Validation of 17 microsatellite markers for parentage testing and individual identification in 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** Corresp., [1,2], **Ping Yan** Corresp. [1,2]

[1] Animal Scienece Department, Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu, China

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

Corresponding Authors: Xian Guo, Ping Yan
Email address: guoxian@caas.cn, pingyan63@126.com

Background. Yak (*Bos grunniens*) is the most important domestic animal for people living at high altitudes, especially on the Qinghai-Tibetan Plateau. Yak ordinarily feed by grazing. Yak grazing has a strong impact on pedigree record accuracy because it is difficult to control mating in grazing yak. Polled yak are less dangerous to herdsman than horned yak. For this reason, polled yak breeding is preferred by herdsmen. **Methods.** A standardized set of microsatellite loci was assessed and applied to parentage testing and individual identification in yak. Seventy-one microsatellite loci were selected from literature databases. Most of them were used for cattle. Thirty-five microsatellite loci generated excellent results from PCR amplification and were evaluated for parentage testing and individual identification in 236 unrelated polled yaks. Seventeen loci had polymorphic information content (PIC) > 0.5 and were in Hardy-Weinberg equilibrium without linkage disequilibrium. Of these 35 markers, the seventeen loci were used in a core set of multiplexed markers suitable for parentage testing and individual identification in polled yak. **Results.** The multiplex PCRs and loading systems provided very high exclusion probabilities (EP) determined from combined core set genotypes. The followed combined exclusion probabilities were obtained: EP1 (exclusion probability for one candidate parent when one confirmed parental genotype is missing) = 0.99718116; EP2 (exclusion probability for one candidate parent when one confirmed parental genotype is available) = 0.99997381; EPP (exclusion probability for one alleged parental genotype pair is available) = 0.99999998. The combined EI (exclusion probability of finding two identical genotypes) and combined ESI (exclusion probability of finding the identity of two siblings) were > 0.99999999 and 0.99999899, respectively. The results of this analysis indicated that this combination of seventeen microsatellite markers could improve process efficiency, reliability, and utility in paternity testing and individual identification. **Discussion.** Fluorescent detection primers can identify many microsatellite alleles that are applicable

in polled yak population genetics analyses. Fluorescence labeling is useful for multiplex amplification and loading systems, decreases detection time, reduces testing costs, and generates copious amounts of microsatellite genotype data. Polymorphic microsatellite loci are suitable for parentage/relatedness analysis in other yak breeds with locus allele frequencies similar to those of the polled yak.

1 **Validation of 17 microsatellite markers for parentage testing and individual identification**

2 **in domestic yak (*Bos grunniens*)**

4 Jie Pei[1,2], Pengjia Bao[1,2], Min Chu[1,2], Chunnian Liang[1,2], Xuezhi Ding[1,2], Hongbo Wang[1,2],

5 Xiaoyun Wu[1,2], Xian Guo[1,2] and Ping Yan[1,2]

7 [1] Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of

8 Agricultural Sciences, Lanzhou, China

9 [2] Key Laboratory for Yak Genetics, Breeding, and Reproduction Engineering of Gansu Province,

10 Lanzhou, China

12 Corresponding authors:

13 Xian Guo, guoxianlz@163.com

14 Ping Yan, pingyan63@126.com

ABSTRACT

Background. Yak (*Bos grunniens*) is the most important domestic animal for people living at high altitudes, especially on the Qinghai-Tibetan Plateau. Yak ordinarily feed by grazing. Yak grazing has a strong impact on pedigree record accuracy because it is difficult to control mating in grazing yak. Polled yak are less dangerous to herdsman than horned yak. For this reason, polled yak breeding is preferred by herdsmen.

Methods. A standardized set of microsatellite loci was assessed and applied to parentage testing and individual identification in yak. Seventy-one microsatellite loci were selected from literature databases. Most of them were used for cattle. Thirty-five microsatellite loci generated excellent results from PCR amplification and were evaluated for parentage testing and individual identification in 236 unrelated polled yaks. Seventeen loci had polymorphic information content (PIC) > 0.5 and were in Hardy-Weinberg equilibrium without linkage disequilibrium. Of these 35 markers, the seventeen loci were used in a core set of multiplexed markers suitable for parentage testing and individual identification in polled yak.

Results. The multiplex PCRs and loading systems provided very high exclusion probabilities (EP) determined from combined core set genotypes. The followed combined exclusion probabilities were obtained: EP1 (exclusion probability for one candidate parent when one confirmed parental genotype is missing) = 0.99718116; EP2 (exclusion probability for one candidate parent when one confirmed parental genotype is available) = 0.99997381; EPP (exclusion probability for one alleged parental genotype pair is available) = 0.99999998. The combined EI (exclusion probability of finding two identical genotypes) and combined ESI (exclusion probability of finding the identity of two siblings) were > 0.99999999 and 0.99999899, respectively. The results of this analysis indicated that this combination of seventeen microsatellite markers could improve process efficiency, reliability, and utility in paternity testing and individual identification.

Discussion. Fluorescent detection primers can identify many microsatellite alleles that are applicable in polled yak population genetics analyses. Fluorescence labeling is useful for

43 multiplex amplification and loading systems, decreases detection time, reduces testing costs, and

44 generates copious amounts of microsatellite genotype data. Polymorphic microsatellite loci are

45 suitable for parentage/relatedness analysis in other yak breeds with locus allele frequencies

46 similar to those of the polled yak.

47

48 **Keywords** Polled yak, Short tandem repeat, Simple sequence repeats, Forensic genetics,

49 Exclusion probabilities, Conservation biology, Genetic diversity

50

51 INTRODUCTION

52 Yak (*Bos grunniens*), a member of the Bovidae, has successfully adapted to the severe cold and

53 low oxygen levels characteristic of high altitude regions (~2,500-5,500 m) such as the Himalayas

54 in south-central Asia, the Qinghai-Tibetan Plateau, Mongolia, and Russia (Wu 2016). These

55 regions are known for their high elevations, pristine natural environments, and frequent seasonal

56 variations (Haynes & Yang 2013; Ishizaki et al. 2005; Mizuno et al. 2015). The ability of yak to

57 survive in such rugged natural environments is the result of numerous behavioral, physiological,

58 and genetic adaptations (Barsila et al. 2014; Ding et al. 2015; Hu et al. 2012; Huang et al. 2012;

59 Qiu et al. 2012; Wang et al. 2017b). For instance, yak have steady daily fasting heat production

60 levels at constant ambient air temperatures even with increasing altitude (Ding et al. 2014; Han

61 et al. 2003). Yak can thrive in an extreme environment such as the Tibetan Plateau where few

62 other animals can survive. Yak have made significant contributions to human life in the Tibetan

63 Plateau by providing meat, milk, fur, leather, transportation, and more (Hu et al. 2016; Lee et al.

64 2017; Medhammar et al. 2012; Wang et al. 2018).

65   The ~13 million domestic yak in China constitute ~90% of the global yak population. There

66 are fourteen yak breeds in China, of which one is artificial (Wu 2016). Polled yak strain have

67 been bred for many years at the foot of the Ashidan Mountain. Polling is very useful in herd

68 management because it reduces the risk of horn-inflicted injury or death among the herdsman.

69 Polled yak breeding is very valuable in pasturing areas. Accurate genealogical records help

70 estimate genetic parameters and improve breeding programs. They ensure efficient and effective

71 breeding progress and avoid excessive inbreeding. When yak graze, however, mating is very

72 difficult to control. Semen samples could be mislabeled in error during preparation, and mating

73 records might be confused as a result of clerical errors made during artificial insemination.

74 Therefore, accurate yak pedigree records compiled by paternity testing and individual

75 identification are essential in polled yak breeding.

76   Molecular markers like microsatellite DNA indicate the degree of genetic relatedness between

77 animals and facilitate paternity verification and individual recognition (Estoup et al. 2002; Zhang

78 et al. 2006). Microsatellites, namely short tandem repeats (STR) or simple sequence repeats

79 (SSR), are tracts of repetitive DNA in which certain motifs ranging in length from 1-10

80 nucleotides are repeated 5-50× (Carneiro Vieira et al. 2016). Microsatellite DNA can be used to

81 develop pedigreed animal populations and evaluate animal breeding to support genetic

82 improvement by selective breeding (Weising et al. 1998). The application of microsatellites as

83 molecular markers for animal identification and parentage verification is highly accurate and

84 effective in both routine applications and forensic casework (Budowle et al. 2005; Iyengar &

85 Hadi 2014; Linacre et al. 2011).

86   Recently, DNA analysis of microsatellite markers has been used to verify parentage in breed

87 registries and identify individual animals linked to a particular database or owner. Microsatellite

88 panels have been well characterized for cattle (Zhao et al. 2017), horses (Kang et al. 2016), sheep

89 (Rosa et al. 2013), dogs (Jeong et al. 2015) and parrots (Coetzer et al. 2017). Many highly

90 polymorphic microsatellites have been identified, characterized, and mapped for bovine breeds.

91 Marker panels have also been developed (Carolino et al. 2009). Microsatellite marker panels

92 have been used to identify and assign individuals to families and populations (Stevanovic et al.

93 2010). The application of microsatellites in the identification of yak relationships has seldom

94 been reported. Moreover, there are no reports on paternity testing for yak breeds. Consequently,

95 there is a strong requirement for yak identity control and parentage verification.

96   A highly informative set of DNA markers whose alleles can be easily scored is critical if it is

97   to be used effectively in paternity testing and individual identification. The objective of the

98   present study was to develop and evaluate multiplex microsatellite systems designed for the

99   purposes of polled yak parentage control and kinship analyses. Many of these microsatellite loci

100   are already commonly used to characterize genetic parameters for cattle. This study aimed to

101   validate a set of highly polymorphic microsatellites for use in parentage testing and individual

102   identification in polled yak. The results of this study can be used to develop powerful, efficient

103   genetic tools for breeders to verify parentage verification and match individual yaks when their

104   identities are ambiguous or missing.

105

MATERIALS AND METHODS

106

Marker selection and primer design

107

108   Microsatellites were selected from cattle breeding literature. The markers had to meet the

109   following criteria: (a) high PIC and heterozygosity; (b) a large number of alleles; (c) no known

110   null alleles; (d) homogeneous repeat motifs (Schnabel et al. 2000). A total of 71 bovine

111   microsatellite markers were selected, of which 65 were derived from cattle references and 6

112   exclusively for yak (Li 2004). All 71 microsatellites and their flanking sequences were found on

113   the cattle genome. Their corresponding sequences on the yak genome were sought. Primers for

114   most of the loci used in earlier studies were adapted to the yak genome because of mutations or

115   low scores. The following 14 loci did not need adjustment: BM1824, BM2113, BMS2533,

116   ETH121, ETH225, ILSTS008, INRA124, RM099, INRA126, UMN0103, UMN0307, UMN0920,

117   UMN2303, UMN3007, and UMN3008. The primers used in this study are listed in Table S1.

118

Sample collection

119

120   The polled yaks were selected from the herds on Ashidan Mountain in Qinghai province. To

121   avoid consanguinity, only samples without genetic relationship were selected in pedigree. Blood

122   was drawn from the jugular veins of 236 unrelated individuals, including 38 sires and 198 dams.

123   All yaks were handled in strict accordance with good animal practice according to *the Animal*

124   *Ethics Procedures and Guidelines of the People's Republic of China*, and the study was approved

125   by *the Animal Administration and Ethics Committee of Lanzhou Institute of Husbandry and*

126   *Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences* (Permit No. SYXK-

127   2016-0039). The blood was mixed with preservation buffer (1.5 mg mL^{-1} EDTA and 137 mmol

128   L^{-1} NaCl) at a 5:1 ratio. The blood was stored at -80 ℃ in an ultra-cold freezer prior to DNA

129   extraction.

130

131   **DNA extraction and quantification**

132   Genomic DNA was extracted from white blood cells digested with proteinase-K. The samples

133   were centrifuged at 5,000 g for 2 min. The clear aqueous layer was then transferred to a new test

134   tube and 0.5 mL of 10 mg mL^{-1} RNase A was added to it. The sample was then extracted with

135   phenol:chloroform:isoamyl alcohol (25:24:1) followed by chloroform. The DNA was

136   precipitated with ethanol then resuspended in 50 mL TE buffer (10 mM Tris-HCl and 1 mM

137   EDTA, pH 8.0). The extracted DNA was quantified using a NanoDrop 2000 fluorometer

138   (Thermo Fisher Scientific, Waltham. MA, USA). OD$_{260:280}$ = 1.7-1.9 and OD$_{260:230}$ = 2.0-2.2

139   (data not shown).

140

141   **Preliminary primer screening**

142   Unlabeled primer pairs (Table S1) were ordered for each microsatellite marker and sample

143   fragments were amplified. The PCRs were conducted in 20-μL aliquots consisting of 20-50 ng

144   genomic DNA, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM MgCl$_2$, 2.0 mM of each dNTP,

145   5 ng BSA, and 1.0 U *Taq*^{TM} Hot Start Version polymerase (TaKaRa Bio Inc., Kusatsu, Shiga,

146   Japan). Primer concentrations ranged from 1.0-5.5 μM. The thermal cycle parameters were 5 min

147   at 95 ºC followed by 35 cycles of 30 s at 95 ºC, 30 s at 55-58 ºC (Table S1), and 20 s at 72 ºC,

148   then a final extension step of 5 min at 72 ºC in a Verity PCR Veriti 96-Well Thermal Cycler

149   (Applied Biosystems Corp., Foster City, CA, USA). After amplification, the PCR products were

150   visualized on 2% agarose gel (Gene Tech Co. Ltd., Chai Wan, Hong Kong). A 5-μL sample was

151   mixed with 1 μL loading dye (TaKaRa Bio Inc., Kusatsu, Shiga, Japan). Electrophoresis was run

152   for 35 min at 85 V and the samples were visualized with an GelDoc™ XR+ gel imaging system

153   (Bio-Rad Corp., Hercules, CA, USA).

154   The amplification fragments with high specificity, high amplification efficiency, and loci that

155   were either non-syntenic or separated by >10 cM (to avoid genetic linkage) (Table S2) were

156   manually cut out of the agarose gel and send to the biological companies for sequencing. The

157   samples were forwarded to Invitrogen (Carlsbad, CA, USA) and Thermo Fisher Scientific

158   (Waltham, MA, USA) without prior purification. Fragment analysis was conducted on an ABI

159   3730xl automated sequencer (Applied Biosystems Corp., Foster City, CA, USA). Only forward

160   primers were used to sequence the targets. Repeat marker sequences were revealed with Sanger

161   sequencing (Invitrogen, Carlsbad, CA, USA, and Thermo Fisher Scientific, Waltham, MA, USA).

162

163

164   **Genetic information acquisition**

165   Confirmed primer sequences flanking microsatellite loci were synthesized with a fluorescent

166   label attached to the 5' end of each forward primer. The forward primers set was fluorescently

167   labeled with FAM™, HEX™, or TAMRA™ dyes (Thermo Fisher Scientific, Waltham, MA,

168   USA). Microsatellites were separately amplified by PCR to identify loci with high levels of

169   allelic polymorphism. The amplification systems and conditions were similar to those used with

170   the unlabeled primers described above.

171   After amplification, 1 μL amplified fragment mix was added to 0.5 μL loading buffer (blue

172   dextran, 50 mg mL^{-1}; EDTA, 25 mM) and 4 μL deionized formamide then denatured by

173   incubation for 5 min at 95 ºC. Then, 0.5 μL internal size standard (Thermo ABI 4322682;

174   Thermo Fisher Scientific, Waltham, MA, USA) was added to each sample. Fluorescently labeled

175   PCR products were identified by capillary electrophoresis (ABI3730xl Genetic Analyzer,

176   Applied Biosystems Corp., Foster City, CA, USA). The fluorescently labeled DNA fragments

177   were analyzed with GENESCAN v. 3.7 (Applied Biosystems Corp., Foster City, CA, USA) then

178   with GENOTYPER v. 3.7 NT (Applied Biosystems Corp., Foster City, CA, USA) and

179   automatically sorted according to the internal size standard. The numbers of the microsatellite

180   repeat motifs were calculated on the basis of the fragment lengths of the PCR product (Table S3).

181

182   **Genetic information analysis**

183   Genotypic data were initially processed in Microsoft Office Excel 2007, manually checked for

184   errors, and transformed into other input files for various analyses. Performance characteristics

185   [observed heterozygosity (HO); expected heterozygosity (HE); polymorphism information

186   content (PIC); estimated null allele frequency (F(null)); Hardy-Weinberg equilibrium (HWE);

187   and genotypic disequilibrium] were measured with GENEPOP v. 4.6 (Raymond & Rousset 1995;

188   Rousset 2008) and CERVUS v. 3.0.7 (Kalinowski et al. 2007; Kalinowski et al. 2010; Slate et al.

189   2000).

190

191   **Multiplex PCR conditions**

192   Genotyping 236 yak for 35 loci (Table 1) produced a core set of 17 loci with high PICs (Table 2).

193   Four multiplex PCRs were assembled, each of which contained four or five microsatellite

194   markers. A list of the primer sequences used in the multiplex appears in Table 2. The multiplex

195   PCR amplification was performed in 15-μL volumes per sample. Each of these consisted of ~25

196   ng genomic DNA, 10 mM Tris-HCl (pH 8.3), 35 mM KCl, 1.8 mM $MgCl_2$, 5.0 mM of each

197   dNTPs, and 2.5 U *Taq*™ Hot Start Version polymerase (TaKaRa Bio Inc., Kusatsu, Shiga,

198   Japan). The primer concentrations are shown in Table 2. For PCR amplification, a thermal cycler

199   (Veriti 96-Well, Applied Biosystems Corp., Foster City, CA, USA) was run at 95 ℃ for 5 min

200 followed by 25 cycles of 95 ℃ for 30 s, 55-58 ℃ for 30 s and 72 ℃ for 30 s, then 10 cycles of

201 95 ℃ for 30 s, 53 ℃ for 30 s, and 72 ℃ for 30 s, finishing with a final extension at 72 ℃ for 5

202 min. Genetic typing of these microsatellite loci was performed in a manner similar to that

203 described above.

204

205 **Paternity test and individual identification**

206 Different types of exclusion probability were determined with CERVUS v. 3.0.7. These included

207 (a) EP1, the exclusion probability for one candidate parent when one confirmed parental

208 genotype was missing; (b) EP2, the exclusion probability for one candidate parent given the

209 genotype of a known parent of the opposite sex; (c) EPP, the exclusion probability for one

210 candidate parent pair; (d) EI, the exclusion probability for the identity of two unrelated

211 individuals; and (e) ESI, the exclusion probability of finding the identity of two siblings

212 (Kalinowski et al. 2007; Kalinowski et al. 2010; Slate et al. 2000).

213   Three parental exclusion probabilities were calculated corresponding to different scenarios.

214 EP1 assumes that genotypes are known for the offspring and one putative parent but not the other.

215 EP2 assumes that genotypes are known for the offspring, one confirmed parent, and one putative

216 parent (both parents genotyped). EPP assumes that genotypes are known for the offspring and

217 one parent pair; it is the probability of a misattributed progeny (mismatched with both parents)

218 from two genotyped parents (Jamieson & Taylor 1997).

219

220 # RESULTS

221 **Microsatellite loci genetic characteristics**

222 The numbers of alleles, allele size ranges, HO, HE, PIC, F(null), and HWE for the 35

223 microsatellite loci for the polled yak are presented in Table 1. A total of 214 alleles were

224 identified in the polled yak population. Significant ($P < 0.001$) heterozygote deficits were

225   detected at loci BM2943, INRA035, and RM099 because they were monomorphic. For the 32

226   polymorphic loci left, the number of alleles per locus ranged from 3 (CSSM013 and CSSM033)

227   to 12 (SPS115) and the PIC varied from 0.084 (MM12) to 0.815 (ILSTS028). Four of the 32

228   polymorphic loci had PIC < 0.25 (low). Eleven of them had moderate PIC values ranging from

229   0.25 to 0.50. The remaining 17 loci had high PIC (>0.50) (Table 1).

230   The 17 genetic markers with high PIC were reproducible, informative, and robust for

231   parentage testing and identification program. The numbers of alleles ranged from 5 to 12 within

232   this subset of 17 polled yak loci. Heterozygosities ranged from 0.573 (POTCHA) to 0.806

233   (ILSTS028) with an overall average of 0.662. Deviations between observed and estimated

234   heterozygosity ranged from 0.003 (INRA005) to 0.103 (TGLA126). The 17 markers had an

235   average PIC of 0.636 ranging from 0.529 (INRA005) to 0.815 (ILSTS028). F(null) alleles were

236   estimated to be in the range of -0.0218 for YAK08 to +0.0819 for TGLA126 (Table 1).

237

238   **Hardy-Weinberg and genotypic disequilibrium tests**

239   Exclusion probabilities were calculated from the allele frequencies based on the underlying

240   assumptions of HWE. The HWE for the 17 microsatellite loci were not significant (P > 0.05).

241   The 17 loci with high PIC (>0.5) were selected for yak paternity testing. Non-random gamete

242   association to form genotypes also influences the allele frequencies used to calculate genotype

243   frequencies. In natural populations, this effect probably occurs because of population

244   substructuring. Genotypic disequilibrium within the polled yak resulted in 136 comparisons.

245   There were no interlocus disequilibria.

246

247   **Multiplex amplification and loading**

248   Multiplexes were organized such that four groups of four or five loci with high PIC were co-

249   amplified by PCR. Non-overlapping allele lengths were chosen in such a way that they would

250   have the same fluorescent color label. Typical fluorescence signals of detections for the core 17

251 microsatellite loci are shown in Figure 1. Using the three available fluorescent colors, we

252 multiplex-loaded and scored two groups of eight or nine loci in each single gel lane.

253

254 **Parentage inference**

255 The EP1 of the 17 markers ranged from 0.189 (BM720) to 0.503 (ILSTS028) and their average

256 value was 0.285. The average EP2 was higher than that for the EP1. The average for all markers

257 was 0.451 and the range was from 0.328 (INRA005) to 0.673 (ILSTS028). For all 17 markers,

258 the combined EP1 and EP2 were 0.99718116 and 0.99997381, respectively. For the alleged

259 parents, the combined EPP was 0.99999998. The combined EI and ESI were > 0.99999999 and

260 0.99999899, respectively (Table 2). Therefore, the identity and parentage exclusion probabilities

261 based on the 17 microsatellite loci were highly discriminant.

262

263 # DISCUSSION

264   Previous studies implied that 3-4% of the losses in genetic progress were caused by 10%

265 increases in pedigree errors per year in dairy breeding (Israel & Weller 2000). Nevertheless,

266 other studies indicated that these dairy breeding pedigree errors have increased to 22% since the

267 1980s (Ron et al. 1996; Visscher et al. 2002). In fact, the yak pedigree error rate may be > 22%

268 since yak feed primarily by grazing. On the other hand, wrong insemination records caused by

269 clerical- and insemination errors and by stain disappearance may also influence pedigree error

270 rates. Therefore, it is necessary to identify and correct the pedigree through paternity testing and

271 individual identification. These practices are essential for reliable breeding programs aimed at

272 the genetic improvement of yak. Several reports have been published on the use of microsatellite

273 markers for cattle identification (Sharma et al. 2015; Zhao et al. 2017), but the performance

274 characteristics of a yak identification panel have not yet been characterized. Paternity testing in

275 yak breeding would increase profitability by improving selective breeding program efficiency.

276   The method of choice for livestock identification and parentage verification has been based on

277   microsatellites (Jan & Fumagalli 2016; Jeong et al. 2015; Wang et al. 2017a). For swine (*Sus*

278   *scrofa*) and cattle (*Bos taurus*), single nucleotide polymorphisms (SNPs) have been applied in

279   identification and parentage verification (Eggen 2012; Rohrer et al. 2007). A recent study

280   debated whether SNPs rather than microsatellites should be used for parentage verification

281   (Kaiser et al. 2017). Nevertheless, the interpretation of complex parentage cases may require

282   >700 SNP markers. The simplified international exchange of microsatellite profiles involves

283   only 32 genotypes as opposed to the 1,400 genotypes for SNPs (Strucken et al. 2016). For these

284   reasons, microsatellites are preferred over SNPs for parentage testing.

285   The cattle identification and parentage verification panel recommended by the ISAG

286   (International Society for Animal Genetics) consists mainly of dimeric microsatellite repeat

287   motifs. Similarly, all of the microsatellite markers applied in the present study were dimeric. The

288   various repeat patterns have both advantages and disadvantages. Dinucleotide microsatellites

289   may have mutations or stutter bands which can introduce allele interpretation errors (Walsh et al.

290   1996). For trimeric-, tetrameric-, and pentameric repeat motifs, stutter slippage efficiency during

291   PCR is even lower (Gill et al. 2005), and ensures clear peak discrimination. On the other hand,

292   an appropriate mutation rate is beneficial for multigenerational parentage verification. For

293   trimeric-, tetrameric-, pentameric-, and hexameric microsatellites, gaps during sequence variant

294   visualization within repeat units result in larger bin sizes (Gill et al. 2000). When sequence

295   variants span too widely, however, it becomes increasingly difficult to confine the microsatellite

296   marker within the same fluorochrome to a single multiplex PCR system. Multimeric repeats can

297   also be compound. The microsatellite markers used in the present study were all dimeric with

298   homogeneous repeat motifs, so the genotyping was unambiguous.

299   Microsatellite loci must have few null alleles, Hardy-Weinberg balance, and no linkage

300   disequilibria if they are to be effectively used in exclusion probability determinations for

301   paternity testing and individual verification. Null alleles are loci that fail to amplify to detectable

302   levels by PCR because of primer binding site mutations and other errors (Dakin & Avise 2004;

303   Kline et al. 2011). Null alleles are estimated by comparing the observed and expected numbers

304   of homozygotes for a locus (Dąbrowski et al. 2015; Dakin & Avise 2004). There may be an

305   increased risk of null alleles for certain markers. This fact must be considered when genotyping

306   individual yak for paternity testing. For example, when there is a mismatch of an allele for which

307   both parent and offspring are homozygous, the allele should be sequenced to screen for null

308   alleles. The risks of null allele occurrence were highest for HEL10 (F(Null)) = 0.703) and

309   MGTG4B (F(Null)) = 0.163) in this study. Therefore, these loci were excluded from the core set

310   used in parental identification.

311   The formulae for estimating exclusion probabilities assume random mating and non-random

312   association among alleles and loci. The allele frequencies used in the calculation of genotype

313   frequencies and exclusion probabilities must be consistent with Hardy-Weinberg assumptions.

314   The 17 loci selected as core microsatellite markers were all in Hardy-Weinberg equilibrium. We

315   tested for linkage disequilibrium because linked loci are relatively ineffective at determining

316   paternity or individual exclusion and the yak locus locations on chromosomes are unknown.

317   Genotypic disequilibrium tests showed no consistent deviations in populations not expected *a*

318   *priori* to have them. None of the 17 loci was in disequilibrium, which means that they could be

319   used in polled yak exclusion probability (PE) determinations.

320   The microsatellites used in cattle parentage testing could not be applied to yak identification

321   directly because it was uncertain as to whether the primers used for cattle would produce the

322   desired results for yak. In addition, allelic frequencies differ between cattle and yak. In the

323   present study, the primers were designed according to yak gene sequences then tested by PCR

324   amplification and electrophoresis. There is a strong correlation between PIC and exclusion

325   probability. Even among various cattle breeds, the PIC of the same microsatellites have different

326   values (Mao et al. 2008; Vohra et al. 2017), and the same applied to yak breeds (Zhang et al.

327   2008). For this reason, when determining the exclusion probabilities of parentage testing and

328   individual identification for a new breed, the allele frequencies and PIC should be calculated *de*

329   *novo*. In the current study, the polled yak population was repeatedly tested on the proposed

330   microsatellite markers. Since some markers have similar allele frequencies and PIC among yak

331 breeds, they are expected to generate accurate results for other yak populations.

332   In the present study, microsatellite loci with high PIC (> 0.5) were selected for the

333 identification panels. Nevertheless, the PIC of these loci were still slightly lower than those used

334 for cattle research (Stevanovic et al. 2010; Vohra et al. 2017). Therefore, a greater number of

335 microsatellite markers were selected for the present study. The set of 17 microsatellites sufficed

336 for PE determinations. The combined probabilities of excluding a wrongly assigned sire were

337 99.718116%, 99.997381%, and 99.999998% for EP1, EP2, and EPP, respectively. These results

338 were similar to those reported for multiplexing systems in Angeln dairy cattle. They had 16

339 microsatellite markers and nearly 99.9% PE1 (Sanders et al. 2006).

340   We used several mapped cattle microsatellites to develop sets of yak loci apt for multiplex

341 PCR amplification and multiplex loading into a single gel lane. PCR-based methods, highly

342 variable microsatellite loci, and fluorescent-based genotyping establish a new standard for

343 parentage testing. Furthermore, multiplexing reduces microsatellite typing costs and time.

344 Nevertheless, selecting the markers comprising a universal panel depends on a balance among

345 the required panel accuracy, amplicon length, and ability to undergo a successful multiplex

346 reaction. Multiplex PCR amplifications are technically more difficult than their single-locus

347 counterparts. Multiplex amplifications are also less likely to transfer across species than single-

348 locus amplifications. There is also a relatively higher risk of introducing genotyping errors with

349 multiplex amplifications due to allelic drop-out or false allele production (Taberlet & Luikart

350 1999). Many multiplex PCR and loading optimization methods have been investigated for cattle

351 paternity testing. Nevertheless, these methods cannot be directly applied to yak identification.

352 Therefore, we developed four multiplex amplifications, each of which contained four or five loci

353 and two multiplex loads each containing eight or nine loci. They were run in two gel lanes. The

354 main advantage of this system is that the allele length ranges do not overlap within the same

355 fluorochrome. Multiplex primer sets to amplify 17 microsatellite markers and fluorescent dye

356 technology combined form a rapid and powerful yak parentage testing method.

357   We propose that a combination of 17 genetic markers could yield a polled yak panel with

358 enhanced processing efficiency, reliability, and utility. This system exploits the advancements

359 made in fragment analysis technology. These markers combined with likelihood-based parentage

360 testing will help improve breeding programs and accurately determine polled yak pedigrees. If

361 some of the polled yaks identified by this system were not descendants of the expected breeding

362 male yak, they would be eliminated from a breeding group by breeders to ensure genetic purity

363 and improvement. On the other hand, when semen samples of high-grade male yak are mixed

364 with those of other males, the detection system in this study could identify the semen using the

365 blood sample of the high-grade male yak. Multiplex systems can also be used for the rapid

366 assessment of breed population history, structure, and diversity, and for reconstructing

367 relationships among breeds. These multiplexes may also work in other yak breeds with gene

368 frequencies similar to that of the population tested in the present study.

369

370 **CONCLUSIONS**

371 The set of 17 microsatellite markers can be amplified in four multiplex PCR reaction systems

372 and genotyped in two multiplex loading systems. The high variability displayed by these

373 microsatellite loci demonstrates that highly precise genotyping panels could potentially be used

374 in individual genotyping, parentage verification, and individual identification. The

375 microsatellites authenticated in this study could also serve in studies of yak population structure,

376 history, and diversity. They are important resources for the genetic improvement of domestic yak.

377

378

379 **ACKNOWLEDGEMENTS**

380 We thank Hongli Jiang from Beijing UBioLab Genetics Technology Co., Ltd. for technical

381 advice on genotype judging and multiplex systems development. We also thank the staff of

382 Editage who helped improve the written English in this manuscript. We thank the anonymous

383 reviewers for their informative comments on the manuscript.

384

REFERENCES

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

Budowle B, Garofano P, Hellman A, Ketchum M, Kanthaswamy S, Parson W, van Haeringen W, Fain S, and Broad T. 2005. Recommendations for animal DNA forensic and identity testing. *International Journal of Legal Medicine* 119:295-302. 10.1007/s00414-005-0545-9

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.

Dakin EE, and Avise JC. 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93:504-509. 10.1038/sj.hdy.6800545

Ding LM, Chen JQ, Long RJ, Gibb MJ, Wang L, Sang C, Mi JD, Zhou JW, Liu PP, Shang ZH, Guo XS, Qiu Q, and Marquardt S. 2015. Blood hormonal and metabolite levels in grazing yak steers undergoing compensatory growth. *Animal Feed Science and Technology* 209:30-39. 10.1016/j.anifeedsci.2015.07.024

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

Han XT, Xie AY, Bi XC, Liu SJ, and Hu LH. 2003. Effects of altitude, ambient temperature and solar radiation on fasting heat production in yellow cattle (Bos taurus). *British Journal of Nutrition* 89:399-407. 10.1079/bjn2003783

424   Haynes MA, and Yang Y. 2013. Adapting to change: transitions in traditional rangeland management of Tibetan yak
425   herders in northwest Yunnan. *Environment, Development and Sustainability* 15:1065-1077.
426   10.1007/s10668-012-9426-9
427   Hu Q, Ma T, Wang K, Xu T, Liu J, and Qiu Q. 2012. The Yak genome database: an integrative database for studying
428   yak biology and high-altitude adaption. *BMC Genomics* 13:(7 November 2012)-(2017 November 2012).
429   Hu T, Cao J, Lee S, Ho K, Li X, Liu S, and Chen J. 2016. Physiochemical characteristics of indoor PM2.5 with
430   combustion of dried yak dung as biofuel in Tibetan Plateau, China. *Indoor and Built Environment* 25:737-
431   747. 10.1177/1420326x15586584
432   Huang XD, Tan HY, Long R, Liang JB, and Wright A-DG. 2012. Comparison of methanogen diversity of yak (Bos
433   grunniens) and cattle (Bos taurus) from the Qinghai-Tibetan plateau, China. *Bmc Microbiology* 12.
434   10.1186/1471-2180-12-237
435   Ishizaki T, Kolzumi T, Ruan Z, Wang ZG, Chen QH, and Sakai A. 2005. Nitric oxide inhibitor altitude-dependently
436   elevates pulmonary arterial pressure in high-altitude adapted yaks. *Respir Physiol Neurobiol* 146:225-230.
437   10.1016/j.resp.2004.12.002
438   Israel C, and Weller J. 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy
439   cattle populations. *J Dairy Sci* 83:181-187.
440   Iyengar A, and Hadi S. 2014. Use of non-human DNA analysis in forensic science: A mini review. *Medicine Science*
441   *and the Law* 54:41-50. 10.1177/0025802413487522
442   Jamieson A, and Taylor S. 1997. Comparisons of three probability formulae for parentage exclusion. *Anim Genet*
443   28:397-400.
444   Jan C, and Fumagalli L. 2016. Polymorphic DNA microsatellite markers for forensic individual identification and
445   parentage analyses of seven threatened species of parrots (family Psittacidae). *Peerj* 4:e2416.
446   Jeong H, Choi B-H, Lee H-E, Gim J-A, and Kim H-S. 2015. Microsatellite analysis of genetic variation and structure in
447   Korean and exotic dog breeds. *Genes & Genomics* 37:819-827. 10.1007/s13258-015-0313-2
448   Kaiser S, Taylor S, Chen N, Sillett T, Bondra E, and Webster M. 2017. A comparative assessment of SNP and
449   microsatellite markers for assigning parentage in a socially monogamous bird. *Mol Ecol Resour* 17:183-193.
450
451   Kalinowski ST, Taper ML, and Marshall TC. 2007. Revising how the computer program CERVUS accommodates
452   genotyping error increases success in paternity assignment. *Molecular Ecology* 16:1099-1106.
453   10.1111/j.1365-294X.2007.03089.x
454   Kalinowski ST, Taper ML, and Marshall TC. 2010. Revising how the computer program CERVUS accommodates
455   genotyping error increases success in paternity assignment (vol 16, pg 1099, 2007). *Molecular Ecology*
456   19:1512-1512. 10.1111/j.1365-294X.2010.04544.x
457   Kang SW, Lee SY, Chio DH, Kang HJ, Hu MB, and Yang YJ. 2016. Statistical analysis of alleles in 4703 thoroughbred
458   racing horses using fifteen microsatellite DNA markers. *Journal of Animal Science* 94:88-88.
459   10.2134/jas2016.94supplement488x
460   Kline MC, Hill CR, Decker AE, and Butler JM. 2011. STR sequence analysis for characterizing normal, variant, and
461   null alleles. *Forensic Science International-Genetics* 5:329-332. 10.1016/j.fsigen.2010.09.005
462   Lee JN, Wang Y, Xu YO, Li YC, Tian F, and Jiang MF. 2017. Characterisation of gene expression related to milk fat
463   synthesis in the mammary tissue of lactating yaks. *Journal of Dairy Research* 84:283-288.
464   10.1017/s0022029917000413

465   Li Q. 2004. Molecular phylogenetic analysis of yak in ChinaPh.D. Nanjing Agricultural University.

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

469   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
470   and differentiation of six Chinese cattle populations using microsatellite markers. *J Genet Genomics* 35:25-
471   32.

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

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

478   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,
479   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,
480   Zhang J, Wang Z, Zhang Y, Zhang D, Yonezawa T, Hasegawa M, Zhong Y, Liu W, Zhang Y, Huang Z, Zhang S,
481   Long R, Yang H, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, and Liu J. 2012. The yak genome and
482   adaptation to life at high altitude. *Nature Genetics* 44:946-+. 10.1038/ng.2343

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

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

487   Ron M, Blanc Y, Band M, Ezra E, and Weller J. 1996. Misidentification rate in the Israeli dairy cattle population and
488   its implications for genetic improvement. *J Dairy Sci* 79:676-681.

489   Rosa AJM, Sardina MT, Mastrangelo S, Tolone M, and Portolano B. 2013. Parentage verification of Valle del Belice
490   dairy sheep using multiplex microsatellite panel. *Small Ruminant Research* 113:62-65.
491   10.1016/j.smallrumres.2013.03.021

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

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

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

498   Sharma R, Kishore A, Mukesh M, Ahlawat S, Maitra A, Pandey A, and Tantia M. 2015. Genetic diversity and
499   relationship of Indian cattle inferred from microsatellite and mitochondrial DNA markers. *BMC Genet
500   16:73.

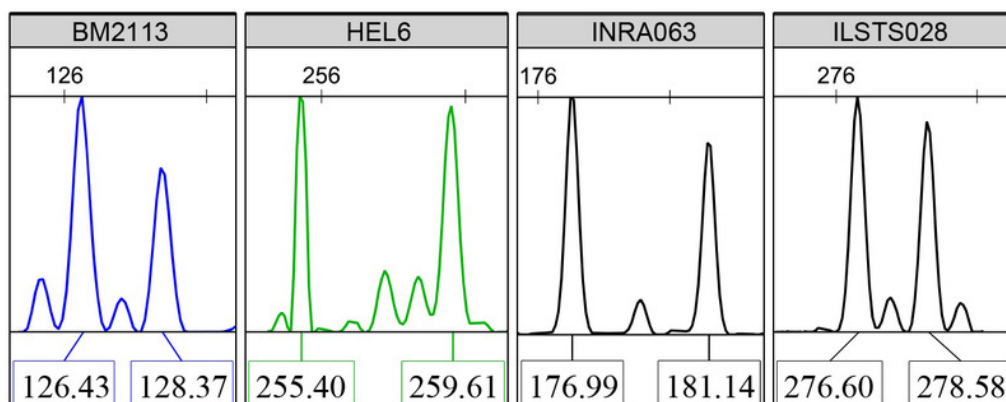
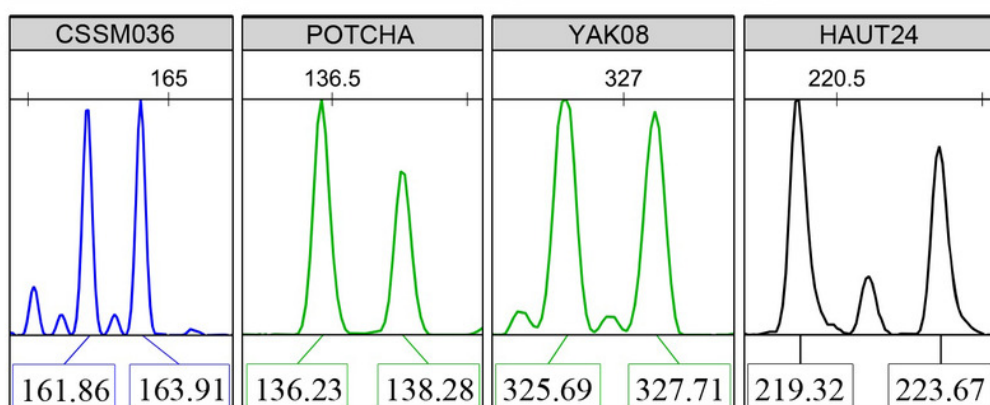
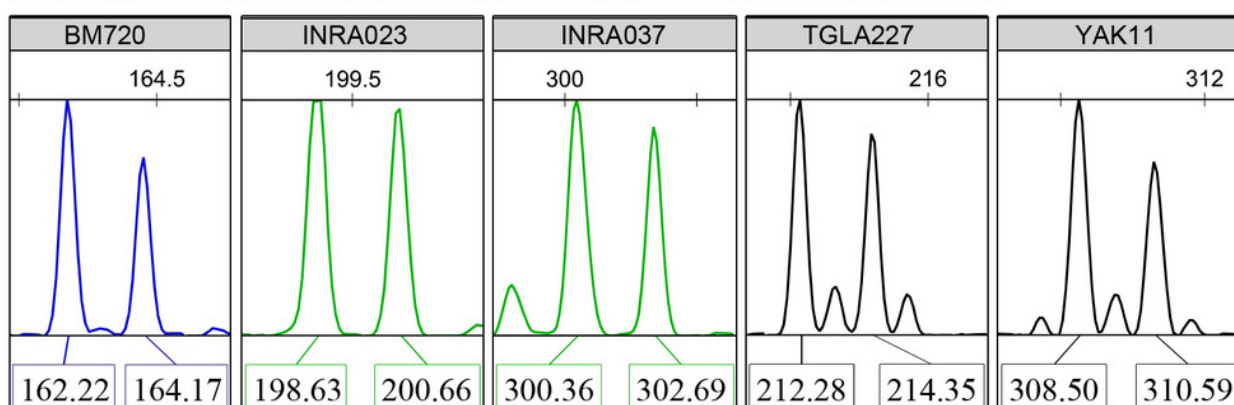
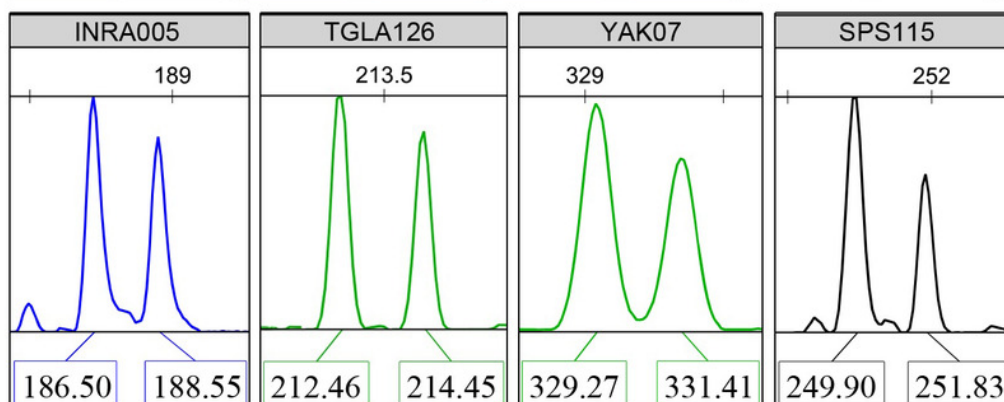
501   Slate J, Marshall T, and Pemberton J. 2000. A retrospective assessment of the accuracy of the paternity inference
502   program CERVUS. *Molecular Ecology* 9:801-808. 10.1046/j.1365-294x.2000.00930.x

503   Stevanovic J, Stanimirovic Z, Dimitrijevic V, and Maletic M. 2010. Evaluation of 11 microsatellite loci for their use in
504   paternity testing in Yugoslav Pied cattle (YU Simmental cattle). *Czech Journal of Animal Science* 55:221-
505   226.

506 Strucken EM, Lee SH, Lee HK, Song KD, Gibson JP, and Gondro C. 2016. How many markers are enough? Factors
507   influencing parentage testing in different livestock populations. *Journal of Animal Breeding and Genetics*
508   133:13-23. 10.1111/jbg.12179
509 Taberlet P, and Luikart G. 1999. Non-invasive genetic sampling and individual identification. *Biological Journal of*
510   *the Linnean Society* 68:41-55. 10.1111/j.1095-8312.1999.tb01157.x
511 Visscher P, Woolliams J, Smith D, and Williams J. 2002. Estimation of pedigree errors in the UK dairy population
512   using microsatellite markers and the impact on selection. *J Dairy Sci* 85:2368-2375.
513 Vohra V, Sodhi M, Niranjan SK, Mishra AK, Chopra A, Kumar M, and Joshi BK. 2017. Characterization of rare
514   migratory cattle and evaluation of its phylogeny using short-tandem-repeat-based markers. *Journal of*
515   *Applied Animal Research* 45:355-363. 10.1080/09712119.2016.1194843
516 Walsh PS, Fildes NJ, and Reynolds R. 1996. Sequence analysis and characterization of stutter products at the
517   tetranucleotide repeat locus vWA. *Nucleic Acids Res* 24:2807-2812. 10.1093/nar/24.14.2807
518 Wang G, Chen S, Chao T, Ji Z, Hou L, Qin Z, and Wang J. 2017a. Analysis of genetic diversity of Chinese dairy goats
519   via microsatellite markers. *J Anim Sci* 95:2304-2313.
520 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
521   species-mediated oxidative stress on the activation of mitochondrial apoptosis and the tenderness of yak
522   meat. *Food Chemistry* 244:394-402. 10.1016/j.foodchem.2017.10.034
523 Wang L-z, Wang Z-s, Xue B, Wu D, and Peng Q-h. 2017b. Comparison of rumen archaeal diversity in adult and
524   elderly yaks (Bos grunniens) using 16S rRNA gene high-throughput sequencing. *Journal of Integrative*
525   *Agriculture* 16:1130-1137. 10.1016/s2095-3119(16)61454-5
526 Weising K, Winter P, Huttel B, and Kahl G. 1998. Microsatellite markers for molecular breeding. *Journal of Crop*
527   *Production* 1:113-143.
528 Wu J. 2016. The distributions of Chinese yak breeds in response to climate change over the past 50 years. *Animal*
529   *Science Journal* 87:947-958. 10.1111/asj.12526
530 Zhang G, Chen W, Xue M, Wang Z, Chang H, Han X, Liao X, and Wang D. 2008. Analysis of genetic diversity and
531   population structure of Chinese yak breeds (Bos grunniens) using microsatellite markers. *J Genet*
532   *Genomics* 35:233-238.
533 Zhang X, Tong J, and Xiong B. 2006. Applications of microsatellite markers on studies of genetics and breeding of
534   fishes. *Journal of Agricultural Biotechnology* 14:117-121.
535 Zhao J, Zhu C, Xu Z, Jiang X, Yang S, and Chen A. 2017. Microsatellite markers for animal identification and meat
536   traceability of six beef cattle breeds in the Chinese market. *Food Control* 78:469-475.
537   10.1016/j.foodcont.2017.03.017
538
539

Figure 1

Typical fluorescence signal s of detections for the core 17 microsatellite loci.

The "Panels" stand for the multiplex PCR groups. The colors of fluorochrome 6-FAM, HEX, and TAMARA are showed by blue, green, and black respectively. The numbers above the sharp peaks represent length scales based on the internal size standard (bp). The numbers below the sharp peaks represent fragment lengths of PCR amplifications (bp).

Genetic information of the 35 polymorphic microsatellite loci labeled by fluorescence.

1   **Table 1   Genetic information of the 35 polymorphic microsatellite loci labeled by fluorescence.**

Locus ID	Range (bp)	Repeat motif	N_A	H_O	H_E	PIC	HW	F (Null)
BM720	160-168	AC	5	0.574	0.588	0.541	NS	0.0127
BM1818	239-251	AC	5	0.325	0.329	0.302	NS	-0.0031
BM2113	115-133	TG	8	0.673	0.662	0.623	NS	-0.0067
BM2943	253	AC	1	0	0	0	ND	ND
CSSM013	156-160	TG	3	0.294	0.309	0.272	NS	0.0180
CSSM029	182-194	AC	6	0.517	0.466	0.417	NS	-0.0623
CSSM033	309-317	TG	3	0.137	0.130	0.126	ND	-0.0261
CSSM036	157-179	TG	10	0.677	0.693	0.646	NS	0.0116
CSSM041	128-134	TG	4	0.457	0.428	0.402	NS	-0.0427
CSSME070	249-259	AC	5	0.513	0.526	0.412	NS	0.0093
HAUT24	217-231	AC	8	0.628	0.633	0.572	NS	-0.0059
HEL5	214-228	TG	6	0.182	0.181	0.174	ND	-0.0055
HEL6	247-265	GT	8	0.686	0.754	0.721	NS	0.0427
HEL10	171-177	TG	4	0.091	0.502	0.417	***	0.7025
ILSTS006	162-170	GT	4	0.453	0.448	0.374	NS	-0.0095
ILSTS008	172-184	AC	6	0.577	0.550	0.446	NS	-0.0257
ILSTS028	261-293	GT	11	0.806	0.838	0.815	NS	0.0194
ILSTS030	150-156	GT	4	0.106	0.129	0.123	ND	0.1033
INRA005	186-202	CA	5	0.607	0.604	0.529	NS	-0.0058
INRA023	184-204	TG	10	0.753	0.780	0.752	NS	0.0162
INRA035	104	TG	1	0	0	0	ND	ND
INRA037	296-312	CA	8	0.688	0.770	0.730	NS	0.0576
INRA063	175-187	TG	7	0.604	0.642	0.585	NS	0.0317
MGTG4B	245-255	AC	5	0.196	0.273	0.254	**	0.1631
MM12	109-121	GT	4	0.085	0.087	0.084	ND	0.0242
POTCHA	128-148	CA	8	0.573	0.585	0.540	NS	0.0227
RM099	233	CA	1	0	0	0	ND	ND
SPS115	231-261	CA	12	0.782	0.806	0.780	NS	0.0146
TGLA57	253-263	GT	6	0.348	0.340	0.322	NS	-0.0120
TGLA126	209-223	TG	8	0.587	0.690	0.639	NS	0.0819
TGLA227	210-222	AC	5	0.611	0.617	0.538	NS	0.0015
YAK07	323-339	TG	9	0.63	0.621	0.589	NS	-0.0104
YAK08	321-343	CA	8	0.698	0.676	0.612	NS	-0.0218
YAK11	306-314	GT	7	0.679	0.666	0.597	NS	-0.0096
YAK12	259-279	GT	8	0.549	0.522	0.406	NS	-0.0269

24   Note:

25   Range, range of allele sizes; Repeat motif, repeat motif of microsatellites; N_A, number of alleles found; H_O, observed

26   heterozygosity; H_E, expected heterozygosity; PIC, polymorphism information content; HWE, departure from Hardy-

27   Weinberg equilibrium; NS, not significant; ND, not done; **, significant at the 1% level; ***, significant at the 0.1%

28   level; F (Null), F value for null allel frequency.

Detail primer information of the core 17 microsatellite loci for multiplexes.

Table 2   Detail primer information of the core 17 microsatellite loci for multiplexes.

Locus name	Forward primer (5'→3')	Reverse primer (5'→3')	C [mM]	Ta (℃)	PMP	PML	Fluoro	RM	EP1	EP2	EPP	EI	ESI	CL
BM720	GAAATCAACAAGTTCCAATCCTG	ATCTCATTCTTGTGTCATGGATGA	3.5	56	3	2	6-FAM	(AC)	0.189	0.353	0.532	0.783	0.489	13
BM2113	GCTGCCTTCTACCAAATACCC	CTTCCTGAGAGAAGCAACACC	1.9	55	1	1	6-FAM	(TG)	0.261	0.439	0.636	0.847	0.542	2
CSSM036	GATAACTCAACCACACGTCT	AAGAAGTACTGGTTGCCAATCG	2.8	56	2	1	6-FAM	(TG)	0.283	0.456	0.644	0.859	0.561	27
HAUT24	CTCTGCCTTTGTCCCTGTCT	CCAAACCCCCTACCCACA	5.7	56	2	1	TAMARA	(AC)	0.221	0.377	0.552	0.805	0.517	22
HEL6	GGACACGACTGAGCAAGTAACA	GCTTTGGCAGGCAGATACAT	4.0	56	1	1	HEX	(GT)	0.367	0.549	0.743	0.907	0.603	1
ILSTS028	AGAAGAGTGTACCTCCTCCCAC	TCCAGATTTTGTACCAGACCAT	4.6	56	1	1	TAMARA	(GT)	0.503	0.673	0.847	0.953	0.656	11
INRA005	CTTCAGGCATACCCTACACCA	GGGGAATCTGTGGAGGAGTT	8.3	56	4	2	6-FAM	(CA)	0.190	0.328	0.483	0.768	0.494	12
INRA023	ATTTCCCTTCTGACTGGTACTTC	GTGTCCCTCCTCTAATCCCTAA	3.0	55	3	2	HEX	(TG)	0.408	0.589	0.782	0.924	0.620	3
INRA037	GCTACAATCCAGACTGAGCACG	GACACGGCTTAGCGACTGAA	3.1	57	3	2	HEX	(CA)	0.370	0.549	0.731	0.909	0.611	10
INRA063	AAACCACAGAAATGCTTGGAAG	ATTTGCACAAGCTAAATCTAACAA	3.6	56	1	1	TAMARA	(TG)	0.228	0.390	0.566	0.816	0.524	18
POTCHA	ATGCCAACTTTTCCCATCACT	GTAAACACAGTTCCCTGGAGAGA	3.5	56	2	1	HEX	(CA)	0.192	0.357	0.540	0.783	0.488	15
SPS115	AAAGTGACACAACAGCTTCACC	ACCGAGTGTCCTAGTTTGGC	4.6	55	4	2	TAMARA	(CA)	0.452	0.628	0.814	0.938	0.636	15
TGLA126	ATGAGAGAGGCTTCTGGGATG	CTTCACCATTGGACCACGAG	3.7	56	4	2	HEX	(TG)	0.272	0.444	0.625	0.854	0.558	20
TGLA227	CAAAGGAGCATAACTTTACAGCA	AGCCCTAACCATTGGACAGC	4.9	57	3	2	TAMARA	(AC)	0.200	0.337	0.494	0.775	0.501	18
YAK07	TAACAAAGCTGCTGGGAACAT	CGGAGTCACTTTCCTCACCTAT	2.4	56	4	2	HEX	(TG)	0.230	0.412	0.615	0.825	0.516	1
YAK08	ACTGGAGTAGGTTGCCCTGC	CCTGGCTTGGTCCTGTCTCT	3.8	56	2	1	HEX	(CA)	0.247	0.405	0.572	0.831	0.545	6
YAK11	TCCCCTCACTCCTCATTGGT	TGCAGGCAGTTTCTTACCAGT	4.4	56	3	2	TAMARA	(GT)	0.233	0.387	0.548	0.820	0.537	1
Combined									0.9971 8116	0.9999 7381	0.9999 9998	>0.999 99999	0.9999 9899	

Note:

C, concentration of primers; Ta, annealing temperature; PMP, panel numbers for multiplex PCR; PML, panel numbers for multiplex loading; Fluoro, fluorescent dye; RM, repeat motif of microsatellites; EP1, exclusion probability for one confirmed parent genotype missing; EP2, exclusion probability for one confirmed parent genotype available ; EPP, exclusion probability for one alleged parent pair genotype available ; EI, exclusion probability for finding two identical genotypes; ESI, exclusion probability for finding sib; CL, locations of microsatellite loci on cattle chromosoms.