International Journal of Legal Medicine (2022) 136:1–12
https://doi.org/10.1007/s00414-021-02711-y

ORIGINAL ARTICLE

Development and validation of simultaneous identification of 26 mammalian and poultry species by a multiplex assay

Chikahiro Mori[1,2] · **Shuichi Matsumura**[3]

Received: 9 July 2021 / Accepted: 24 September 2021 / Published online: 9 October 2021

Abstract

A multiplex PCR assay was developed to simultaneously identify 22 mammalian species (alpaca, Asiatic black bear, Bactrian camel, brown rat, cat, cattle, common raccoon, dog, European rabbit, goat, horse, house mouse, human, Japanese badger, Japanese wild boar, masked palm civet, pig, raccoon dog, red fox, sheep, Siberian weasel, and sika deer) and four poultry species (chicken, domestic turkey, Japanese quail, and mallard), even from a biological sample containing a DNA mixture of multiple species. The assay was designed to identify species through multiplex PCR and capillary electrophoresis, with a combination of amplification of a partial region of the mitochondrial D-loop by universal primer sets and a partial region of the cytochrome *b* (cyt *b*) gene by species-specific primer sets. The assay was highly sensitive, with a detection limit of 100 copies of mitochondrial DNA. The assay's ability to identify species from complex DNA mixtures was demonstrated using an experimental sample consisting of 10 species. Efficacy, accuracy, and reliability of the assay were validated for use in forensic analysis with the guidelines of Scientific Working Group on DNA Analysis Methods (SWGDAM). The multiplex PCR assay developed in this study enables cost-effective, highly sensitive, and simultaneous species identification without massively parallel sequencing (MPS) platforms. Thus, the technique described is straightforward and suitable for routine forensic investigations.

Keywords Species identification · Forensic science · DNA mixtures · Multiplex PCR · DNA degradation

Introduction

Non-human biological samples are frequently encountered in routine forensic investigations. Forensic scientists are required to identify mammalian species in many forensic cases; for example, traffic accidents involving animals [1], animal cruelty [2–5], livestock robbery [6, 7], animal attacks [5, 7, 8], murder cases [9], and postmortem investigations [10]. In addition, species identification has become important in wildlife forensics [11–15], conservation of endangered animals [16–19], and the detection of food fraud [20–22].

✉ Chikahiro Mori
cmori2010gp@gmail.com

[1] The United Graduate School of Agricultural Science, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan

[2] Forensic Science Laboratory, Gifu Prefectural Police Headquarters, 2-1-1 Yabutaminami, Gifu 500-8501, Japan

[3] Faculty of Applied Biological Sciences, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan

Currently, mitochondrial DNA (mtDNA) markers are commonly used for species identification because the high copy number of mtDNA per cell [23–25] allows for successful identification, even in samples that contain insufficient amounts of nuclear DNA (e.g., partial bone fragments and hair shafts). The cytochrome *b* (cyt *b*) gene, the cytochrome *c* oxidase I (COI) gene, and the D-loop region are typical mtDNA markers in forensic, phylogenetic, and biodiversity studies [7, 13, 15, 26–32]. Other genes, such as the 12S and 16S ribosomal RNA genes, have also been studied for species identification [33–35].

In forensic investigations, non-human biological samples recovered from crime scenes are often contaminated with human DNA [25]. In other cases, samples may contain DNA from more than one non-human species. Typical DNA mixture samples include feces containing materials from ingested species, mixed meat or meat products, and saliva collected from a bite mark on a victim's body. Species identification from these mixtures is a challenge for forensic scientists. This is mainly because conventional Sanger sequencing using universal primers, which

has been widely applied for species identification, produces mixed DNA profiles which are difficult to interpret and lead to database matching of low reliability if it is applied to such DNA mixture samples [36]. In recent years, DNA metabarcoding using massively parallel sequencing (MPS), also called next-generation sequencing (NGS), technologies have been developed to identify multiple species in mixed DNA samples [5, 18, 22, 37]. However, DNA metabarcoding is not suitable for small-scale analysis, which is common in local forensic laboratories, mainly because of its costly and time-consuming procedures including data analysis in comparison with other techniques [36]. By contrast, fluorescent multiplex PCR involving capillary electrophoresis or real-time PCR allows for cost-effective species identification from such mixed samples. The target species of these assays, however, are limited by their primer sets [7, 31, 32, 38]. Therefore, the usefulness of the assays critically depends on the number of species they can analyze simultaneously because forensic laboratories often encounter cases with little information on the source species of the recovered DNA samples. Furthermore, negative results from the assays with a large number of identifiable species are also useful to eliminate the necessity of further investigation. Therefore, there is still a need to develop a more cost-effective assay for simultaneous identification of a wide range of mammalian species.

In this study, we developed a new multiplex PCR assay that can simultaneously identify 22 mammalian and four poultry species, even from a biological sample containing DNA from multiple species. The target species are composed of common domestic animals and wild animal forensic laboratories in Japan often encounter. Partial regions of the D-loop and cyt *b* loci were used for the assay. Universal primer sets were designed to amplify the hypervariable region (HV) of the D-loop of most of the target species. The region is equivalent to the HV1 of human mtDNA and shows interspecific size variation, giving size differences in PCR products between species. Additional species-specific cyt *b* primer sets were also designed to identify the species of which the HV regions cannot be amplified by the universal primer sets or are too long to amplify from degraded mtDNA. Each PCR product was analyzed by capillary electrophoresis using an Applied Biosystems 3500xL Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA). The combination of the primer sets of the D-loop and cyt *b* loci enabled an increase of the number of identifiable species, while the number of species-specific primer sets remained small. Developmental validation, which is a standard procedure to evaluate conditions and limitations of a new DNA methodology for forensic analysis, was performed to demonstrate efficacy, reliability, and robustness of the assay in accordance with the guidelines of Scientific Working Group on DNA Analysis Methods (SWGDAM) [39].

Materials and methods

Primer design

MtDNA sequences of the target species were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/). For each species, the NCBI reference sequence was used, if available. The species and the accession numbers used in this study are listed in Table 1. The sequences were aligned using MEGA7 software [40] to identify variable and conserved regions of the D-loop and the cyt *b* gene. Two universal primer sets, one targeting mammals and the other targeting poultry, were used for D-loop. Another primer set, which consisted of one universal forward primer and 10 species-specific reverse primers, was used for the cyt *b* in 12 mammals. All the primer sets were designed to amplify fragments of different sizes among species below ~ 500 bp (Tables 1, 2). DL_UniF and CYTB_UniF were modified from Nakamura et al. [7] and Pääbo et al. [41], respectively, in order to increase the number of the target species. CYTB_PigR and CYTB_GoatR were modified from Tobe et al. [31] and Matsunaga et al. [42], respectively. CYTB_RabbitR was identical with that in Tobe et al. [31]. The remaining primers were newly designed for this study. All forward primers were labeled with 6-FAM, VIC, or NED dyes and purified by HPLC. The fluorescent primers were ordered from Thermo Fisher Scientific, and the other primers were ordered from Rikaken Co., Ltd. (Nagoya, Japan).

Sample collection and DNA extraction

Biological samples (blood, buccal cells, tissues, hair, and nails) were obtained from 28 animal species, including 12 domestic mammals, 12 wild mammals, and four poultry species (Table 1). Most of the species are often encountered in forensic investigations in Japan. Total DNA was extracted from each sample using the QIAamp DNA Investigator Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocols. The following commercial DNA samples were used as human samples: AmpFLSTR Control DNA 9947A (Thermo Fisher Scientific), AmpFLSTR Control DNA 007 (Thermo Fisher Scientific), 2800 M Control DNA (Promega, Madison, WI), K562 Genomic DNA (Promega), and human genomic DNA (HeLa) (Takara Bio Inc., Shiga, Japan).

The mtDNA copy number of all the DNA samples tested in this study was quantified using a real-time quantitative PCR (qPCR) analysis described below because the copy number varies depending on the tissue type. A universal primer set (Forward: 5'-TACGACCTCGATGTTGGA

Table 1 List of the species and the GenBank accession number used in this study

Common name	Scientific name	Number of samples	Expected amplicon length (bp)		Observed Mean size (bp)		Accession number
			cyt *b*	D-loop	cyt *b* (SD)	D-loop (SD)	
Cattle	*Bos taurus*	5	118	481	110.5 (0.04)	476.3 (0.06)	NC_006853
European rabbit	*Oryctolagus cuniculus*	8	187	503	181.3 (0.03)	496.8 (0.05)	NC_001913
Human	*Homo sapiens*	5	197	485	190.8 (0.02)	479.5 (0.06)	NC_012920
Sheep	*Ovis aries*	5	205	–	199.5 (0.005)	–	NC_001941
Pig	*Sus scrofa domesticus*	5	225	431	217.0 (0.04)	426.1 (0.06)	NC_000845
Japanese wild boar	*Sus scrofa leucomystax*	5	N/A	N/A	217.0 (0.02)	425.0 (0.08)	AB015070, AB015085
Horse	*Equus caballus*	5	246	–	240.8 (0.04)	–	NC_001640
Goat	*Capra hircus*	5	261	–	256.0 (0.05)	–	NC_005044
Cat	*Felis catus*	5	283	–	278.3 (0.05)	–	NC_001700
Sika deer	*Cervus nippon*	5	429	–	424.4 (0.02)	–	NC_006993
Common raccoon	*Procyon lotor*	10	–	278	–	273.7 (0.08)	NC_009126
Japanese badger	*Meles anakuma*	5	–	285	–	281.2 (0.05)	NC_009677
Siberian weasel	*Mustela sibirica*	5	–	295	–	291.6 (0.04)	NC_020637
Red fox	*Vulpes vulpes*	5	–	313	–	309.1 (0.05)	NC_008434
Raccoon dog	*Nyctereutes procyonoides*	7	–	326	–	322.0 (0.10)	NC_013700
Dog	*Canis lupus familiaris*	5	–	336	–	332.6 (0.10)	NC_002008
Masked palm civet	*Paguma larvata*	5	–	356	–	351.2–352.2 (0.02–0.06)	NC_029403
Bactrian camel	*Camelus bactrianus*	5	370	359	365.9 (0.04)	352.0 (0.03)	NC_009628
Alpaca	*Vicugna pacos*	4	370	360	366.1 (0.23)	354.2 (0.07)	KU168760
Asiatic black bear	*Ursus thibetanus*	12	–	367	–	361.6–364.5 (0.06–0.1)	NC_009971
Brown rat	*Rattus norvegicus*	5	–	373	–	369.8 (0.04)	NC_001665
House mouse	*Mus musculus*	5	–	374	–	370.4 (0.03)	NC_005089
Rhesus macaque	*Macaca mulatta*	6	–	487	–	481.6–483.7 (0.005–0.16)	NC_005943
Japanese macaque	*Macaca fuscata*	6	–	488	–	481.6 (0.07)	NC_025513
Japanese quail	*Coturnix japonica*	5	–	280	–	276.5 (0.05)	NC_003408
Domestic turkey	*Meleagris gallopavo domesticus*	5	–	285	–	281.0 (0)	NC_010195
Mallard	*Anas platyrhynchos*	5	–	317	–	311.3 (0.04)	NC_009684
Chicken	*Gallus gallus*	5	–	344	–	340.2 (0.05)	NC_001323

N/A: The length is not available because the registered sequences (AB015070 and AB015085) do not cover the primer binding sites

TCA-3'; Reverse: 5'-AGATAGAAACCGACCTGGATT-3') was newly designed to amplify a conserved region of the 16S rRNA gene of vertebrates (Supplementary Table S1). As a qPCR standard DNA, the sequence of the conserved region corresponding to the human mtDNA position 2980–3104 of the revised Cambridge Reference Sequence (rCRS, NC_012920) [43] was inserted into the plasmid pUCFA and cloned (commercially available from Fasmac Co., Ltd., Kanagawa, Japan). The concentration of the purified plasmid DNA solution was quantified by measuring the absorbance at 260 nm (A260) using Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific). The copy number (copies/μL) of the plasmid DNA in the solution was then calculated by dividing the concentration by the average molecular weight. The solution was serially diluted to the following concentrations and subsequently used for the standard curve analysis: 10^7, 10^6, 10^5, 10^4, 10^3, 10^2, and 10^1 copies/μL. The qPCR analysis was performed using a SmartCycler II system (Cepheid, Sunnyvale, CA). Each 25-μL reaction contained 12.5 μL of SYBR Premix Ex Taq (Takara Bio Inc.), 0.2 μM of the primers, and 2 μL of sample DNA or standard DNA. The amplification conditions were as follows: 95 °C for 30 s followed by 40 cycles of 95 °C for 5 s and 60 °C for 20 s. Each standard DNA was quantified in triplicate.

Multiplex PCR

Multiplex PCR was performed using Multiplex PCR Assay Kit Ver. 2 (Takara Bio Inc.) in a 25-μL reaction mixture containing 5000 copies of sample DNA, 12.5 μL of

Table 2 List of primers used in the multiplex PCR

Primer name	Sequence (5'-3')	Conc. (μM)	Target region	5' label
DL_UniF	CACCATCAGCACCCAAAGCT[a]	0.10	D-loop	VIC (Green)
DL_UniR	ATGGGCCCGGAGCGAGAAGAG	0.10		–
DL_Bird_UniF	TCGTGCATACATTTATATTCCACATA	0.50	D-loop	NED (Yellow)
DL_Bird_UniR1	GTGTACGATTAATAAATCCATCTGGTAC	0.20		–
DL_Bird_UniR2	GTGGACGATCAATAAATCCATCTGATAC	0.40		–
CYTB_UniF	GACCAATGATATGAAAAATCATCGTTGT[b]	0.80	cyt *b*	6-FAM (Blue)
CYTB_CattleR	GGCTGGAAGGTCGATGAATGTA	0.15		–
CYTB_RabbitR	GTGAAAATTTGAATTATAAGGCACAG[c]	0.20		–
CYTB_HumanR	ATAGTCCTGTGGTGATTTGGAGGATC	0.10		–
CYTB_SheepR	TGCTAGGAATAGGTCTGTTGGAATC	0.20		–
CYTB_PigR	GTCTGATGTGTAATGTATTGCTAAGAAC[d]	0.15		–
CYTB_HorseR	ACGGATGAGAAGGCAGTTGTC	0.06		–
CYTB_GoatR	CGACAAATGTGAGTTACAGAGGGA[e]	0.08		–
CYTB_CatR	TGATTCAGCCATAATTAACGTCG	0.20		–
CYTB_CamelR	GTAGGAGCCGTAGTAAAGCCCA	0.10		–
CYTB_SikaR	GCTGTGGCTATAACTGTAAATAGGACA	1.60		–

[a]Modified from Nakamura et al.[7]

[b]Modified from Pääbo et al.[41]

[c]Same as Tobe et al.[31]

[d]Modified from Tobe et al.[31]

[e]Modified from Matsunaga et al.[42]

2 × Multiplex PCR Buffer (Mg^{2+}, dNTP plus), 0.125 μL of Multiplex Enzyme Mix, and the primer mix listed in Table 2. Amplification was carried out on an iCycler (Bio-Rad Laboratories, Inc., Hercules, CA) using the following conditions: initial denaturation at 94 °C for 1 min followed by 27 cycles of 94 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s, and a final extension at 72 °C for 10 min.

Allelic ladder

To construct an allelic ladder for electrophoretic analysis using the 3500xL Genetic Analyzer, fragments of each target species except for house mouse and rhesus macaque were amplified by singleplex PCR. The fragments were sequenced using unlabeled primers and BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) to confirm target-specific amplification. The amplified fragments by the singleplex PCR were mixed together so that all peak heights detected by the 3500xL Genetic Analyzer were balanced. The mix was subsequently purified using Performa DTR Gel Filtration Cartridges (Edge Biosystems, Gaithersburg, MD). The constructed allelic ladder was stored at – 20 °C until use.

Electrophoresis and data analysis

All PCR products were analyzed on the 3500xL Genetic Analyzer with a 36-cm array and POP-4 polymer (Thermo Fisher Scientific). Spectral calibration was initially performed using DS-36 Matrix Standard (Dye Set J6). 1 μL of the PCR product or the allelic ladder was added to 9.5 μL of Hi-Di Formamide and 0.5 μL of GeneScan 600 LIZ dye Size Standard v2.0, and subsequently denatured at 95 °C for 3 min. The samples and the allelic ladder were injected at 1.2 kV for 24 s and electrophoresed at 13 kV for 1600 s. Electrophoresis results were analyzed using GeneMapper ID-X Software v1.4 with a peak amplitude threshold of 175 RFU and customized panel and bin sets.

Species specificity, sensitivity, and repeatability

Species specificity and the intraspecific variability of the assay were assessed by analyzing all samples for each species listed in Table 1. Sensitivity of the multiplex assay was evaluated by testing serial dilutions (10,000, 1000, 500, 100, 10 copies of mtDNA) from cattle, human, pig, dog, and chicken samples. Each test was repeated in triplicate. Repeatability of the assay was also evaluated by testing one

individual of each species in triplicate. In order to assess the sizing precision of capillary electrophoresis on the 3500xL Genetic Analyzer, one microliter of the allelic ladder was injected onto each of the 24 capillaries and electrophoresed. The mean peak size (bp) and standard deviation within a single run (i.e., $n = 24$) were calculated for each peak, and this was repeated five times.

Mixture studies

To evaluate the ability to identify species from DNA mixture, three sets of DNA mixture (human with cattle, human with dog, and human with chicken) were tested with a total input of 5000 copies of mtDNA and one of the following mixture ratios: 50:1, 20:1, 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:10, 1:20, and 1:50. An additional two sets of DNA mixture (brown rat with house mouse and Japanese wild boar with pig) were analyzed in the same manner to evaluate whether two adjacent peaks were detectable by the GeneMapper ID-X software. An experimental mixture sample, which contained 5000 mtDNA copies of each 10 species (cattle, European rabbit, sheep, pig, horse, goat, sika deer, domestic turkey, mallard, and chicken), was also tested.

Case studies

To demonstrate the utility of the assay in forensic casework, a total of 31 unknown samples collected from crime or accident scenes were tested. These samples were either bloodstains, a saliva stain, tissues, partial bones, hairs, or feces (Supplementary Table S2).

A 1 bp-InDel and phylogenetic relationship of wild boars and domestic pigs

To investigate the relationship between the presence of a 1 bp-insertion/deletion (InDel) in the D-loop region (NC_000845.1:m.137delC) and the phylogeny among European and Asian wild boars and domestic pigs, complete mtDNA sequences of 307 Eurasian wild boars and domestic pigs were downloaded from GenBank. A complete mtDNA sequence of desert warthog (*Phacochoerus africanus*) was also downloaded and used as an outgroup of the phylogenetic analysis. Multiple sequence alignment and construction of a neighbor-joining phylogenetic tree were carried out using MAFFT online version 7 (https://mafft.cbrc.jp/alignment/server/large.html) [44]. The generated tree was visualized using iTOL v6.1.1 (https://itol.embl.de/) [45].

Results

Species specificity

The target mtDNA regions of 24 mammalian and four poultry species were successfully amplified in all tested samples using the designed primers. Some differences were observed between the expected and observed size of the products, most likely because of differences in electrophoretic mobility. Nevertheless, the PCR products of all species were distinguished from each other through the capillary electrophoresis, except for Japanese macaque and rhesus macaque (Fig. 1). Although the sizes of the peaks at the cyt *b* locus in the observed electropherograms were the same between Bactrian camel and alpaca, or pig and Japanese wild boar, each species was identifiable based on the D-loop results. The source species of each PCR product was confirmed by sequencing and subsequent NCBI BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Intraspecific length variation of the PCR product was observed in several species. Three types of peak size at the D-loop locus in the electropherograms were observed in rhesus macaque, and the shortest peak was the same size as the peak of Japanese macaque (Fig. 2). Asiatic black bear had four types of peak size at the D-loop locus, at most with 3 bp difference, although these peaks were easily distinguished from those of other species. Some individuals of masked palm civet had two peaks at the D-loop locus with 1 bp difference. Sequencing analysis indicated that the two peaks were caused by length heteroplasmy. The longer peak was the same size as Bactrian camel. However, the two species were distinguishable by the presence or absence of the camel peak at the cyt *b* locus in the electropherogram. No intraspecific variation was observed in the remaining species.

A low non-targeted peak was observed at the D-loop locus in the electropherogram only when horse DNA was amplified. Because the size of the peak differed from those of the other species, it did not affect species identification.

Sensitivity

The sensitivity study showed successful identification from 10,000 copies down to 100 copies of input mtDNA for all the tested species except for chicken (from 10,000 copies down to 500 copies). No peak was observed from 100 copies of input chicken mtDNA. Based on a calculation using the human control DNA 9947A (0.1 ng/μL), 100 copies of mtDNA is equivalent to approximately 0.15 pg of total DNA.

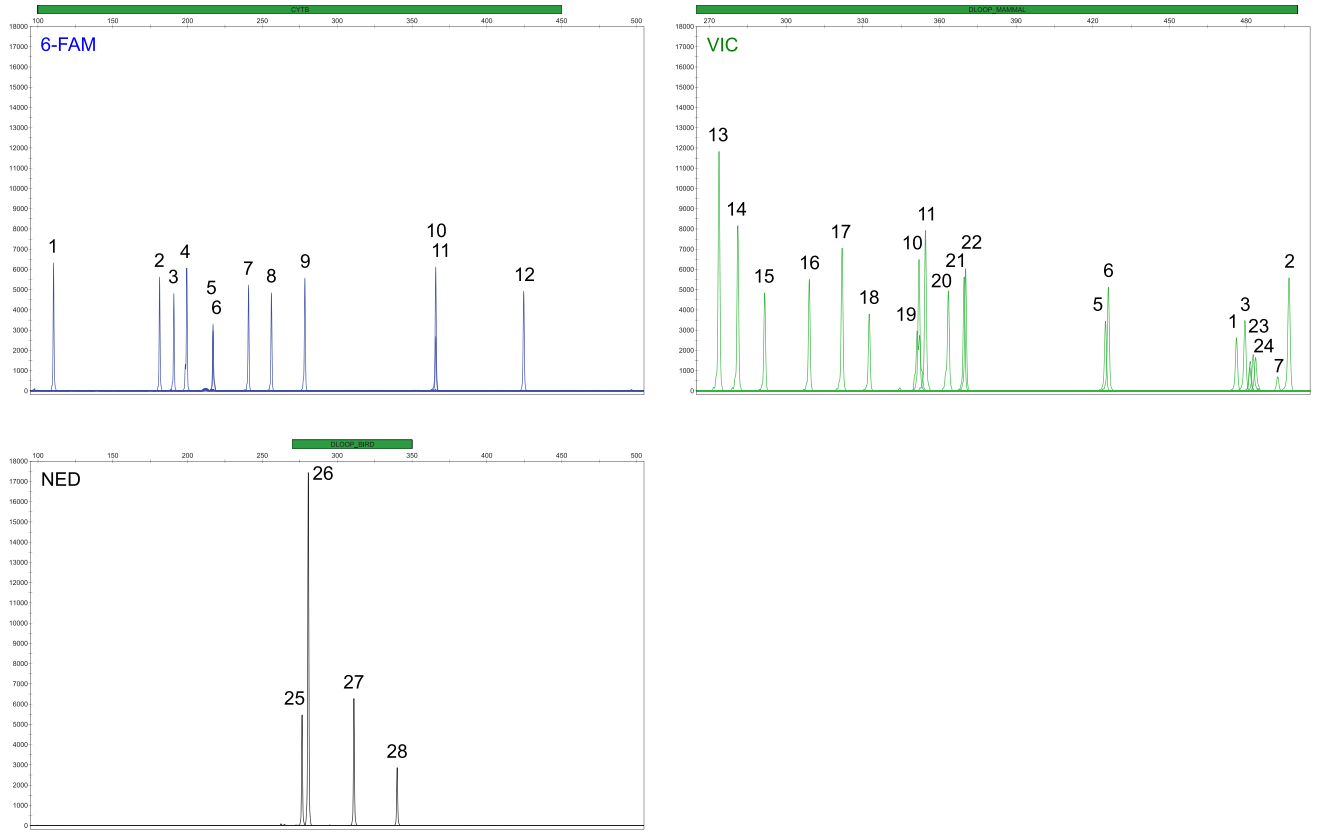

Fig. 1 Electropherograms of each tested species shown by "Overlay All" plots using GeneMapper ID-X Software v1.4. Peaks are (1) cattle, (2) European rabbit, (3) human, (4) sheep, (5) Japanese wild boar, (6) pig, (7) horse, (8) goat, (9) cat, (10) Bactrian camel, (11) alpaca, (12) sika deer, (13) common raccoon, (14) Japanese badger, (15) Siberian weasel, (16) red fox, (17) raccoon dog, (18) dog, (19) masked palm civet, (20) Asiatic black bear, (21) brown rat, (22) house mouse, (23) Japanese macaque, (24) Rhesus macaque, (25) Japanese quail, (26) domestic turkey, (27) mallard, and (28) Chicken

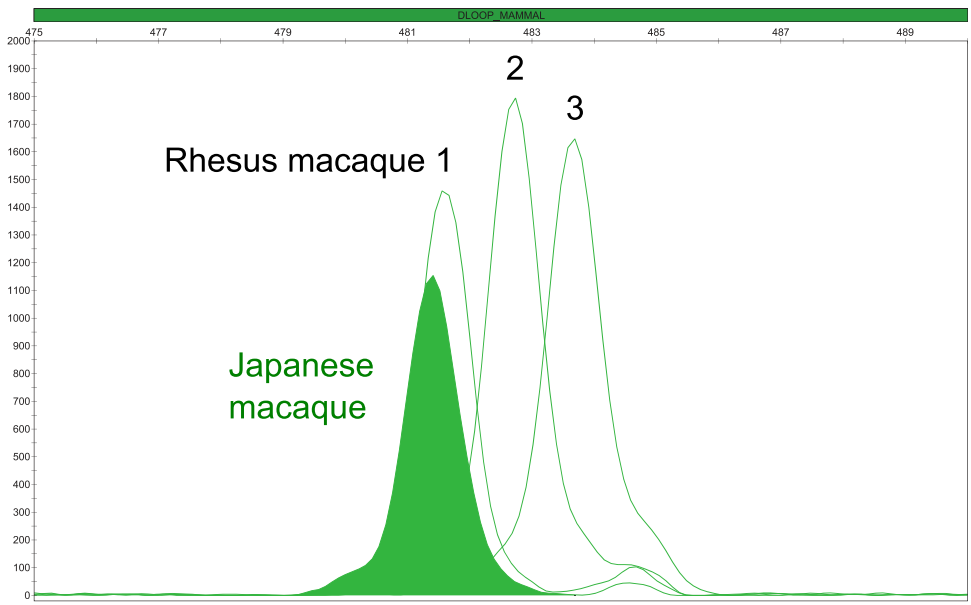

Fig. 2 Peak positions of Japanese macaque and individuals of rhesus macaque (1–3)

Repeatability

The accuracy of species identification based on the peak size detected by capillary electrophoresis depends on the precision of the peak size measurement as well as species specificity of the peak size. To increase the precision, the assay used a high-density internal size standard, 600 LIZ Size Standard v2.0. The peak size data was collected five times by injections of the constructed allelic ladder (Fig. 3) for all 24 capillaries on the 3500xL Genetic Analyzer. The sizing precision of the fragment analysis was assessed by calculating the standard deviations (SDs) of the size measurement for each peak of the allelic ladder among the 24 capillaries in a single run. The maximum SD of a peak was 0.09 bp (Fig. 4) and was lower than the threshold (0.15 bp) of the sizing precision described in the performance check section of the User Guide of the 3500xL Genetic Analyzer [46]. To detect a 1-bp difference in the peak size between species correctly, peak size measurements should lie within a ± 0.5 bp "window" around the size obtained for each peak in the allelic ladder with a sufficiently high probability [47]. The observed SD in this study (< 0.1 bp) indicates that the probability is over 99.99% because the five SDs were still less than 0.5 bp.

The repeatability was evaluated by testing one sample from each species multiple times. The maximum SD of the peak size among the target species was 0.06 bp. In addition, testing all samples once showed that the SDs of the observed peaks for each species were less than 0.5 bp (max 0.23 bp).

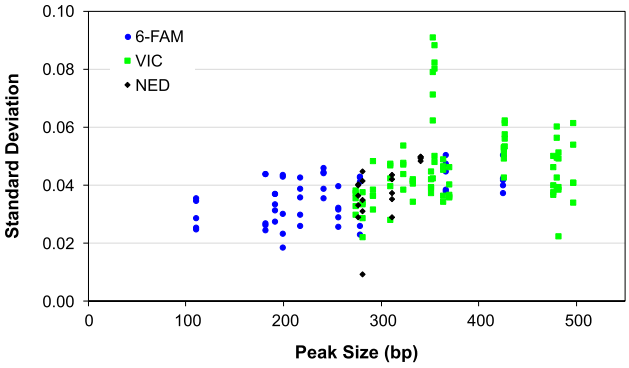

Fig. 4 Sizing precision of the allelic ladder across five runs on a 3500xL Genetic Analyzer. Each dot indicates the standard deviation of the observed peak sizes on 24 capillaries within a single run

Mixture studies

It is important to identify species from mixtures of DNA because casework samples are often contaminated with human or environmental DNA. Three pairs (human with cattle, human with dog, and human with chicken) of DNA mixture with ratios ranging from 50:1 to 1: 50 (50:1, 20:1, 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:10, 1:20, and 1:50) were tested to evaluate the assay's ability to identify both major and minor contributors correctly. All major contributors were identifiable in all cases. The peak of each minor contributor was detected and its source species was identified correctly in the mixture ratios between 20:1 and 1:20 (≧240 copies

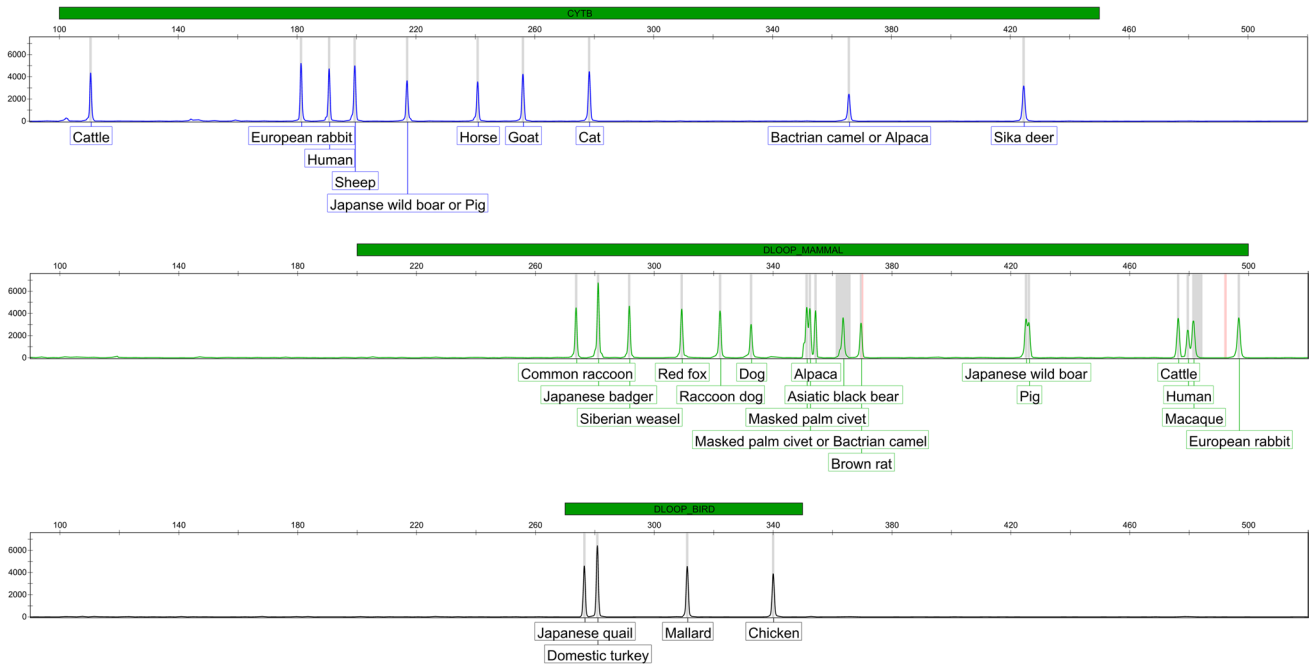

Fig. 3 An electropherogram of the allelic ladder. The three panels correspond to 6-FAM, VIC, and NED dye-labeled peaks from top to bottom

minor contributor) for the human–chicken DNA mixture, and in the ratios between 50:1 and 1:20 for the human–cattle and human–dog mixtures. Another two pairs (brown rat with house mouse, and Japanese wild boar with pig) of DNA mixture, which have 1 bp-adjacent peaks, were tested in a similar way. The peak of each minor contributor was not detected in any ratios for the brown rat and house mouse DNA mixture because it was masked by the tail of the major contributor peak. Similarly, the peak of each minor contributor was not detected in any ratios for the DNA mixture of Japanese wild boar and pig, except for the case when pig and Japanese wild boar were mixed in the mixture ratios of 1:1, 1:2, or 1:5.

The assay's ability to identify multiple species in a sample was also demonstrated using the experimental DNA mixture sample composed of ten species. Each species in the sample was correctly identified (Fig. 5).

Case studies

In forensic casework, samples may be degraded by environmental exposure and/or contaminated with human or other environmental DNA. The utility of the developed assay was evaluated using a total of 31 forensic casework samples. The assay unambiguously identified the source of DNA in each sample (Supplementary Table S2): 6 of them were derived from sika deer; 5 from cat; 5 from common raccoon; 5 from wild boar; 4 from chicken; 3 from red fox; 1 from cattle; 1 from macaque; and 1 from raccoon dog. The identification results were consistent with the situations of the casework from which the samples were obtained.

A 1 bp-InDel and phylogenetic relationship of wild boars and domestic pigs

The assay distinguishes Japanese wild boar and pig based on a 1-bp difference of the fragment length of the D-loop region (Fig. 1). To evaluate the assay's ability to identify Japanese wild boar and pig correctly, the phylogenetic relationship of Eurasian wild boar and pig was analyzed. The complete mtDNA sequences of the 307 Eurasian wild boars and pigs were aligned with that of desert warthog (*Phacochoerus africanus*). The phylogenetic analysis showed that they were mainly grouped into two lineage groups: the European and Asian lineages (Supplementary Fig. S1). An InDel was found in the part of the D-loop region amplified by the assay. The presence or absence of the 1-bp deletion (NC_000845.1:m.137delC) corresponded well to the two lineage groups. Almost all the East Asian wild boars and indigenous pig breeds, as well as Japanese wild boars tested in this study, had the 1-bp deletion. There were some exceptions to this lineage–InDel relationship. Six individuals (Yorkshire breed, Chinese Meishan breed, or native breed of the Andaman and Nicobar Islands in India) were phylogenetically grouped into the Asian lineage although they did not have the deletion. There were also some exceptional individuals whose geographic origin did not match the phylogenetic position. Eight individuals of European pig breeds (Berkshire, Yorkshire, or Large White) were grouped into the Asian lineage, while a Chinese Northeast wild boar was grouped into the European lineage. Besides, the multiple sequence alignment showed that 11 individuals had other InDels in the D-loop region amplified by the assay.

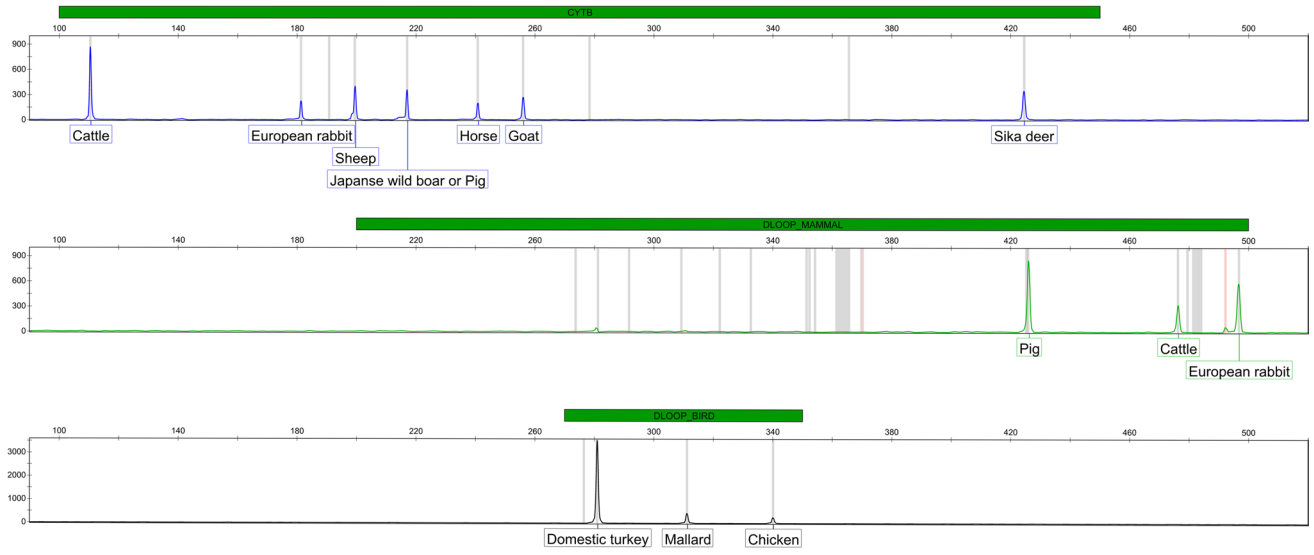

Fig. 5 Amplification of an experimental mixture sample consisting of 10 species (cattle, European rabbit, sheep, pig, horse, goat, sika deer, domestic turkey, mallard, and chicken)

Discussion

The developed multiplex PCR assay enables simultaneous identification of 26 mammalian and poultry species from biological samples recovered from crime scenes. The assay is cost-effective because it does not require specific fluorescently labeled primers for each species and only three forward primers were labeled with fluorescent dyes. The species specificity test and the repeatability test demonstrated that the assay identifies species accurately based on peak detection with 1 bp-precision using the internal size standard and the allelic ladder. The sensitivity study indicated that the assay was highly sensitive, with a detection limit of 100 copies of mtDNA. This is sufficient to detect one mammalian cell because somatic cells have a high copy number of mtDNA ranging from hundreds to thousands of copies depending on tissue type [23–25]. The sensitivity level was similar to that in a previous study [32] but was achieved using fewer PCR cycles (27 cycles compared with 30 cycles). It is important to note that this fewer PCR cycles leave room for further sensitivity improvement, because additional cycles might enable us to identify species from ultra-low-level samples containing less than 100 copies of mtDNA. The high sensitivity coupled with any PCR product below ~ 500 bp in size made the assay robust for degraded samples. The assay was able to identify species even from degraded casework samples such as partial bones, hairs, and feces. The assay also enables the analysis of DNA mixtures. The peaks of different species contained in a DNA mixture were distinguishable as long as they were separated by > 1 bp (Fig. 5). Although species identification was difficult in the mixture test using brown rat and house mouse DNA or Japanese wild boar and pig DNA, forensic scientists rarely encounter mixtures of these species in casework.

The assay requires careful consideration for some species due to their intraspecific variation. The species specificity tests in the present study showed that rhesus macaque (*Macaca mulatta*) had intraspecific length variation in the amplified D-loop region. Some individuals of rhesus macaque had the same PCR product size as Japanese macaque (*Macaca fuscata*) (Fig. 2). According to NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), there are intraspecific length variations in crab-eating macaque (*Macaca fascicularis*) as well as rhesus macaque (data not shown) and some individuals from crab-eating macaque and rhesus macaque are expected to show the same PCR product size as Japanese macaque. Therefore, we cannot conclude firmly that a sample is derived from Japanese macaque based on the assay result. However, there were clear differences in the PCR product size between these macaque species and the other species tested in this study. This indicates that these macaque species can be identified at least as one of the *Macaca* species.

No intraspecific variation of the D-loop region was observed in the human control DNA samples used in this study. However, the targeted D-loop region contains the HV1 in which intraspecific length variation called "C-Stretch" has been reported in humans [48–50]. Therefore, some human samples can be identified as *Macaca* species based on the PCR product size of the D-loop. Even in this case, the assay can correctly identify the sample as human by the presence of the cyt *b* human-specific peak because no cross-amplification of the cyt *b* region in *Macaca* species by the human specific primer was observed.

The tested domestic pig (*Sus scrofa domesticus*) samples, which were obtained from European breeds (Yorkshire and Duroc) reared in Japan, were successfully distinguished from Japanese wild boar (*Sus scrofa leucomystax*) samples even though they are phylogenetically closely related to each other. An InDel made the size of the D-loop peak derived from Japanese wild boar 1 bp shorter than that of the pig samples. However, identification of pig and Japanese wild boar requires careful consideration. Firstly, the multiple sequence alignment and the phylogenetic tree showed that almost all the East Asian wild boars and pig breeds had the 1 bp-deletion (Supplementary Fig. S1), and therefore the D-loop peak derived from these East Asian breeds may be the same size as that of Japanese wild boar. Secondly, other InDels in the amplified region may affect the assignment of samples to pig or Japanese wild boar although the present analysis showed such InDels rarely occur. Thirdly, crossbreeding between European pigs and East Asian pigs complicates the interpretation of electrophoresis results. These crossbred European pigs such as Berkshire [51] may have the 1-bp deletion if they inherit mtDNA from the Asian lineage. Therefore, even if the peak of Japanese wild boar is detected by the assay, the possibility that the DNA sample is derived from a pig cannot be excluded. Fourthly, genetic introgression of European domestic pig breeds into Japanese wild boars [52, 53] might also affect the conclusion from the assay. The hybrid individuals, called "Inobuta" in Japanese, will be identified as either of the two species, depending on their maternal lineage. To avoid this misidentification, forensic scientists should take the possibility that an analyzed sample originates from a hybrid individual into consideration, as proposed by Amorim et al. [54].

In this study a cost-effective, highly sensitive, and straightforward assay for simultaneous species identification was developed. The validation studies demonstrated the efficacy, accuracy, and reliability of the assay. We markedly reduced the species-specific primers in the multiplex by using the D-loop universal primers while keeping the

number of identifiable species high. In addition, we minimized the number of fluorescent dyes used in the assay: the multiplex reaction requires only three fluorescent dyes for the peak detection on a 3500xL Genetic Analyzer. These will facilitate further development of the assay, for example, depending on the geographical origin of samples, by adding new primers to identify species untested in this study. One of the advantages of the assay is that it can identify numerous species even from mixed biological samples for which the analysis based on the conventional sequencing is inappropriate. This advantage makes it easy for forensic laboratories to conduct species identification tests on the evidence left unanalyzed because it contains DNA mixture from multiple source species. Although DNA metabarcoding using MPS technologies is now an option for researchers handling mixed DNA samples, the assay is much more cost-effective than DNA metabarcoding. Furthermore, the procedure of the assay is time-saving: the time required from multiplex PCR to the interpretation of results is only 2 h. The assay is therefore suitable for local forensic laboratories that have no MPS platform. Previous methods based on real-time PCR coupled with species-specific probes are indeed simple, cost-effective, and timesaving, but the number of target species is limited by the number of detection channels of the instruments [38]. Another type of real-time PCR methods using melting curve analysis [15, 35, 55] can increase the number of target species compared with the probe-based methods. However, this type of methods is not appropriate for DNA mixture because the melting curve profile of a particular species can change depending on mixing of another species and its mixture ratios. In the case of forensic casework samples that are suspected of DNA mixture, researchers may not conclude whether the profile was obtained from single source or not. In comparison with such real-time PCR methods, the developed assay can identify numerous species in any mixture ratios as long as their peaks are detectable.

The results presented in this study also indicate that the developed assay has a potential for application not only to routine forensic investigations but to other fields such as food authentication and textile authentication.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00414-021-02711-y.

Acknowledgements The authors gratefully acknowledge the following individuals or organizations for their help in obtaining samples: Dr. Makoto Asano, Dr. Takashi Hayakawa, Dr. Satoshi Nagaoka, and Dr. Ryo Tadano (Gifu University); Dr. Hiroo Imai (Kyoto University); Kyoto City Zoo (Japan); Shunan Tokuyama Zoo (Japan); Takaoka Kojyo Zoo (Japan); KE'KEN Textile Testing & Certification Center (Japan). The authors would also like to thank Dr. Kohei Nakamura (Gifu University) for the use of the Nanodrop 1000 Spectrophotometer.

This work was supported in part by the Cooperative Research Program of the Primate Research Institute, Kyoto University (No. 2016-D-19).

Authors' contributions **Chikahiro Mori:** Conceptualization, methodology, validation, formal analysis, investigation, resources, data curation, writing (original draft), visualization, project administration. **Shuichi Matsumura:** conceptualization, methodology, resources, writing (review and editing), supervision.

Funding The authors did not receive support from any organization for the submitted work.

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare they have no competing interests.

References

1. Schneider PM, Seo Y, Rittner C (1999) Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. Int J Legal Med 112:315–316. https://doi.org/10.1007/s004140050257
2. Lorenzini R (2005) DNA forensics and the poaching of wildlife in Italy: A case study. Forensic Sci Int 153:218–221. https://doi.org/10.1016/j.forsciint.2005.04.032
3. Sato I, Nakaki S, Murata K et al (2010) Forensic hair analysis to identify animal species on a case of pet animal abuse. Int J Legal Med 124:249–256. https://doi.org/10.1007/s00414-009-0383-2
4. Stern AW, Lamm CG (2011) Utilization of paw prints for species identification in the Canidae family. J Forensic Sci 56:1041–1043. https://doi.org/10.1111/j.1556-4029.2011.01768.x
5. Tillmar AO, Dell'Amico B, Welander J, Holmlund G (2013) A universal method for species identification of mammals utilizing next generation sequencing for the analysis of DNA mixtures. PLoS ONE 8:e83761. https://doi.org/10.1371/journal.pone.0083761
6. Bravi CM, Lirón JP, Mirol PM et al (2004) A simple method for domestic animal identification in Argentina using PCR-RFLP analysis of cytochrome *b* gene. Leg Med 6:246–251
7. Nakamura H, Muro T, Imamura S, Yuasa I (2009) Forensic species identification based on size variation of mitochondrial DNA hypervariable regions. Int J Legal Med 123:177–184. https://doi.org/10.1007/s00414-008-0306-7
8. Naue J, Lutz-Bonengel S, Pietsch K et al (2012) Bite through the tent. Int J Legal Med 126:483–488. https://doi.org/10.1007/s00414-012-0674-x
9. Savolainen P, Lundeberg J (1999) Forensic evidence based on mtDNA from dog and wolf hairs. J Forensic Sci 44:77–81. https://doi.org/10.1016/S1353-1131(99)90078-0
10. Schulz I, Schneider PM, Olek K et al (2006) Examination of postmortem animal interference to human remains using cross-species multiplex PCR. Forensic Sci Med Pathol 2:95–101. https://doi.org/10.1385/FSMP:2:2:95

11. Alacs EA, Georges A, FitzSimmons NN, Robertson J (2010) DNA detective: A review of molecular approaches to wildlife forensics. Forensic Sci Med Pathol 6:180–194. https://doi.org/10.1007/s12024-009-9131-7
12. Johnson RN, Wilson-Wilde L, Linacre A (2014) Current and future directions of DNA in wildlife forensic science. Forensic Sci Int Genet 10:1–11. https://doi.org/10.1016/j.fsigen.2013.12.007
13. Ewart KM, Frankham GJ, McEwing R et al (2018) An internationally standardized species identification test for use on suspected seized rhinoceros horn in the illegal wildlife trade. Forensic Sci Int Genet 32:33–39. https://doi.org/10.1016/j.fsigen.2017.10.003
14. Summerell AE, Frankham GJ, Gunn P, Johnson RN (2019) DNA based method for determining source country of the short beaked echidna (*Tachyglossus aculeatus*) in the illegal wildlife trade. Forensic Sci Int 295:46–53. https://doi.org/10.1016/j.forsciint.2018.11.019
15. Ouso DO, Otiende MY, Jeneby MM et al (2020) Three-gene PCR and high-resolution melting analysis for differentiating vertebrate species mitochondrial DNA for biodiversity research and complementing forensic surveillance. Sci Rep 10:1–13. https://doi.org/10.1038/s41598-020-61600-3
16. Hsieh HM, Chiang HL, Tsai LC et al (2001) Cytochrome *b* gene for species identification of the conservation animals. Forensic Sci Int 122:7–18. https://doi.org/10.1016/S0379-0738(01)00403-0
17. Iyengar A (2014) Forensic DNA analysis for animal protection and biodiversity conservation: A review. J Nat Conserv 22:195–205. https://doi.org/10.1016/j.jnc.2013.12.001
18. Arulandhu AJ, Staats M, Hagelaar R, et al (2017) Development and validation of a multi-locus DNA metabarcoding method to identify endangered species in complex samples. Gigascience 6. https://doi.org/10.1093/gigascience/gix080
19. Conte J, Potoczniak MJ, Mower C, Tobe SS (2019) ELEquant: a developmental framework and validation of forensic and conservation real-time PCR assays. Mol Biol Rep 46:2093–2100. https://doi.org/10.1007/s11033-019-04660-7
20. Mafra I, Ferreira IMPLVO, Oliveira MBPP (2008) Food authentication by PCR-based methods. Eur Food Res Technol 227:649–665. https://doi.org/10.1007/s00217-007-0782-x
21. Galimberti A, De Mattia F, Losa A et al (2013) DNA barcoding as a new tool for food traceability. Food Res Int 50:55–63. https://doi.org/10.1016/j.foodres.2012.09.036
22. Dobrovolny S, Blaschitz M, Weinmaier T et al (2019) Development of a DNA metabarcoding method for the identification of fifteen mammalian and six poultry species in food. Food Chem 272:354–361. https://doi.org/10.1016/j.foodchem.2018.08.032
23. Robin ED, Wong R (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol 136:507–513. https://doi.org/10.1002/jcp.1041360316
24. Holland MM, Parsons TJ (1999) Mitochondrial DNA Sequence Analysis - Validation and Use for Forensic Casework. Forensic Sci Rev 11:21–50
25. Tobe SS, Linacre AMT (2008) A technique for the quantification of human and non-human mammalian mitochondrial DNA copy number in forensic and other mixtures. Forensic Sci Int Genet 2:249–256. https://doi.org/10.1016/J.FSIGEN.2008.03.002
26. Kocher TD, Thomas WK, Meyer A et al (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci 86:6196–6200. https://doi.org/10.1073/pnas.86.16.6196
27. Johns GC, Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Mol Biol Evol 15:1481–1490. https://doi.org/10.1093/oxfordjournals.molbev.a025875
28. Parson W, Pegoraro K, Niederstätter H et al (2000) Species identification by means of the cytochrome b gene. Int J Legal Med 114:23–28. https://doi.org/10.1007/s004140000134
29. Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System: Barcoding. Mol Ecol Notes 7:355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x
30. Dawnay N, Ogden R, McEwing R et al (2007) Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Sci Int 173:1–6. https://doi.org/10.1016/j.forsciint.2006.09.013
31. Tobe SS, Linacre AMT (2008) A multiplex assay to identify 18 European mammal species from mixtures using the mitochondrial cytochrome *b* gene. Electrophoresis 29:340–347. https://doi.org/10.1002/elps.200700706
32. Ramón-Laca A, Linacre AMT, Gleeson DM, Tobe SS (2013) Identification multiplex assay of 19 terrestrial mammal species present in New Zealand. Electrophoresis 34:3370–3376. https://doi.org/10.1002/elps.201300324
33. Kitano T, Umetsu K, Tian W, Osawa M (2007) Two universal primer sets for species identification among vertebrates. Int J Legal Med 121:423–427. https://doi.org/10.1007/s00414-006-0113-y
34. Pereira F, Carneiro J, Matthiesen R et al (2010) Identification of species by multiplex analysis of variable-length sequences. Nucleic Acids Res 38:e203–e203. https://doi.org/10.1093/nar/gkq865
35. Ishida N, Sakurada M, Kusunoki H, Ueno Y (2018) Development of a simultaneous identification method for 13 animal species using two multiplex real-time PCR assays and melting curve analysis. Leg Med 30:64–71. https://doi.org/10.1016/J.LEGALMED.2017.11.007
36. Mori C, Matsumura S (2021) Current issues for mammalian species identification in forensic science: a review. Int J Legal Med 135:3–12. https://doi.org/10.1007/s00414-020-02341-w
37. Coghlan ML, Haile J, Houston J et al (2012) Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genet 8:e1002657. https://doi.org/10.1371/journal.pgen.1002657
38. Köppel R, van Velsen F, Ganeshan A et al (2020) Multiplex real-time PCR for the detection and quantification of DNA from chamois, roe, deer, pork and beef. Eur Food Res Technol 246:1007–1015. https://doi.org/10.1007/s00217-020-03468-1
39. Scientific Working Group on DNA Analysis Methods (SWGDAM) (2016) Validation Guidelines for DNA Analysis Methods. https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf. Accessed 1 Jul 2020
40. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054
41. Pääbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. Nucleic Acids Res 16:9775–9787. https://doi.org/10.1093/nar/16.20.9775
42. Matsunaga T, Chikuni K, Tanabe R et al (1999) A quick and simple method for the identification of meat species and meat products by PCR assay. Meat Sci 51:143–148. https://doi.org/10.1016/S0309-1740(98)00112-0
43. Andrews RM, Kubacka I, Chinnery PF et al (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147–147. https://doi.org/10.1038/13779
44. Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166. https://doi.org/10.1093/bib/bbx108

45. Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res gkab301. https://doi.org/10.1093/nar/gkab301
46. ThermoFisher Scientific (2018) User Guide: 3500/3500xL Genetic Analyzer with 3500 Series Data Collection Software 3.1 Revision C. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100031809_3500_3500xL_Software_v3_1_UG.pdf. Accessed 30 Apr 2021
47. Smith RN (1995) Accurate size comparison of short tandem repeat alleles amplified by PCR. Biotechniques 18:122–128
48. Bendall KE, Sykes BC (1995) Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet 57:248–256
49. Sekiguchi K, Imaizumi K, Fujii K et al (2008) Mitochondrial DNA population data of HV1 and HV2 sequences from Japanese individuals. Leg Med 10:284–286. https://doi.org/10.1016/j.legalmed.2008.02.002
50. Chen F, Dang YH, Yan CX et al (2009) Sequence-length variation of mtDNA HVS-I C-stretch in Chinese ethnic groups. J Zhejiang Univ Sci B 10:711–720. https://doi.org/10.1631/jzus.B0920140
51. Li M, Tian S, Yeung CKL et al (2014) Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication. Sci Rep 4:1–7. https://doi.org/10.1038/srep04678
52. Takahashi R (2018) Detection of inobuta from wild boar population in Japan by genetic analysis. Rev Agric Sci 6:61–71. https://doi.org/10.7831/ras.6.61
53. Anderson D, Toma R, Negishi Y et al (2019) Mating of escaped domestic pigs with wild boar and possibility of their offspring migration after the Fukushima Daiichi Nuclear Power Plant accident. Sci Rep 9:1–6. https://doi.org/10.1038/s41598-019-47982-z
54. Amorim A, Pereira F, Alves C, García O (2020) Species assignment in forensics and the challenge of hybrids. Forensic Sci Int Genet 48:102333. https://doi.org/10.1016/j.fsigen.2020.102333
55. Lopez-Oceja A, Nuñez C, Baeta M et al (2017) Species identification in meat products: A new screening method based on high resolution melting analysis of *cyt b* gene. Food Chem 237:701–706. https://doi.org/10.1016/j.foodchem.2017.06.004