Evaluation of 17 microsatellite markers for parentage testing and individual identification of domestic yak (*Bos grunniens*)

Jie Pei[1,2], Pengjia Bao[1,2], Min Chu[1,2], Chunnian Liang[1,2], Xuezhi Ding[1,2], Hongbo Wang[1,2], Xiaoyun Wu[1,2], Xian Guo[1,2] and Ping Yan[1,2]

[1] Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou, China

[2] Key Laboratory for Yak Genetics, Breeding, and Reproduction Engineering of Gansu Province, Lanzhou, China

Corresponding authors:
Xian Guo, guoxianlz@163.com
Ping Yan, pingyan63@126.com

ABSTRACT

Background. Yak (*Bos grunniens*) is the most important domestic animal for people living at high altitudes. Yak ordinarily feeds by grazing, and this behavior impacts the accuracy of the pedigree record ~~accuracy~~ because it is difficult to control mating in grazing yak. This study aimed to evaluate the ~~system of~~ pedigree system and individual identification in polled yak.

Methods. Seventy-one microsatellite loci were selected from the literature, mostly from the studies on cattle. Thirty-five microsatellite loci generated excellent ~~results in~~ PCR results and were evaluated for the parentage testing and individual identification of 236 unrelated polled yaks. Seventeen of these 35 microsatellite loci had polymorphic information content (PIC) values greater than 0.5, and these loci were in Hardy-Weinberg equilibrium without linkage disequilibrium~~Of these, 17 microsatellite loci had polymorphic information content (PIC) of >0.5, and they were included in Hardy-Weinberg equilibrium without linkage disequilibrium~~.

Results. Using multiplex PCR~~s~~, capillary electrophoresis, and genotyping, very high exclusion probabilities were obtained for the combined core set of 17 loci~~:~~. ~~the~~ The exclusion probability (PE) for one candidate parent when the genotype of the other parent is not known ~~=~~ was 0.99718116~~;~~. PE for one candidate parent when the genotype of the other parent is known ~~=~~ was 0.99997381~~;~~. ~~and~~ PE for a known candidate parent pair ~~=~~ was 0.99999998. The combined PEI (exclusion probability for identity of two unrelated individuals) and PESI (exclusion probability for identity of two siblings) were > 0.99999999 and 0.99999899, respectively. These findings indicated that the combination of 17 microsatellite markers could be useful for efficient and reliable parentage testing and individual identification in polled yak.

Discussion. Many microsatellite loci have been investigated for cattle paternity testing. Nevertheless, these loci cannot be directly applied to yak identification because the two bovid species have different genomic~~e~~ sequences and organization. Seventeen loci were selected from 71 microsatellite loci based on efficient amplification, unambiguous genotyping, and high PIC values for polled yaks, and were suitable for parentage analysis in polled yak

populations~~The 17 loci selected were deemed suitable for yak parentage testing based on their efficient amplification, unambiguous genotyping, and high PIC. Thus, this set of microsatellite loci is suitable for parentage analysis in polled yak populations~~.

INTRODUCTION

Yak (*Bos grunniens*), a member of Bovidae family, has successfully been adapted to the severe cold and low oxygen levels, ~~which is a~~ characteristics of high altitude regions (~2,500–5,500 m), such as the Himalayas in South-Central Asia, the Qinghai-Tibetan Plateau ~~in ,~~ Mongolia, and Russia (Wu 2016). These regions are known for their high elevations, pristine natural environments, and extreme seasonal variations (Mizuno et al. 2015). The ability of yak to survive in such rugged natural environment is due to its ~~various~~ varied behavioral, physiological, and genetic adaptations (Barsila et al. 2014; Hu et al. 2012; Qiu et al. 2012). For instance, remarkable reduction of heat production at night when not grazing and increased energy consumption when grazing ~~in~~ under free-range conditions enables yaks to save more energy and resist the extremely harsh conditions than other cattle under ~~the same~~ similar environmental conditions ~~in order to resist the extremely harsh conditions~~ (Ding et al. 2014). Yak can thrive ~~in~~ under extreme environmental conditions, such as the Tibetan Plateau~~,~~ where few other animals ~~also~~ can survive. In this region, yak have significantly contributed to human life by providing meat, milk, fur, leather, and transportation ~~when~~ compared to other animals (Medhammar et al. 2012; Wang et al. 2018).

There are about 13 million domestic yak in China, accounting for~~Of~~ approximately 90% of the global yak population~~, domestic yak constituted about 13 million in China~~. Although there are 18 yak breeds in China, only one breed (Datong) has been included in a breeding program (Wu 2016). However, polled yak ~~have~~ has been bred for many years at the foot of Ashidan Mountain, as polling reduces the risk of horn-inflicted injury or death among the herdsmen. Accurate genealogical records can help estimate ~~the~~ genetic parameters and improve ~~the~~ breeding programs, ensuring efficient and effective breeding progression to avoid excessive

inbreeding. However, ~~while grazing,~~ it is difficult to control mating among yaks while grazing. Furthermore, semen samples could be erroneously mislabeled during preparation, and mating records might be misinterpreted because of clerical errors made during artificial insemination. Therefore, accurate yak pedigree records that are compiled by parentage testing and individual identification are essential ~~in~~ to the yak breeding process ~~of yak~~.

Molecular markers can indicate the degree of genetic relatedness between animals, facilitating ~~the~~ parentage verification and individual identification (Estoup et al. 2002). Microsatellites are ~~also known~~refer to ~~as~~ short tandem repeats (STRs) or simple sequence repeats (SSRs), and are considered as tracts of DNA motifs ranging from ~~1~~ one to 10 nucleotides in length with repeats of 5–50 times (Carneiro Vieira et al. 2016). ~~The~~ ~~m~~Microsatellites can be used to develop pedigree~~d~~ animal populations and ~~to~~ evaluate animal breeding, ~~which in turn~~ supporting ~~the~~ genetic improvement by selective breeding (Weising et al. 1998). The application~~s~~ of microsatellites as molecular markers for animal identification and parentage verification produced highly accurate and effective results in both breeding and forensics (Linacre et al. 2011).

Microsatellite marker analysis has been used to verify the parentage in breeding registries and identify individual animals that are linked to a particular database or owner. Microsatellite panels of cattle (Zhao et al. 2017), horse (Kang et al. 2016), sheep (Rosa et al. 2013), dog (Jeong et al. 2015), and parrot (Coetzer et al. 2017) have been well characterized. Parentage control in the beef cattle breeds, Charolais, Limousin, and Preta, in Portugal was assessed using 10 microsatellite markers, and the results revealed a combined exclusion probability (PE) ~~of~~ above 0.9995, indicating their ability to exclude a random parent pair (Carolino et al. 2009). The application of 11 microsatellite loci in paternity testing ~~regarding~~ ~~the~~in Yugoslav Pied cattle breed in Serbia revealed a combined PE of 0.999 (Stevanovic et al. 2010). Sixteen specific microsatellite markers were used to develop a genetic system ~~for~~ of meat traceability ~~of~~ for several beef cattle breeds, including Japanese Black, Anduo yak, Limousin, Jiaxian Red, Nanyang Yellow, and Luxi Yellow (Zhao et al. 2017). Previous studies ~~have~~ reported that microsatellite genotyping was used for ~~kinship relationship~~

~~identification~~population genetics analysis and parentage testing in yak. However, these loci and their primers were originally developed for cattle, and then used directly in yak (Li et al. 2013; Nguyen et al. 2005). Therefore, there is a need for explorative and application-based studies on microsatellite markers or panels of microsatellite markers that are suitable for ~~yak have to be explored and applied for~~ accurate individual identification and parentage testing in yak.

Hence, in the present study, we aimed to establish a paternity test and individual identification system for the polled yak. ~~So~~Therefore, the study ~~planned~~ was expected to (1) ~~to~~ calculate the genetic parameters of polled yak microsatellite loci, which have been commonly reported by ~~studies on kinship relationships~~population genetics studies ~~in~~ of cattle and yak; (2) ~~to~~ evaluate the application values of the loci~~,~~ with high polymorphic information content (PIC)~~,~~ for parentage testing and individual identification; and (3) ~~to~~ explore a multi-loci combination test system for parentage testing and individual identification.

MATERIALS AND METHODS

Marker selection and primer design

The microsatellites used in the present study were selected from the previous reports on cattle breeding based on the following criteria: (a) high PIC; (b) large number of alleles; (c) relatively infrequent null alleles; and (d) homogeneous or approximately homogeneous repeat motifs (Schnabel et al. 2000). Among the 71 bovine microsatellite markers selected, 65 were derived from ~~the~~ cattle, ~~(Tian et al. 2008; Zhang et al. 2008b)~~ and ~~six~~ 6 were exclusive to yak (Supplemental Table S1) ~~(Li 2004)~~. All ~~the~~ 71 microsatellites and their flanking sequences were found in the cattle genome, and were searched for in the yak genome. The primers for most ~~of the~~ loci used in ~~the~~ previous studies were not ~~adapted~~ suitable ~~to~~ for the yak genome ~~due to~~because of low scores. However, the primers used to amplify ~~the loci~~ BM1824, BM2113, BMS2533, ETH121, ETH225, ILSTS008, INRA124, RM099, INRA126, UMN0103, UMN0307, UMN0920, UMN2303, UMN3007, and UMN3008 loci presented

relatively high scores, ~~and therefore,~~so these loci were not ~~adjusted~~ redesigned for yak. The primers used in the present study are listed in Table S1.

Sample collection

Polled ~~Yaks~~ yaks ~~polled~~ bred ~~by~~via selective breeding were selected from ~~the~~ herds in Ashidan Mountain~~,~~ region in Qinghai Province, China. To avoid consanguinity, samples ~~from animals with no genetic relationship~~ were ~~selected~~taken from ~~aninals~~animals with no genetic relationship. All yaks were handled in strict accordance with good animal practices~~,~~ by following the *Animal Ethics Procedures and Guidelines of the People's Republic of China*. The present study was approved by the *Animal Administration and Ethics Committee of Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences* (Permit No. SYXK-2016-0039). Blood was drawn from the jugular veins of 236 unrelated individuals, including 38 sires and 198 dams, and samples were mixed with ~~the~~ preservation buffer (containing 1.5 mg mL^{-1} EDTA and 137 mmol L^{-1} NaCl) at a ratio of 5:1. The blood samples were stored at -80 ºC in an ultra-cold freezer until DNA extraction.

DNA extraction and quantification

~~The~~ gGenomic DNA was extracted from white blood cells, ~~which were~~ separated from ~~the~~ whole blood, and digested with ~~Proteinase~~ proteinase K. After digestion, the samples were centrifuged at 5000 × *g* for 2 min, and the resulting supernatant (clear aqueous layer) was transferred to a new test tube. After the addition of 0.5 mL of 10 mg mL^{-1} RNase A, ~~the~~ DNA was individually extracted with a phenol:chloroform:isoamyl alcohol mixture (25:24:1) followed by chloroform, precipitated with ethanol, and ~~then~~ resuspended in 50 mL TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH 8.0). The extracted DNA was then quantified using a NanoDrop 2000 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

Preliminary primer screening

Unlabeled primer pairs (Table S1) for each microsatellite marker were used for ~~amplifying~~ amplification of ~~the~~ DNA fragments. PCR was performed with a reaction mixture at a total volume of ~~total volume~~ 20 µL, ~~and~~ comprising of 20–50 ng ~~of~~ genomic DNA, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 2.0 mM each dNTP, 5 ng of bovine serum albumin, and 1.0 U of *Taq*™ Hot Start Version polymerase (TaKaRa Bio Inc., Kusatsu, Shiga, Japan). Primer ~~The~~ concentrations ~~of Primers~~ ranged from 1.0 to 5.5 µM. The thermal cycle parameters were as follows: 5 min at 95 ºC; 35 cycles of 30 s at 95 ºC, 30 s at 55–58 ºC (Table S1), and 20 s at 72 ºC; and a final extension step for 5 min at 72 ºC. Amplifications were performed on a Veriti 96-Well Thermal Cycler (Applied Biosystems Corp., Foster City, CA, USA). The amplicons were visualized on 2% agarose gels (Gene Tech Co. Ltd., Chai Wan, Hong Kong, PRC). Five microliters of PCR product were mixed with 1 µL of loading dye (TaKaRa Bio Inc.) and electrophoresed for 35 min at 85 V~~;~~. ~~the~~ The samples were then visualized using the GelDoc™ XR+ gel imaging system (Bio-Rad Corp., Hercules, CA, USA).

The amplicons with high specificity~~,~~ and ~~high~~ amplification efficiency, ~~and~~ whose loci ~~that~~ were either non-syntenic or separated by > 10 cM (to avoid strong genetic linkage) (Table S2)~~.~~, were manually cut out of agarose gels and forwarded to Invitrogen (Carlsbad, CA, USA) and Thermo Fisher Scientific~~,~~ without prior purification~~,~~ for Sanger sequencing on an ABI3730xl automated sequencer (Applied Biosystems Corp.) to identify repeat markers. Only the forward primers were used to sequence the target markers.

Genetic information acquisition

The confirmed primer sequences flanking the microsatellite loci were synthesized with a fluorescent label (FAM™, HEX™, or TAMRA™; Thermo Fisher Scientific) attached to the 5' end of each forward primer. The microsatellites were separately amplified by PCR to identify the highly polymorphic loci. The amplification systems and conditions were similar

to those described above for unlabeled primers.

After amplification, 1 μL of amplified fragment mix was added to 0.5 μL of loading buffer (blue dextran, 50 mg mL^{-1}; EDTA, 25 mM) and 4 μL of deionized formamide, and it was then denatured by incubation for 5 min at 95 ºC. An internal size standard (0.5 μL of Thermo ABI 4322682; Thermo Fisher Scientific) was added to each sample, and the fluorescently labeled PCR products were then separated by capillary electrophoresis (ABI3730xl Genetic Analyzer; Applied Biosystems Corp.). The fluorescently labeled DNA fragments were first analyzed with GENESCAN v. 3.7 (Applied Biosystems Corp.) ~~and then with~~followed by GENOTYPER v. 3.7 NT (Applied Biosystems Corp.). ~~Then it was~~The ~~fragement~~fragments were then automatically sorted according to the internal size standard. The numbers of microsatellite repeat motifs were calculated based on the amplicon lengths ~~of the amplicons~~ (Table S3).

Genetic information analysis

The genotypic data were initially processed in Microsoft Office Excel 2007, manually checked for errors, and then transformed into input files that were required for ~~the following~~subsequent analyses. The performance characteristics [observed heterozygosity (HO~~);~~), expected heterozygosity (HE~~);~~), PIC,~~:~~ estimated null allele frequency (F(null~~));~~)), Hardy-Weinberg equilibrium (HWE~~);~~), and ~~genotypic~~ linkage disequilibrium] were measured with GENEPOP v. 4.6 (Raymond & Rousset 1995; Rousset 2008) and CERVUS v. 3.0.7 (Kalinowski et al. 2010; Slate et al. 2000).

Multiplex PCR conditions

Genotyping of 236 yaks for 35 loci (Table 1) produced a core set of 17 loci with high PIC values (Table 2). Four multiplex PCR reactions were assembled, each containing four or five microsatellite markers. The primer sequences and concentrations used in the multiplex PCR reactions are shown in Table 2. The multiplex PCR reactions were performed with reaction mixtures ~~of~~ at a total volume of 15 μL, ~~comprising of~~ containing 25 ng of genomic DNA, 10

mM Tris-HCl (pH 8.3), 35 mM KCl, 1.8 mM $MgCl_2$, 5.0 mM each dNTPs, and 2.5 U of *Taq*™ Hot Start Version polymerase (TaKaRa Bio Inc.). Amplifications were performed in a Veriti 96-Well thermal cycler (Applied Biosystems Corp.) under the following conditions: 95 ºC for 5 min; 25 cycles of 95 ºC for 30 s, 55–58 ºC for 30 s, and 72 ºC for 30 s; 10 cycles of 95 ºC for 30 s, 53 ºC for 30 s, and 72 ºC for 30 s; and final extension at 72 ºC for 5 min. Genotyping of these microsatellite loci was performed as described above.

Parentage testing and individual identification

Exclusion probability values for parentage testing ~~was~~ were obtained using a likelihood-based method ~~with~~ based on ~~the~~ genotypic information. Five types of PEs were calculated for the set of loci in CERVUS v. 3.0.7. PE1 was defined as the average probability of excluding an unrelated candidate parent of an arbitrary offspring when the genotype of the other parent is unknown. PE2 was defined as the average probability of excluding an unrelated candidate parent of an arbitrary offspring when the genotype of the other parent is known. PEP was defined as the average probability of excluding a pair of unrelated candidate parents of an arbitrary offspring. PEI was defined as the average probability of differentiating two randomly-selected individuals. PESI was defined as the average probability of differentiating two randomly-selected full siblings (Kalinowski et al. 2010; Slate et al. 2000).

RESULTS

Microsatellite loci characteristics

After preliminary screening, a total of 35 microsatellite loci with the highest primer specificity ~~and allele numbers~~ among the 71 microsatellite loci were selected for further analysis. The number of alleles, allele size ranged, HO, HE, PIC, F(null), and HWE of the 35 microsatellite loci of polled yaks are presented in Table 1. Three monomorphic loci, ~~viz.,~~ namely BM2943, INRA035, and RM099, were identified. The remaining 32 loci were

polymorphic. The~~ir~~ number of alleles per locus ranged from ~~3~~ three (CSSM013 and CSSM033) to 12 (SPS115), and ~~the~~ PIC values ranged from 0.084 (MM12) to 0.815 (ILSTS028). Four of these 32 polymorphic loci had low PIC values (i.e., < 0.25), 11 had moderate PIC values (0.25 ~~to~~ 0.50), and 17 loci had high PIC values (~~i.e.,~~ > 0.50) (Table 1).

The 17 genetic markers with high PIC values were reproducible, informative, and locus-specific for parentage testing and individual identification. The number of alleles ranged from ~~5~~five (BM720, INRA005, TGLA227) to 12 (SPS115). Heterozygosity~~ies~~ values ranged from 0.573 (POTCHA) to 0.806 (ILSTS028), with an average of 0.662. Deviations between HO and HE ranged from 0.003 (INRA005) to 0.103 (TGLA126). The 17 loci had an average PIC values of 0.636, ranging from 0.529 (INRA005) to 0.815 (ILSTS028). The null allele frequency ranged from 0.0015 (TGLA227) to 0.0819 (TGLA126) (Table 1).

Hardy-Weinberg and ~~genotypic~~ linkage disequilibrium tests

The results of HWE tests of the 17 microsatellite loci ~~showed~~ indicated no ~~significance~~ significant differences (P > 0.05) (Table 1). Therefore, the 17 loci with high PIC values (> 0.5) were selected for yak paternity testing. ~~Genotypic~~ Linkage disequilibrium within ~~the~~ polled yaks resulted in 136 comparisons~~. There were~~, and no ~~disequilibrium found in the~~ interlocus disequilibrium was detected between ~~each locus~~loci ~~within~~ on the same chromosome (Table S6).

Multiplex amplification and loading

~~The m~~Multiplex PCR reactions were organized to co-amplify ~~the~~ four groups of 4–5~~the four or five~~ loci with high PIC values. Non-overlapping allele lengths that presented the same fluorescent color label were selected. Typical fluorescence signals of the core 17 microsatellite loci are shown in Figure 1. Using the three available fluorescent colors, we multiplex-loaded and scored two groups of eight or nine loci in each run.

Parentage inference

The exclusion probabilities were calculated from the allele frequencies based on the underlying assumptions of HWE. The PE1 values of the 17 core loci ranged from 0.189 (BM720) to 0.503 (ILSTS028), with an average of 0.285. The average PE2 values of the 17 markers was 0.451, and ~~it~~ values ranged from 0.328 (INRA005) to 0.673 (ILSTS028). ~~For~~ Regarding the 17 core loci, the combined PE1 and PE2 values were 0.99718116 and 0.99997381, respectively. For the ~~alleged~~ putative parents, the combined PEP value was 0.99999998. ~~The c~~Combined PEI and PESI values were > 0.99999999 and 0.99999899, respectively (Table 2). Therefore, the PE values indicated that the discriminatory power of the 17 loci was high~~the identity and parentage PEs based on the 17 core loci were highly different~~.

DISCUSSION

Previous studies indicated that 4.3% of annual losses with regard to genetic gain during dairy breeding were caused by pedigree errors (10%), compared to simulation analysis of accurate paternity determination data~~Previous studies have indicated that 4.3% of the annual losses in genetic progress during dairy breeding was caused by 10%~~ (Israel & Weller 2000) ~~to 22% pedigree errors in some farms due to inappropriate management systems (Ron et al. 1996; Visscher et al. 2002)~~. In fact, the pedigree error rate in yaks was high due to incorrect paternity as ~~the~~ yaks feed primarily by grazing, thwarting ~~kinship~~ parentage attribution. Additionally, clerical and insemination errors and fading ink on artificial insemination records and labels might contribute to sample mixing, thereby, leading to pedigree errors. Therefore, it is necessary to identify and correct the pedigree through parentage testing and individual identification. These practices that aim at genetic improvement of yak are essential for generating reliable breeding programs ~~that aim at genetic improvement of yak~~. Several reports have been published on the use of microsatellite markers for cattle identification (Sharma et al. 2015; Zhao et al. 2017), but the performance characteristics of yak identification panels

have not yet been established. Parentage testing ~~in~~ for yak breeding ~~will~~ increases the profitability by improving the efficiency of selective breeding programs.

The most commonly used methods of livestock identification and parentage verification rely on microsatellites (Jan & Fumagalli 2016; Jeong et al. 2015; Wang et al. 2017). However, single nucleotide polymorphisms (SNPs) have been applied in the identification and parentage verification of swine (*Sus scrofa*) and cattle (*Bos taurus*) (Eggen 2012; Rohrer et al. 2007). A recent study debated on ~~whether~~ the use of SNPs ~~rather than~~ instead of microsatellites ~~should be used~~ for parentage verification (Kaiser et al. 2017). ~~But~~ However, at least 200 SNPs should be used for parentage testing to reduce false-negative results, and at least 700 SNPs are required to completely eliminate ~~the~~ false positives (Strucken et al. 2016). In addition, ~~a kinship~~ parentage analysis based on SNPs has few predictable statistical problems that must be ~~carefully~~ considered carefully and ~~appropriately~~ evaluated appropriately before substituting the classical STRs approach (Amorim & Pereira 2005). For these reasons, microsatellites are preferred over SNPs for parentage testing.

Most of the microsatellite markers used for cattle identification and parentage verification are dimeric repeat motifs (Carolino et al. 2009; Stevanovic et al. 2010; van de Goor et al. 2009; Zhao et al. 2017). Similarly, all microsatellite markers used in the present study were dimeric. Repeat patterns have advantages as well as disadvantages. Dimeric microsatellites might have mutations or stutter bands that ~~-~~present allele interpretation errors (Walsh et al. 1996). On the contrary, ~~D~~during PCR conduction, trimeric, tetrameric, and pentameric repeat motifs demonstrated lower stutter slippage efficiency than dimeric microsatellites (Gill et al. 2005), and they ~~they~~ ensured clear peak discrimination. On the ~~contrary~~other hand, an appropriate mutation rate might be beneficial for the verification of multigenerational parentage. This is because ~~the~~ mutations might occur between an ancestor and a descendant's assumed father/mother, but not between an assumed father/mother and offspring, ~~which thereby~~ thus ~~allows~~ allowing the identification of ~~the~~ true father/mother. For trimeric, tetrameric, pentameric, and hexameric microsatellites, ~~the~~ gaps occurring during the sequence variant visualization within the repeat units can result in larger bin sizes ~~larger~~ than ~~that~~ those

of dimeric microsatellites (Gill et al. 2000). However, ~~when~~ if the span of the sequence variants ~~was~~ is too wide, it becomes increasingly difficult to confine the microsatellite markers within a fluorochrome to a single multiplex PCR system. Furthermore, multimeric repeats can also be compound. All of ~~The~~ the microsatellite markers used in the present study were ~~all~~ dimeric containing homogeneous repeat motifs, ~~and~~ therefore, some markers can be labeled as one fluorochrome in a single multiplex PCR with unambiguous genotyping~~the genotyping remained unambiguous~~.

In order to use them in paternity testing and individual identification, ~~the~~ microsatellite loci must have rare null alleles, ~~to~~ and be ~~involved in~~at HWE, and the gametic association (linkage disequilibrium) should be absent. Null alleles are not amplified to ~~the~~ detectable levels ~~by~~ via PCR because of mutations ~~of~~ at primer binding sites (Kline et al. 2011). The frequency of null alleles is mainly estimated by Mendelian incompatibilities (Strucken et al. 2016) and by comparing the observed and expected number~~s~~ of homozygotes ~~in~~ at a locus (Dąbrowski et al. 2015). This fact must be considered when performing genotyping for parentage testing~~,~~ and when there is an apparent opposite homozygosity between parent and offspring. In the present study, the estimated null allele frequencies remained the highest for HEL10 (F(Null) = 0.703) and MGTG4B (F(Null) = 0.163). Therefore, these loci were excluded from the core set used for parental identification.

The formulae used to estimate the exclusion probabilities assume random mating, random association between alleles of different loci, and allele frequencies consistent with HWE. The 17 loci that ~~are~~ were selected as core microsatellite markers were ~~included in~~at HWE (Table 1), indicating that they can be used to calculate ~~the~~ PE values.

The microsatellites used in cattle parentage testing could not be directly applied to yak identification due to uncertainty ~~as to~~of whether the primers used for cattle would produce the desired results in yak. In the present study, the microsatellite primers were designed ~~according to~~based on the yak gene sequences, and ~~then~~ were then tested by PCR and electrophoresis. In addition, the allelic frequencies differed between the cattle and yak. Even among ~~the~~ cattle breeds, the microsatellite PIC values ~~have different~~ differed ~~values~~ (Mao et al. 2008), and the same might be applied to ~~the~~ yak breeds ~~of yak~~ (Zhang et al. 2008). Thus, while determining

the PE values of parentage testing and individual identification for a new breed, allele frequencies and PIC values should be calculated *de novo*.

The microsatellite loci with PIC value > 0.5 were selected for the identification panel. Nevertheless, the PIC values of these loci were still moderately lower than those used for cattle testing (Stevanovic et al. 2010; Vohra et al. 2017). Therefore, a high number of microsatellite markers were screened in the present study, and a set of 17 microsatellite markers proved to be sufficient ~~to~~ for ~~determine~~ determining ~~the~~ PE values. The combined exclusion probabilities of wrongly assigned sires were 99.718116%, 99.997381%, and 99.999998% for PE1, PE2, and PEP, respectively. Similar results were reported in Angeln dairy cattle, with 16 microsatellites and approximately 99.9% PE1 (Sanders et al. 2006). The PEP value calculated for the set of 17 microsatellites for parentage testing in the present study was 99.997381%, which was marginally higher than that obtained for Swiss yaks (99.5%) using 13 STR markers (Nguyen et al. 2005). Therefore, the set of microsatellite loci used in the present study ~~appeared to~~ significantly contributed ~~for~~ to parental identification in the polled yak population.

We used several mapped cattle microsatellites to develop the sets of yak loci that ~~are~~ were suitable for multiplex PCR amplification, and multiplex loading was conducted in a single run ~~in order~~ to reduce human errors, typing cost, and time. Nevertheless, selecting markers ~~comprising of a~~for a universal panel depends on the balance among the required panel accuracy, amplicon length, and ability to undergo a successful multiplex reaction. Multiplex PCR amplifications are technically more difficult than their single-locus counterparts, but they are less likely to transfer across species than single-locus amplifications. Several multiplex PCR and loading optimization methods have been investigated for ~~cattle~~ parentage testing of cattle. Nevertheless, these methods cannot be directly applied to yak identification. Therefore, we developed four multiplex amplifications~~,~~ (each containing four or five loci~~,~~) and two multiplex loads~~,~~ (each containing eight or nine loci~~,~~) ~~which were~~ running in two gel lanes. The main advantage of this system is that the ~~ranges of~~ allele length ranges do not overlap within the same fluorochrome.

Hence, we proposed that a combination of 17 microsatellites can yield a polled yak panel with enhanced processing efficiency, reliability, and utility. ~~Also~~Moreover, this system uses the standard genotyping methods of DNA fragment analysis technology. Combined with ~~the~~ likelihood-based parentage testing, these 17 markers will help ~~to~~ improve ~~the~~ breeding programs and accurately determine ~~the~~ polled yak pedigrees. If this system is used to identify polled yaks that are not the descendants of the expected breeding male yak, breeders can eliminate them from the breeding group to ensure genetic purity and breed improvement. On the other hand, if the semen samples of high-grade male yaks are mixed with those of other males, ~~then~~ the detection system can ~~be used to identify it by using~~ use DNA from semen and ~~the~~ blood ~~sample~~ of the high-grade male yak for identification. Multiplex systems can also be used to rapidly assess the ~~breeding population~~ history, structure, and diversity of the breeding population, and ~~for~~ these systems can reconstruct~~ing the~~ relationships among breeds. Furthermore, ~~T~~these ~~multiplexe~~multiplex systems might also be applied to other yak breeds ~~of the yak~~ with ~~similar~~ gene frequencies that are similar to that of the population tested in the present study.

Although the core set of microsatellite loci presented here was meaningful for yak parentage testing, this methodology ~~has~~ still has ~~some~~the following limitations: (a) ~~as~~ Since different yak breeds have different microsatellite genotypes, this core set was only suitable for parentage testing of polled yak; (b) ~~the~~ The PIC values of the microsatellite loci are not sufficiently high to reduce the number of loci, thus avoiding low PE values; and (c) ~~the~~ The number of panels for multiplex PCR reactions and multiplex loading ~~was~~ still remained too high, ~~which was~~leading to time-consuming and high cost. Therefore, a large number of microsatellites with high PIC values should be obtained for different yak breeds to develop efficient parentage test systems, with higher PE values and fewer markers. ~~Also~~ Furthermore, microsatellites should be suitable to each specific yak breed, and markers should be ~~possibly~~ screened ~~in a~~using single multiplex PCR reactions or multiplex loading.

CONCLUSIONS

A set of 17 microsatellite markers, ~~that can be~~which were assembled in~~to~~ ~~the~~ four multiplex PCR reaction systems and genotyped in ~~the~~ two multiplex loading systems, were identified and evaluated. The high variability displayed by these microsatellite loci demonstrated that ~~the~~ highly precise genotyping panels might be used ~~in~~ for individual genotyping, parentage verification, and individual identification. The microsatellites reported ~~herein~~ in this study ~~can~~ could also ~~assist in~~ be used ~~to studying~~ evaluate to evaluate ~~the~~ yak population structure, history, and diversity, which subsequently aids ~~for~~ the genetic improvement of domestic yak.

ACKNOWLEDGEMENTS

We thank Hongli Jiang from Beijing UBioLab Genetics Technology Co., Ltd. for technical advice on genotype evaluation and multiplex system~~s~~ development. We also thank Editage for improving the language of this manuscript. We thank the E~~e~~ditor and the anonymous Reviewers for their constructive comments on the manuscript.

REFERENCES

Amorim A, and Pereira L. 2005. Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs. *Forensic Sci Int* 150:17-21. 10.1016/j.forsciint.2004.06.018

Barsila SR, Kreuzer M, Devkota NR, Ding L, and Marquardt S. 2014. Adaptation to Himalayan high altitude pasture sites by yaks and different types of hybrids of yaks with cattle. *Livestock Science* 169:125-136. 10.1016/j.livsci.2014.09.004

Carneiro Vieira ML, Santini L, Diniz AL, and Munhoz CdF. 2016. Microsatellite markers: what they mean and why they are so useful. *Genetics and Molecular Biology* 39:312-328. 10.1590/1678-4685-gmb-2016-0027

Carolino I, Sousa CO, Ferreira S, Carolino N, Silva FS, and Gama LT. 2009. Implementation of a parentage control system in Portuguese beef-cattle with a panel of microsatellite markers. *Genetics and Molecular Biology* 32:306-311. 10.1590/s1415-47572009005000026

Coetzer WG, Downs CT, Perrin MR, and Willows-Munro S. 2017. Testing of microsatellite multiplexes for individual identification of Cape Parrots (Poicephalus robustus): paternity testing and monitoring trade. *Peerj* 5. 10.7717/peerj.2900

Dąbrowski M, Bornelöv S, Kruczyk M, Baltzer N, and Komorowski J. 2015. 'True' null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties and survey of possible improvements. *Mol Ecol Resour* 15:477-488.

Ding LM, Wang YP, Brosh A, Chen JQ, Gibb MJ, Shang ZH, Guo XS, Mi JD, Zhou JW, Wang HC, Qiu Q, and Long RJ. 2014. Seasonal heat production and energy balance of grazing yaks on the Qinghai-Tibetan plateau. *Animal Feed Science and Technology* 198:83-93. 10.1016/j.anifeedsci.2014.09.022

Eggen A. 2012. The development and application of genomic selection as a new breeding paradigm. *Animal Frontiers* 2:10-15. 10.2527/af.2011-0027

Estoup A, Jarne P, and Cornuet JM. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11:1591-1604. 10.1046/j.1365-294X.2002.01576.x

Gill P, Curran J, and Elliot K. 2005. A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. *Nucleic Acids Res* 33:632-643. 10.1093/nar/gki205

Gill P, Whitaker J, Flaxman C, Brown N, and Buckleton J. 2000. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci Int* 112:17-40. 10.1016/s0379-0738(00)00158-4

Hu Q, Ma T, Wang K, Xu T, Liu J, and Qiu Q. 2012. The Yak genome database: an integrative database for studying yak biology and high-altitude adaption. *BMC Genomics* 13:(7 November 2012)-(2017 November 2012).

Israel C, and Weller J. 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy cattle populations. *J Dairy Sci* 83:181-187.

Jan C, and Fumagalli L. 2016. Polymorphic DNA microsatellite markers for forensic individual identification and parentage analyses of seven threatened species of parrots (family Psittacidae). *Peerj* 4:e2416.

Jeong H, Choi B-H, Lee H-E, Gim J-A, and Kim H-S. 2015. Microsatellite analysis of genetic variation and structure in Korean and exotic dog breeds. *Genes & Genomics* 37:819-827. 10.1007/s13258-015-0313-2

Kaiser S, Taylor S, Chen N, Sillett T, Bondra E, and Webster M. 2017. A comparative assessment of SNP and microsatellite markers for assigning parentage in a socially monogamous bird. *Mol Ecol Resour* 17:183-193.

Formatted: Portuguese (Portugal)

Field Code Changed

Formatted: Portuguese (Portugal)

Formatted: Portuguese (Portugal)

Formatted: Portuguese (Portugal)

Kalinowski ST, Taper ML, and Marshall TC. 2010. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment (vol 16, pg 1099, 2007). *Molecular Ecology* 19:1512-1512. 10.1111/j.1365-294X.2010.04544.x

Kang SW, Lee SY, Chio DH, Kang HJ, Hu MB, and Yang YJ. 2016. Statistical analysis of alleles in 4703 thoroughbred racing horses using fifteen microsatellite DNA markers. *Journal of Animal Science* 94:88-88. 10.2134/jas2016.94supplement488x

Kline MC, Hill CR, Decker AE, and Butler JM. 2011. STR sequence analysis for characterizing normal, variant, and null alleles. *Forensic Science International-Genetics* 5:329-332. 10.1016/j.fsigen.2010.09.005

Li D, Chai Z, Ji Q, Zhang C, Xin J, and Zhong J. 2013. Genetic diversity of DNA microsatellite for Tibetan yak. *Hereditas (Beijing)* 35:175-184.

Linacre A, Gusmao L, Hecht W, Hellmann AP, Mayr WR, Parson W, Prinz M, Schneider PM, and Morling N. 2011. ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Science International-Genetics* 5:501-505. 10.1016/j.fsigen.2010.10.017

Mao Y, Chang H, Yang Z, Zhang L, Xu M, Chang G, Sun W, Song G, and Ji D. 2008. The analysis of genetic diversity and differentiation of six Chinese cattle populations using microsatellite markers. *J Genet Genomics* 35:25-32.

Medhammar E, Wijesinha-Bettoni R, Stadlmayr B, Nilsson E, Charrondiere UR, and Burlingame B. 2012. Composition of milk from minor dairy animals and buffalo breeds: a biodiversity perspective. *Journal of the Science of Food and Agriculture* 92:445-474. 10.1002/jsfa.4690

Mizuno S, Ishizaki T, Toga H, Sakai A, Isakova J, Taalaibekova E, Baiserkeev Z, Kojonazarov B, and Aldashev A. 2015. Endogenous Asymmetric Dimethylarginine Pathway in High Altitude Adapted Yaks. *Biomed Research International*. 10.1155/2015/196904

Nguyen T, Genini S, Ménétrey F, Malek M, Vögeli P, Goe M, and Stranzinger G. 2005. Application of bovine microsatellite markers for genetic diversity analysis of Swiss yak (Poephagus grunniens). *Anim Genet* 36:484-489.

Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM, Auvil L, Capitanu B, Ma J, Lewin HA, Qian X, Lang Y, Zhou R, Wang L, Wang K, Xia J, Liao S, Pan S, Lu X, Hou H, Wang Y, Zang X, Yin Y, Ma H, Zhang J, Wang Z, Zhang Y, Zhang D, Yonezawa T, Hasegawa M, Zhong Y, Liu W, Zhang Y, Huang Z, Zhang S, Long R, Yang H, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, and Liu J. 2012. The yak genome and adaptation to life at high altitude. *Nature Genetics* 44:946-+. 10.1038/ng.2343

Raymond M, and Rousset F. 1995. GENEPOP (VERSION-1.2) - POPULATION-GENETICS SOFTWARE FOR EXACT TESTS AND ECUMENICISM. *Journal of Heredity* 86:248-249.

Rohrer GA, Freking BA, and Nonneman D. 2007. Single nucleotide polymorphisms for pig identification and parentage exclusion. *Animal Genetics* 38:253-258. 10.1111/j.1365-2052.2007.01593.x

Rosa AJM, Sardina MT, Mastrangelo S, Tolone M, and Portolano B. 2013. Parentage verification of Valle del Belice dairy sheep using multiplex microsatellite panel. *Small Ruminant Research* 113:62-65. 10.1016/j.smallrumres.2013.03.021

Rousset F. 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8:103-106. 10.1111/j.1471-8286.2007.01931.x

Sanders K, Bennewitz J, and Kalm E. 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. *J Dairy Sci* 89:315-321.

Schnabel RD, Ward TJ, and Derr JN. 2000. Validation of 15 microsatellites for parentage testing in North American bison, Bison bison and domestic cattle. *Animal Genetics* 31:360-366. 10.1046/j.1365-2052.2000.00685.x

Sharma R, Kishore A, Mukesh M, Ahlawat S, Maitra A, Pandey A, and Tantia M. 2015. Genetic diversity and relationship of Indian cattle inferred from microsatellite and mitochondrial DNA markers. *BMC Genet* 16:73.

Slate J, Marshall T, and Pemberton J. 2000. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology* 9:801-808. 10.1046/j.1365-294x.2000.00930.x

Stevanovic J, Stanimirovic Z, Dimitrijevic V, and Maletic M. 2010. Evaluation of 11 microsatellite loci for their use in paternity testing in Yugoslav Pied cattle (YU Simmental cattle). *Czech Journal of Animal Science* 55:221-226.

Strucken EM, Lee SH, Lee HK, Song KD, Gibson JP, and Gondro C. 2016. How many markers are enough? Factors influencing parentage testing in different livestock populations. *Journal of Animal Breeding and Genetics* 133:13-23. 10.1111/jbg.12179

van de Goor LHP, Panneman H, and van Haeringen WA. 2009. A proposal for standardization in forensic bovine DNA typing: allele nomenclature of 16 cattle-specific short tandem repeat loci. *Animal Genetics* 40:630-636. 10.1111/j.1365-2052.2009.01891.x

Vohra V, Sodhi M, Niranjan SK, Mishra AK, Chopra A, Kumar M, and Joshi BK. 2017. Characterization of rare migratory cattle and evaluation of its phylogeny using short-tandem-repeat-based markers. *Journal of Applied Animal Research* 45:355-363. 10.1080/09712119.2016.1194843

Walsh PS, Fildes NJ, and Reynolds R. 1996. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res* 24:2807-2812. 10.1093/nar/24.14.2807

Wang G, Chen S, Chao T, Ji Z, Hou L, Qin Z, and Wang J. 2017. Analysis of genetic diversity of Chinese dairy goats via microsatellite markers. *J Anim Sci* 95:2304-2313.

Wang L-L, Yu Q-L, Han L, Ma X-L, Song R-D, Zhao S-N, and Zhang W-H. 2018. Study on the effect of reactive oxygen species-mediated oxidative stress on the activation of mitochondrial apoptosis and the tenderness of yak meat. *Food Chemistry* 244:394-402. 10.1016/j.foodchem.2017.10.034

Weising K, Winter P, Huttel B, and Kahl G. 1998. Microsatellite markers for molecular breeding. *Journal of Crop Production* 1:113-143.

Wu J. 2016. The distributions of Chinese yak breeds in response to climate change over the past 50 years. *Animal Science Journal* 87:947-958. 10.1111/asj.12526

Zhang G, Chen W, Xue M, Wang Z, Chang H, Han X, Liao X, and Wang D. 2008. Analysis of genetic diversity and population structure of Chinese yak breeds (Bos grunniens) using microsatellite markers. *J Genet Genomics* 35:233-238.

Zhao J, Zhu C, Xu Z, Jiang X, Yang S, and Chen A. 2017. Microsatellite markers for animal identification and meat traceability of six beef cattle breeds in the Chinese market. *Food Control* 78:469-475. 10.1016/j.foodcont.2017.03.017