Received: 4 January 2024 | Accepted: 25 March 2024

DOI: 10.1111/age.13428

RESEARCH ARTICLE

ANIMAL GENETICS WILEY

A panel of tetranucleotide STR markers as an alternative approach to forensic DNA identification of wolf and dog

Aliaksandra E. Hrebianchuk[1,2] | **Nastassia S. Parfionava**[1] | **Tatsiana V. Zabauskaya**[1] | **Iosif S. Tsybovsky**[3]

[1]Scientific and Practical Center of the State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus

[2]State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus

[3]BelJurZabespjachjenne, Minsk, Belarus

Correspondence
Aliaksandra E. Hrebianchuk, State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus.
Email: iamsanya94@mail.ru

Abstract

Commercial panels of microsatellite (STR) loci are focused on the use of DNA of the domestic dog (*Canis lupus familiaris*) and are often inapplicable for genotyping the DNA of the gray wolf (*Canis lupus lupus*). We propose a CPlex test system, including one hexa- and 12 tetranucleotide autosomal STR loci, as well as two sex loci, that is equally efficient in DNA identification of biological samples of the wolf and the dog. Analysis of molecular variance between samples revealed significant differentiation values (F_{ST}=0.0784, p<0.001), which allows to use the panel to differentiate wolf and dog samples. Population subdivision coefficients (θ-values) were calculated for each of the 13 STR loci of the developed test system. It was shown that the values of the genotype frequency for dogs and wolves, without and with considering the θ-value, differ by three orders of magnitude (for dogs 8.9×10^{-16} and 2.1×10^{-14} and for wolves 1.9×10^{-15} and 4.5×10^{-14}, respectively). The use of population subdivision coefficients will allow to identify the most reliable results of an expert identification study and the power of exclusion provided by the STR loci of the CPlex test system makes it possible to achieve a reliable level of evidence in forensic DNA analysis of both wolves and dogs. The test system has been validated for use in forensic identification of the dog and wolf based on biological traces found at crime scenes, as well as for individual identification and establishing biological relationship of animals of these species.

KEYWORDS
dog, identification, microsatellites, test system, wildlife forensics, wolf

INTRODUCTION

Wildlife has accompanied humanity throughout its existence and comprises an irreplaceable resource. Poaching of wolves (*Canis lupus lupus*) and other species not only harms the economic interests of the country, but also causes environmental problems as it leads to uncontrolled changes in natural biocenoses. A domestic dog (*Canis lupus familiaris*) can be implicated in such crimes as animal abuse, dog attacks on people and livestock, theft, fraud in the trade of purebred animals, as well as illegal hunting. In this regard, reliable differentiation and identification of biological samples of animal origin is of great importance, since offenses against animals are a common occurrence. For this reason, forensic DNA analysis of animals continues to develop rapidly, as evidenced by the appearance of multiple scientific publications about the selection of DNA markers for genotyping various animal species (Hill et al., 2022; Hrebianchuk et al., 2023; Rębała et al., 2022; Wostenberg & Burnham-Curtis, 2023).

The experience in human identification has shown that, for the purposes of forensic identification, the most reliable DNA markers are microsatellites with

tetranucleotide tandem repeats because of the greatest stability of amplification with low-copy or degraded DNA and in the presence of various admixtures (Allendorf, n.d.; Coyle, n.d.; Linacre et al., 2011). Most dog STR markers described in the literature have di- and tetranucleotide repeats (Clarke, 2010; Dayton, 2009; Ogden, 2012; van Asch, 2012; Zenke, 2011). However, the available panels for dog DNA analysis are, in most cases, inapplicable to the study of wolf DNA due to deviations from the Hardy–Weinberg equilibrium and presence of null alleles in DNA markers, which prevents the use of these markers for identification of both wolves and dogs in wildlife forensic research. This fact was once again shown in our previous study (Hrebianchuk & Tsybovsky, 2024).

International Society for Animal Genetics (ISAG) has proposed a panel of microsatellite loci to identify individuals and establish biological relationships in dogs. The panel consists of two multiplex systems, which allow obtaining information about 23 polymorphic dinucleotide loci in the dog genome. However, tetra-, penta-, or hexanucleotide microsatellite loci are easier to evaluate and show fewer occurrences and lower intensity of the stutter effect. Therefore, the use of tetranucleotide STR loci significantly improves the consistency and reproducibility of genotyping. This issue was extensively discussed at the Animal Forensic Genetics session at the 39th International Society for Animal Genetics conference (Cape Town, South Africa, ISAG 2023), and the participants agreed that tetranucleotide STRs are optimal markers in forensic genetics of animal specimens.

The aim of this study was to develop a panel of tetra- and hexanucleotide STR loci that have required polymorphism and sufficient values of random match probability for the forensic DNA analysis of both the gray wolf and the domestic dog.

MATERIALS AND METHODS

Sample collection and DNA extraction

To develop a panel for forensic DNA identification, we genotyped representative samples unrelated individuals of the gray wolf ($n=103$) and the domestic dog ($n=198$). The wolf sample consisted of fragments of muscle and cartilage tissue (53 from males and 50 from females), while the dog sample included blood, hair, and buccal epithelium from purebred, mixed-breed, and mongrel dogs (90 from males, 108 from females). All the samples were legally procured: the wolf samples were obtained during scheduled shootings and from animals killed as a result of road accidents, or provided by zoos; the dog samples were collected with the consent of the owners (in most cases, by veterinary institutions in Belarus).

DNA was extracted by incubating biological material in the lysis buffer consisting of 20 mM Tris–HCl, pH 8.0, 2% SDS, 100 mM NaCl, 20 mM EDTA, proteinase K, and 0.01 mM dithiothreitol. The lysate of biological samples was subjected to the purification procedure on silica (Boom et al., 1990). Total DNA extracts were quantified using a DS-11 spectrophotometer for microvolumes (DeNovix, USA) following the manufacturer's recommendations.

Test system design, DNA amplification, and genotyping

The design of the CPlex test system for forensic DNA identification of wolves and dogs was a single multiplex reaction and included one hexanucleotide (*vWF.X*), 12 tetranucleotide autosomal STR loci, and two sex loci (Table 1), selected based on analysis of research articles. The source species for all used loci was the domestic dog.

To optimize the amplification conditions, we evaluated the specificity and intensity of allele detection depending on the concentration of Mg^{2+} ions, annealing temperature and primer concentration, quantity and quality of DNA, and use of different polymerase stabilizers (Triton X-100, Tween-20, BSA, TMGNa, DMSO) at various concentrations, as well as of DNA polymerases of different types and manufacturers. As a result, the 10-μL PCR volume contained 10 mm Tris–HCl, pH 8.6; 25 mM KCl; 2.0 mM $MgCl_2$; 0.2 mM of each dNTP; 0.2–1.0 μm of each pair of primers; 0.15 U of DNA polymerase activity; 1.5 ng/μL BSA; 0.02% Tween-20, and 1–20 ng of DNA to be analyzed.

During the development of the CPlex test system, the PCR was conducted in a thermocycler C1000 (BioRad, USA) using the following conditions: initial denaturation step at 95°C for 10 min followed by 30 cycles of denaturation at 95°C (for 30 s), annealing at 60°C (for 40 s) and elongation at 72°C (for 60 s), with a final elongation at 72°C for 30 min.

The combination of alleles of each of the samples was detected by electrophoretic separation of PCR products on a 3500 Genetic Analyzer (ThermoFisher Scientific, USA). The estimated size of the alleles (in bp) in the studied loci, detected by electrophoresis, was determined using the Orange 500-bp internal size standards (NimaGen®, The Netherlands) and GeneScan-600 LIZ™ SizeStandard v2.0 (ThermoFisher Scientific, USA). Genotyping was evaluated with GeneMapper ID-X v1.6 software package (ThermoFisher Scientific, USA).

Alleles sequence determination

To identify possible iso-alleles and microvariants of the sequence, as well as to perform tandem determination of alleles in the test system CPlex, the primary structure of alleles was determined by Sanger dideoxy sequencing. Nucleotide sequences of alleles of each

TABLE 1 Characteristics of STR loci.

Loci	Chromosome	Tandem repeat	Dye	GenBank accession number	Reference
FH2096	11	(AATG)n	6-FAM	OQ216492 OQ216493	Caniglia et al. (2010)
vWF.X	27	(AGGAAT)n	6-FAM	OQ216494 OQ216495	DeNise et al. (2004)
FH2079	24	(TGGA)n	6-FAM	OQ216498 OQ216499	Radko and Podbielska (2021)
FH2361	33	(CTTT)n	6-FAM	OQ216518 OQ216519	van Asch (2012)
		$(TCTT)_{11\text{-}18}TC$		OQ216520 OQ216521	
PEZ17	7	(TTTC)n	R6G	OQ216504 OQ216505	van Asch (2012)
FH2004	11	(TTCT)n	R6G	OQ216508 OQ216509	Caniglia et al. (2010)
FH2054	12	(ATCT)n	TMR	OQ216510 OQ216511	Radko and Podbielska (2021)
FH2010	24	(GAAT)n	TMR	OQ216502 OQ216503	Eichmann et al. (2004)
FH2016	1	(CTTT)n	TMR	OQ216525 OQ216526	Fan et al. (2016)
		$(CTTT)_{19\text{-}21}CT$		OQ216527 OQ216528	
FH2001	23	(ATCT)n	ROX	OQ216514 OQ216515	Verardi et al. (2006)
		$(ATCT)_{10\text{-}14}CAACTC$		OQ216516 OQ216517	
FH2328	8	(AAAG)n	ROX	OQ216522 OQ216523	van Asch (2012)
PEZ16	27	(GAAA)n	ROX	OQ216506 OQ216507	Magory Cohen et al. (2013)
VGL3438	34	(AAAG)n	TAMRA	OQ216512 OQ216513	Magory Cohen et al. (2013)
DBX6	X	–	6-FAM	OQ216490	Seddon (2005)
DBY7	Y	–	R6G	OQ216491	Seddon (2005)

STR locus and sex loci of the wolf and the dog were determined in the forward and reverse directions. Sequencing was performed on a 3500 Genetic Analyzer. Cycle sequencing was conducted with the BigDye® Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific, USA) according to the manufacturers' protocols. The results were analyzed using the Sequencing Analysis Software v5.4 (ThermoFisher Scientific, USA). A comparative analysis of the allele sequences of the studied loci in DNA of the wolf and the dog was carried out using the resources of the GenBank database (Benson et al., 2005) and BioEdit v7.0.5.3 (Hall et al., 2011). The sequences of the smallest and largest molecular size of alleles and microvariants were deposited to the GenBank database in our previous study (Hrebianchuk & Tsybovsky, 2024).

Statistical analysis of results

The main parameters of genetic diversity, including the level of polymorphism, allele frequencies, values of observed and expected heterozygosity, compliance with the Hardy–Weinberg equilibrium and linkage disequilibrium were assessed using CERVUS v 3.0.7 (Kalinowski et al., 2007) Possible errors in interpretation of genetic profiles caused by absence of alleles, allele dropouts or other possible PCR artifacts were detected using MICRO-CHECKER v.2.2.1 (van Oosterhout et al., 2004) and CERVUS v.3.0.7.

Since incorrect species identification distorts calculations based on the analysis of genetic diversity indexes (Galinskayaa et al., 2019), a cluster analysis of genotyping data for wolves and dogs was first carried out for

the studied loci. The population structure was determined using the Monte Carlo algorithm according to the Markov chain method using STRUCTURE v.2.3.4 software and Admixture model (Pritchard et al., 2000) with further determination of the true number of clusters by the method of Evanno et al. (2005). The burn-in period included 500 000 iterations, followed by the construction of the Markov chain for 1 000 000 iterations for the expected number of groups in the sample, *K*, from 1 to 5, with six repeats for each value of *K*. Values of K were selected taking into account a previous study in which results of 32-locus genotypes of dogs with *K* from 1 to 10 formed only two separate clusters.

Analysis of molecular variance and estimation of inbreeding coefficients were performed using ARLEQUIN v.3.5.1.3 software (Excoffier et al., 2005). Analysis of assignment of the individual to the sample pool (Assignment Test) based on selected loci was performed using GenAlEx v.6.5 (Peakall & Smouse, 2006, 2012).

The forensic parameters of the selected microsatellite loci (*PIC*, probabilities of identity for unrelated individuals and for siblings, and genotype frequency) were calculated using GenAlEx v.6.5 and CERVUS v. 3.0.7. The genotypes of the studied samples of wolves and dogs were used as a reference set for calculating the match probability and likelihood ratios, which are used in the forensic genetics of objects of animal origin.

Validation

The conditions for using the test system were tested and optimized to achieve a high level of specificity of amplification, optimal intensity and uniformity of detection of marker fluorophores. We performed testing and optimization of primer annealing temperature, primer concentrations, primers ratio, as well as the amount of DNA in the sample.

Validation was carried out in accordance with the protocol of the Scientific Working Group on DNA Analysis Methods. The validation process and results are detailed in the Supplementary Material S1.

RESULTS AND DISCUSSION

Alleles sequence determination

Sequencing confirmed the presence of simple tandem repeats for nine tetranucleotide (*FH2096*, *FH2079*, *PEZ17*, *FH2004*, *FH2054*, *FH2010*, *FH2328*, *PEZ16*, and *VGL3438*) and one hexanucleotide (*vWF.X*) autosomal loci, with identical allele sequences for the wolf and the dog and identical allele sequences for the sex loci. Alleles of the *FH2016* and *FH2361* loci included incomplete tandem repeats. Specifically, in the *FH2016* locus, incomplete tandems were identified in samples of both wolves and dogs. In the *FH2361* locus, microvariants were detected only in the sample of dogs.

Sequencing the alleles of the *FH2001* locus produced an unexpected result (Figure 1). A 6-bp insertion was found in the non-tandem region of the locus in the samples of wolves and dogs. This insertion was detected only in long alleles (with 10 or more tandem repeats), and there were no alleles of 11 or more repeats that did not contain it. For this reason, locus *FH2001* was not excluded from the final forensic panel, and special names were assigned to alleles with the insertion (10in–14in alleles).

To unify the panel to the instrumentation, we introduced a tandem allele designation based on nucleotide sequences of all identified alleles. Alleles with minimum and maximum molecular sizes, as well as the identified microvariants, were deposited in the GenBank database with corresponding accession numbers (Table 1).

Statistical analysis of the results of using the test system

A total of 166 alleles were identified in the domestic dog and gray wolf samples using the proposed test system, of which 150 and 115 alleles were detected in the dog and wolf sample, respectively (Figure 2a,b).

The analysis by MICRO-CHECKER v.2.2.1 and CERVUS v.3.0.7 revealed absence of null alleles or genotyping errors in all studied loci. Analysis of linkage

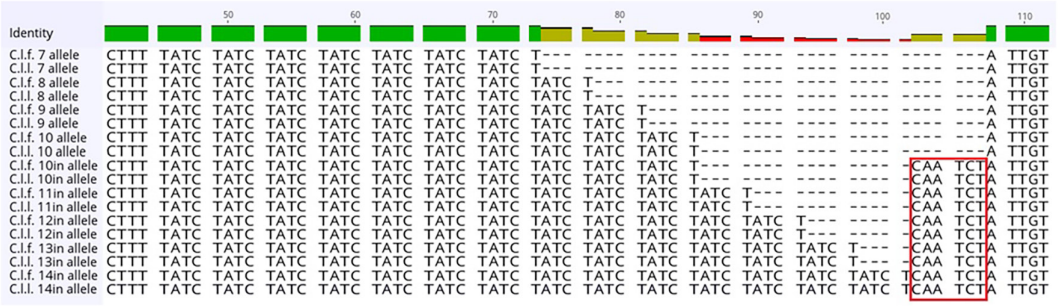

FIGURE 1 Part of the sequence of the *FH2001* locus in wolf and dog DNA (C.l.f., domestic dog; C.l.l., gray wolf; red rectangle, 6-bp insertion in the non-tandem region of the locus).

disequilibrium of the loci using ARLEQUIN v.3.5.1.3 showed independent distribution of alleles in both samples of the test system loci.

The highest observed and expected heterozygosity rates were obtained for the wolf sample, at 0.861 and 0.821 (for *FH2361* locus), respectively. In dog samples, lower values of expected heterozygosity compared to observed heterozygosity are a potent indicator of the presence of inbreeding resulting in synthetic selection and genetic drift, which can irreversibly remove alleles from the population, leading to significantly reduced diversity (Galinskayaa et al., 2019). Although mutations counteract the genetic drift, it is difficult to achieve a balance of genetic processes in dog breeding due to exclusion of individuals with identified mutations in a particular trait. The highest values of heterozygosity and the highest effective number of alleles were found in the wolf sample, which indicates natural development and presence of mutation-drift balance in the natural population.

Eleven loci in the wolf and dog samples conformed to the Hardy–Weinberg equilibrium ($p>0.05$): *FH2001*, *FH2004*, *FH2010*, *FH2016*, *FH2096*, *FH2328*, *FH2361*, *PEZ16*, *PEZ17*, *VGL3438*, and *vWF.X*. Two loci, *FH2054* and *FH2079*, deviated from the equilibrium ($p<0.005$) in all studied samples; however, this deviation lost statistical significance when the Bonferroni correction was used ($p>0.0038$).

According to Botstein et al. (Botstein et al., 1980), *PIC* for a highly informative locus should exceed 0.5, for a sufficiently informative locus it can be less than 0.5 but must exceed 0.25, whereas for a low-informative locus it is <0.25. In the proposed test system, the *FH2096* locus had the smallest *PIC* values in the samples of both the dog and the wolf (0.472 and 0.552, respectively, Table 2).

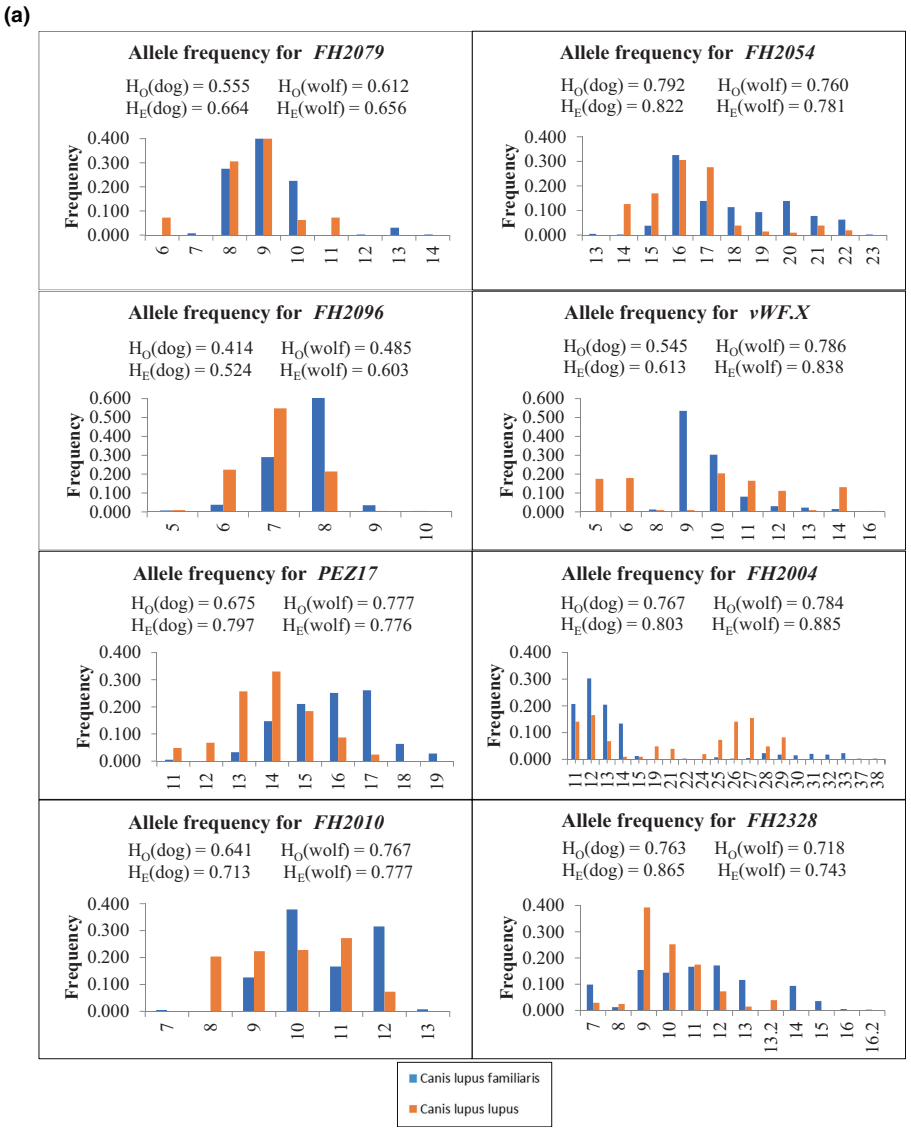

FIGURE 2 (a) Allele frequencies and values of observed and expected heterozygosity of the CPlex test system loci in the wolf and dog samples. (b) Allele frequencies and values of observed and expected heterozygosity of the CPlex test system loci in the wolf and dog samples.

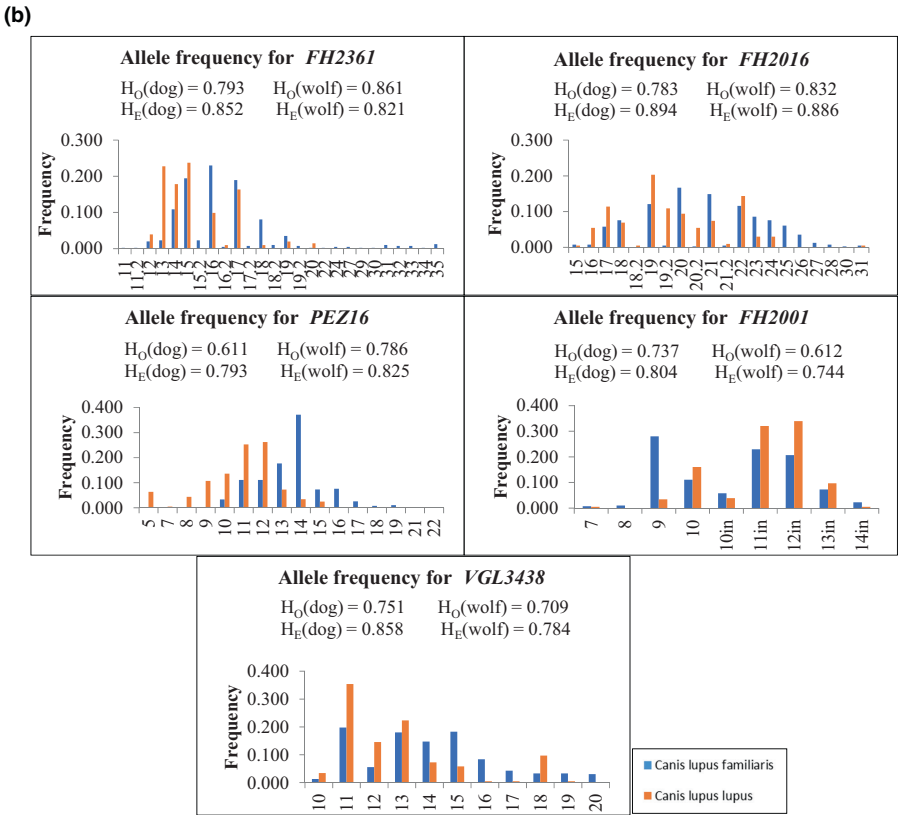

FIGURE 2 (Continued)

The next smallest values were found for the *vWF.X* locus in the sample of the dog (0.541) and for the *FH2079* locus in the wolf sample (0.579). Such *PIC* values are due to the strong expression of the major alleles and (or) a small number of identified alleles at the locus. The highest *PIC* values were found for the *FH2016* locus (0.882 in the wolf sample and 0.885 in the dog sample). The mean *PIC* values were 0.720 for the dog sample and 0.742 for the wolf sample, indicating that the loci selection is suitable for the interpretation of results in forensic genetic testing.

A posteriori analysis of the STRUCTURE results for the combined pool of the wolf and dog genotypes revealed the maximum value of the test statistics ΔK at $K=2$, which indicates the presence of two genetic clusters in the analyzed sample of animals: the gray wolf (green cluster) and the domestic dog (red cluster; Figure 3a, Table 3). Analysis of the population structure of wolves and dogs showed a strong genetic differentiation between them with the average values of the cluster membership coefficient *Q* of 0.974 and 0.981, respectively. The data obtained on differentiation of the wolf and the dog are in good agreement with the data of Korablev et al. (2021). It should be noted that Lorenzini et al. (2022) used 22 markers (of which 17 dinucleotide), and the results of cluster analysis of wolves and dogs showed the cluster membership coefficient *Q* of 0.990 and 0.995, respectively. Taking into account the fact that 22 loci were involved in this study, and dinucleotide loci are considered more polymorphic than tetranucleotide loci, the CPlex test system provides sufficient to reliable differentiation values of *Q*.

The results of the cluster analysis are consistent with the results of analysis of molecular variance (AMOVA; Table 4). AMOVA was performed to assess the level of differentiation between a wolf and a dog using selected microsatellite loci. This analysis revealed a significant (Hedrick, 2000) genetic differentiation for the set of selected autosomal STR loci. The AMOVA results showed that the percentage of variation between the wolf and dog samples was 7.84% ($F_{ST}=0.0784$; $p<0.001$), compared to 92.16% within the samples. Variance components between the samples were significant for all the studied loci, which indicates the presence of differentiation between the wolf and dog samples. The *vWF.X* and *FH2096* loci accounted for 21.11% and 19.33% of inter-sample genetic variability, respectively, while the *FH2079* and *FH2016* loci showed the lowest inter-population variability (2.45% and 2.20%, respectively). The commercial dog panels have not been tested on samples of wolves or have not been examined simultaneously on two subspecies, which does not allow for comparative analysis of AMOVA results.

The significance of the differentiation of the wolf and the dog using the selected loci can be illustrated by the analysis of assignment of a definite sample (Assignment test). The analysis is based on the calculation of the

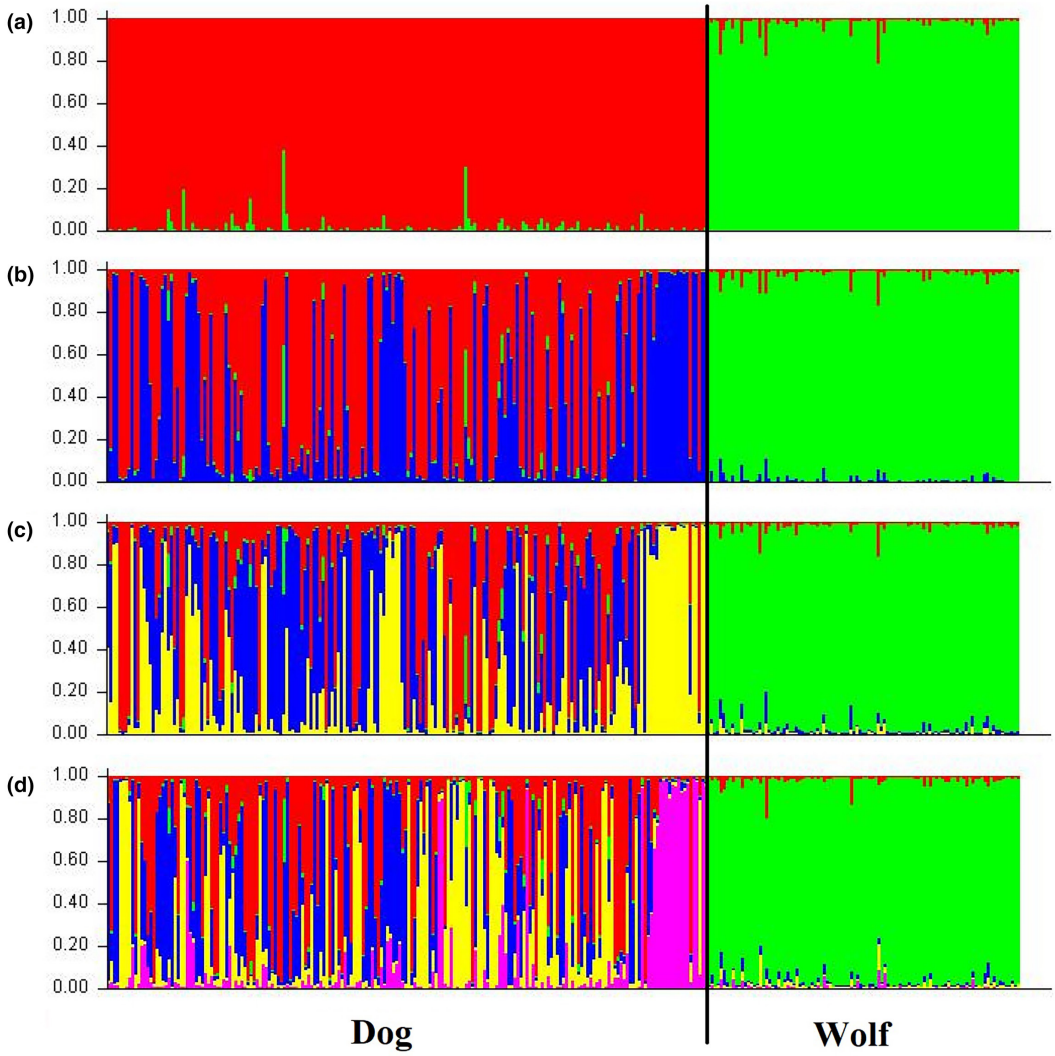

FIGURE 3 (a–d) Results of cluster analysis of wolf and dog samples, performed in STRUCTURE software, for the most probable value of the number of genetic clusters $K=2$–5, sorted by samples.

probability value of the presence of the genotype of a certain individual in the sample from which it was selected, and its comparison with the probability value of the same genotype in another sample. Based on these calculations, an individual belongs to the sample for which it has the highest probability (Figure 4).

The calculation of true genetic affiliation to the sample showed a high consolidation of wolves and dogs, with 100% of all studied animals genetically assigned to their own cluster. At the same time, we observed a large difference in the moduli of natural logarithms of expected frequencies of the genotypes when they belonged to their own vs. an alternative sample; for wolves, the difference was on average 7.413 and, for the dog, it was 7.337. The high values of the difference between the moduli of the logarithms of the expected genotype frequencies confirm successful differentiation of the gray wolf and the domestic dog using the proposed loci. A common method of species identification in wildlife forensics is mtDNA sequencing. The main difficulty in the genetic differentiation of the wolf and the dog is that DNA markers unique to both the wolf and the dog have not been found. A comparison of the dog and wolf genomes showed a similarity of 99%, which once again indicates their group origin (Freedman & Wayne, 2017). The results of Assignment test of the CPlex panel show that, despite the close genetic relationship, the gene pools of the wolf and the dog remain significantly different, and based on the analysis of the frequency characteristics of the alleles, it is possible to conclude that the animal belongs to the gray wolf or the domestic dog.

Since the test system was developed for genotyping wolves and dogs, whose samples are significantly differentiated, we calculated the coefficient of subdivision (θ-value). The θ-value is necessary for calculating the match probability of genotypes in populations with identified subdivision (Morf et al., 2021). The θ-value was calculated using the formula recommended by NRCII (Buckleton et al., 2016; *National Research Council*, 1996). The mean θ-value was 0.088 for the developed test

system. It is recommended that θ-values are used for all loci when calculating the match probability in forensic expertise for samples of unknown origin (dog vs. wolf).

To assess the applicability of the Cplex test system in forensic identification, we calculated the match probability of genotypes of two individuals (for unrelated individuals and for siblings), as well as the likelihood ratio for the total set of genotypes (Taberlet & Luikart, 1999; Waits et al., 2001), since both these calculation approaches are used in forensic expertise of objects of animal origin (Morf et al., 2021). Calculations were performed using the frequencies of genotypes in the samples of wolves, dogs, and the combined sample used for developing the forensic panel, with or without taking into account the θ-value (Buckleton et al., 2016). The results are shown in Table 5.

The match probability without using the θ-value in the dog sample varied from 1.6×10^{-25} to 6.0×10^{-14} with a mean value of 8.9×10^{-16}, whereas for the wolf sample, the range of values was from 2.3×10^{-21} to 9.3×10^{-14} (mean: 1.9×10^{-15}). The mean LR values for sibs of wolves and dogs were 9.2×10^{6} and 7.9×10^{6}, respectively. The obtained data are significantly higher than DeNise et al. (2004) values, obtained using tetranucleotide loci recommended by American Kennel Club (match probability [MP] value for dogs did not exceed 3.2×10^{-8}). Similar values to CPlex were obtained from Berger et al. (2019) using the CaDNAP panel (Dayton, 2009) recommended by The Canine DNA Profiling Group (https://gerichtsmedizin.at/cadnap.html) and Kunita et al. (n.d.), who used Canine Genotypes™ Panel 2.1 Kit. Both panels include mainly tetranucleotide STRs and were used in three breed groups of dogs. The MP values for dog samples depending on breed using the CaDNAP panel ranged from 4.55×10^{-14} to 1.05×10^{-13}, and in the case of the commercial Canine Genotypes™ Panel 2.1 kit – from 2.11×10^{-18} to 3.26×10^{-16}.

Higher values were obtained for the dinucleotide panel recommended by ISAG. Thus, the MP value for mixed-breed dog samples in the study by Kanthaswamy et al. (2018) was 1.25×10^{-21}, and, in the work of Polish colleagues (Radko et al., 2022), it ranged from 1.25×10^{-20} (for a standard panel) to 3.5×10^{-23} (for an extended panel). High MP values are easily explained by the high allelic

TABLE 2 Values of the *PIC* and the coefficient of subdivision (θ-value) using CPlex for samples of the wolf and dog.

	PIC		
Locus	**Domestic dog**	**Gray wolf**	**θ-Value**
FH2096	0.472	0.552	0.131
vWF.X	0.541	0.813	0.192
FH2079	0.620	0.579	0.145
FH2010	0.657	0.747	0.070
PEZ17	0.766	0.746	0.080
PEZ16	0.769	0.827	0.170
FH2001	0.774	0.701	0.031
FH2004	0.778	0.876	0.090
FH2054	0.801	0.759	0.081
FH2361	0.830	0.788	0.035
VGL3438	0.845	0.755	0.020
FH2328	0.848	0.713	0.065
FH2016	0.885	0.882	0.039
Average	0.720	0.742	0.088

Abbreviation: *PIC*, polymorphism information content.

TABLE 3 Evanno's resulting table for the total sample of wolves ($n=103$) and dogs ($n=198$).

K	Reps	Mean LnP(*K*)	Stdev LnP(*K*)	Ln′(*K*)	|ln″(*K*)|	Δ*K*
1	6	−14956.067	0.266	–	–	–
2	**6**	**−13876.667**	**0.186**	**1079.400**	**864.767**	**4644.542**
3	6	−13662.033	0.585	214.633	107.217	183.158
4	6	−13554.616	1.693	107.417	9.700	5.730
5	6	−13437.500	3.412	117.117	–	–

Note: The bold values are values for the most probable number of clusters.

TABLE 4 Analysis of molecular variance of samples of the gray wolf and the domestic dog using CPlex.

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage variation, %
Between populations	1	121.471	0.42744	7.84
Within populations	600	3012.745	5.02326	92.16
Total	601	3134.216	5.45070	
Fixation indices				
F_{IS}	0.120 37			
F_{ST}	0.078 42			
F_{IT}	0.189 35			

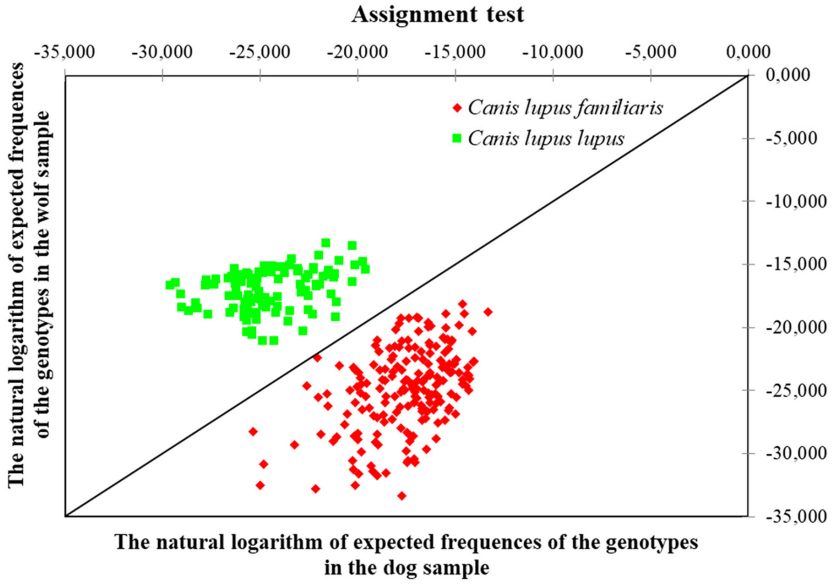

FIGURE 4 Assignment graph (above the diagonal line are individuals assigned to the sample of the gray wolf; under the line, individuals assigned to the sample of the domestic dog).

		Min	Max	Mean	SD
Dog sample without θ-value	MP	1.6×10^{-25}	6.0×10^{-14}	8.9×10^{-16}	4.6×10^{-15}
	LR	1.7×10^{13}	6.1×10^{24}	3.8×10^{23}	4.4×10^{23}
Dog sample with θ-value	MP	3.8×10^{-24}	1.4×10^{-12}	2.1×10^{-14}	1.1×10^{-13}
	LR	7.1×10^{11}	2.6×10^{23}	1.6×10^{21}	1.9×10^{22}
Wolf sample without θ-value	MP	2.3×10^{-21}	9.3×10^{-14}	1.9×10^{-15}	1.1×10^{-14}
	LR	1.1×10^{13}	4.3×10^{20}	8.9×10^{18}	4.6×10^{19}
Wolf sample with θ-value	MP	5.4×10^{-20}	2.2×10^{-12}	4.5×10^{-14}	2.6×10^{-13}
	LR	4.6×10^{11}	1.9×10^{19}	3.8×10^{17}	2.0×10^{18}
Total sample without θ-value	MP	9.7×10^{-26}	1.7×10^{-14}	1.1×10^{-16}	1.0×10^{-15}
	LR	5.8×10^{13}	1.0×10^{25}	4.7×10^{22}	6.1×10^{23}
Total sample with θ-value	MP	2.3×10^{-24}	4.0×10^{-13}	2.6×10^{-15}	2.5×10^{-14}
	LR	2.5×10^{12}	4.4×10^{23}	2.0×10^{21}	2.6×10^{22}
Dog sibs	MP	4.7×10^{-7}	4.8×10^{-7}	1.5×10^{-7}	6.6×10^{-8}
	LR	2.1×10^{6}	2.2×10^{7}	7.9×10^{6}	$3.6\times\times10^{6}$
Wolf sibs	MP	4.2×10^{-8}	1.8×10^{-6}	1.5×10^{-7}	2.3×10^{-7}
	LR	5.6×10^{5}	2.0×10^{7}	9.2×10^{6}	3.7×10^{6}

TABLE 5 Match probabilities and likelihood ratios for the total set of loci using the developed CPlex test system.

Abbreviations: LR, likelihood ratio for the total set of loci; max, maximum value; mean, mean value; min, minimum value; MP, match probability; SD, standard deviation.

diversity of dinucleotide markers compared to tetranucleotide repeats. It should also be noted that all currently developed panels for forensic DNA analysis have been studied exclusively on dog samples. However, as our previous study showed, the loci recommended by ISAG are not applicable for their use, at least in the Belarusian wolf population (Hrebianchuk & Tsybovsky, 2024).

The use of a dinucleotide panel with high MP values increases the information content of studies in parentage testing. However, in forensic identification, the determining factor is reliability, since the researcher very often has DNA of poor quality and in small quantity, which determines the undoubted priority of tetranucleotide DNA markers with their consistently clear interpretation of genetic profiles, confirmed by all participants of the ISAG 2023 conference. In addition, it should be noted that considering that the wolf population in the Republic of Belarus as of 2022 was estimated at 1630–1890 individuals, while the number of domestic and stray dogs reached 100 000 due to some recent increases, the power of exclusion provided by the STR loci of the CPlex test system makes it possible to achieve a reliable level of evidence in forensic DNA analysis of both wolves and dogs.

Based on Table 5, the values of match probability and likelihood ratio for dogs, wolves, as well as for the total sample, differ by one to two orders of magnitude when calculated without and with using the θ-value. Therefore, committing the θ-value in an identification study can

lead to overestimation or underestimation of the frequency of genotype occurrence, which will result in an overestimation of the reliability of the forensic expertise.

CONCLUSION

We propose 13 autosomal STR-loci with tetra- and hexanucleotide tandems and two sex loci, which have sufficient polymorphism both in the study of DNA of gray wolves and in the study of DNA of domestic dogs, for use in forensic science. This study assessed polymorphism and forensic parameters of the loci included in the test system and confirmed the tandem structure of the alleles of the loci, and studied the genetic structure of the Belarusian populations. Results of the cluster analysis and AMOVA between samples revealed significant differentiation values, which allows to use the panel to differentiate wolf and dog samples. According to the results of genetic and statistical analysis of a dataset of wolf and dog genotypes, all the loci are inherited independently and conform to Hardy–Weinberg equilibrium. The coefficients of subdivision of the population for each STR locus of the test system were calculated, and the effectiveness of their use was proven. The developed CPlex test system has high repeatability and 100% precision, high analytical specificity and low limit of detection of genetic material. The developed CPlex test system can be used in forensic investigation of cases of fraud in the trade of purebred dogs, animal attacks, illegal hunting, and other cases involving dogs or wolves where biological samples are available. The test system exhibits cross-activity on DNA matrices of other canines, and therefore the identification study must be preceded by an expert identification of the species of animal from which the test sample originated.

Based on these development efforts, we have created a practical guideline for DNA identification of biological samples of the gray wolf and the domestic dog that is intended for practicing forensic experts of the State Forensic Examination Committee of the Republic of Belarus. This methodological document has been included in the Registry of Forensic Methods and other methodological materials of the State Forensic Examination Committee of the Republic of Belarus, thereby incorporating the developed test system into the national legal system.

AUTHOR CONTRIBUTIONS

Aliaksandra E. Hrebianchuk: Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft; writing – review and editing. **Nastassia S. Parfionava:** Investigation; validation; writing – review and editing. **Tatsiana V. Zabauskaya:** Investigation; validation; writing – review and editing. **Iosif S. Tsybovsky:** Funding acquisition; project administration; resources; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

The authors express their gratitude personally to Dr Vadim E. Sidorovich and Irena A. Solovej and to the personnel of Republican State Public Association "Belarusian Society of Hunters and Fishermen", veterinary center "Drug" and the State Institution "Minsk City Veterinary Station" for their help in procuring a collection of biological samples. The authors also thank the anonymous reviewer who made this article much better.

FUNDING INFORMATION

This research was carried out at the Scientific and Practical Center of the State Forensic Examination Committee of the Republic of Belarus and funded by the State Committee of the Science and Technology of the Republic of Belarus (Project No. 20190195).

All applicable international, national and/or institutional guidelines for the care and use of animals have been followed.

CONFLICT OF INTEREST STATEMENT

The authors have declared that no competing interest exists.

DATA AVAILABILITY STATEMENT

Detailed information about the primer sequences used and the alleles of the studied loci, as well as accession numbers of allele sequences in GenBank, are presented in Table 1 in this article.

Validation was carried out in accordance with the protocol of the Scientific Working Group on DNA Analysis Methods. The validation process and results are detailed in the Supplementary Material S1.

ORCID

Aliaksandra E. Hrebianchuk https://orcid.org/0000-0002-1224-3275
Nastassia S. Parfionava https://orcid.org/0000-0003-3573-3039
Tatsiana V. Zabauskaya https://orcid.org/0000-0003-0716-4035
Iosif S. Tsybovsky https://orcid.org/0000-0002-8611-8215

REFERENCES

Allendorf, F.W. (2007) *Conservation and the genetics of populations*. UK: Blackwell Publishing, p. 642.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J. et al. (2005) GenBank. *Nucleic Acids Res*, 34, D16–D20. Available from: https://doi.org/10.1093/nar/gkj157

Berger, B., Heinrich, J., Niederstätter, H. et al. (2019) Forensic characterization and statistical considerations of the CaDNAP 13-STR panel in 1,184 donestic dogs from Germany, Austria, and Switzerland. *Forensic Science International: Genetics*, 42, 90–98. Available from: https://doi.org/10.1016/j.fsigen.2019.06.017

Boom, R.C.J.A., Sol, C.J., Salimans, M.M. et al. (1990) Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology*, 28(3), 495–503. Available from: https://doi.org/10.1128/jcm.28.3.495-503.1990

Botstein, D., White, R.L., Skolnick, M. & Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32(3), 314.

Buckleton, J.S., Bright, J.-A. & Taylor, D. (2016) *Forensic DNA evidence interpretation*, 2nd edition. Boca Raton, FL: CRC Press, p. 486.

Caniglia, R., Fabbri, E., Greco, C., Galaverni, M. & Randi, E. (2010) Forensic DNA against wildlife poaching: identification of a serial wolf killing in Italy. *Forensic Science International: Genetics*, 4(5), 334–338. Available from: https://doi.org/10.1016/j.fsigen.2009.10.012

Clarke, M. (2010) Dog attack: the application of canine DNA profiling in forensic casework. *Forensic Science, Medicine, and Pathology*, 6(3), 151–157. Available from: https://doi.org/10.1007/s12024-009-9114-8

Coyle, H.M. (Ed.). (2008) *Nonhuman DNA typing: theory and casework applications*. USA: CRC Press, p. 254.

DeNise, S., Johnston, E., Halverson, J. et al. (2004) Power of exclusion for parentage verification and probability of match for identity in American kennel club breeds using 17 canine microsatellite markers. *Animal Genetics*, 35(1), 14–17. Available from: https://doi.org/10.1046/j.1365-2052.2003.01074.x

Dayton, M. (2009) Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological materials. *Croatian Medical Journal*, 50(3), 268–285. Available from: https://doi.org/10.3325/cmj.2009.50.268

Eichmann, C., Berger, B., Reinhold, M. et al. (2004) Canine-specific STR typing of saliva traces on dog bite wounds. *International Journal of Legal Medicine*, 118(6), 337–342. Available from: https://doi.org/10.1007/s00414-004-0479-7

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611–2620. Available from: https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Laval, G. & Schneider, S. (2005) An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1(4), 47–50. Available from: 19325852

Fan, Z., Silva, P., Gronau, I. et al. (2016) Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Research*, 26(2), 163–173. Available from: https://doi.org/10.1101/gr.197517.115

Freedman, A.H. & Wayne, R.K. (2017) Deciphering the origin of dogs: from fossils to genomes. *Annual Review of Animal Biosciences*, 5, 281–307. Available from: https://doi.org/10.1146/annurev-animal-535022114-110937

Galinskayaa, T.V., Schepetov, D.M. & Lysenkov, S.N. (2019) Prejudices against microsatellite studies and how to resist them. *Russian Journal of Genetics*, 55(6), 657–671. Available from: https://doi.org/10.1134/S0016675819060043

Hall, T., Biosciences, I. & Carlsbad, C. (2011) BioEdit: an important software for molecular biology. *GERF Bull Biosci*, 2(1), 60–61.

Hedrick, P.W. (2000) *Genetics of populations*. Boston: Jones and Bartlett, p. 675.

Hill, E. et al. (2022) Development of an STR panel for a non-native population of an endangered species. *Molecular Biology Reports*, 49(1), 839–845. Available from: https://doi.org/10.1007/s11033-021-06905-w

Hrebianchuk, A.E. & Tsybovsky, I.S. (2024) Universal panel of STR-loci for the study of polymorphism of the species *Canis lupus* and forensic identification of dog and wolf. *Vavilov Journal of Genetics and Breeding*, 28(1), 98–107. Available from: https://doi.org/10.18699/vjgb-24-12

Hrebianchuk, A.E., Lukashkova, O.N., Kotava, S.A. & Tsybovsky, I.S. (2023) A robust test system for DNA identification of raccoon dogs. *Russian Journal of Genetics*, 59(5), 503–517. Available from: https://doi.org/10.1134/S102279542305006X

Kalinowski, S.T., Taper, M.L. & Marshall, T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16(5), 1099–1106. Available from: https://doi.org/10.1111/j.1365-294x.2007.03089.x

Kanthaswamy, S., Oldt, R.F., Montes, M. & Falak, A. (2018) Comparing two commercial domestic dog (*Canis familiaris*) STR genotyping kits for forensic identity calculations in a mixed-breed dog population sample. *Animal Genetics*, 50(1), 105–111. Available from: https://doi.org/10.1111/age.12758

Korablev, M.P., Korablev, N.P. & Korablev, P.N. (2021) Genetic diversity and population structure of the grey wolf (*Canis lupus* Linnaeus, 1758) and evidence of wolf × dog hybridisation in the centre of European Russia. *Mammalian Biology*, 101(1), 91–104. Available from: https://doi.org/10.1007/s42991-020-00074-2

Kunita, F., Udagawa, C., Inagaki, T. et al. Allele frequencies and the probability of identity for 18 STR loci generated via the Canine Genotypes™ Panel 2.1 kit in Golden retriever, Miniature dachshund, and Shiba Inu groups for forensic applications. https://doi.org/10.20944/preprints202312.0089.v1

Linacre, A. et al. (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Science International: Genetics*, 5(5), 501–505. Available from: https://doi.org/10.1016/j.fsigen.2010.10.017

Lorenzini, R., Attili, L., Tancredi, C. et al. (2022) A validated molecular protocol to differentiate pure wolves, dogs and wolf×dog hybrids through a panel of multiplexed canine STR markers. *Diversity*, 14, 511. Available from: https://doi.org/10.3390/d14070511

Magory Cohen, T., King, R., Dolev, A., Boldo, A., Lichter-Peled, A. & Kahila Bar-Gal, G. (2013) Genetic characterization of populations of the golden jackal and the red fox in Israel. *Conservation Genetics*, 14(1), 55–63. Available from: https://doi.org/10.1007/s10592-012-0423-1

Morf, N.V., Kopps, A.M., Nater, A. et al. (2021) STRoe deer: a validated forensic STR profiling system for the European roe deer (*Capreolus capreolus*). *Forensic Science International*, 1, 100023. Available from: https://doi.org/10.1016/j.fsiae.2021.100023

National Research Council. (1996) *The evaluation of forensic DNA evidencelcommittee on DNA forensic science: an update*. Washington D.C: National Academy Press, p. 272.

Ogden, R. (2012) Genetic data from 15 STR loci for forensic individual identification and parentage analyses in UK domestic dogs (*Canis lupus familiaris*). *Forensic Science International: Genetics*, 6(2), e63–e65. Available from: https://doi.org/10.1016/j.fsigen.2011.04.015

Peakall, R. & Smouse, P.E. (2006) GenAlEx 6: genetic analysis in excel. population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295. Available from: https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in excel. population genetic software for teaching and research—an update. *Bioinformatics*, 28, 2537–2539. Available from: https://doi.org/10.1093/bioinformatics/bts460

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. Available from: https://doi.org/10.1093/genetics/155.2.945

Radko, A. & Podbielska, A. (2021) Microsatellite DNA analysis of genetic diversity and parentage testing in the popular dog breeds in Poland. *Genes*, 12(4), 485. Available from: https://doi.org/10.3390/genes12040485

Radko, A., Podbielska, A. & Wierzbowska, I.A. (2022) The assessment of the efficacy of STRs panels recommended by the ISAG for canine pedigrees analysis for forensic casework. *Annals of Animal Science*, 22(3), 885–889. Available from: https://doi.org/10.2478/aoas-2021-0085

Rębała, K. et al. (2022) STR typing of European elk (moose) and European roe deer with novel forensic assays reveals contrasting patterns of genetic structure of the two cervids in Belarus. *Russian Journal of Genetics*, 58(12), 1493–1503. Available from: https://doi.org/10.1134/S1022795422120109

Seddon, J.M. (2005) Canid-specific primers for molecular sexing using tissue or non-invasive samples. *Conservation Genetics*, 6(1), 147–149. Available from: https://doi.org/10.1007/s10592-004-7734-9

Taberlet, P. & Luikart, G. (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, 68, 41–55. Available from: https://doi.org/10.1016/S0169-5347(99)01637-7

Verardi, A., Lucchini, V. & Randi, E. (2006) Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molecular Ecology*, 15(10), 2845–2855. Available from: https://doi.org/10.1111/j.1365-294X.2006.02995.x

van Asch, B. (2012) State-of-the-art and future prospects of canine STR-based genotyping. *Open Forensic Science Journal*, 3(1), 45–52. Available from: https://doi.org/10.2174/1874402801003010045

van Oosterhout, C., Hutchinson, W.F., Wills, D.P. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), 535–538. Available from: https://doi.org/10.1111/j.1471-8286.2004.00684.x

Waits, L.P. et al. (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, 10, 249–256. Available from: https://doi.org/10.1046/j.1365-294X.2001.01185.x

Wostenberg, D.J. & Burnham-Curtis, M.K. (2023) The development of multiplex STR panels for the identification of bald eagles (*Haliaeetus leucocephalus*) and golden eagles (*Aquila chrysaetos*). *Forensic Science International: Animals and Environments*, 3, 100062. Available from: https://doi.org/10.1016/j.fsiae.2022.100062

Zenke, P. (2011) Population genetic study in Hungarian canine populations using forensically informative STR loci. *Forensic Science International: Genetics*, 5(1), e31–e36. Available from: https://doi.org/10.1016/j.fsigen.2010.03.013

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hrebianchuk, A.E., Parfionava, N.S., Zabauskaya, T.V. & Tsybovsky, I.S. (2024) A panel of tetranucleotide STR markers as an alternative approach to forensic DNA identification of wolf and dog. *Animal Genetics*, 55, 440–451. Available from: https://doi.org/10.1111/age.13428