Non-specific peaks generated by animal DNA during human STR analysis: peak characteristics and a novel analysis method for mixed human/animal samples

Shota Inokuchi[a,b,c*], Natsuko Mizuno[b], Hiroaki Nakanishi[a], Kazuyuki Saito[a], Tetsushi Kitayama[b], Koji Fujii[b], Hiroaki Nakahara[b], and Kazumasa Sekiguchi[b]

[a]*Department of Forensic Medicine, Graduate School of Medicine, Juntendo University, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan*

[b]*National Research Institute of Police Science, 6-3-1 Kashiwanoha, Kashiwa 277-0882, Japan*

[c]*Forensic Science Laboratory, Tokyo Metropolitan Police Department, 3-35-21, Shakujiidai, Nerima, Tokyo 177-0045, Japan*

*Corresponding author: Shota Inokuchi

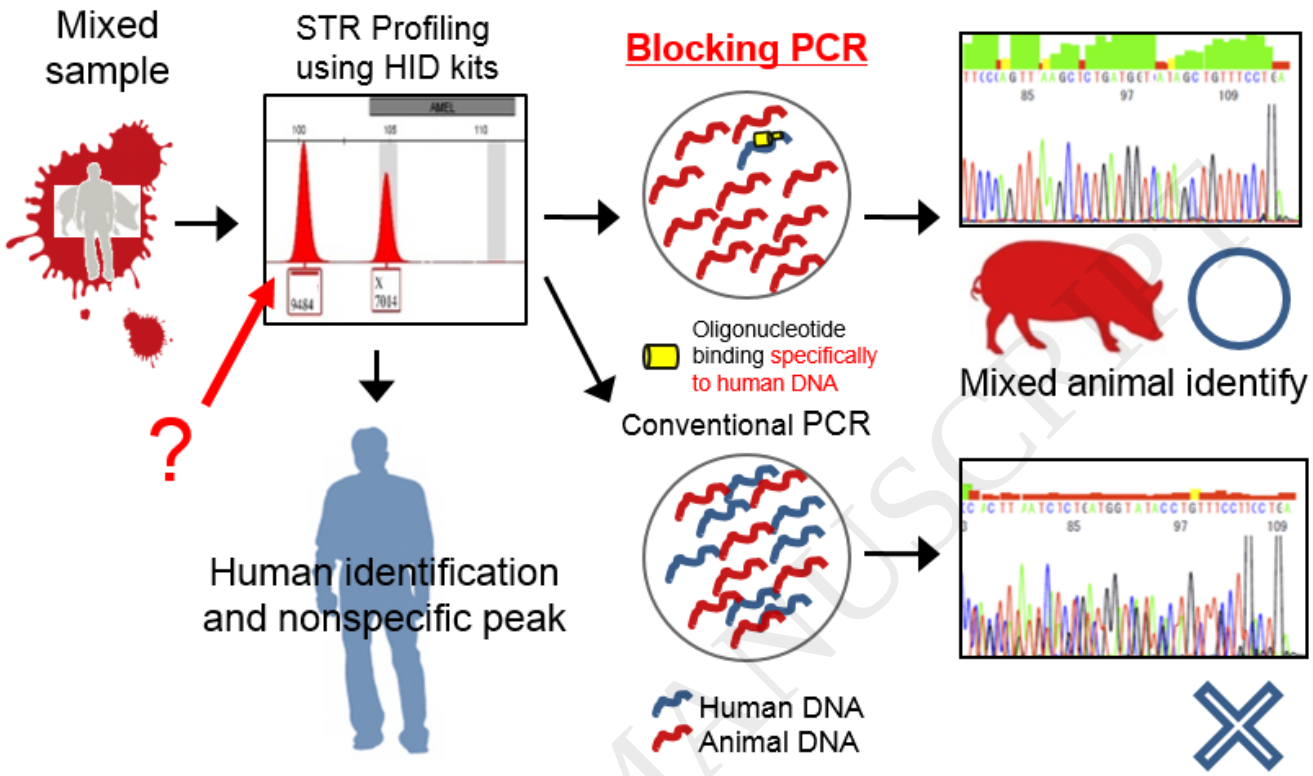
Department of Forensic Medicine, Graduate School of Medicine, Juntendo University, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan

E-mail: s-inoku@juntendo.ac.jp

Telephone: +81-3-3813-3111

Fax: +81-3-3813-3622

Graphical abstract



Highlights

- DNA of animals at the crime scene is amplified by HID kit amelogenin primers
- Non-specific peaks generated by animal DNA differ among animal families
- Blocking PCR was developed for analysis of only animal DNA from mixed samples
- This method could simultaneously perform HID and animal family discrimination

Forensic human identification (HID) laboratories occasionally encounter non-specific peaks generated by non-human DNA. Casework samples for human short tandem repeat (STR) profiling may be contaminated by animal DNA because of the specific environment or situation from which they were obtained. Validation studies for HID kits have reported that non-specific peaks generated from some animals are observed near the human amelogenin peak. In this study, we first revealed that DNA sequences associated with the non-specific peaks generated from animal DNA differ from one animal family to the other. However, non-specific peaks cannot be analyzed using the remainder of polymerase chain reaction (PCR) products left over from conventional HID kits when human and animal DNA are mixed. To overcome this issue, we have developed a

novel analysis method of using non-specific peaks generated from animal DNA in human STR profiling to identify the source of contaminating animal DNA at the family level. The method applied here is termed as blocking PCR, which involves selective animal DNA re-amplification by blocking nontarget human amelogenin DNA amplification using an oligonucleotide probe that specifically binds to human amelogenin using the remaining PCR product from the HID kit. Our data demonstrated that HID and family discrimination among animals that are often encountered in forensic contexts could be performed simultaneously. This study enabled recovery of more information from limited quantities of casework samples contaminated with animal DNA, which would be useful for forensic HID scientists.

Keywords: blocking PCR; mixture sample; non-specific peak; STR

Introduction

Short tandem repeat (STR) profiling using multiplex polymerase chain reaction (PCR) amplification has been widely used in applications involving human identification (HID), including forensic analysis and relationship testing [1]. Casework samples for HID may be contaminated by animal (vertebrate animals) cells because humans are frequently in contact with pets, livestock, wildlife, and animal products. For example, a casework sample such as a human blood stain caused by a dog bite may include human DNA (from human blood) and dog DNA (from dog saliva). Similarly, saliva taken from cutlery may contain human DNA (from human saliva) and contaminating animal DNA (from food). We have encountered forensic casework scenarios that required human identification and animal species identification of a source contaminating the casework sample. For example, a glove was left behind by a suspect at a crime scene. A partial STR profile that did not exclude the suspect was obtained from the surface of the glove. Additionally, a non-specific peak was observed near the amelogenin locus. The suspect had been employed at a pig farm. Police dogs were also utilized to

investigate the suspect using the glove before performing human STR profiling. In this case, it was important to determine whether the non-specific peak was because of pig DNA or dog DNA from a casework sample for human identification. In actual casework, proof indicating human involvement is often not obtained because of degradation and/or a complex mixture; hence, additional information about casework samples is helpful. Therefore, the ability to identify the contaminating animal source in addition to performing HID from such casework samples would be a benefit to trial and criminal investigations in some casework scenarios.

In previous studies involving sequence analysis of 12S or 16S ribosomal RNA (rRNA) [2,3], methods using species-specific oligonucleotides (SSOs) [4] and classification by restriction fragment length polymorphisms (RFLPs) [5] were developed for identifying animal species in forensic settings. Currently, in wildlife forensic science, identification of animal species is commonly performed using cytochrome b (*Cytb*) or cytochrome c oxidase subunit I (*COI*) in mitochondrial DNA [6–8]. These methods are powerful tools when identification of animal species associated with a particular sample is the primary purpose. However, because HID is generally the primary goal in criminal investigations, almost all purified DNA obtained from casework samples is used for HID. Additionally, although species analysis tests using mitochondrial DNA, such *COI* and *Cytb*, require only small quantities of DNA, in cases when human DNA and animal DNA are mixed, it is difficult to perform the sequence analysis because almost all universal primers for species analysis tests amplify both DNAs. Therefore, it might not be possible to perform animal discrimination tests on mixed human/animal casework samples, particularly when only limited quantities are available.

Earlier cross-species validation studies reported for the non-specific peaks generated from animal DNA are observed in electropherograms near human amelogenin peaks [9–12]. The shape and observed position of these non-specific peaks are distinct and can be easily distinguished from other artifacts such as dye blobs and primer dimers. However, little is known about the other characteristics of the non-specific peak generated from animal DNA. We, thus, focused on these non-specific peaks for simultaneous HID and animal discrimination. We initially investigated the characteristic of non-specific peaks and the possibility of identifying an animal family using sequence analysis of non-specific peaks generated from animal DNA using samples that are often relevant in the forensic HID practice.

We then sought to develop a novel PCR assay, termed as "blocking PCR" (bPCR), for analyzing non-specific peaks generated from animal DNA. This assay enables identification of the animal source associated with a sample at the family level by analyzing non-specific peaks generated from animal DNA using PCR products left over after using the commercially available HID kit, despite the casework sample comprising small quantities of DNA present as a mixture of human and animal DNA. We also examined bPCR conditions and evaluated whether this assay was useful in the forensic practice by using remaining PCR products from an HID kit for analyzing various animal/human mixed samples and a mock casework sample.

2. Materials and methods

2.1. Samples

Animal samples used included those of cattle (Bovidae), sheep (Bovidae), goat (Bovidae), horse (Equidae), pig (Suidae), dog (Canidae), fox (Canidae), raccoon (Canidae), cat (Felidae), and chicken

(Phasianidae). DNA of goat, fox, and raccoon was extracted from their respective blood samples. DNA of cattle, sheep, horse, pig, and chicken was extracted from their respective muscle samples. Canine DNA was extracted from saliva and feline DNA was extracted from nail clippings. Control DNA samples used included two human males (9948 and 007) and one human female DNA sample (9947A), sourced from Thermo Fisher Scientific (Waltham, MA, USA). Origins of animal samples were verified using sequence analysis of 16S ribosomal RNA [13]. All procedures performed in the study were in accordance with ethical standards of National Research Institute of Police Science, Japan.

2.2. DNA extraction and quantification

DNA was extracted and purified using the EZ1 DNA Investigator Kit (Qiagen, Venlo, Netherlands) on the EZ1 Advanced XL platform (Qiagen). Animal DNA was quantified by optical absorbance measurement using a NanoVue Plus spectrophotometer (GE Healthcare, Little Chalfont, UK). Human DNA was quantified by human-specific real-time PCR using a human DNA quantification kit that targets 207 bp of the D17Z1 region (Takara Biomedicals, Otsu, Japan) with the SmartCycler II system (Cephid, Sunnyvale, CA, USA).

2.3. PCR amplification, electrophoresis, and data analysis

PCR amplification was performed on the GeneAmp® PCR System 9700 (Thermo Fisher Scientific). PCR products for fragment analysis were separated by capillary electrophoresis on a 3500xL Genetic Analyzer (Thermo Fisher Scientific) using POP-4® Polymer (Thermo Fisher Scientific), under injection conditions of 1.2 kV for 22 s. Alternatively, a 3130*xl* Genetic Analyzer (Thermo Fisher Scientific) was employed, using POP-4® Polymer under injection conditions of 3 kV and 10 s. GeneScan™ 600 LIZ® Size Standard v2.0 (Thermo Fisher

Scientific) was used for sizing. DNA sequencing runs were performed by capillary electrophoresis on the 3130*xl* Genetic Analyzer (Thermo Fisher Scientific) using POP-4® Polymer. PCR products were purified using the QuickStep™ 2 PCR Purification Kit (Edge BioSystems, Gaithersburg, MD, USA).

2.3.1 Multiplex PCR amplification using commercially available HID kits

Multiplex PCR amplification was conducted using the AmpFℓSTR® Identifiler® PCR Amplification Kit (Thermo Fisher Scientific) (Identifiler) or the AmpFℓSTR® Identifiler®Plus PCR Amplification Kit (Thermo Fisher Scientific) (IdentifilerPlus), in accordance with the manufacturer's protocols. All samples were analyzed using GeneMapper® Software *ID* v3.2.1 (Thermo Fisher Scientific) or GeneMapper® *ID-X* Software v.1.4 (Thermo Fisher Scientific).

2.3.2 Monoplex PCR amplification

Monoplex PCR amplification was performed in 25-µL reaction volumes with 10 µL of the AmpFlSTR® Identifiler® Plus Master Mix and 10-pmol of each amelogenin primer included in the PowerPlex 16 System. The cycling conditions for PCRs were set in accordance with the IdentifilerPlus protocol with 35 cycles.

2.3.3 Sequence analysis

Direct sequencing was conducted using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Thermo Fisher Scientific). All sequencing runs were performed for both DNA template strands using the same primers as those used for monoplex PCR amplification. Sequencing run data were analyzed using Sequencing Analysis Software v.5.2 (Thermo Fisher Scientific) and aligned using Sequencher® v.4.7 (Gene Codes, Ann Arbor, USA).

2.4. Non-specific peaks generated from animal DNA when using HID kits

One nanogram of horse and pig DNA; 10 ng of cattle, dog, and fox DNA; 50 ng of sheep and raccoon DNA; or 100 ng of goat, cat, and chicken DNA was amplified using IdentifilerPlus respectively. PCR amplifications were performed in duplicate. Three electrokinetic injections and subsequent electrophoresis were performed for each PCR product.

2.5. Sequence analysis of non-specific peaks generated from animal DNA

The sequences of the amelogenin primers in the AmpFℓSTR® Series (Thermo Fisher Scientific) and the PowerPlex 16 System (Promega, Madison, WI, USA) HID kits have been reported previously (Fig 1) [14–16].

Monoplex PCR amplification for analyzing non-specific peaks generated from animal DNA was performed with animal DNA (horse and pig: 1 ng; cattle, dog, and fox: 10 ng; and sheep and raccoon 50 ng). Agarose electrophoresis was performed for verifying amplification. Direct sequencing was then conducted using purified monoplex PCR products.

2.6. Blocking PCR

2.6.1. Principle

When human and animal DNA coexist in the same casework sample, the amelogenin primer pair amplifies both types of DNA. For direct sequence analysis of only the non-specific peaks generated from animal DNA in such mixed samples, only the animal DNA (rather than the human DNA) needs to be amplified. To overcome this problem, we synthesized an oligonucleotide probe (human amelogenin amplification blocker: HAAB) that specifically binds only to the human DNA (Fig. 1) so as to block PCR amplification of the human

DNA. HAAB approximately matches a part of the human amelogenin sequence, except for a few mismatched, noncomplementary bases at the 3′ terminus. Thus, HAAB binds to the complementary sequence of the human amelogenin gene. However, extension of HAAB itself is prevented because of mismatches of a few bases at the 3′ end. HAAB was designed to overlap certain bases of the amelogenin reverse primer binding regions. Thus, the amelogenin reverse primer is prevented from binding to the human amelogenin site because of the presence of HAAB. Conversely, because there are multiple mismatches to animal DNA involving the binding site, HAAB cannot to bind to it. Animal DNA is amplified by the amelogenin primers even though HAAB is present in the PCR mixture. Figure 2 presents a schematic of the principle behind bPCR with HAAB. Although there are some differences between the sequences of the X and Y chromosomes at the HAAB binding region, HAAB was designed to bind to both X and Y chromosomes of human DNA (Fig. 1).

2.6.2. Validation of bPCR conditions for analyzing non-specific peaks from mixed human/animal samples

When HAAB does not overlap with the human amelogenin reverse primer binding region, it would be expected to be degraded by the inherent DNA polymerase 5′ -3′ exonuclease activity. Hence, we examined the correlation between the blocking effect and number of bases overlapping with the human amelogenin reverse primer binding region (Table 1). A mixture of 1 ng of the 9947A control DNA and 1 ng of pig DNA was amplified by monoplex PCR using the AmpFℓSTR® Series amelogenin primer pair which was labeled with 6-carboxyfluorescein (6-FAM) in the presence of 100 pmol HAAB with various extents of overlap with the human amelogenin reverse primer binding region.

For further investigation of appropriate concentrations of HAAB, a mixture of 1 ng of the 007 control DNA and 1 ng of pig DNA was amplified using the IdentifilerPlus with 0, 10, 25, 50, and 100 pmol HAAB with 14 overlapping nucleotides. Then, direct sequencing was performed for PCR products obtained with bPCR using 50-pmol 14-base-overlapping HAAB and those obtained without HAAB. For subsequent experiments, bPCR was performed with 50-pmol 14-base-overlapping HAAB.

To investigate whether HAAB blocks both the human X and Y chromosome ameliogenin sites, a mixture of 0.5 ng of the 007 male control DNA and 0.5 ng of the 9947A female control DNA was amplified using IdentifilerPlus with HAAB.

To verify human DNA specificity of the designed HAAB, a mixture of 1 ng of human DNA (9947A) and 1 ng of horse or pig DNA; 10 ng of cattle, dog, or fox DNA; and 50 ng of sheep or raccoon DNA was amplified using IdentifilerPlus in the presence of HAAB.

2.7. Forensic application

For simulating various contaminated forensic samples, 1 ng of horse or pig; 10 ng of cattle, dog, or fox; and 50 ng of sheep or raccoon DNAs were mixed with 0.5 ng of human male DNA (007) respectively. Chopsticks that had been used to eat pork by a male volunteer were employed as a mock casework sample. Standard STR profiling on DNA of these simulation samples was performed using IdentifilerPlus. After electrophoresis on a 3500xL Genetic Analyzer, bPCR was performed using 1 μL of purified first-round PCR products with or without HAAB. Direct sequencing was then performed on the purified bPCR products.

2.8. In silico analysis of animal amelogenin gene

A search for animal amelogenin genes was performed using the nucleotide database of the National Center for Biotechnology Information (NCBI). Basic Local Alignment Search Tool (Blast) was used to determine whether a sequence is similar or identical to the human amelogenin gene (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The human and animal amelogenin DNA sequences were aligned using ClustalW multiple sequence alignment algorithm within Molecular Evolutionary Genetics Analysis v.10.0.2.

3. Results

3.1. Non-specific peaks generated from animal DNA when using HID kits

Cattle, sheep, horse, pig, dog, fox, and raccoon DNA generated non-specific peaks near the amelogenin peak (Fig. S1). Goat, cat, and chicken DNA templates did not result in similar non-specific peaks. It was not statistically possible to distinguish between animal families using fragment sizes of these non-specific peaks (Table S1). Furthermore, peak heights were not correlated with the amount of template DNA. Horse and pig generated a stronger non-specific peak even with a small quantity of input DNA.

3.2. Sequence analysis of non-specific peaks generated from animal DNA when using HID kits

The amelogenin DNA sequences obtained for each animal are shown in Table 2. All non-specific peaks obtained from each animal were 59 bases in length, following removal of the primer binding regions. The non-specific peaks generated from dog, fox, and raccoon (all Canidae) DNA displayed identical nucleotide sequences. Although sheep and cattle are from same animal group (both being Bovidae), DNA sequences of their non-specific peaks were different. They also differed from the human amelogenin sequences.

3.3. Validation of bPCR conditions for analyzing non-specific peaks from mixed human/animal samples

A mixture of 1 ng of the 9947A control DNA and 1 ng of pig DNA was amplified using various HAAB bases added to the oligonucleotide probe that overlapped with the amelogenin primer binding site and specifically binding to human DNA. Although a greater number of overlapping bases reduced the efficiency of human amelogenin PCR amplification in a dose-dependent manner, overlap with the complete primer binding site blocked PCR amplification of human and animal DNA. Our data showed that the use of HAAB with 14 bases overlapping with the complete primer binding site was the most efficient scheme for blocking human DNA amplification without affecting animal DNA amplification.

A mixture of 1 ng of the 007 control DNA and 1 ng of pig DNA was amplified using the IdentifilerPlus with various amounts of HAAB with a 14-nucleotide overlap. The results showed that HAAB at 50 pmol was sufficient for blocking human amelogenin amplification. Figure 3 illustrates electropherograms of DNA sequence analysis of a mixed human/pig sample that was subjected to bPCR with 50-pmol HAAB with 14 overlapping bases and one for which bPCR was not performed. The electropherogram obtained without performing bPCR could not be analyzed because of contamination with human DNA (Fig. 3A). Conversely, the pig sequence was readily and accurately analyzed when bPCR was performed (Fig. 3B).

IdentifilerPlus electropherograms generated from mixtures of human male and female samples with HAAB are shown in Figure 4. Peaks corresponding to PCR amplification from the human amelogenin locus were not evident when HAAB was present. On the other hand, all human STR peaks other than that for the

amelogenin locus were detected. It was confirmed that HAAB blocked both human X and Y chromosome amelogenin loci amplifications.

When a mixture of human DNA and DNA from each animal was amplified using IdentifilerPlus with the inclusion of HAAB, only the human-specific amelogenin peak did not amplify (Fig. S2). An example of the specificity of HAAB for human DNA is shown in Figure 5: HAAB specifically blocked human DNA without affecting PCR amplification of animal DNA.

3.4. Forensic application

In mixed forensic samples, approximately 2,000 RFU for the human amelogenin peaks (derived from X and Y chromosomes) were observed in all electropherograms. Intensities of the non-specific peaks generated from each animal significantly differed. Peak heights were 6,405, 234, 6,149, 3,775, 7,003, 7548 and 384 RFU for the cattle, sheep, horse, pig, dog, fox, and raccoon DNA, respectively. According to the results of direct sequence analysis after performing bPCR using HAAB, only animal DNA sequences were obtained from the samples even when the intensity of the animal-specific peak was low.

In the mock casework sample involving chopsticks, a full male DNA profile was derived from the saliva, and the animal-specific peak with an intensity of approximately 7170 RFU and a calculated size of 101.24 bp was observed adjacent to the X and Y chromosome amelogenin peaks. The DNA sequence associated with the non-specific peak was determined after performing bPCR with the remaining IdentifilerPlus PCR product (Fig. S3). The obtained DNA sequence was identical to that of a pig. These results confirmed that in a mock forensic

casework sample, animal families from which DNA was mixed in the saliva originated had been successfully identified.

3.5 In silico analysis of animal amelogenin gene

DNA sequences of some animal species were found in the nucleotide database of NCBI and using the Blast search. The alignment of DNA sequences performed using MEGA-X is shown in Fig. S4. The data generated by in silico analysis suggested that the DNA sequences differed among animal families. For cat (Felidae family), there were some mismatches near the 3′ -terminal end of the forward primer. For chicken, no hits were identified by the Blast search. There were some mismatches at the HAAB binding site for almost all animals except primates.

4. Discussion

We investigated non-specific peaks generated from animal DNA on human STR analysis and developed a novel analysis method for identifying such non-specific peaks in mixtures of human/animal samples. The sequences of the non-specific peaks generated with amelogenin primers from HID kits differed among animal families. Hence, the animal families in this study could be identified by sequence analysis of the non-specific peaks. In the present study, we developed a novel PCR assay for appropriate analysis of non-specific peaks, even when analyzing a mixture of human and animal DNA. The human STR profile and contaminating animal DNA at the family level could be simultaneously identified using the remainder PCR products of the HID kit left over after human STR profiling. Because this method does not require additional

DNA from casework samples, further information (i.e., identifying the family of the contaminating animal) can be obtained from limited quantities of DNA obtained from the casework sample.

Sequences of the non-specific peaks observed near the amelogenin locus generated by the amelogenin primer pair differed among animal families in this study. Data of the in silico analysis showed that primate have DNA sequences similar to that of humans. For Bovidae, differences existed even at the species level. If an unknown sequence is obtained by analyzing a non-specific peak, the animal family can, thus, be identified by performing a search of a DNA database, such as GenBank. Analyzing the non-specific peak can, thus, lead to discrimination of the animal family contaminating a casework sample for HID. The intensity of the non-specific peaks generated from animal DNA significantly differed among each animal group even if the same amount of DNA was amplified. In case of a very small amount of human DNA (e.g., touch DNA), amplification is performed using the quantification value calculated by human-specific real-time PCR; however, a large amount of contaminated animal DNA might also be amplified simultaneously. In this study, mismatches were observed at the binding site of the amelogenin primer of each animal by in silico analysis. This resulted in different amplification efficiencies among animals. It is difficult to estimate the animal family which is the source of contamination with casework samples using intensities of the non-specific peak. Hence, analysis of the sequence associated with the non-specific peak was required for identifying the animal at the family level. Here, the non-specific peak was not generated because of cat or chicken DNA. It was considered that cat DNA could not be amplified because of some mismatches near the 3′ -terminal end of the forward primer. Moreover, for chicken, no hits were identified after a Blast search. The amelogenin gene of birds is

less homologous to that of humans; thus, it was believed that chicken DNA would not be amplified with the amelogenin primer pairs used in this study.

It is expected that improved results of electropherograms will be obtained during direct sequence analysis by performing bPCR when the animal DNA in a sample is more abundant than human DNA. However, our study showed that although the intensity of the human amelogenin peaks (X and Y chromosomes) was ≥7-fold greater than that of the non-specific peak (234 RFU on 3500xL platform), only the animal DNA sequence was obtained by performing bPCR with HAAB. The assay developed by us could accurately analyze very low-intensity non-specific peaks (approximately analytical threshold) even if abundant human DNA was present.

PCR amplification has been considered as an assay for amplifying only target sequences. Our developed assay used the opposite approach which deconvolutes target sequences using PCR amplification. In forensic settings, STR profiling using common primer pairs, including those of commercially available kits, improves reliability. Unexpected non-specific peaks are often observed in electropherograms when STR profiling is performed for casework samples. The principle of our assay is simple and uncomplicated; it is considered to be applicable to popular PCR amplification systems. Hence, with our assay, we may be able to analyze other non-specific peaks using PCR products of the latest HID kits, such as the GlobalFiler® PCR Amplification Kit (Thermo Fisher Scientific) and the PowerPlex® Fusion 6C System (Promega). The bPCR is expected to be applicable to various situations in forensic settings. For example, when STR profiling is performed using HID kits on human remains contaminated with large amounts of horse and pig DNA, analysis is not possible because of non-specific peaks. By performing bPCR with oligonucleotide probes that specifically

bind to only the animal DNA, human STR profiles may be recovered from casework samples contaminated with large amounts of animal DNA.

In conclusion, we investigated the impact of animal DNA on human STR analysis and developed a novel method for identifying the source of non-specific peaks from mixed human/animal samples. For forensic applications, our developed method could facilitate simultaneous human identification and animal family discrimination in mixed human/animal samples. This approach makes it possible to obtain further information from limited and precious casework samples.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

[1] J.M. Butler, Forensic DNA Typing, second ed., 2005.

[2] K. Imaizumi, T. Akutsu, S. Miyasaka, M. Yoshino, Development of species identification tests targeting the 16S ribosomal RNA coding region in mitochondrial DNA, Int. J. Legal Med. 121 (2007) 184–191.

[3] A.O. Karlsson, G. Holmlund, Identification of mammal species using species-specific DNA pyrosequencing, Forensic Sci. Int. 173 (2007) 16–20.

[4] H.D. Marshall, K.A. Johnstone, S.M. Carr, Species-specific oligonucleotides and multiplex PCR for forensic discrimination of two species of scallops, *Placopecten magellanicus* and *Chlamys islandica*, Forensic Sci. Int. 167 (2007) 1–7.

[5] Y.S. El-Sayed, O.I. Mohamed, K.M. Ashry, S.M. Abd El-Rahman, Using species-specific repeat and PCR-RFLP in typing of DNA derived from blood of human and animal species, Forensic Sci. Med. Pathol. 6 (2010) 158–64.

[6] L. Sinclair, M. Merck, R. Lockwood, Forensic investigation of animal cruelty : a guide for veterinary and law enforcement professionals, Humane Society Press, 2006. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950118/ (accessed 26 March 2017).

[7] N. Dawnay, R. Ogden, R. McEwing, G.R. Carvalho, R.S. Thorpe, Validation of the barcoding gene COI for use in forensic genetic species identification, Forensic Sci. Int. 173 (2007) 1–6.

[8] A. Lopez-Oceja, D. Gamarra, S. Borragan, S. Jiménez-Moreno, M.M. de Pancorbo, New cyt b gene universal primer set for forensic analysis, Forensic Sci. Int. Genet. 23 (2016) 159–165.

[9] P.J. Collins, L.K. Hennessy, C.S. Leibelt, R.K. Roby, D.J. Reeder, P.A. Foxall, Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR Identifiler PCR Amplification Kit, J. Forensic Sci. 49 (2004) 1265–1277.

[10] D.Y. Wang, C.W. Chang, R.E. Lagacé, L.M. Calandro, L.K. Hennessy, Developmental validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit: an established multiplex assay with improved performance, J. Forensic Sci. 57 (2012) 453–465.

[11] M.G. Ensenberger, J. Thompson, B. Hill, K. Homick, V. Kearney, K.A. Mayntz-Press, et al., Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex, Forensic Sci. Int. Genet. 4 (2010) 257–264.

[12] S. Inokuchi, T. Kitayama, K. Fujii, H. Nakahara, N. Mizuno, K. Kasai, et al., Forensic Validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit and Comparison of Performance with AmpFℓSTR® Identifiler® PCR Amplification Kit, Japanese J. Forensic Sci. Technol. 18 (2013) 45–56.

[13] S.R. Palumbi, A. Martin, S. Romano, W.O. McMillan, L. Stice, G. Grabowski, The simple fool's guide to PCR version 2, Univ. Hawaii, 96822 (2002) 1–45.

[14] PCR primers and conditions, (n.d.). http://www.cstl.nist.gov/strbase/primer.htm (accessed 23 September 2014).

[15] J.M. Butler, P.M. Vallone, J.M. Devaney, M.A. Marino, AAFS Feb 2001 Talk, Comparison of primer sequences used in commercial STR Kits AAFS Feb 2001 talk analysis of released promega primer sequences overlapping loci between promega, (2001) 1–7.

[16] E.A. Cotton, R.F. Allsop, J.L. Guest, R.R.E. Frazier, P. Koumi, I.P. Callow, et al., Validation of the AMP Fl STR ® SGM Plus E system for use in forensic casework, 112 (2000) 151–161.

Figure Captions

Figure 1. Human X and Y chromosome sequences amplified by the amelogenin primer pair

The binding region of the human amelogenin amplification blocker (HAAB) was designed to overlap the reverse primer. There is a difference of three bases in the HAAB binding region between the X and Y chromosomes. The two bases at the A position are complementary to the Y chromosome and one base at the B position is complementary to the X chromosome so as to efficiently bind both X and Y chromosomes.

Figure 2. Schematic diagram of the principle of blocking PCR using the human-specific oligonucleotide probe (HAAB)

HAAB specifically binds only to human DNA. A; Mismatches at the 3′ -end of HAAB prevent extension. Overlap with some bases of the amelogenin reverse primer binding region prevents binding of the amelogenin reverse primer to human amelogenin. B; Only animal DNA is amplified because the human DNA-binding site is occupied by HAAB.

Figure 3. Comparison between electropherogram with bPCR and that without bPCR using Sequencing Analysis Software ver.5.4

A; Electropherogram showing the sequencing results without performing bPCR. B; Electropherogram showing the results after performing bPCR with HAAB.

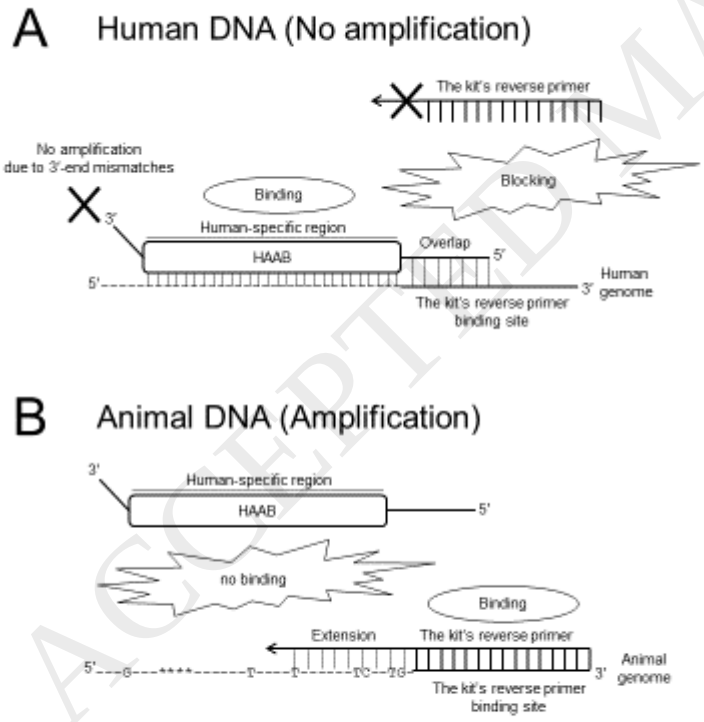
Figure 4. Electropherogram of the IdentifilerPlus profile for PCR products from a mixture of human male and female DNA (007 and 9947A, respectively) amplified with HAAB

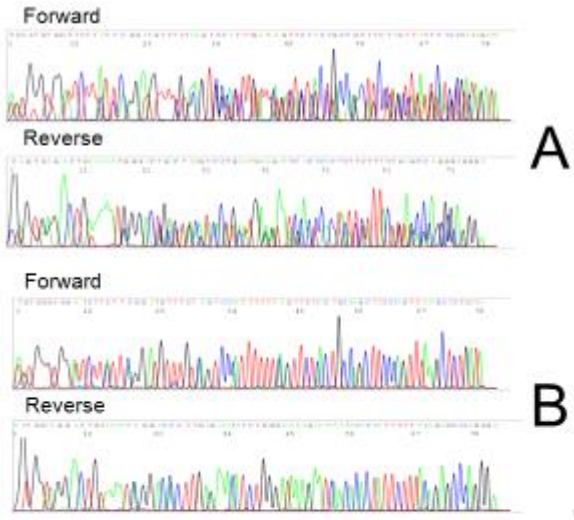
Amelogenin peaks were absent when HAAB was present in the PCR mixture (black arrow).

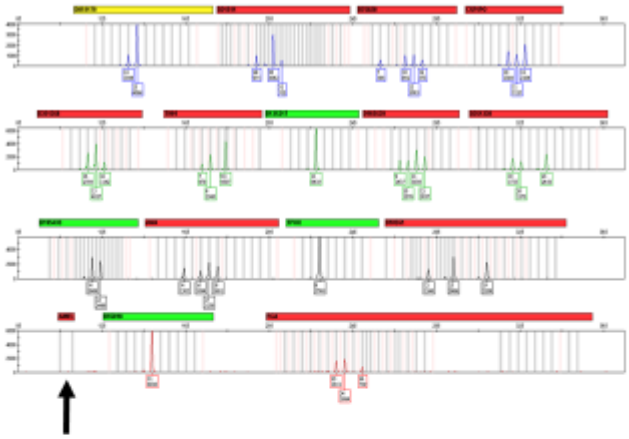
Figure 5. Specificity of the oligonucleotide probe for blocking human amelogenin amplification

A; A mixture of human and pig DNA was amplified using IdentifilerPlus without HAAB. B; A mixture of human and pig DNA was amplified using IdentifilerPlus with HAAB in the PCR mixture. Only the human amelogenin peak was not amplified (black arrow).

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                          10         20         30         40         50
Amelogenin X    CCCTGGGCTC TGTAAAGAAT AGTGTGTTGA TTCTTTATCC CAGAT~~~~~
Amelogenin Y    CCCTGGGCTC TGTAAAGAAT AGTGGGTGGA TTCTTCATCC CAAATAAAGT
                F-Primer PowerPlex 16
                F-Primer AmpFlSTR
                ....|....| ....|....| ....|....| ....|....| ....|....|
                          60        A70         80        B 90        100
Amelogenin X    ~GTTTCTCAA GTGGTCCTGA TTTTACAGTT CCTACCACCA GCTTCCCAGT
Amelogenin Y    GGTTTCTCAA GTGGTCCCAA TTTTACAGTT CCTACCATCA GCTTCCCAGT
                                                                    <--
                          ------------------------------------------
                                          HAAB
                ....|....| ..
                          110
Amelogenin X    TTAAGCTCTG AT
Amelogenin Y    TTAAGCTCTG AT
                Identical R-primer region
```







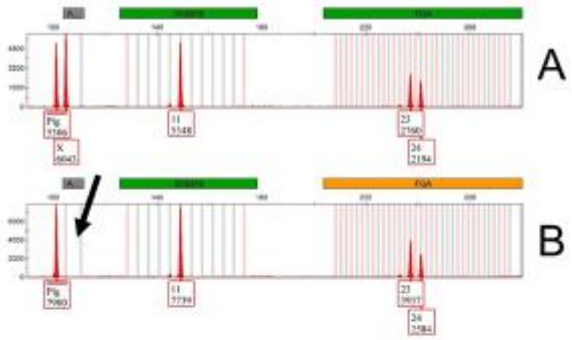


Table 1 Sequences of oligonucleotide probes with various numbers of bases overlapping with the human amelogenin reverse primer site

Number of bases overlapping with the reverse primer binding site	Sequence (5′ -3′)[1,2)]	Melting temperature (°C)
No overlap	GTGGTAGGAACTGTAAAATTGGGACCACT*CTC*	60.1
3 bases	CTGGTGGTAGGAACTGTAAAATTGGGACCACT*CTC*	63.1
7 bases	GAAGCTGGTGGTAGGAACTGTAAAATTGGGACCACT*CTC*	65.6
14 bases	AAACTGGGAAGCTGGTGGTAGGAACTGTAAAATTGGGACCACT*CTC*	68.3
24 bases[3)]	ATCAGAGCTTAAACTGGGAAGCTGGTGGTAGGAACTGTAAAATTGGGACCACT*CTC*	70.8

1) Bases overlapping with the amelogenin reverse primer are underlined.

2) Noncomplementary bases added at the 3'-end to prevent amplification by HAAB are shown in *italics*.

3) Complete overlap with the amelogenin reverse primer binding site.

Table 2 DNA sequences obtained from monoplex PCR amplification using PowerPlex16 System amelogenin primer pairs

Common name (*species*)	**DNA sequences except for primer[1)]**
Human X chromosome (*Homo sapiens*)	TAGTGTGTTGATTCTTTATCCCAGATGTTTCTCAAGTGGTCCTGATTTTACAGTTCCTACCAC
Cattle (*Bos taurus*)	----------------C--T-A------------****----CAC----T------TC--T--
Sheep (*Ovis ammon*)	----------------C--T-G------------****----CAC----T------TC--T--
Horse (*Equus caballus*)	----------------C--T-A--------T---****----CA-----T-------C--T--
Pig (*Sus scrofa*)	--------G-------C--T-AG-------G---****----CAT----T------TC--TG-
Dog (*Canis familiaris*)	-------------------T-A------------****------C----T-------C-TT--
Fox (*Vulpes vulpes*)	-------------------T-A------------****------C----T-------C-TT--
Raccoon (*Nyctereutes procyonoides*)	-------------------T-A------------****------C----T-------C-TT--

1) Hyphens indicate bases that are identical to those in the human X chromosome. Asterisks indicate a deletion compared with the human X chromosome.