of horses in wars throughout Chinese history drove the gene flow and mixture of horses from different regions (Xie, 1987), resulting in reduced genetic diversity among horses in different regions.

More than half number of the Guanzhong horse, Kazakh horse and Yanji horse displayed Y chromosome diversity. And the Kazakh horse and Balikun horse breeds displayed the highest haplotype and nucleotide diversity. The Balikun horse, Kazakh horse and Yanji horse breeds distribute in Xinjiang region having vast and widespread grasslands, where large-scale animal husbandry is still sustained. Thus it appears natural that higher genetic diversity is found in horses from Xinjiang. In the distribution region for Kazakh horse, breeding strategy was adopted to reform Kazakh horse performance traits due to the decreasing quality, so that some samples we selected might have undergone the breed proceeding and could be classified as a cultivated breed called Yili horse. Also it is worth noticing that 64% of Yanji horse exhibited CHT3 and 42.67% of the total genetic variation for CHT3 was attributable to Yanji breed which was obtained from Hejing county, Xinjiang region. In 1920s, 11 stallions were imported from Soviet Union's to ameliorate the local breed quality and their posterities were

still distributed in Hejing county, resulting in the high genetic variation in Yanji breed. The Guanzhong horse belongs to cultivated breed which was formed by crossing local mares with imported Carabai and Ardennes stallions primarily (Xie, 1987). This may be an explanation for the higher mutation frequency displayed in Guanzhong horse. On the contrary, no genetic variation was found in Ningqiang horse which displayed the lowest nucleotide and haplotype diversity, consisting with previous mtDNA results (Lei et al., 2009), this could be explained by the recent bottleneck they underwent that resulted in a decrease of its population number and genetic diversity.

From our results, it is obvious that horses from Northwestern China revealed high Y chromosome diversity. Moreover, the mtDNA and Y-STR researches have proved that Chinese native horses preserved much of genetic diversity, particularly in the Southwestern regions (Lei et al., 2009; Ling et al., 2010; 2011). Combining these results, it is credible that Chinese horses exhibit high genetic variation. Archaeological and osteological evidence have, thus far, identified Central Asia as the region where the earliest horse domestication took place, but a lot of efforts on mtDNA found that the domestication of horses was not a single discrete event (McGahern et al., 2006; Cai et al., 2009) and multiple horse domestication events may have occurred later in other regions including China (Lei et al., 2009). Based on our data, we cannot exclude horses possibly to be independently domesticated in China. Meanwhile, this study is not enough to comprehensively understand the paternal origins of Chinese horses. Therefore, more studies on Y chromosome of ancient and modern horses in China are needed in the future.

Table 6. Haplotype diversity and nucleotide diversity in 13 horse breeds

Breed	Samples	K	Haplotype diversity (mean±SD)	Nucleotide diversity (mean±SD)	Type
CDM	25	4	0.500±0.048	0.00053±0.00005	Mongolia h = 0.530, π = 5.9×10 ⁻⁴
CKY	16	3	0.342±0.140	0.00038±0.00017	
DT	10	3	0.378±0.181	0.00043±0.00022	
YJ	50	3	0.484±0.047	0.00054±0.00006	Southwest h = 0.190, π = 2.0×10 ⁻⁴
BS	40	3	0.142±0.071	0.00015±0.00008	
DB	19	2	0.000±0.000	0.00000±0.00000	
GUZ	7	3	0.286±0.196	0.00030±0.00021	Kazakh h = 0.516, π = 5.5×10 ⁻⁴
GZ	22	2	0.091±0.081	0.00010±0.00009	
NQ	7	1	0.000±0.000	0.00000±0.00000	
BLK	14	2	0.527±0.064	0.00056±0.00007	Tibet
KZK	18	2	0.523±0.048	0.00056±0.00005	
YS	51	3	0.184±0.070	0.00020±0.00008	Hequ
HQ	25	3	0.153±0.092	0.00016±0.00001	
Total	304	5	0.402±0.026	0.00044±0.00003	

CDM, Chaidamu; CKY, Chakouyi; DT, Datong; YJ, Yanji; BS, Baise; DB, Debao pony; GUZ, Guanzhong; GZ, Guizhou; NQ, Ningqiang; BLK, Balikun; KZK, Kazakh; YS, Yushu; HQ, Hequ.

K, haplotype numbers; h, haplotype diversity; π, nucleotide diversity.

CONCLUSION

In conclusion, 33 Y chromosome loci were investigated and first detected two Y-SNPs and one Y-indel in Chinese horses, identifying five Y-SNP haplotypes, which suggested that Chinese horse exhibited high genetic diversity and different pattern of genetic variation from European and American horse populations. In order to clarify paternal origins of horse domestication, further studies are required to evaluate the relationship between Chinese horses and horses outside China.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (31072001).

REFERENCES

- Agulnik, A. I., C. E. Bishop, J. L. Lerner, S. I. Agulnik, and V. V. Solovyev. 1997. Analysis of mutation rates in the *SMCY/SMCX* genes shows that mammalian evolution is male driven. *Mamm. Genome* 8:134-138.
- Brandariz-Fontes, C., J. A. Leonard, J. L. Vega-Pla, N. Backström, G. Lindgren, S. Lippold, and C. Rico. 2013. Y-chromosome analysis in reuertas horses. *PLoS One* 8(5):e64985.
- Cai, D., Z. Tang, L. Han, C. F. Speller, D. Y. Yang, X. Ma, J. Cao, H. Zhu, and H. Zhou. 2009. Ancient DNA provides new insights into the origin of the Chinese domestic horse. *J. Archaeol. Sci.* 36:835-842.
- Cieslak, M., M. Pruvost, N. Benecke, M. Hofreiter, A. Morales, M. Reissmann, and A. Ludwig. 2010. Origin and history of mitochondrial DNA lineages in domestic horses. *PLoS One* 5(12):e15311.
- Deng, T. 2000. Phylogenetic relationship of the Chinese miniature pony to *Equus przewalskii* (Perissodactyla, Equidae). *Acta Veterinaria et Zootechnica Sinica* 31:28-33 (in Chinese).
- Diamond, J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418:700-707.
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47-50.
- Goto, H., O. A. Ryder, A. R. Fisher, B. Schultz, S. L. K. Pond, A. Nekrutenko, and K. D. Makova. 2011. A massively parallel sequencing approach uncovers ancient origins and high genetic variability of endangered Przewalski's horses. *Genome Biol. Evol.* 3:1096-1106.
- Guo, S., P. Savolainen, J. Su, Q. Zhang, D. Qi, J. Zhou, Y. Zhong, X. Zhao, and J. Liu. 2006. Origin of mitochondrial DNA diversity of domestic yaks. *BMC Evol. Biol.* 6:73.
- Gupta, A. K. 2004. Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Curr. Sci.* 87:54-59.
- Hellborg, L. and H. Ellegren. 2003. Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. *Mol. Ecol.* 12:283-291.
- Ishida, N., T. Oyunsuren, S. Mashima, H. Mukoyama, and N. Saitou. 1995. Mitochondrial DNA sequences of various species of the genus *Equus* with special reference to the phylogenetic relationship between Przewalski's wild horse and domestic horse. *J. Mol. Evol.* 41:180-188.
- Larson, G., K. Dobney, U. Albarella, M. Fang, E. Matisoo-Smith, J. Robins, S. Lowden, H. Finlayson, T. Brand, and E. Willerslev. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307(5715):1618-1621.
- Larson, G., R. Liu, X. Zhao, J. Yuan, D. Fuller, L. Barton, K. Dobney, Q. Fan, Z. Gu, X. H. Liu, Y. Luo, P. Lv, L. Andersson, and N. Li. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc. Natl. Acad. Sci. USA.* 107:7686-7691.
- Lau, A. N., L. Peng, H. Goto, L. Chemnick, O. A. Ryder, and K. D. Makova. 2009. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. *Mol. Biol. Evol.* 26:199-208.
- Lei, C. Z., R. Su, M. A. Bower, C. J. Edwards, X. B. Wang, S. Weining, L. Liu, W. M. Xie, F. Li, R. Y. Liu, Y. S. Zhang, C. M. Zhang, and H. Chen. 2009. Multiple maternal origins of native modern and ancient horse populations in China. *Anim. Genet.* 40:933-944.
- Lei, C. Z., W. Zhang, H. Chen, F. Lu, R. Y. Liu, X. Y. Yang, H. C. Zhang, Z. G. Liu, L. B. Yao, Z. F. Lu, and Z. L. Zhao. 2007. Independent maternal origin of Chinese swamp buffalo (*Bubalus bubalis*). *Anim. Genet.* 38:97-102.
- Librado, P. and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Lindgren, G., N. Backström, J. Swinburne, L. Hellborg, A. Einarsson, K. Sandberg, G. Cothran, C. Vilà, M. Binns, and H. Ellegren. 2004. Limited number of patrilineages in horse domestication. *Nat. Genet.* 36:335-336.
- Ling, Y., Y. Ma, W. Guan, Y. Cheng, Y. Wang, J. Han, D. Jin, L. Mang, and H. Mahmut. 2010. Identification of Y chromosome genetic variations in Chinese indigenous horse breeds. *J. Hered.* 101:639-643.
- Ling, Y. H., Y. H. Ma, W. J. Guan, Y. J. Cheng, Y. P. Wang, J. L. Han, L. Mang, Q. J. Zhao, X. H. He, Y. B. Pu, and B. L. Fu. 2011. Evaluation of the genetic diversity and population structure of Chinese indigenous horse breeds using 27 microsatellite markers. *Anim. Genet.* 42:56-65.
- Liu, Y. P., G. S. Wu, Y. G. Yao, Y. W. Miao, G. Luikart, M. Baig, A. Beja-Pereira, Z. L. Ding, M. G. Palanichamy, and Y. P. Zhang. 2006. Multiple maternal origins of chickens: out of the Asian jungles. *Mol. Phylogenet. Evol.* 38:12-19.
- Ludwig, A., M. Pruvost, M. Reissmann, N. Benecke, G. A. Brockmann, P. Castaños, M. Cieslak, S. Lippold, L. Llorente, A. S. Malaspinas, M. Slatkin, and M. Hofreiter. 2009. Coat color variation at the beginning of horse domestication. *Science* 324(5926):485-485.
- McGahern, A., M. A. M. Bower, C. J. Edwards, P. O. Brophy, G. Sulimova, I. Zakharov, M. Vizuete-Forster, M. Levine, S. Li, D. E. MacHugh, and E. W. Hill. 2006. Evidence for biogeographic patterning of mitochondrial DNA sequences in Eastern horse populations. *Anim. Genet.* 37:494-497.

- Nei, M. 1972. Genetic distance between populations. *Am. Nat.* 106:283-292.
- Oakenfull, E. A. and O. A. Ryder. 1998. Mitochondrial control region and 12S rRNA variation in Przewalski's horse (*Equus przewalskii*). *Anim. Genet.* 29:456-459.
- Oakenfull, E. A., H. N. Lim, and O. A. Ryder. 2000. A survey of equid mitochondrial DNA: Implications for the evolution, genetic diversity and conservation of *Equus*. *Conserv. Genet.* 1:341-355.
- Outram, A. K., N. A. Stear, R. Bendrey, S. Olsen, A. Kasparov, V. Zaibert, N. Thorpe, and R. P. Evershed. 2009. The earliest horse harnessing and milking. *Science* 323(5919):1332-1335.
- Raudsepp, T., A. Santani, B. Wallner, S. R. Kata, C. Ren, H. B. Zhang, J. E. Womack, L. C. Skow, and B. P. Chowdhary. 2004. A detailed physical map of the horse Y chromosome. *Proc. Natl. Acad. Sci. USA.* 101:9321-9326.
- Paria, N., T. Raudsepp, A. J. P. Wilkerson, P. C. M. O'Brien, M. A. Ferguson-Smith, C. C. Love, C. Arnold, P. Rakestraw, W. J. Murphy, and B. P. Chowdhary. 2011. A gene catalogue of the euchromatic male-specific region of the horse Y chromosome: comparison with human and other mammals. *PLoS One* 6(7):e21374.
- Sambrook, J. and D. W. Russell. 2002. *Molecular cloning: A laboratory manual*. Translated by P. T. Huang. Beijing: Science Press, Beijing, China.
- Savolainen, P., Y. P. Zhang, J. Luo, J. Lundeberg, and T. Leitner. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* 298(5598):1610-1613.
- Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, T. Pyntikova, J. Ali, and T. Bieri et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825-837.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28:2731-2739.
- Vilà, C., J. A. Leonard, A. Götherström, S. Marklund, K. Sandberg, K. Lidén, R. K. Wayne, and H. Ellegren. 2001. Widespread origins of domestic horse lineages. *Science* 291(5503):474-477.
- Wallner, B., G. Brem, M. Müller, and R. Achmann. 2003. Fixed nucleotide differences on the Y chromosome indicate clear divergence between *Equus przewalskii* and *Equus caballus*. *Anim. Genet.* 34:453-456.
- Wallner, B., F. Piumi, G. Brem, M. Müller, and R. Achmann. 2004. Isolation of Y chromosome-specific microsatellites in the horse and cross-species amplification in the genus *Equus*. *J. Hered.* 95:158-164.
- Wallner, B., C. Vogl, P. Shukla, J. P. Burgstaller, T. Druml, and G. Brem. 2013. Identification of genetic variation on the horse Y chromosome and the tracing of male founder lineages in modern breeds. *PLoS One.* 8(4):e60015.
- Xie, C. 1987. *Chinese Horse Donkey Breeds*. Shanghai Science and Technology Press, Shanghai, China (in Chinese).
- Zhang, C. S. 2004. Wild horses, domestic horses, and center for Eastern Asia horse raising. *Agric. Archaeol.* 1:252-254 (in Chinese).