*Animal Science Papers and Reports* vol. 28 (2010), no. 4, 347-354
Institute of Genetics and Animal Breeding, Jastrzębiec, Poland

# The indel polymorphism in cattle amelogenin gene (*AMEL*) and its significance for the identification and evolutionary studies

**Beata Prusak*, Grzegorz Grzybowski**

Polish Academy of Sciences Institute of Genetics and Animal Breeding,
Jastrzębiec, 05-552 Wólka Kosowska, Poland

*(Received April 15, 2010; accepted July 1, 2010)*

**Nucleotide and amino acid sequences of cattle amelogenin, the major protein of forming enamel, were analysed by screening 1547 samples collected from cattle of various breeds. Sequenced was the part of cattle exon six of the *AMEL* gene (codons from 107 to 186) encompassing so-called hot spot of mutation in the mammals amelogenin gene. The analysis showed the presence of characteristic repeats of 9-nucleotide motifs (triplet PXQ) that makes its structure similar to STR sequences. In Polish Red (PR) cattle (37 cows and 6 bulls) a novel variant of 271 bp was detected on chromosome X. The alignment of sequences obtained showed that a novel variant of the amelogenin gene in PR cattle – *AMEL-X(271)* – is caused by deletion of one 9-nucleotide motif in position 485-493 (numbering according to Gibson *et al.* 1991). Irrespectively of the use of length polymorphism in the *AMEL* gene in early sex determination, the new possibilities in evolutionary and identificatory research are created by analysis of sequence structure. In the amelogenin gene the insertion-deletion of 9-nucleotide motifs can be localized in different sequence positions. As a result, sequences can differ by the number of tandem repeats as well as by the position of indels. Comparison of the *AMEL-X* gene sequences of *Bos/Bison* and *Bubalus* species revealed that distinguishing between the two is feasible not only based on the size of the PCR amplicon but also by the presence of additional 9-nucleotide motif in certain repeat region.**

**KEYWORDS: amelogenin / bovine species / cattle / gene polymorphism / sequence variation / sex determination**

*Corresponding author: b.prusak@ighz.pl

Amelogenin (AMEL), the origin of which has been traced back to the Precambrian period [Delgado *et al.* 2001] is the major component of enamel matrix [Termine *et al.* 1980, Fincham *et al.* 1992, Sasaki and Shimokawa 1995] and plays crucial role in enamel mineralization. In human mutations in the amelogenin gene (*AMEL*) sequence have been reported to be responsible for various hypoplastic and imperfect enamel formation called *amelogenesis imperfecta* [Wright *et al*. 2001, Hart *et al.* 2002]. The amino acid sequence of amelogenin protein is similar even in evolutionary distant lineages, e.g. in eutherians, monotremes, amphibians and reptiles [Toyosava *et al.* 1998, Sire *et al*. 2005]. Amelogenin *loci* span an ancient pseudoautosomal boundary in diverse mammalian species [Iwase *et al.* 2003] that allows for applying the analyses of polymorphism of the *AMEL* gene in phylogenetic studies on mammals. In the cattle genome, similar to other eutherian species, two amelogenin genes located on chromosome X (*AMEL-X*) and chromosome Y (*AMEL-Y*) are present [Iwase *et al.* 2001]. The *AMEL* gene is composed of seven exons, but two extra exons exist at least in humans and rodents [Li *et al.* 1998]. The central region of exon six is the most significant in identification and phylogenetic analysis. This region is considered a hot spot of mutation for the mammalian amelogenin [Delgado *et al.* 2005]. In cattle the 63 bp deletion in sequence of Y-linked gene has been used to distinguish samples of homogametic from heterogametic sex [Ennis and Galagher 1994, Lechniak and Cumming 1997]. High sequence similarity of the *AMEL* gene in the annealing regions of cattle PCR starters was also successively used to cross-amplification and sex determination in sheep *(Ovis aries)*, European red deer (*Cervus elaphus*) – Pfeiffer and Brenig [2005] – and Japanese black bear – Yamamoto *et al.* [2002]. Basing on the size of amplicon of the amelogenin gene on chromosome X in bovine species (*Bos taurus, Bos javanicus, Bos frontalis*), American bison, European wisent (amplicon 280 bp) and buffalo species (*Syncerus caffer, Bubalus bubalis* and *B. depressicornis*) – amplicon 289 bp – the distinguishing between taxon *Bos/Bison* and other *Bovidae* species has been done by Weikard *et al.* [2006]. In the present study, using primers described by Ennis and Galagher [1994] a new, hitherto not reported variant of the *AMEL-X* gene in cattle of 271 bp length is reported. In view of the research done so far the use of the *AMEL* gene in sex determination and distinguishing *Bos/Bison* species from *Bubalus* species only basing on amplicon size could be ambiguous. The aim of the present study was to determine the distribution of a novel variant of the *AMEL-X* gene in various cattle breeds, determine molecular background of its polymorphism, and to analyse potential implications of polymorphism discovered for the sex determination and phylogenetic analyses of *Bovidae*.

## Material and methods

### Samples

The material consisted of blood and semen samples collected from cattle with typical karyotypes of various breeds. In total tested were 1547 cows and bulls of the

following breeds: Polish Red (PR – 223 cows and 75 bulls – included in the Polish National Rare Livestock Breeds Preservation Programme), Polish Black-and-White (PBW with Holstein-Friesian blood and purebred Holstein-Friesian (1041 bulls), Red-and-White (RW (98 bulls), Polish Whitebacked (79 cows), Simmental (26 bulls) and Piedmontese (5 bulls).

**Amelogenin PCR assay and sex determination**

Genomic DNA from blood samples was extracted using Wizard® Genomic DNA Purification Kit (PROMEGA, USA), while the genomic DNA from semen samples with NucleoSpin Tissue Kit (MARCHERY NAGEL, Germany) according to the manufacturers' support protocol. The polymerase chain reaction (PCR) primers were those described by Ennis and Gallagher [1994] with the forward primer end labelled with FAM. PCR reaction was performed in a volume of 10 µl using 50 ng of template DNA, 1 U of AmpliTaqGold™ (APPLIED BIOSYSTEMS, USA), 1.5 mM MgCl2, 200 µM each dNTP (APPLIED BIOSYSTEMS, USA) and 0.05 µM of each primer. The PCR products were separated in 5% Long Ranger gel (FMC BIOPRODUCTS, USA) on an ABI PRISM 377 DNA Sequencer (APPLIED BIOSYSTEMS, USA) using GeneScan-500 internal size standard. Fragment sizes were determined using GeneScan v.3.1 software (APPLIED BIOSYSTEMS, USA).

**Sequencing of amelogenin amplicons**

The PCR products for *AMEL-X(280)* were purified through ultrafiltration using Microcon 100 microconcentrators (AMICON, USA). The PCR products for *AMEL-X(271)* and *AMEL-Y(217)* were excised from the 2% agarose gel and purified using NucleoSpin Extract (MARCHERY NAGEL, Germany). Purified PCR products were sequenced with BigDye® Terminator v1.1 Cycle Sequencing Kit (APPLIED BIOSYSTEMS, USA) according to the user's manual in a GenAmp PCR System 9600 Thermal Cycler (APPLIED BIOSYSTEMS, USA). The sequencing products were separated in APPLIED BIOSYSTEMS 3130 DNA Analyser. The electrophoretic data were collected by the Data Collection v.2.1. software and analysed by the Sequencing Analysis v.3.0. software (APPLIED BIOSYSTEMS, USA).

## Results and discussion

Typical pattern for the homogametic sex (XX karyotype) consisted of two fragments of 280 bp generating single detection signal while for the heterogametic sex (XY karyotype) comprised one fragment of 217 bp and one of 280 bp. This was observed in 1504 cows and bulls. Differently from that pattern, exclusively in PR cattle amplicon for the *AMEL-X* copy had size of 271 bp. It was observed in 37 cows (271 bp/280 bp – 35 cows and 271 bp/271 bp – 2 cows) and in 6 bulls. The frequency of the *AMEL-X* genotypes and alleles observed in PR cows is given in Table 1.

**Table 1**. Allele and genotype frequencies of the *AMEL-X* gene copy in Polish Red cows

| Cows pc (n) | Allele frequency | | Genotype frequency | Genotypes | | |
|---|---|---|---|---|---|---|
| | 280 | 271 | | 280/280 | 280/271 | 271/271 |
| 223 | 0.912 | 0.088 | observed | 0.838 | 0.1486 | 0.0135 |
| | | | expected | 0.8317 | 0.160 | 0.0077 |

Aiming at characterizing the sequence polymorphism, PCR products obtained for the X-271 bp and X-280 bp and Y-linked amelogenin genes were sequenced and deposited in GenBank (accession numbers: EU569298, EU569299, EU569300 respectively). The PCR product amplified covered the nucleotide sequence encoding the amelogenin protein between codon 107 and 186 where the main, conservative alternative splicing site is located. This intra-exonic splicing gives rise to the leucine-rich amelogenin peptide (LRAP) – Gibson *et al.* [1991]. The splicing consists in cutting PMQPLP motive localized in the acceptor site 3' after the codon encoding glutamine (Q = codon 171). Comparative analysis of the sequences obtained for the *AMEL-X(280* and *271)* in cattle showed high similarity to buffalo *AMEL-X(289)* variant (Fig. 1). Starting from the codon 171 all nucleotides and corresponding amino acids covering the LRAP protein are identical. LRAP proteins were found in several mammals and are suspected to play a role as regulatory molecules promoting bone and cartilage growth [Delgado *et al.* 2005]. Analyzing the sequence from the acceptor splicing site (codon 171) downstream to the codon 107, the high similarity reflecting evolutionary significance of the amelogenin structure in mammals was observed. The amelogenin is characterized by high content of proline (P) and glutamine (Q). Given the role these amino acids play in the peptide-peptide interaction [Williamson 1994], their high content could be important for improving the enamel resistance [Delgado *et al.* 2005]. Comparing the sequences of cattle *AMEL-X(280)*, *AMEL-X(271)*, *AMEL-Y(217)* variants and buffalo *AMEL-X(289)* variant with regard to codons for proline and glutamine the repetitive motif of nine nucleotides is observed (aa triplet PXQ). According to the analysis conducted in the present study the authors assume that the lack of 63 bp in amelogenin gene on chromosome Y in cattle is the deletion of seven triplet motifs of amino acids, while the newly identified in PR cattle variant of the amelogenin gene on chromosome X – *AMEL-X(271)* – is caused by a deletion of one triplet (Fig. 1). The *AMEL-X(289)* variant identified in buffalo [Weikard *et al.* 2006] has an additional nine-nucleotide motif CCC TTG CAG (triplet PLQ) comparing to the cattle variant *AMEL-X(280)* – Figure 1. The region of the amelogenin gene where the triplets PXQ are localized is called an *insertion locus* (according to numbering given by Delgado *et al*. [2005] *insertion locus* covers positions 133-159). To date, the *insertion locus* was observed in six species: tree shrew, guinea pig, goat, bovine, hedgehog, and opossum. The number of triplets (PIQ, PLQ or PMQ) ranges from 2 (tree shrew) to nine (opossum). After each round of repeat, proline (P) and glutamine (Q) are conserved in almost all sequences, while the central amino acid is sometimes

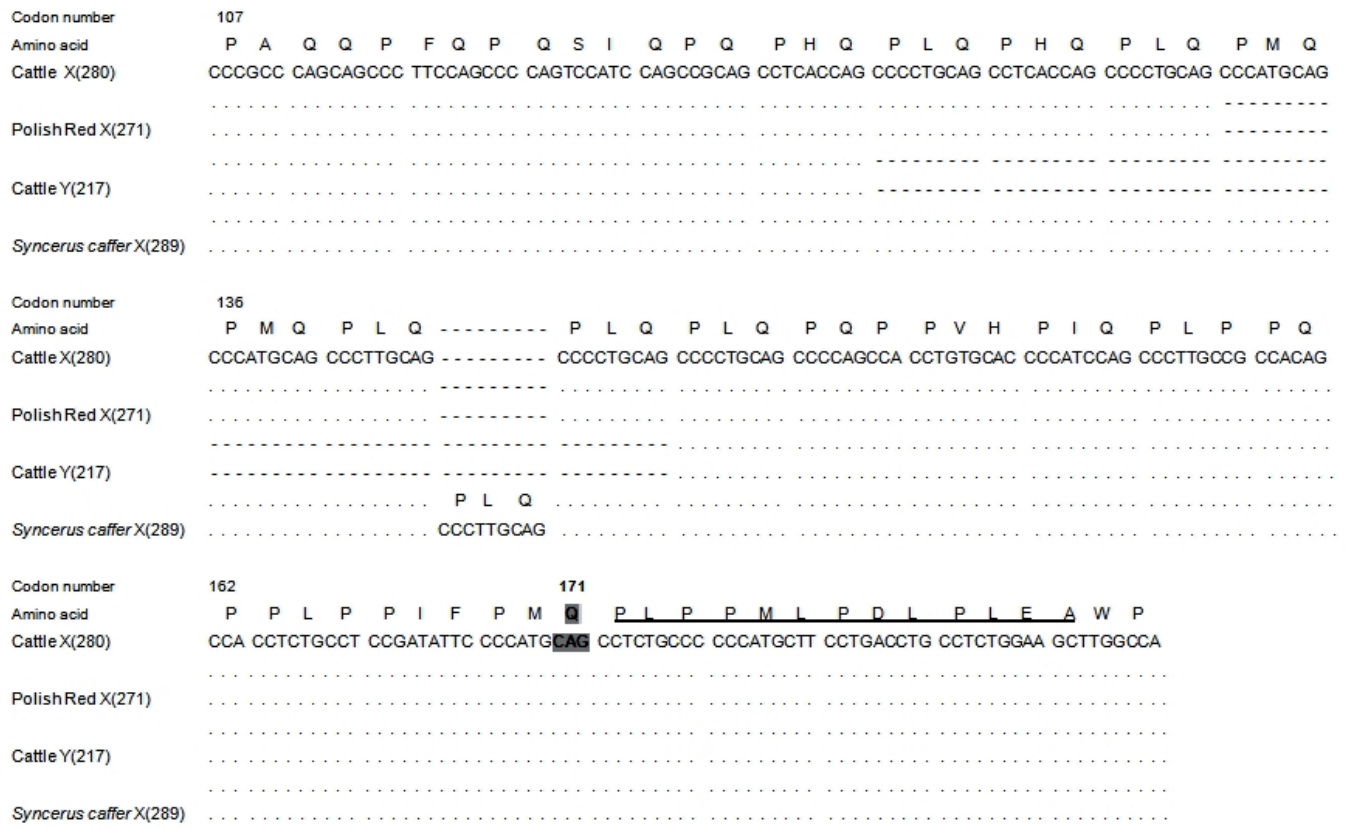


**Fig. 1.** Alignment of the variable region of exon 6 of the *AMEL* gene in Polish Red cattle (this study) and African buffalo *AMEL-X(289)* variant (*Syncerus caffer*, GenBank accession number DQ469604). (.) – identical residue, (-) – lacking residue, underlined – LRAP, shadowed – codon 171 (alternative splice acceptor site). Numbering according to the amelogenin sequence by Gibson *et al.* [1991].

substituted, suggesting that the central residue in these triplets is less constrained than P and Q [Delgado *et al.* 2005]. The triplets are always inserted between a Q and a P. The deletion of 63 bp in *AMEL*-Y gene copy, the deletion of 9 bp in *AMEL-X(271)* variant, and the insertion of nine nucleotide motifs in *AMEL-X(280)* in buffalo, all of them are localized exactly between Q and P. In the *insertion locus* 133-159 there are several possible positions of insertion/deletion (Fig. 1). It is thought that the repetitive character of this variable region in amelogenin, similarly to tandem repeats has been generated during evolution through the mechanism of slipped-strand mispairing (SSM). Domestication of several species has made the *Bovinae* one of the numerically most important taxonomic groups of mammals with 11 species: ox (*Bos taurus*), zebu (*Bos indicus*), gayal (*Bos frontalis*), gaur (*Bos gaurus*) banteng (*Bos javanicus*), Bison (*Bison bison*), wisent or European bison (*Bison bonasus*), yak (*Bos grunniens*), water buffalo (*Bubalus bubalis*), African buffalo (*Syncerus caffer*) and anoa (*Bubalus depressicornis*) – Lenstra and Bradley [1999]. Within this taxon, the earliest divergence 5 to 10 MYA separated water buffalo (*Bubalus bubalis*), African buffalo (*Syncerus caffer*) and anoa (*Bubalus depressicornis*) from the *Bos* and *Bison* species [Janecek *et al.* 1996]. In amelogenin evolution the trend to decrease the proline and glutamine content has been described [Delgado *et al.* 2005]. The *AMEL*-X gene variant identified in PR cattle (amplicon 271 bp) is the shortest known amelogenin sequence. It seems that the novel variant of *AMEL*-X (271 bp) can be considered as a characteristic of PR cattle. The results presented and GenBank database screening for amelogenin polymorphisms revealed that the *AMEL-X(271)* was not hitherto reported in other cattle breeds. Thus, it can be assumed that the presence of the *AMEL-X(271)* variant in PR cattle is not due to gene flow between breeds. The uniqueness of the PR cattle included in the PNRLBPP has also been confirmed earlier on the basis of STR *loci* [Grzybowski and Prusak 2004].

The presence of the novel *AMEL-X* variant in cattle throws a new light on the use of polymorphism of the amelogenin gene in gender identification and evolutionary analysis of *Bovidae.* Gender identification based on size differences in the bovine *AMEL* X- linked and *AMEL* Y-linked copies has, thus several limitations. The occurrence of the *AMEL*-X(271bp) variant requires special approach to sex identifying in cattle. It is possible, the atypical heterozygous (280bp/271bp) and atypical homozygous (271bp/271bp) genotype to occur for homogametic and atypical genotype (271bp/217bp) for heterogametic sex. Thangaraj *et al.* [2002] showed that the amelogenin gene is not a reliable method for gender assignment in forensic casework and prenatal diagnosis and can lead to false sex determination. Thus, the additional Y chromosome unambiguous markers should be included in the test to avoid confusion. The repetitive character of amelogenin sequence and mechanism of slipped-strand mispairing lead to insertion or deletion of the repeat motifs. Thus, the amplicon size can vary in species or breed (like in PR cattle) or be identical in distant lineages. In the *insertion locus* mentioned there are numerous positions where insertions and deletions can potentially occur. In consequence, obtaining the amplicon of typical size for certain species or lineage

could be insufficient for its correct identification. The position polymorphism, *i.e*. the position in sequence where the motif is inserted or deleted could be the deciding factor. It is also necessary to consider the introgression between species after domestication. For example, in Africa the introgression of Indian zebu bulls in taurine herds has occurred [Bradley *et al.* 1996]. The Chinese *yakow* is a yak-ox hybrid, which is kept at altitudes between the habitat ranges of parent species [Felius 1985]. The South-East Asian Selembu dairy and beef cattle result from gayal with zebu crossing [Felius 1985]. Taurine mitochondria have been found in wild North American bison populations [Ward *et al.* 1999] while the *beefalo* may be regarded as a domestic hybrid breed of bison and taurine cattle. The random mode of polymerase slippage mechanism which explains the production of indel polymorphism in the *insertion locus* in the *AMEL* gene in cattle leads to insertion/deletion of nine nucleotide motifs in variable positions in the sequence and in variable time and species. Moreover, the knowledge on the indel polymorphism in the amelogenin gene in *Bovidae* is limited. In view of introgression events in domesticated species and their wild ancestors and in view of gene flow that occurred in populations after domestication, the recommendation of the *AMEL* gene for evolutionary studies in family *Bovidae* as suggested by Brinkmann [2002] seems to be poorly justified.

**REFERENCES**


1. BRADLEY D.G., MACHUGH D.E., CUNNINGHAM E.P., LOFTUS R.T., 1996 – Mitochondrial diversity and the origins of African and European cattle. ***Proceedings of the National Academy of Sciences of the USA*** 93, 5131-5135.
2. BRINKMANN B., 2002 – Is the amelogenin sex test valid? ***International Journal of Legal Medicine*** 116(2), 63.
3. DELGADO S., CASANE D., BONNAUD L., LAURIN M., SIRE J-Y., GIRONDOT M., 2001 – Molecular evidence for Precambrian origin of amelogenin, the major protein of vertebrate enamel. ***Molecular Biology and Evolution*** 18(2), 2146-2153.
4. DELGADO S., GIRONDOT M., SIRE J-Y., 2005 – Molecular evolution of amelogenin in mammals. ***Journal of Molecular Evolution*** 60, 12-30.
5. ENNIS S., GALLAGHER T.F., 1994 – A PCR-based sex-determination assays in cattle based on the bovine amelogenin locus. ***Animal Genetics*** 25, 425-427.
6. FELIUS M., 1985 – Cattle Breeds: An Encyclopedia. Misset: Doetinchem, The Netherlands.
7. FINCHAM A.G., LAU E.C., SIMMER J. AND ZEICHNER D.M., 1992 – Amelogenin biochemistry – form and function. In: Chemistry and Biology of Mineralized Tissues (H. Slavkin and P. Price, Eds.), Elsevier, Amsterdam, 187-201.
8. GIBSON C., GOLUB E., HEROLD R., RISSER M., DING W., SHIMOKAWA H., YOUNG M., TERMINE J., ROSENBLOOM J., 1991 – Structure and expression of the bovine amelogenin gene. ***Biochemistry*** 30(4), 1075-1079.
9. GRZYBOWSKI G., PRUSAK B., 2004 – Genetic variation in nine European cattle breeds as determined on the basis of microsatellite markers. III. Genetic integrity of the Polish Red cattle included in the breeds preservation programme. ***Animal Science Papers and Reports*** 22 (1), 45-56.
10. HART P.S., HART T.C., SIMMER J.P., WRIGHT J.T., 2002 – A nomenclature for X-linked amelogenesis imperfecta. ***Archives of Oral Biology*** 47(4), 255-260.

11. IWASE M., SATTA Y., HIRAI Y., HIRAI H., IMAI H., TAKAHATA M., 2003 – The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. ***Proceedings of the National Academy of Sciences of the USA*** 100, 5258-5263.
12. IWASE M., SATTA Y., TAKAHATA N., 2001 – Sex-chromosomal differentiation and amelogenin genes in mammals. ***Molecular Biology and Evolution*** 18(8), 1601-1603.
13. JANECEK L.L., HONEYCUTT R.L., ADKINS R.M., DAVIS S.K., 1996 – Mitochondrial gene sequences and the molecular systematics of the artiodactyl subfamily Bovinae. ***Molecular Phylogenetics and Evolution*** 6, 107-119.
14. LENSTRA J.A., BRADLEY D.G., 1999 – In: F. Fries, A. Ruvensky (eds) Systematics and Phylogeny of Cattle, CAB International, Wallingford, Oxon, UK, pp 1-14.
15. LECHNIAK D., CUMMING I.R., 1997 – The use of amelogenin gene polymorphism in PCR embryo sexing in bovine IVF embryos. ***Journal of Applied Genetics*** 38(1), 45-49.
16. LI W., MATHEWS C., GAO C., DENBESTEN P.K., 1998 – Identification of two additional exons at 5' end of the amelogenin gene. ***Archives of Oral Biology*** 43, 497-504.
17. PFEIFFER I., 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 Genetics*** 6, 16.
18. SASAKI S, AND SHIMOKAWA H, 1995 – The amelogenin gene. ***International Journal of Developmental Biology*** 39, 127-133.
19. SIRE J-Y., DELGADO S., FROMENTIN D., GIRONDOT M., 2005 – Amelogenin: lessons from evolution. ***Archives of Oral Biology*** 50, 205-212.
20. THANGARAJ K., REDDY A.G., SINGH L., 2002 – Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? ***International Journal of Legal Medicine*** 116, 121-123.
21. TERMINE J.D., BELCOURT A.B., CHRISTNER P.V., CONN K.M., NYLEN M.U., 1980 – Properties of dissociatively extracted fetal tooth matrix proteins. I. Principal molecular species in developing bovine enamel. ***Journal of Biological Chemistry*** 255, 9760-9768.
22. TOYOSAVA S., O´HUIGIN C., FIGUEROA F., TICHY H., KLEIN J., 1998 – Identification and characterization of amelogenin genes in monotremes, reptiles, and amphibians. ***Proceedings of the National Academy of Sciences of the USA*** 95, 13056-13061.
23. WARD T.J., BIELAWSKI J.P., DAVIS S.K., TEMPLETON J.D., DERR J.N., 1999 – Identification of domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA markers and the parametric bootstrap. ***Animal Conservation*** 2, 51-57.
24. WEIKARD R., PITRA C., KÜHN C., 2006 – Amelogenin cross-amplification in the family bovidae and its application for sex determination. ***Molecular Reproduction and Development*** 73, 1333-1337.
25. WILLIAMSON M.P., 1994 – The structure and function of proline-rich regions in proteins. ***Biochemical Journal*** 297, 249-260.
26. WRIGHT J.T., DECKER S., PIDDINGTON R., HARRISON G., KULKARNI A.B., 2001 – Amelogenin-deficient mice display an amelogenesis imperfecta phenotype. ***Journal of Biological Chemistry*** 276, 31871-31875.
27. YAMAMOTO K., TSUBOTA T., KOMATSU T., KATAYAMA A., MURASE T., KITA I., KUDO T., 2002 – Sex identification of Japanese Black Bear, *Ursus thibetanus japonicus*, by PCR based on amelogenin gene. ***Journal* of *Veterinary Medical Science*** 64(6), 505-508.