Journal of Genetics (2019) 98:7
https://doi.org/10.1007/s12041-019-1061-x

RESEARCH NOTE

Establishing gene *Amelogenin* as sex-specific marker in yak by genomic approach

P. P. DAS[1], G. KRISHNAN[2], J. DOLEY[1], D. BHATTACHARYA[3], S. M. DEB[4], P. CHAKRAVARTY[1] and P. J. DAS[5]*

[1]*Indian Council of Agricultural Research-National Research Centre on Yak, Dirang 790 101, India*
[2]*Indian Council of Agricultural Research -National Institute of Animal Nutrition and Physiology, Bengaluru 560 030, India*
[3]*Indian Council of Agricultural Research -Central Island Agricultural Research Institute, Garacharama, Port Blair 744 101, India*
[4]*Indian Council of Agricultural Research -National Dairy Research Institute, Karnal 132 001, India*
[5]*Indian Council of Agricultural Research-National Research Centre on Pig, Rani, Guwahati 781 131, India*
*For correspondence. E-mail: drpranabjyotidas@gmail.com, Pranab.Das@icar.gov.in.

Received 20 June 2018; revised 3 October 2018; accepted 6 October 2018; published online 12 February 2019

Abstract. Yak, an economically important bovine species considered as lifeline of the Himalaya. Indeed, this gigantic bovine is neglected because of the scientific intervention for its conservation as well as research documentation for a long time. *Amelogenin* is an essential protein for tooth enamel which eutherian mammals contain two copies in both X and Y chromosome each. In bovine, the deletion of a fragment of the nucleotide sequence in Y chromosome copy of exon 6 made *Amelogenin* an excellent sex-specific marker. Thus, an attempt was made to use the gene as an advanced molecular marker of sexing of the yak to improve breeding strategies and reproduction. The present study confirmed that the polymerase chain reaction amplification of the *Amelogenin* gene with a unique primer is useful in sex identification of the yak. The test is further refined with qPCR validation by quantifying the DNA copy number of the *Amelogenin* gene in male and female. We observed a high level of sequence polymorphisms of *AMELX* and *AMELY* in yak considered as novel identification. These tests can be further extended into several other specialized fields including forensics, meat production and processing, and quality control.

Keywords. yak; Y-chromosome; *AMELX* gene; *AMELY* gene; marker; sex determination.

Introduction

Yak (*Bos grunniens*) is a semidomesticated, multifunctional bovid found in high altitudes of the Himalayan region. It provides milk, meat, wool, fuel, shelter and are also used as a pack animals for highlanders (Wang *et al.* 2014). Although similar chromosome number persist among yak and cattle, their survival capacity in extreme cold and harsh environments make yak economically viable animal. Here, we illustrated a reliable method for sex determination in the yak with X and Y chromosome-specific *Amelogenin* gene (*AMELX* and *AMELY*) and found *AMELY* gene as a signature point in the evolution of Y chromosome in the yak. The *Amelogenin* gene has undergone very less recombination throughout its evolution, thereby thought to represent the ancestral pseudoautosomal boundary (Sasaki and Shimokawa 2003) and thereby used in the study of chromosome evolution in mammals (Lahn and Page 1999; Pfeiffer and Brenig 2005; Gokulakrishnan *et al.* 2012). However, precise understanding and annotation of this gene in yak are yet to be studied. Few previous studies proposed that the mammalian sex chromosomes were evolved by four events and a series of inversions on the Y chromosome, and *Amelogenin* loci span ancient pseudoautosomal boundary in diverse mammalian species (Iwase *et al.* 2003; Grzybowski *et al.* 2006). In eutherian mammals, a copy of the *Amelogenin* gene is present in X and Y chromosomes with a significant sequence variation (Girondot and Sire 1998). The two copies of this gene appears due to its independent

duplication from X to Y chromosome (Sire *et al.* 2007), whereas in some other vertebrate a single copy is present only on X chromosome (Lahn and Page 1999). Bovine *Amelogenin* gene consists of six exons and a portion of about 63 nucleotides is deleted in Y linked copy of the gene (Girondot and Sire 1998; Das *et al.* 2009). This deletion appears to be an excellent tool for sex determination marker with considerable distinction in *AMELX* and *AMELY*. Till date, the different choice of methodologies has been developed to determine the sex of animal and animal products by specific gene analysis or hormone assessment (Pfeiffer and Brenig 2005). However, the DNA based techniques of sex determination in both human and animal have turned out to be more prominent in recent times. With the aim of similar hypothesis, this study was conducted to establish *Amelogenin* gene, a reliable sex-specific marker in yak by the genomic approach through PCR, quantitative PCR (qPCR), and comparative sequence analysis.

Materials and methods

Procurement of blood was done in accordance with the approval of Institute Animal Ethics Committee of ICAR-NRC on Yak, Dirang, India. The approved animal use protocol number is 4(17)/NRCY/IAEC-02 dated 01.08.2013.

The blood sample was collected from organized yak farm, and subsequently DNA was isolated using the standard protocol. Eight pairs of primers were designed from bovine genome spanning whole *AMELX* gene and one pair primer from X-linked androgen receptor (*AR*) as a positive control and four pairs from housekeeping gene using Primer3 software (http://www.primer3.ut.ee) (table 1). PCR was performed as per manufacturer protocol (Thermo Scientific) with a slight modification. All the eight amplified PCR products of female yak along with extra band amplified in male yak by primer-5 were extracted from agarose gel and purified with Qiagen gel extraction kit. A total of nine fragments, namely AMELX1-393bp, AMELX2-200bp, AMELX3-298bp, AMELX4-195bp, AMELX5-279bp and 216, AMELX6-94bp, AMELX7-289bp and AMELX8-103bp products size covering 9023 bp of *AMELX* were sequenced by standard double-stranded sequencing reaction and analysed in ABI3700 automated DNA Analyzer (ABI).

All the fragments of *Amelogenin* sequences of female yak and one male-specific *Amelogenin* sequence were submitted to NCBI databank (accession numbers KT826760 and KT835052), and these sequences were aligned with available cattle sequences (http://www.ncbi.nlm.nih.gov/). Phylogenetic and molecular evolutionary analyses were conducted using MEGA v6.0 (Thompson *et al.* 1994). The qPCR was performed using the Applied Biosystems StepOne Plus Real-Time PCR System (Life Technologies) to construct standard curve and estimation of ΔΔCT

Table 1. Detail of primer sequences, annealing temperature and product size of AMELX (different fragments), AMLEY and *AR*.

	Gene symbol	Forward primer	Reverse primer	T_m (°C)	Product size (bp)
1	*AMELX-1*	GCCAGTAATTTGCACCACAA	AGAACCAGGAGGCACAGAAA	60	393
2	*AMELX-2*	AACAAAAGTCTGCCCCACAG	CTCCTGGGGTGCTTGTTAAA	60	200
3	*AMELX-3*	GCCAACCAGAATTCCACAAT	CATGAGCAGCAACCAAAGAA	60	298
4	*AMELX-4*	CCAGTACCCTTCCTATGGTT	GGGTTGGAGTCATGGAGTGT	60	195
5	*AMELX-5*	CAGCCAAACCTCCCTCTGC	CCCGCTTGGCTTGTCTGTTGC	60	279, 216
6	*AMELX-6*	CACCTCTGCCTCCGATATTC	CCGCTTGGTCTTGTCTGTTG	60	94
7	*AMELX-7*	CCAGCAACAGACAAGACCAA	CTTTACGACCCCATGAATCG	60	289
8	*AMELX-8*	GGATTAAAAGATCAGAAAATGAGAAGA	AAGTAAGGTGGTGGTAGGAAAAA	60	103
9	*AR*	AGCAGCAACAGGAGACCAGT	TGCTTAAGCCTGGGAAAGTG	58	273
10	*TUBB2B*	TGTCCCTCGTGCTATCTTGGT	TGCTTAAGCCTGGGAAAGTG	58	180
11	*ACTB*	CGGCATCGAGGACAGGAT	CATCGTACTCCTGCTTGCTGAT	58	169
12	*TBP*	CAGAGAGCTCCGGGATCGT	CACCATCTTCCCAGAACTGAATAT	58	194
13	*GAPDH*	CCTGCCCGTTCGACAGATA	GGCGACGATGTCCACTTTG	58	150

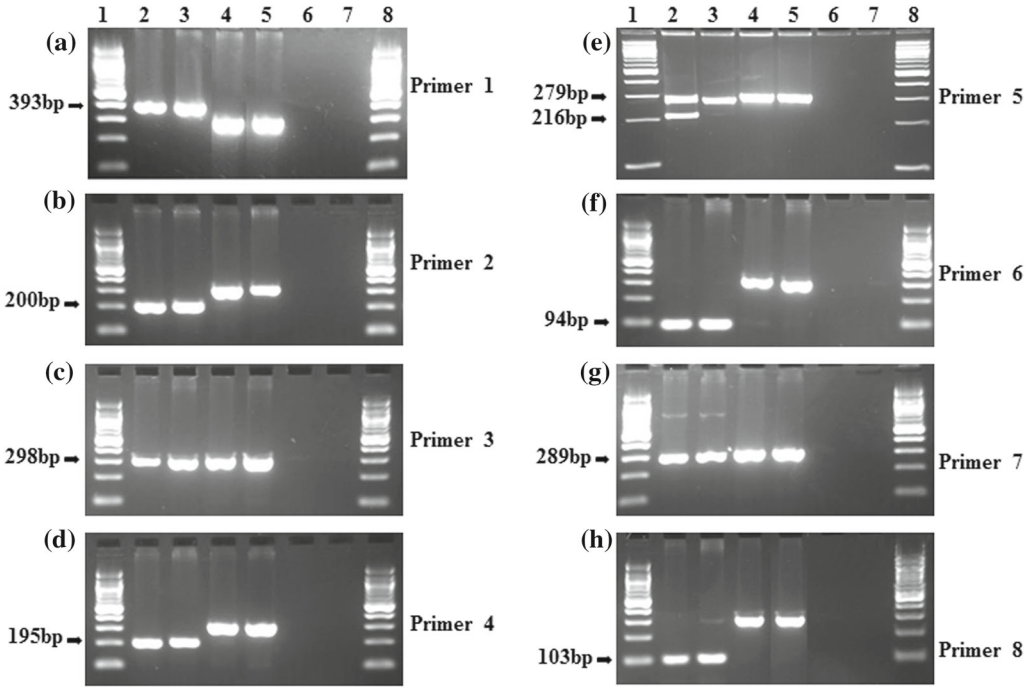

Figure 1. Agarose gel images showing PCR results. Amplification of product of (a) 393 bp with primer-1 (AMELX1) in lanes 2 and 3; (b) 200 bp with primer-2 (AMELX2) in lanes 2 and 3; (c) 298 bp with primer-3 (AMELX3) in lanes 2 and 3; (d) 195 bp with primer-4 (AMELX4) in lanes 2 and 3; (e) 279 and 216 bp with primer-5 (AMELX5) in lanes 2 and 3; (f) 94 bp with primer-6 (AMELX6) of in lanes 2 and 3; (g) 289 bp with primer-7 (AMELX7) in lanes 2 and 3 (h) 103 bp with primer-8 (AMELX8) in lanes 2 and 3. Lanes: 1 and 8, 100-bp marker; 2, female yak genomic DNA; 3, male yak genomic DNA; 4, *AR* gene as the positive control in female yak genomic DNA; 5, *AR* gene as the positive control in male yak genomic DNA; 6, negative control; 7, blank.

for quantification of copy number of *AMELX* gene. The DNAs from five male and five female yaks were amplified by primers 1, 2, 3, 4, 5, 6, 7 and 8 (table 1). Each reaction was run on triplicate at a final volume of 20 μL per reaction with 2 μL of 10-fold diluted DNA, 100 nM of each primer and 10 μL of SYBR Select Master Mix in 58°C, 2 min; 95°C, 2 min; 40 cycles of 95°C,15 s and 60°C, 1 min; followed by a melting curve program. The copy number variation of the gene and comparative results were analysed in StepOne plus software.

Results

Except for primer-5 (additional amplicon of 216 bp is amplified), all other primers produced similar product size in the male and female yaks along with positive control *AR*. Primer-5 gave two distinct amplified products in male but a single amplicon in the female. The *AR* gene amplified in both male and female genomic DNA of yak (figure 1, a–h). The sequencing of all the eight amplicons of *AMELX* (AMELX1, AMELX2, AMELX3, AMELX4, AMELX5, AMELX6, AMELX7 and AMELX8) from female yak and male-specific AMELX5 fragments of the male were successfully retrieved with expected product size. The qPCR with primers 2, 3, 4, 6, 7 and 8 in five male and five female genomic DNA distinguish the copy number in X and Y chromosomes. The result showed a significant higher copy number, almost equivalent to double in the female, compared to male ($P > 0.5$), in all the primers except primer-6.

A total 1851-bp sequence size covering 9023-bp cattle *AMELX* at different intervals were sequenced. Further, the sequencing data of yak *AMELX* was aligned with the cattle and other closely related eutherian mammals sequences using NCBI blast tools. Although uniformity prevails after alignments of *AMELX* with cattle for the fragments AMELX1 and AMELX8, few mismatches were observed in the fragments AMELX2, AMELX3, AMELX4 and AMELX5 at positions 2106.A>C, 3255.A>G, 6217.A>T, 6235.C>A, G.6657.C respectively. A deletion was found at position G. 6657.C in AMELX5 amplicon in the female yak, but not in cattle. A 63-bp deletion in male-specific fragments of AMELX5 at position C.6464-6525 is a novel finding in the yak. Further, eight sequence mismatches at positions 6416.T>C, 6431.C>T, 6442.A>G, C.6462-6525.A, 6555.G>C, 6562.A>G, 6583.A>G, 6598.A>G, 6610.TG>CA, 6637.C>T also observed when compared among Y and X specific fragment of AMELX5 in yak (figure 2, a&b).

The phylogeny was analysed with yak *AMELX* fragment5 from both male and female. Sequence alignment of *AMELX* isoform (fragment5) of yak with bison, buffalo, cattle, goat and sheep showed a *Bos* specific deletion at

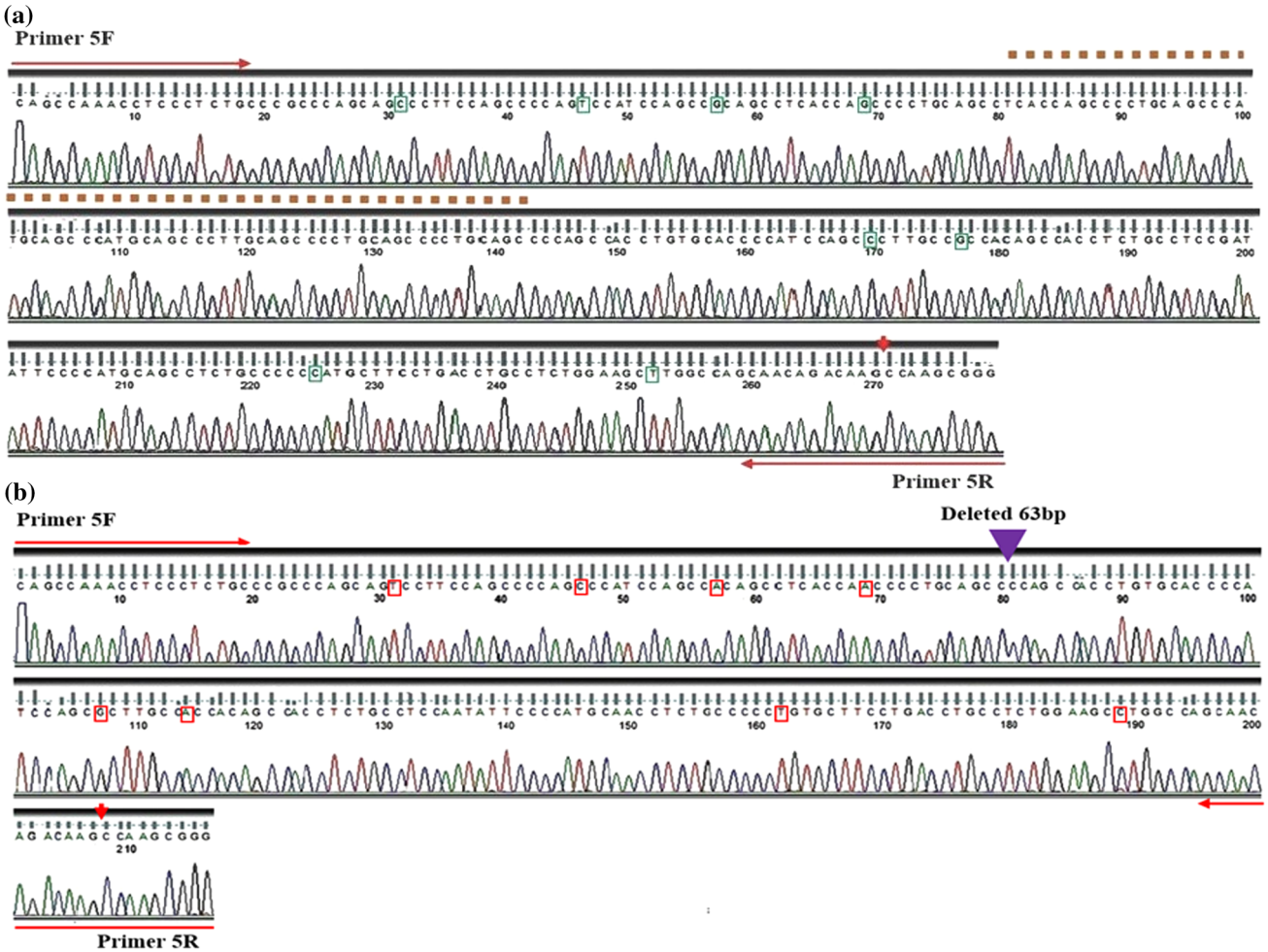

Figure 2. Chromatogram of sequenced yak. (a) *AMELX* isoform; (b) *AMELY* isoform confirmed that a 63 bp deletion in male-specific fragments at position C.81-143.C as well as eight sequence mismatch at positions G.31G>T, G.46T>C, C.57G>A, A.69G>A, C.170/106C>G, C.177/114G>A, C.225/163G>T, C.253/189T>C comparing between Y and X specific fragment of *AMEL* of with primer5. Vertical red arrows are indicating primer sitting site of both forward (F) and reverse primer (R), horizontal red arrows indicate deletion of single nucleotide comparing to cattle, brown dotted line indicated deleted sequences in Y specific *AMEL*, purple arrows indicates position of deletion in Y specific *AMEL*; green and red squares indicated mismatch position in Y and X specific fragment of *AMEL*, respectively.

position 126.C-135.C and yak specific deletion at position 278.A-280.C. The present study also revealed some distinctive deletions and substitutions in yak *Amelogenin* gene sequence when compared with its closely related eukaryotic species. Finally, a composite schematic diagram was constructed using all the available information on yak *Amelogenin* gene which confirms that the *Amelogenin* gene can be used as a molecular marker to determine the sex differences in yak (figure 3, a–f).

Discussion

Use of sophisticated tools for sex determination in an animal is always benefiting the breeding strategies and various scientific approaches. The present study revealed a sequence variation in *Amelogenin* gene could access as a phenomenal sex assurance marker in yak at the molecular level based on a sequence insertion and deletion attributes of X and Y-specific *Amelogenin* gene. It correlates with the previous study carried out in other vertebrates like cattle (Grzybowski *et al.* 2006), sheep (Pfeiffer and Brenig 2005), horse (Hasegawa *et al.* 2001), red deer (Gurgul *et al.* 2010) and human (Salido *et al.* 1992; Dutta *et al.* 2017). The differences in sequence length of X and Y copy of *Amelogenin* gene in yak were confirmed by the primer (AMELX-5), is first such report in the yak. Moreover, qPCR validation of DNA copy number in either sex enhances trials of the hypothesis. Double the copy number in female indicates X specific amplification. The sequencing data of yak *AMELX* may further facilitate the understanding of its biological importance, which provides the sequence differences with cattle *AMELX* genes. Sequencing results of few fragments of yak i.e.,

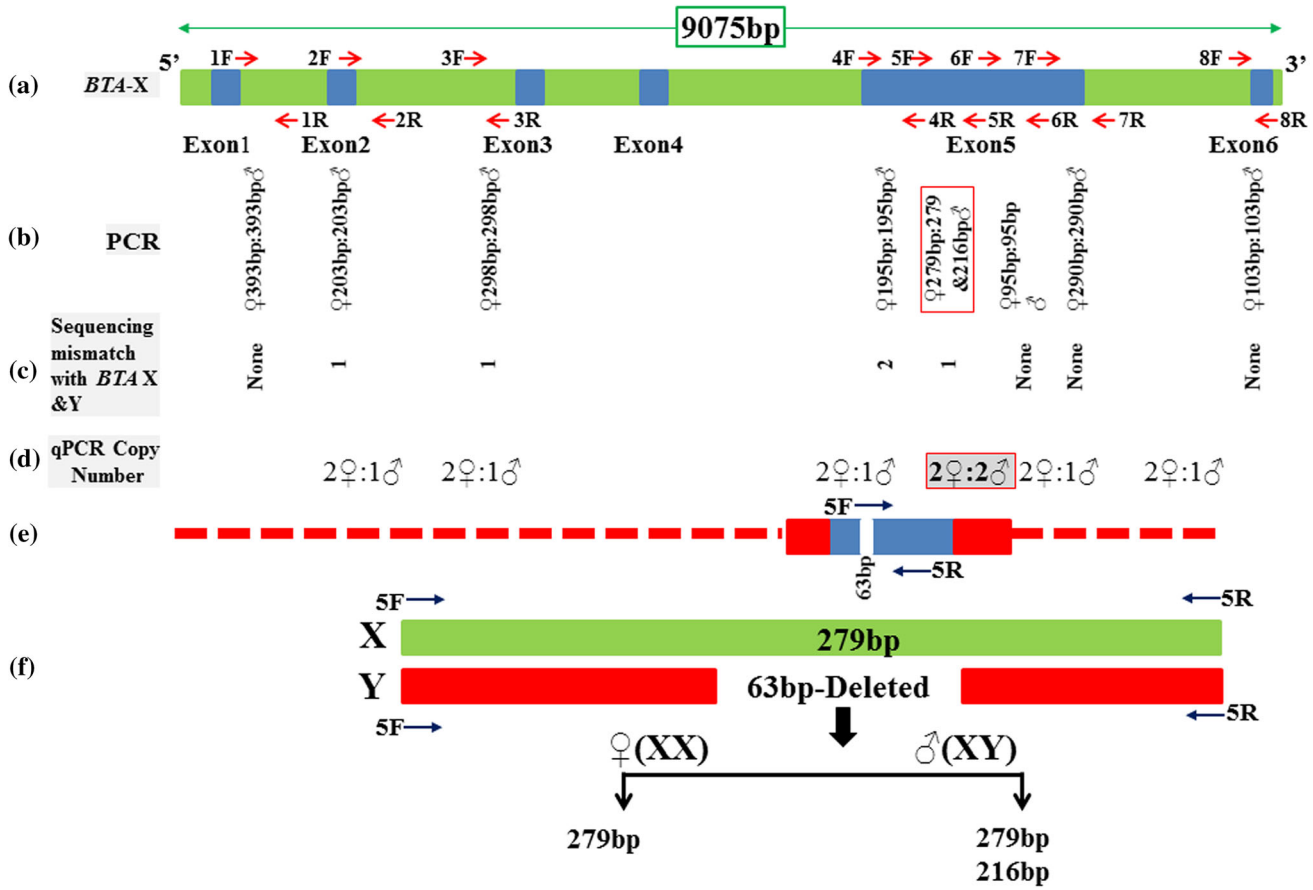

Figure 3. A schematic diagram of the *Amelogenin* gene as a sex-specific marker in the yak. (a) Schematic diagram of complete *AMELX* gene of cattle showing different exons in blue squares and introns in the green square. Vertical arrows in both direction showing forward (F) and reverse (R) primer sitting site of primers 1, 2, 3, 4, 5, 6, 7 and 8. Exact start and end sites of all primers. (b) Showing PCR product size in male and female yak genomic DNA amplified by primers 1, 2, 3, 4, 5, 6, 7 and 8. The red square box indicates sex-specific PCR product size in male and female genomic DNA by primer5. (c) Indicating sequence mismatches between yak and cattle *AMEL* gene. Numbers indicate mismatches observed under each fragment of *AMEL* sequences. (d) Showing copy number variation in male and female genomic DNA using primers 2, 3, 4, 7, 8 confirming 2:1 ratio in male and female. Primer6 in the red box is presenting equal copy number in male and female genomic DNA. (e) Schematic representation showing 63 bp deletion in exon 5 in Y chromosome-specific *AMEL* detected by primer5. (f) Schematic representation of Y (red) and X (green) chromosome-specific fragment of *AMEL* gene, showing deletion of 63 bp in male-specific *AMEL* gene and exact sequenced product size 279 bp in female and 279 and 216 bp in male genomic DNA.

AMELX1 and AMELX8 showed uniformity with cattle *AMELX* whereas, others i.e., AMELX2, AMELX3, AMELX4, AMELX5, AMELX6 and AMELX7 of yak witnessed few mismatches with cattle seem to be unique, as of Amelogenin protein found highly conserved in placental mammals (figure 2, a&b). We also observed sequence variations within male and female as well as between species in the yak. A 63-bp deletion in male-specific fragments at position C.81-143.C as well as eight sequence mismatches at positions G.31G>T, G.46T>C, C.57G>A, A.69G>A, C.170/106C>G, C.177/114G>A, C.225/163G>T, C.253/189T>C in between Y and X specific fragments of AMELX5 reflects a unique identity of yak evolutionary lineage (figure 2, a&b). This deletion may occur in the Y chromosome, which subsequently lost most of its gene sequences in the course of evolution. The *Amelogenin*, on other hands, could not be completely lost or else may owe a transposon-based partial duplication to other chromosomes. The qPCR results of primer-6 showed same DNA copy number in both male and female yak since primer-6 was designed from the identical sequence in X and Y chromosomes (figure 2, a&b). Further, the DNA copy number variation in *Amelogenin* gene for sex determination is established for the first time in yak along with its mapping and unique sequence information on the yak is considered as a novel phenomenon.

Acknowledgements

The authors are thankful to the Director, ICAR-NRCY for granting permission to carry out this project. The study was

supported by ICAR-NRC on Yak Institutional Project Grant 2013–2015.

References

Das P. J., Chowdhary B. P. and Raudsepp T. 2009 Characterization of the bovine pseudoautosomal region (PAR) and comparison with sheep, goat, and other mammalian PARs. *Cytogenet. Genome Res.* **126**, 139–147.

Dutta P., Bhosale S., Singh R., Gubrellay P., Patil J., Sehdev B. *et al.* 2017 *Amelogenin* gene-the pioneer in gender determination from forensic dental samples. *J. Clin. Diagn. Res.* **11**, ZC56.

Girondot M. and Sire J. Y. 1998 Evolution of the *Amelogenin* gene in toothed and toothless vertebrates. *Eur. J. Oral Sci.* **106**, 501–508.

Gokulakrishnan P., Kumar R., Sharma B., Mendiratta S., Sharm D. and Malav O. 2012 Sex determination of cattle meat by polymerase chain reaction amplification of the *Amelogenin* (*AMELX*/*AMELY*) gene. *Vet. World* **5**, 526–529.

Grzybowski G., Prusak B. and Romaniuk B. 2006 A novel variant of the *Amelogenin* gene (AMEL-X) in cattle and its implications for sex determination. *Anim. Sci. Pap. Rep.* **24**, 111–118.

Gurgul A., Radko A. and Słota E. 2010 Characteristics of X-and Y-chromosome specific regions of the *Amelogenin* gene and a PCR-based method for sex identification in red deer (Cervus elaphus). *Mol. Biol. Rep.* **37**, 2915–2918.

Hasegawa T., Sato F., Ishida N., Fukushima Y. and Mukoyama H. 2001 Sex determination by simultaneous amplification of equine SRY and *Amelogenin* genes. *J. Vet. Med. Sci.* **62**, 1109–1110.

Iwase M., Satta Y., Hirai Y., Hirai H., Imai H. and Takahata N. 2003 The *Amelogenin* loci span an ancient pseudoautosomal boundary in diverse mammalian species. *Proc. Natl. Acad. Sci.* **100**, 5258–5263.

Lahn B. T. and Page D. C. 1999 Four evolutionary strata on the human X chromosome. *Science* **286**, 964–967.

Pfeiffer I. and Brenig B. 2005 X-and Y-chromosome specific variants of the *Amelogenin* gene allow sex determination in sheep (Ovis aries) and European red deer (Cervus elaphus). *BMC Genet.* **6**, 16.

Salido E. C., Yen P., Koprivnikar K., Yu L. and Shapiro L. 1992 The human enamel protein gene *Amelogenin* is expressed from both the X and the Y chromosomes. *Am. J. Hum. Genet.* **50**, 303.

Sasaki S. and Shimokawa H. 2003 The *Amelogenin* gene. *Int. J. Dev. Biol.* **39**, 127–133.

Sire J.-Y., Davit-Béal T., Delgado S. and Gu X. 2007 The origin and evolution of enamel mineralization genes. *Cells Tissues Organs* **186**, 25–48.

Thompson J. D., Higgins D. G. and Gibson T. J. 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Wang K., Hu Q., Ma H., Wang L., Yang Y., Luo W. *et al.* 2014 Genome-wide variation within and between wild and domestic yak. *Mol. Ecol. Resour.* **14**, 794–801.

Corresponding editor: H. A. Ranganath