

Molecular Gender Determination of Ancient Human from Malay Peninsular

¹Z.A. Shahrul Hisham, ¹S. Sahidan, ²M.A.W. Rohaya, ¹M.Y. Siti Afeefah, ¹Z.A. Intan Zarina,
¹J.A.N. Nor Hidayah, ¹R.M.A. Nadiah and ³Z.A. Zaidah

¹School of Biosciences and Biotechnology, Faculty of Science and Technology,
University Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

²Department of Orthodontic, Faculty of Dentistry, University Kebangsaan Malaysia,
Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia

³Department of Microbiology, Faculty of Applied Science, University Technology MARA,
40450 Shah Alam, Selangor, Malaysia

Abstract: Problem statement: DNA samples from fourteen modern human bloods (seven males and seven females) and two ancient skeletal samples excavated from Kalumpang Island and Dungun, Peninsular Malaysia were subjected for molecular genders determination using specific primers of human *AMELX* and *AMELY*. Approach: A standard multiple PCR mixture with forward primer and either a human X-specific reverse primer for *AMELX* or a human Y-specific reverse primer for *AMELY* amplifications were used to assess the presence of these genes in male and female samples. Results: PCR amplification of a modern male sample yielded 329 and 235 bp bands, whereas a modern female sample only 329 bp band. The Kalumpang Island sample produced two positive bands of *AMELY* and *AMELX*. After reamplification of the Dungun sample, only an *AMELY* band was visible. All amplified bands were cloned into TA plasmid and sequenced. BLAST analysis showed that the 329 bp band is *AMELX*, while the 235 bp product is *AMELY*. Conclusion/Recommendations: The skeletal remains of both Kalumpang Island and Dungun samples from west and east of Peninsular Malaysia respectively, are males.

Key words: Amelogenin gene, PCR, ancient bone, modern blood, sex marker

INTRODUCTION

Ancient DNA can be extracted from any formerly living material recovered from either excavation sites or museum specimens. The field of bio-molecular archaeology can include the analysis of any ancient biological material, defined as being older than 75 years. The first sequence of ancient DNA was successfully retrieved in 1984 from a museum specimen of an extinct quagga (a zebra-like animal), aged 140 years^[1]. Careful and well-established molecular biology techniques were used to obtain a 229 base pair (bp) sequence of its mitochondrial DNA. Subsequent analysis showed the quagga to be a very close relative of the zebra. This achievement was quickly followed by nuclear DNA recovery from 2400-year-old tissue collected from an Egyptian mummy^[2].

The main challenge faced by bio-molecular archaeologists is the minute quantity and low purity of DNA samples that can be recovered from ancient

remains for future analysis. As these remains are often exposed to extreme environments for decades, the DNA undergoes hydrolytic degradation followed by fragmentation, leaving only short fragments^[3,4]. The advent of PCR technology has helped to overcome this hurdle as this technique is sensitive enough that, theoretically, sequences from a single molecule of DNA can be amplified to millions of copies in a single optimized reaction.

Sex determination is of interest to this study, as it can help to obtain sex-specific demographic data and to further understand the gender-specific roles in a society. The Amelogenin (*AMEL*) gene is considered to be the best candidate for PCR-based sex determination, as it has variants at the same loci of both the X and Y chromosomes that are measurably different in size from one another^[5-7]. These are referred to as the *AMELX* and *AMELY* genes, respectively. Because only three primers are needed to amplify amelogenin from both of the sex chromosomes, determination of whether both

Corresponding Authors: Z.A. Shahrul Hisham, School of Biosciences and Biotechnology, Faculty of Science and Technology, University Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
Tel: +603-89213245 Fax: +603-89252698

AMEL gene variants are present also can be used as a molecular sex test. We also describe here a simple molecular sex analysis of ancient skeletal samples excavated from Peninsular Malaysia.

MATERIALS AND METHODS

Skeletal remains: The skeletal remains examined in the study were collected from two different locations. The first was from a swampy area of Kalumpang Island, situated in the west Peninsular Malaysia. The second was from the site of the shipwrecked Wanli, which was discovered at an approximate depth of 30 m in sea water around the coastal area of Dungun in the east Peninsular Malaysia. The bones were human samples based on the typical morphological of human bones. The samples were subjected to further analysis to verify the gender of each archeological remains.

Extraction of modern and ancient genomic DNA: The procedure to extract DNA from modern sources is distinct from that used to extract ancient DNA. This is due to the different types of starting material that were used for the protocol, as most modern DNA is extracted from peripheral blood while ancient DNA is extracted from skeletal remains. For the extraction of modern DNA, approximately 3 mL of whole blood samples were drawn from 14 healthy males and females who had given informed consent. Red Cell Lysis Buffer (100 nM Tris HCl, 10 mM MgCl₂ and 1 mM NaCl) was added to the blood with in a 9:1 ratio and the sample was spun at 660 g and 4°C for 5 minutes. This step was repeated until the resultant pellet became white. The pellet was dissolved in 2 mL of STE buffer (200 mM Tris HCl, 100 mM NaCl and 10 mM EDTA, 0.5% (w/v) SDS) containing 2.5 μL of Proteinase K (more than 3 U μL⁻¹) (Sigma) and then incubated overnight at 37°C. Next, 3 mL of 6 M NaCl was added and the samples were vortexed and then spun at 660 g, 4°C for 15 min. To precipitate DNA, 2 volumes of absolute ethanol (EtOH) were added to the supernatant, which was then subjected to centrifugation at 2600 g for 5 min, followed by a precipitation step using 70% (v/v) EtOH. The pellet was air dried and an appropriate volume of TE buffer was added to the sample. An OD reading was obtained to determine the quantity and purity of the extracted sample.

During the handling of ancient skeletal remains, great care was taken to avoid contamination with modern DNA. All materials and disposables were UV irradiated to destroy any remaining contamination that could alter the results. The extraction of DNA from ancient bone was performed as the bone surface was

ground using specialized dental micromotor (Nouvag, Switzerland) in a special dust protective chamber (Bredent, Germany) using specialized coarse plain cut straight hand piece metal bur (Nouvag, Switzerland). The bone was powdered by UV irradiated mortar and pestle. Approximately 5 mg of powdered bone were taken for each DNA extraction. In this study, due to the importance of the respective samples, we were allowed to extract the mtDNA from a total of 5 g of powdered bone. The powdered bone was UV irradiated prior to DNA extraction. The UV irradiated bone powder was dissolved in 8 mL extraction buffer (10 mM Tris HCl, 10 mM EDTA pH8.0, 50 mM NaCl, 2.0% (w/v) SDS and more than 24 U of Proteinase K) and incubated in a shaking water bath at 56°C followed by incubation at 37°C for 24 h. The extraction solution was centrifuged at 2000 g for 5 min. Approximately 1.5 mL aliquots of the supernatant was transferred to 1.5 mL microtube and spun for 5 min at 13500 g using Microfuge® 18 Centrifuge (Beckman Coulter). The supernatant was transferred into another 1.5 mL microtube for two cycles of phenol-chloroform extraction followed by one cycle of a standard 100 and 70% (v/v) ethanol precipitation. The pellet produced during ethanol precipitation was dissolved in 50 μL deionized distilled water.

Amplification of modern and ancient genomic DNA: Both modern and ancient DNA samples were analyzed by PCR. The amelogenin genes were amplified using the following set of primers: forward *AMEL* primer; 5' CTTCCCAGTTAACGCTCTGAT 3', *AMELX*-specific reverse primer; 5' GTCTCTCCTATACCACTTAGTC 3' and *AMELY*-specific reverse primer; 5' TTTGCCAAAGTTAGTAATTAA 3'. Primers were designed using Primer Premier Version 5.0 based on *AMELX* (Accession number: M55418) and *AMELY* (Accession number: M55419) sequences. PCR was performed using an Eppendorf Mastercycler gradient (Germany). The PCR mixture contained 1.25 U AmpliTaq Gold® DNA Polymerase, 1X GeneAmp® PCR Gold Buffer, 8.8 pM forward primer, 4.16 pM X-specific reverse primer, 4.84 pM Y-specific reverse primer, 2.5 mM MgCl₂, 0.1 mM dNTP, 40 ng μL⁻¹ BSA, 300 ng DNA template and sterile dH₂O added to a final volume of 20 μL. The PCR cycle starts with pre-denaturation step at 94°C for 2 min, then 94°C for 1 min, 50°C for 2 min, 72°C for 1 min for 30 cycles followed by a final elongation step at 72°C for 1 min. The samples were held at 4°C until further analysis. Aliquots (10 μL) were analyzed in a 1.0% (w/v) agarose gel using TAE (90 mM Tris-OH, 90 mM acetic acid and 2 mM Na EDTA) and visualized by ethidium

bromide staining under UV light. For the ancient sample from Dungun, the PCR amplification procedure was repeated after gel purification to generate a visible band of interest due to the minute and unpurified starting samples of DNA.

Cloning and sequencing analysis: Prior to cloning, the genes of interest were purified and then ligated with the TA plasmid (TA cloning vector kit™, USA) as per the manufacturer's instructions. The recombinant plasmid was introduced to competent cells by heat-shock. Plasmids isolated from positive clones were digested with EcoR1 to further confirm that the inserts were successfully cloned before being sequenced. The sequencing results were analyzed using Basic Local Alignment Search Tool (BLAST) software, which can be obtained from <http://www.ncbi.nih.gov/web>.

RESULTS

Modern genomic DNA from peripheral blood was used to first test the efficiency of the multiplex PCR amplification approach used in this study. DNA was successfully extracted from whole blood samples drawn from 7 males and 7 females, as shown in Fig. 1A. The DNA bands obtained from modern genomic DNA extraction procedures are intact and most of the samples are spectrophotometrically pure, as measured by their OD readings (approximately 1.8-2.0 using a 260/230 and 260/280 nm ratio).

Ancient DNA samples from bones originating in the swampy area of Kalumpang Island in west Peninsular Malaysia and a coastal area of Dungun in east Peninsular Malaysia was extracted with great care when handling ancient samples to minimize contamination that could lead to the misinterpretation of the results of a PCR analysis. All materials and consumables were therefore UV irradiated to destroy traces of modern DNA. However, no bands could be detected (i.e., that this demonstrates the low amount of starting sample and shows the needs of PCR amplification) in an agarose gel loaded with genomic DNA from the Kalumpang Island and Dungun samples after being stained with ethidium bromide (EtBr) and photographed under UV light (Fig. 2A; Lane 1 and 2). This was also true for the negative control (Fig. 2A; Lane 3), in which the ancient bone powder was replaced by sterile distilled water. Unlike genetic analyses using modern sources, ancient DNA studies are characterized by low starting quantity and quality of DNA. The advent of Polymerase Chain Reaction (PCR), which can amplify minute amount of DNA from samples, makes the field of ancient molecular DNA research possible.

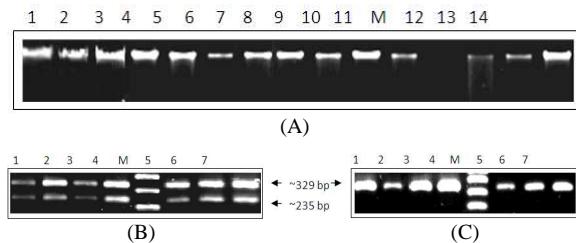


Fig. 1: Genomic DNA extraction from blood and the PCR amplification of *AMELX* and *AMELY* genes. Panel A shows bands of modern genomic DNA isolated from samples of human peripheral blood. Lanes 1-7 contain genomic DNA from male samples and lanes 8-14 contain genomic DNA from female samples. Lane M is a standard 1 kb marker. The 1 kb DNA marker are not appeared in the picture due to smaller in size (maximum size of DNA marker is 10kb) as compared to nuclear genomic DNA. The PCR of male samples produced two bands, as shown in Panel B (Lane 1-7). PCR of female samples produced a single band, as shown in Panel C (Lane 1-7). M is a standard 100 bp marker in Panels B and C

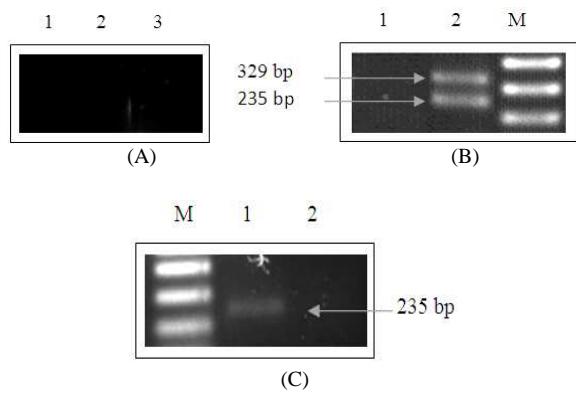


Fig. 2: DNA extractions of ancient bones and PCR amplification of *AMELX* and *AMELY* from Kalumpang Island and Dungun, peninsular Malaysia. Genomic DNA bands from ancient samples were not apparent on the gel, as shown in Panel A Lane 1A is from the Dungun sample and lane 2A is from the Kalumpang Island sample. Panel B shows the PCR amplification of *AMELX* and *AMELY* genes from the Kalumpang Island sample; Lane 2B is the results of the amplification of *AMELX* and *AMELY*. Panel C is from the Dungun sample; Lane 1C is the second PCR amplification of *AMELY*. Lane 1A, 2C and 3B are negative controls and M is a standard 100 bp marker.

Table 1: Results of sequencing *AMELX* and *AMELY* genes from modern and ancient DNA samples. After cloning, the samples were sequenced using M13 forward and reverse primers. All the genes are 100% identical to human amelogenin variant genes of the X and Y chromosomes, respectively

Sample of origin (<i>Homo sapiens</i>)	Gender	Identity of gene	Band size (bp)
Modern (blood)	Male	<i>AMELX</i>	329
		<i>AMELY</i>	235
	Female	<i>AMELX</i>	329
Ancient (bone)	Male	<i>AMELX</i>	329
		<i>AMELY</i>	235
	Male	<i>AMELY</i>	235

Skeletal remains from Kalumpang Island and Dungun are concluded to be male based on the amplification of the male-specific *AMELY* gene variant and are thus marked as such in italics

In this study, amelogenin gene variants were used as a marker for sex determination. Although the genes located on the X and Y chromosomes are different in size and sequence, they share a homology region that enabled us to utilize a forward primer that could anneal to both. Therefore, we used a cocktail of three primers, consisting of the common forward primer and both the human X-specific reverse primer (*AMELX*) and human Y-specific (*AMELY*) reverse primer. PCR will amplify one or both of the respective genes based on their presence in the sex chromosomes, hence reflecting the presence or absence of these chromosomes in the source DNA. In male samples, the amelogenin gene is located on both the X and Y chromosomes and, thus, two bands of PCR products will be produced, i.e., 329 bp (*AMELX*) and 235 bp (*AMELY*). On the other hand, as female samples have two X chromosomes, PCR will generate only a single band of 235 bp, but in higher concentration than the X chromosome band from males. Figure 1B shows that PCR amplification from 7 modern male samples stained with EtBr and photographed under UV light produced the expected two bands, i.e., 329 bp for *AMELX* and 235 bp for *AMELY*. In contrast, amplification from genomic DNA samples of 7 modern females produced a single, more intense band of *AMELX* at 329 bp, as shown in Fig. 1C.

After validation of this technique by PCR amplifications from the genomic DNA of equal numbers of males and females ($n = 14$), the amelogenin gene amplification approach was used to test ancient bone samples from two different locations, Kalumpang Island and Dungun. Figure 2B (Lane 2), shows the results of PCR amplification of *AMELX* and *AMELY* from the Kalumpang Island sample. The amplification

produced both of the expected band sizes, i.e., 329 bp (*AMELX*) and 235 bp (*AMELY*). In contrast, the ancient bone sample from the coastal area of Dungun did not yield any bands of interest after the first PCR amplification. However, upon the second amplification, a band of 235 bp, the expected size for *AMELY*, was detected (Fig. 2C, Lane 1).

After the PCR amplification of *AMELX* and *AMELY* from both modern and ancient DNA samples, the bands of interest for all samples were purified using Wizard® SV gel and PCR Clean-Up System Kit (Promega, USA) and ligated into a TA cloning vector (TA cloning kit™, USA). The resultant plasmids were digested with EcoR1 to further confirm the presence of 329 bp and 235 bp insert for plasmid consisting *AMELX* and *AMELY* genes respectively. The recombinant plasmids were then submitted for DNA sequencing. Analysis of the DNA sequences showed that *AMELX* amplified from modern male and female samples as well as ancient samples were 100% identical to the sequence retrieved from the NCBI database (NCBI accession number: M55418) for the human X chromosome variant of amelogenin. In addition, *AMELY* from modern male and ancient samples are also 100% identical with the sequence of amelogenin on the human Y chromosome retrieved from the NCBI database (NCBI accession number: M55419) (Table 1).

DISCUSSION

Multicopy DNA genomes, such as mitochondrial DNA, that are present in large numbers in cells are more likely to persist in ancient specimens than the nuclear genome. Therefore, most ancient DNA research has focused on retrieving sequences from mitochondria (mtDNA) in mammals and chloroplasts (cpDNA) in plants. Although mtDNA has many advantages for studying evolutionary and population genetics, it is clear that the retrieval of nuclear DNA would allow a much more information to be gained from archaeological remains^[3]. In this study, we concentrate on the extraction of nuclear DNA from ancient sources for further PCR analysis.

Through sex determination, the historical roles of gender identities can be directly articulated^[8]. Thus assessing sex-specific demographic data such as age and stature while allowing one to better understand the gender-specific roles in a society^[9]. The analysis of ancient DNA has the potential to provide answers to archaeological, paleontological and anthropological questions that classical approaches of these disciplines cannot. Radiography and other means of identification

of ancient bone and dental morphology can be done to estimate the ethnic origin, age and gender of the sample. However, the analysis will be difficult if the sample is not in good condition or if the remains are largely incomplete. Moreover, each physical skeleton is unique and incomparable with other archaeological samples found throughout the world. This is because of the population-specific nature of the factors that give rise to the appearance of the physical skeletal. In addition, of the inability to reliably estimate the sex of individuals hinders the analysis of gender and life-course identities in historical populations^[8].

Techniques of molecular analysis for gender determination provide an expedient solution when the use of conventional methods is not possible. Improvements that make the practice easier and faster are most welcomed in the field; thus, PCR is a particularly attractive option for sex identification. This type of molecular analysis generates precise results due to the sensitivity of PCR as a tool. If the procedure goes well, for example if inhibitors of PCR and contaminating modern DNA are excluded, the results are reliable. Even though the extraction of genomic DNA can be considered a destructive procedure (5 mg samples of powdered bone are needed), only a relatively small amount of the sample is needed (in parallel, the amount of ancient samples available is very limited). PCR is capable of amplifying minute amounts of samples to produce bands of interest that can easily be visualized in an agarose gel stained with EtBr and exposed to UV light. Our preliminary results using modern DNA extracted from the peripheral blood showed that the X-Y homologous amelogenin gene was 100% reliable as a sex marker. All of our samples ($n = 14$), consisting of 7 males and 7 females, exhibited positive results, i.e., amplified bands of the expected size of *AMELX* and *AMELY* according to their presence in sex chromosomes (Fig. 1B and 1C).

Amelogenin is one of the major matrix proteins secreted by the ameloblasts during enamel formation. The *AMEL* gene encodes a highly conserved protein and is located on both the X and Y chromosomes in humans. The discrimination of gender using this gene is possible as the variants are measurably different in size and males possess both X and Y chromosomes, whereas females have two X chromosomes. Human females will thus never produce an *AMELY* band following PCR amplification. Both of the genes are similar in their exonic regions but differ in their intronic regions, as there is a deletion of an intronic region of amelogenin from the Y chromosome (*AMELY*) compared to the gene in the X chromosome (*AMELX*)^[5]. As a result, PCR amplification of samples

from males will produce two bands (*AMELX* and *AMELY*), whereas the female samples will only produce a single band. Figure 1B and C show that male donors indeed gave rise to two amelogenin bands, while female donors produced only a single band.

The bands of interest were purified and cloned before being subjected to DNA sequencing using M13 forward and reverse primers. Analysis of the sequencing results showed that both the *AMELX* and *AMELY* bands are identical to the known sequence of amelogenin in the X and Y chromosomes, respectively (Table 1). This result showed that *AMELX* and *AMELY* were amplified from their respective sex chromosomes and thus can act as a distinctive factor for gender determination.

Two ancient samples were subjected to molecular gender determination in this study. Samples from Kalumpang Island were recovered from a swampy area that is high in minerals and salt. It is believed that the area is an ancient graveyard based on the finding of another almost complete skeleton in the same locale. It is estimated by a local archaeologist that the skeletal remains are ~2000 years in age. The skeletal remains from Dungun, on the other hand, had been in the seabed for a very long time (they are approximately 400 years old). Thus, both samples spent a long time exposed to extreme environmental conditions consisting of high salt and other minerals, inhibitors from the soil, or water and high pressure. These conditions can contribute to the process of DNA degradation and the inhibition of PCR amplification. In our study, we had to amplify the PCR products from the Dungun samples twice, as the first amplification did not yield any bands of interest. Notably, only *AMELY* was amplified after the second round (Fig. 2C). The loss of *AMELX* may be explained by the size of the X chromosome itself, as it is 75% larger than Y chromosome^[10] and hence has a higher chance of degradation due to extreme and prolonged environment exposure. However, we can conclude that the gender of the bones is male based on human karyotype, which indicates that the Y chromosome only exists in males. Meanwhile, PCR amplification of the Kalumpang Island sample yielded two bands of interest, i.e., *AMELX* and *AMELY*, from which we can conclude that the skeletal remains are also those of a male.

CONCLUSION

Based on the distinctive features of amelogenin gene variants on the X and Y chromosomes, the respective genes can be used as a molecular sex determinant. The ancient bones from Kalumpang Island

and Dungun have both been determined to be male based on the amplification of partial *AMELX* (329 bp) and *AMELY* (235 bp) genes for the Kalumpang Island sample and the male-specific *AMELY* (235 bp) gene for the Dungun sample.

ACKNOWLEDGEMENT

The researchers would like to thank the University Kebangsaan Malaysia for the financial grant (UKM-GUP-BTK-07-15-197) during this study. The authors also would like to thank Prof Dr Shuhaimi, N.K. that leads the expedition team at Kelumpang Island and Syamsul, B. for expedition of Wanli Shipwreck, Dungun. All of the hard work in this study is well-acknowledged.

REFERENCES

1. Leonard, J.A., N. Rohland, S. Glaberman, R.C. Fleischer, A. Caccone and M. Hofreiter, 2005. A rapid loss of stripes: The evolutionary history of the extinct quagga. *Biol. Lett.*, 1: 291-295. DOI: 10.1098/rsbl.2005.0323.
2. Münnemann, K., T. Böni, G. Colacicco, B. Blümich and F. Rühli, 2007. Noninvasive (¹H and (23) Na nuclear magnetic resonance imaging of ancient Egyptian human mummified tissue. *Magnet. Reson. Imag.*, 25: 1341-1345. DOI: 10.1016/j.mri.2007.03.023
3. Willeslev, E. and A. Cooper, 2005. Ancient DNA. *Proc. Biol. Sci.*, 272: 3-16. DOI: 10.1098/rspb.2004.2813
4. Krause, J., P.H. Dear, J.L. Pollack, M. Slatkin, H. Spriggs and I. Barnes *et al.*, 2006. Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature*, 439: 724-727. DOI: 10.1038/nature04432
5. Sivagami, A.V., A.R. Rao and U. Varshney, 2000. A simple and cost-effective method for preparing DNA from the hard tooth tissue and its use in polymerase chain reaction amplification of amelogenin gene segment for sex determination in an Indian population. *Forensic Sci. Int.*, 110: 107-115. DOI: 10.1016/S0379-0738(00)00155-9
6. Weikard, R., C. Pitra and C. Kühn, 2006. Amelogenin cross-amplification in the family Bovidae and its application for sex determination. *Mol. Reprod. Dev.*, 73: 1333-1337. DOI: 10.1002/mrd.20554
7. Richard, B., S. Delgado, P. Gorry and J.Y. Sire, 2007. A study of polymorphism in human *AMELX*. *Arch. Oral Biol.*, 52: 1026-1031. DOI: 10.1016/j.archoralbio.2007.06.001
8. Knudson, K.J. and C.M. Stojanowski, 2008. New directions in bioarchaeology: Recent contributions to the study of human social identities. *J. Archaeol. Res.*, 16: 397-432. DOI: 10.1007/s10814-008-9024-4.
9. Gibbon, V., M. Paximadis, G. Štrkalj, P. Ruff and C. Penny, 2009. Novel methods of molecular sex identification from skeletal tissue using the amelogenin gene. *Forensic Sci. Int.*, 3: 74-79. DOI: 10.1016/j.fsc.2008.10.007.
10. Ballabio, A., D. Nelson and S. Rozen, 2006. Genetics of disease: The sex chromosomes and human disease. *Curr. Opin. Genet. Dev.*, 16: 209-212. DOI: 10.1016/j.gde.2006.04.016