Int J Legal Med (2007) 121:411–416
DOI 10.1007/s00414-006-0143-5

ORIGINAL ARTICLE

Molecular characterization of the canine mitochondrial DNA control region for forensic applications

Cordula Eichmann · Walther Parson

Received: 12 May 2006 / Accepted: 14 November 2006 / Published online: 12 January 2007
© Springer-Verlag 2007

Abstract The canine mitochondrial DNA (mtDNA) control region of 133 dogs living in the area around Innsbruck, Austria was sequenced. A total of 40 polymorphic sites were observed in the first hypervariable segment and 15 in the second, which resulted in the differentiation of 40 distinct haplotypes. We observed five nucleotide positions that were highly polymorphic within different haplogroups, and they represent good candidates for mtDNA screening. We found five point heteroplasmic positions; all located in HVS-I and a polythymine region in HVS-II, the latter often being associated with length heteroplasmy. In contrast to human mtDNA, the canine control region contains a hypervariable 10 nucleotide repeat region, which is located between the two hypervariable regions. In our population sample, we observed eight different repeat types, which we characterized by direct sequencing and fragment length analysis. The discrimination power of the canine mtDNA control region was 0.93, not taking the polymorphic repeat region into consideration.

Keywords Canine mtDNA · HVS-I and HVS-II · Haplogroups · Hypervariable repeat region · Sequence analysis · Forensics

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00414-006-0143-5 and is accessible for authorized users.

C. Eichmann · W. Parson (✉)
Institute of Legal Medicine, Innsbruck Medical University,
Müllerstrasse 44,
6020 Innsbruck, Austria
e-mail: walther.parson@i-med.ac.at

Introduction

Genetic studies have shown that modern dogs (*Canis familiaris*) originated from the domestication of wild grey wolves (*Canis lupus*) [1–5]. Mitochondrial DNA (mtDNA) data suggest that there were several independent domestication events, which is primarily indicated by the observed patterns of mtDNA clades that arose in the course of a continued genetic exchange between wolves and dogs during coexistence over a wide geographical range [3–9]. Most modern dog breeds are relatively young; the majority having developed within the past 300 years and, many of these, derived from a relatively small number of animals. Therefore, purebred dogs represent a limited genetic pool with genetic characteristics derived from one or a small number of recent genetic founders. Compared to other animals, the population of dogs is characterized by a higher degree of isolation, narrower bottlenecks and much better genealogical records. As the dog is deemed to be our closest companion and most popular pet, it can also be considered one of the most interesting animal species from a forensic point of view. As a consequence of its high abundance, forensically relevant cases involving dogs are observed on a regular basis. Dog hairs or remains of dog saliva can help to link a crime scene to a person involved in the crime, e.g. the owner of the dog or a victim being transported in a dog owner's car. D'Andrea et al. [10] demonstrated that it is almost impossible to enter the house of a pet holder without being contaminated by animal hairs. Moreover, as dogs shed more hairs than humans, the probability of finding hairs originating from a dog can be higher than finding biological material from its owner. It has been shown previously that the analysis of canine nuclear DNA (STRs [11]) and mtDNA [12–15] including species identification by use of the cytochrome b gene [16]

complement the human molecular forensic armamentarium very effectively and open the door for enhanced applied animal DNA testing in forensic casework [17].

In the current study, we investigated the canine mtDNA control region (CR), which is—similar to the situation in humans—often the last resort for DNA typing when nuclear DNA analysis fails to give a useful result. We characterized the observed polymorphisms within the two hypervariable regions, hypervariable segment I (HVS-I) and HVS-II, of the mtDNA CR and add more data to the emerging canine mtDNA databases required for establishing frequency estimates of unknown sequences. Furthermore, we draw attention to the polymorphic repeat region located between HVS-I and HVS-II.

Materials and methods

Buccal swabs were taken from 133 dogs living in the area surrounding Innsbruck, Austria as described in [18]. DNA was extracted from the swabs using the Chelex method [19].

All primers used for the amplification and sequencing of the mtDNA CR were designed using Primer Express software (version 5.0). The denomination of the primers corresponds with the 5′-end of the reference sequence according to [20]. The sequencing strategy of the CR is shown in Fig. 1. The CR of the mitochondrial genome was amplified in a 50-µl assay including 1× PCR buffer II, 1.5 mM $MgCl_2$, 200 µM each deoxynucleotide triphosphate (dNTP), 4 U AmpliTaq Gold polymerase (Applied Biosystems [AB], Foster City, CA), 0.25 mg/ml BSA (Serva, Heidelberg, Germany) and 50 nM each primer (Table S1). Amplification was performed on a Gene Amp PCR System 9600 (Perkin Elmer [PE], Norwalk, CT) comprising initial denaturation at 95°C for 11 min followed by 36 cycles at 95°C for 20 s, 60°C for 30 s and 72°C for 90 s. PCR products were purified using ExoSAP-IT (Amersham, Bucks, UK) and sequenced using BigDye Terminator sequencing reagents (version 2.0 and version 1.1, AB). Sequencing was performed on a Gene Amp PCR System 9600 (PE) comprising initial denaturation at 96°C for 30 s followed by 25 cycles at 96°C for 15 s, 59°C for 30 s (F15416, R56, R16095), 66°C for 30 s (F16010), 50°C for 30 s (F16527, F15960, R16643, R35) and 60°C for 4 min. Sequencing reaction products were purified from residual dye terminators using Sephadex G-50 Fine (Amersham), all according to the manufacturer’s manual. DNA sequencing was carried out on an AB Prism 3700 Genetic Analyzer using POP 6, 50-cm capillary arrays, and an AB 3100 Genetic Analyzer (AB) using POP 6, 36-cm capillary arrays, and default instrument settings as recommended by the manufacturer. The data were analysed using Sequencing Analysis Version 3.7 (AB) and Sequencher version 4.1.4 (GeneCodes, Ann Arbor, MI).

The primers used for fragment length analysis of the repeat region were designed using the primer3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). The total reaction volume was 25 µl including 1× PCR buffer II, 1.5 mM $MgCl_2$, 200 µM each dNTP, 2 U AmpliTaq Gold polymerase (AB), 0.25 mg/ml BSA (Serva) and 200 nM of each primer (Table S1). Amplification was performed on a Multicycler PTC2 40 Tetrad 2 (Biozyme Diagnostik Hessisch Oldendorf, Germany) comprising initial denaturation at 95°C for 11 min followed by 29 cycles of 95°C for 30 s, 54°C for 45 s, 72°C for 90 s and a final incubation at 72°C for 60 min. Aliquots of 2 µl of the amplification products were combined with 20-µl deionized formamide, including 0.4-µl internal size standard (Genescan-500 TAMRA, AB), heat-denatured at 95°C for 3 min, snap-cooled on ice and subjected to capillary electrophoresis on an ABI Prism 3100 Genetic Analyzer using POP 6, 36-cm capillary arrays, and default instrument settings as recommended by the manufacturer. The data were analysed using GeneScan Analysis version 3.7 and Genotyper version 2.5 (both AB). Extraction blanks and negative controls were carried through the complete amplification and sequencing processes.

The discrimination power of the CR haplotypes was calculated as described in [12, 14]

Results and discussion

Polymorphic positions/haplotypes

We determined complete mtDNA CR sequences from 133 dogs by amplification of the entire CR and full double-strand sequencing coverage using eight primers. Sequence polymorphisms were reported as differences from a canine mtDNA sequence (accession number: U96639) [20], which is also recommended as reference in [21]. In our study, we found a total of 55 polymorphic positions in the two hypervariable segments, of which 40 sites were located in HVS-I and 15 in HVS-II (Fig. 1, Table S2). Of these positions, 43 (78%) showed a transition, of which 31 (72%) constituted a C–T and 12 (28%) an A–G exchange; position 16507 showed a transversion (T–A) and position 15639 was found to be a ternary site (T–G, T–A). Furthermore, we observed six insertions (in a total of 30 individuals) at positions 15464.1, 16663.1, 16663.2, 16663.3, 16671.1 and 16674.1 and three deletions (in a total of eight individuals) at positions 15931, 15938 and 16674.

We identified 40 different haplotypes (HT1–HT40) among the 133 samples, not taking the repeat region [21] into consideration (Table S2); 22 of these 40 haplotypes

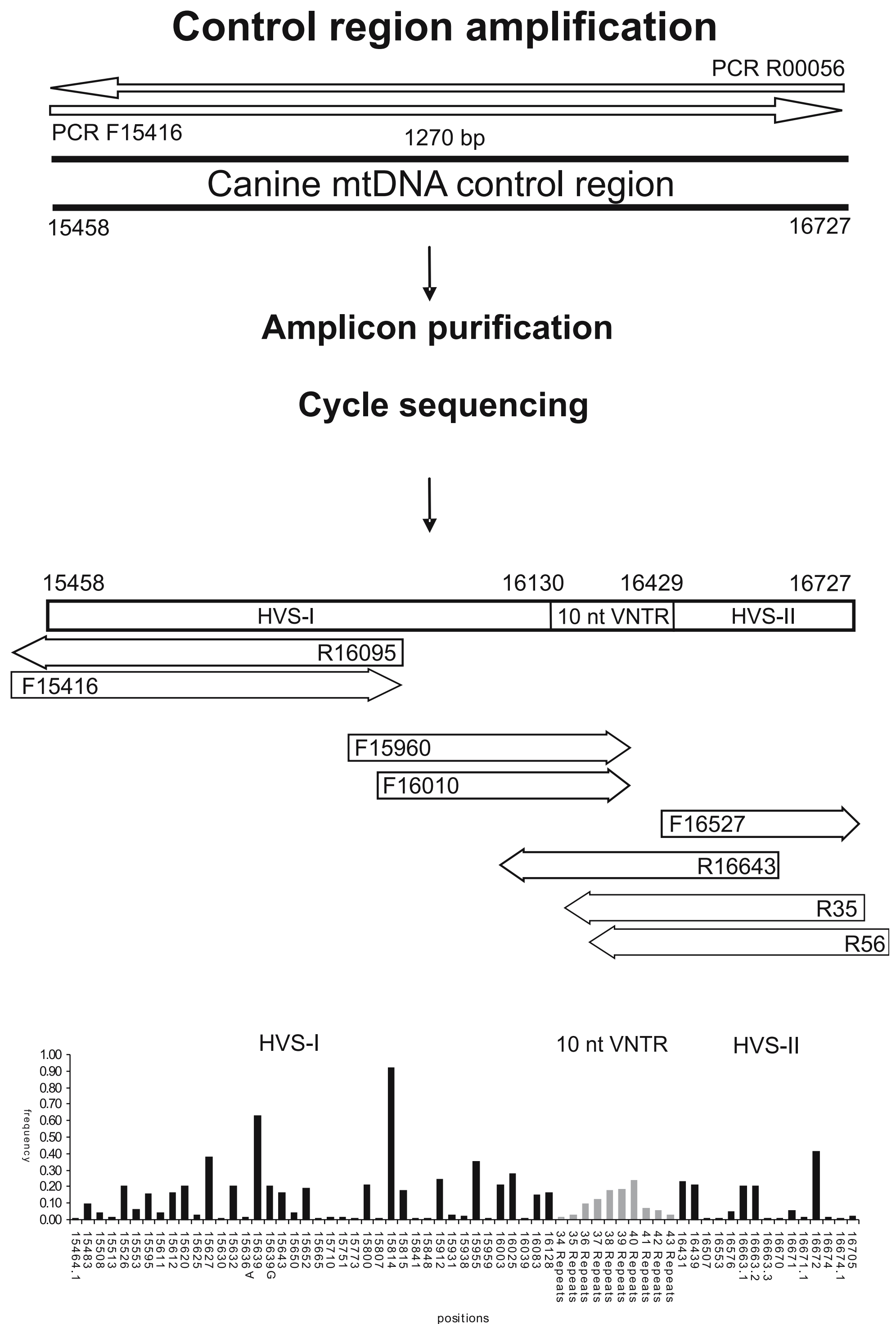

Fig. 1 Schematic representation of the amplification and sequencing strategy for the analysis of the canine mtDNA control region. Primers were named according to their 5′-position with respect to the reference sequence [20]; *arrows* indicate sequencing direction and reading length. Number and distribution of all 55 polymorphic sites found in the canine mtDNA control region is given below (133 individuals). *X*-axis: nucleotide position in the control region according to [20]. *Y*-axis: frequency of observed substitutions. The 10 nt repeat region located between HVS-I and HVS-II is indicated by *repeat numbers*, sequence clusters are depicted in Table S5. 15639 is a ternary position (T–A, T–G) and is therefore listed twice

were observed only once in our data set (point heteroplasmy disregarded), and the most common mtDNA sequence (HT6) was shared by 20 individuals (15%). We found clusters of identical HVS-I sequences (HT2, 3; HT6, 8, 9, 12; HT22, 25 and HT38, 39), which were further distinguished by analysing HVS-II. Therefore, the additional analysis of HVS-II increased the total number of observed haplotypes from 34 to 40 and increased the exclusion capacity from 0.92 (HVS-I) to 0.93 (HVS-I and HVS-II), which is in accordance with Angleby and Savolainen [14] (range between 0.86 and 0.95).

Although there was no direct correlation between dog breed and haplotypes [14], we found that some mtDNA sequences were overrepresented (Table S2). In ten German

shepherds, six showed HT1, and in seven rottweilers, we found five individuals with HT6. Because the dog samples were collected by convenience, the presented data set does not necessarily represent a typical set from the local area. This should be taken into consideration when forensic applications are performed.

Hot spots

We identified five positions: 15627, 15639 and 15955 (HVS-I), and 16431 and 16672 (HVS-II), which displayed a high degree of polymorphism among different haplogroups (hgs, Fig. 1, Table S2). Based on the hg definition in [14], positions 15627, 15955, 16431 and 16672 showed a substituted variant in 14 (44%), 12 (38%), 12 (38%) and 16 (50%) hgs, respectively. The ternary site 15639 brought mutations in a total of 28 hgs (87.5%), where a transversion to A was observed in 20 hgs (62.5%) and a transversion to G in 8 hgs (25%). The positions located in HVS-I were also found to be highly polymorphic in previous studies [12, 14, 22–24], partly using other nomenclatures for base designation (Table S3). In Angleby and Savolainen [14], position 15627 showed a mutation in 49 hgs (39.5%), position 15639 in 109 hgs (88%; 73 hgs [59%] harboured the variant A, 36 hgs [29%] showed a G) and position 15955 displayed a mutation in 48 hgs (39%); in Savolainen et al. [24], position 15627 showed a G in 75 hgs (48%) and even one transversion (A–T, W11), position 15639 displayed a mutation in 139 hgs (90%), where 106 hgs (68%) showed an A and 33 hgs showed a G (21%), of which two had a deletion at this position, and position 15955 showed a mutation in 51 hgs (33%). These five positions alone allow for the differentiation of 12 clusters within the 133 individuals of our study; the exclusion capacity amounts to 0.86 (Table S4), which seem to represent good candidates for mtDNA screening.

Heteroplasmy

High copy number and high mutation rate lead to mtDNA sequence heteroplasmy, a phenomenon that describes the co-existence of more than one population of mtDNA genomes per cell/individual. Heteroplasmy has forensic relevance—as described in the human field—because the co-occurrence of the same heteroplasmic event in two otherwise identical sequences may increase the significance of a match [25]. It has been demonstrated that the detection of point heteroplasmy depends on the reaction conditions [26] and is likely to differ between tissues. In our study, we found five point heteroplasmic positions among the 133 individuals (Table S2), all located in HVS-I. The heteroplasmic mixtures were confirmed by double-strand sequencing. Following the guidelines established in the human field [27], point heteroplasmy was designated with an International Union of Biochemistry code (mixture of purine bases R, a mixture of pyrimidine bases Y). At positions 15635 and 16083, an A–G mixture was observed and a C–T mixture at positions 15620 and 15814. Position 15931 harboured a mixture, including the nucleotide A, and a deletion and was designated 15931 A/DEL according to [21] (Fig. S2). Length heteroplasmy was observed in HVS-II, as this region includes a polycytosine region (16661–16663) directly followed by a polythymine region (16664–16671). We observed both a transition of C to T and an insertion of T in this T-tract, which led to an extension of the polythymine region. This was observed in 45% of the examined individuals and then mostly associated with length heteroplasmy.

Notation of the canine CR

In the literature, the two hypervariable regions have been denominated in different ways using the terms CBS I and II (conserved sequence blocks I and II) [20], HV1 and HV2 (hypervariable regions 1 and 2) [12], HVR (hypervariable region) [23], hypervariable region and conserved region [28]. According to the convention in human mtDNA, we refer to the hypervariable segments as HVS-I and HVS-II.

As already stated in [21], most laboratories analyse HVS-I as a segment ranging between positions 15458 and 16039, with the exception of [22, 28] who also sequenced HVS-II. As the repeat region starts at position 16130, sequence information within a string of 90 nucleotides is missed when the commonly used sequencing strategy is applied. Among others, this would involve positions 16083 (A–G) and 16128 (G–A), which showed a high degree of variability (Fig. 1, Table S2 and [22]).

In agreement with [21], the importance of a unique standardized reference sequence is also emphasized here, as the application of different reference sequences confuses the reporting of mtDNA haplotypes [12, 15, 20, 22–24, 28]. Table S3 presents some different notations of already established data in a comparative way [12, 20, 22, 23] to help harmonization of results.

A 10 nt VNTR in the canine mtDNA CR

In contrast to humans, the canine mtDNA CR harbours a 10 nucleotide (nt) variable number of tandem repeat (VNTR) between HVS-I and HVS-II (Fig. 1, Table S2). Each repeat unit consists of a constant (first 6 nts) and a variable sequence (last 4 nts), in which point mutations discriminate

between different repeat types. In this study, we characterized eight different repeat types (I–VIII, Table S5), whereas earlier studies described only repeat types I, II and IV in detail [20, 22, 28–32]. We found seven different motifs consisting of mixtures of various repeat types that we classified as "a" through "g" suffixing roman numbers for the repeat types involved, e.g. $\mathbf{a}_{I,\ II}$ (GTACACGT**R**C) consists of type I (GTACACGT**A**C) and type II (GTA CACGT**G**C, Table S5). Table S2 shows that all but one individual shared the $\mathbf{a}_{I,\ II}$ motif, some of them displayed additional mixture motifs.

Based on fragment length analysis, we found substantial length heteroplasmy within the VNTR (Fig. S3). In most individuals, we determined a single dominant length variant; in 21 (18.4%) dogs, however, we found two or even three prominent length variants (Fig. S3, Table S6). We analysed the 10 nt repeat region in six different tissues (blood, hair roots, muscle, liver, bone and tooth) that we obtained from six different dog breeds (bearded collie, husky, cocker spaniel, dachshund, German shepherd, rottweiler). The observed pattern of length variation in the 10 nts repeat region was similar between the different tissues within an individual; however, muscle tissues gave distinct patterns compared with the other tissues in the collie, the spaniel, the rottweiler and the German shepherd (Fig. S3).

Whether or not the VNTR polymorphism can be used for forensic identification has been dealt with in earlier studies. Fridez et al. [29] observed high intra-individual fragment length variation of the VNTR and concluded that this marker cannot be used for forensic identification. Savolainen et al. [31] stated that "...fragment length analysis (...) can be used to distinguish between individuals even within the same family when DNA is obtained from blood." Further, they conclude that "...(fragment length analysis) is useful for the comparison of blood stains and blood samples from dogs, but not for the comparison of single hairs to each other or to blood samples." For the comparison of hairs, the authors suggest to analyse the sequence of the VNTR by cloning as they found that "...if only a sufficient number of molecules are analysed from a hair, it should always be possible to find array types (repeat types) that are common with those in a blood sample." Based on our findings, at least muscle tissue might also be critical for the comparison of results obtained from other tissues within one individual. Our data further show a high level of sequence diversity within the repeat region, which will give rise to a more detailed study on tissue-specific intra-individual variation.

Acknowledgement The authors would like to thank Volker Büchele, Armin Anranter, Helena and Paul Scheithauer for collecting buccal samples from the individual dogs and Ina Pfeiffer for preparing the tissue samples. We also thank an anonymous reviewer, Harald Niederstätter, Burkhard Berger and Anna König for helpful assistance, comments and discussion.

References

1. Olsen SJ (1985) Origins of the domestic dog: the fossil record. University of Arizona Press, Tucson, AZ
2. Wayne RK (1993) Molecular evolution of the dog family. Trends Genet 9:218–224
3. Vila C, Savolainen P, Maldonado JE et al (1997) Multiple and ancient origins of the domestic dog. Science 276:1687–1689
4. Leonard JA, Wayne RK, Wheeler J, Valadez R, Guillen S, Vila C (2002) Ancient DNA evidence for Old World origin of New World dogs. Science 298:1613–1616
5. Sutter NB, Ostrander EA (2004) Dog star rising: the canine genetic system. Nat Rev Genet 5:900–910
6. Vila C, Maldonado JE, Wayne RK (1999) Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. J Heredity 90:71–77
7. Wayne RK, Ostrander EA (1999) Origin, genetic diversity, and genome structure of the domestic dog. Bioessays 21:247–257
8. Flagstad O, Walker CW, Vila C et al (2003) Two centuries of the Scandinavian wolf population: patterns of genetic variability and migration during an era of dramatic decline. Mol Ecol 12:869–880
9. Vila C, Walker C, Sundqvist AK et al (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. Heredity 90:17–24
10. D'Andrea F, Fridez F, Coquoz R (1998) Preliminary experiments on the transfer of animal hair during simulated criminal behavior. J Forensic Sci 43:1257–1258
11. Eichmann C, Berger B, Reinhold M, Lutz M, Parson W (2004) Canine-specific STR typing of saliva traces on dog bite wounds. Int J Leg Med 118:337–342
12. Savolainen P, Rosen B, Holmberg A, Leitner T, Uhlen M, Lundeberg J (1997) Sequence analysis of domestic dog mitochondrial DNA for forensic use. J Forensic Sci 42:593–600
13. Savolainen P, Lundeberg J (1999) Forensic evidence based on mtDNA from dog and wolf hairs. J Forensic Sci 44:77–81
14. Angleby H, Savolainen P (2005) Forensic informativity of domestic dog mtDNA control region sequences. Forensic Sci Int 154:99–110
15. Schneider PM, Seo Y, Rittner C (1999) Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. Int J Leg Med 112:315–316
16. Parson W, Pegoraro K, Niederstätter H, Föger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. Int J Leg Med 114:23–28
17. Budowle B, Garofano P, Hellman A et al (2005) Recommendations for animal DNA forensic and identity testing. Int J Leg Med 119:295–302
18. Eichmann C, Berger B, Parson W (2004) A proposed nomenclature for 15 canine-specific polymorphic STR loci for forensic purposes. Int J Leg Med 118:249–266
19. Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10:506–513
20. Kim KS, Lee SE, Jeong HW, Ha JH (1998) The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Mol Phylogenet Evol 10:210–220
21. Pereira L, Van Asch B, Amorim A (2004) Standardisation of nomenclature for dog mtDNA D-loop: a prerequisite for launching a Canis familiaris database. Forensic Sci Int 141:99–108
22. Okumura N, Ishiguro N, Nakano M, Matsui A, Sahara M (1996) Intra- and interbreed genetic variations of mitochondrial DNA

major non-coding regions in Japanese native dog breeds (Canis familiaris). Anim Genet 27:397–405

23. Wetton JH, Higgs JE, Spriggs AC, Roney CA, Tsang CS, Foster AP (2003) Mitochondrial profiling of dog hairs. Forensic Sci Int 133:235–241
24. Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E, Lundeberg J (2004) A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. Proc Natl Acad Sci USA 101:12387–12390
25. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ (1996) Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat Genet 12:417–420
26. Brandstatter A, Parson W (2003) Mitochondrial DNA heteroplasmy or artefacts—a matter of the amplification strategy? Int J Leg Med 117:180–184
27. Tully G, Bar W, Brinkmann B et al (2001) Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci Int 124:83–91
28. Tsuda K, Kikkawa Y, Yonekawa H, Tanabe Y (1997) Extensive interbreeding occurred among multiple matriarchal ancestors during the domestication of dogs: evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. Genes Genet Syst 72:229–238
29. Fridez F, Rochat S, Coquoz R (1999) Individual identification of cats and dogs using mitochondrial DNA tandem repeats? Sci Justice 39:167–171
30. Hoelzel AR, Lopez JV, Dover GA, O'Brien SJ (1994) Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. J Mol Evol 39:191–199
31. Savolainen P, Arvestad L, Lundeberg J (2000) A novel method for forensic DNA investigations: repeat-type sequence analysis of tandemly repeated mtDNA in domestic dogs. J Forensic Sci 45:990–999
32. Savolainen P, Arvestad L, Lundeberg J (2000) mtDNA tandem repeats in domestic dogs and wolves: mutation mechanism studied by analysis of the sequence of imperfect repeats. Mol Biol Evol 17:474–488