Forensic Science International: Animals and Environments 6 (2024) 100092

Contents lists available at ScienceDirect

Forensic Science International: Animals and Environments

journal homepage: www.sciencedirect.com/journal/forensic-science-international-animals-and-environments

Original research

A DNA profiling system for conservation management of Kamchatka brown bear (*Ursus arctos piscator*); population data and system performance from 16 autosomal STRs

Rune Andreassen [a,b,*], Berit K. Hansen [a,b], Liya Pokrovskaya [c], Vladimir Zhakov [d], Daniel Kling [e,f], Cornelya F.C. Klütsch [b], Ida Fløystad [b], Hans G. Eiken [b], Snorre B. Hagen [b]

[a] Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway
[b] Norwegian Institute of Bioeconomy Research (NIBIO), Division of Environment and Natural Resources, Svanhovd, Svanvik N-9925, Norway
[c] Laboratory of Animal Behaviour, Department of Vertebrate Zoology, Faculty of Biology, Lomonosov Moscow State University, 119234 Leninskiye Gory, 1/12, Moscow, Russia
[d] Kamchatka Branch of Pacific Geographical Institute (KB PGI) FEB RAS, Petropavlovsk-Kamchatsky, Russia
[e] Department of Forensic Sciences, Oslo University Hospital, Oslo, Norway
[f] Biostatistics (BIAS), Norwegian University of Life Sciences, Ås, Norway

ARTICLE INFO

Keywords:
Wildlife forensics
STRs
Brown bear
Non-invasive genetic sampling
Population genetics

ABSTRACT

Despite the high density of brown bears (*Ursus arctos piscator*) on the Kamchatka peninsula their genetic variation has not been studied by STR analysis. Our aim was, therefore, to provide population data from the Kamchatka brown bear population applying a validated DNA profiling system. Twelve dinucleotide STRs commonly used in Western-European (WE) populations and four additional ones (G10C, G10J, G10O, G10X), were included. Template input $\geq$ 0.2 ng was successfully amplified. Measurements of precision, stutter and heterozygous balance showed that markers could be reliably genotyped applying the thresholds used for genotyping WE brown bears. However, locus G10X revealed an ancient allele-specific polymorphism that led to suboptimal amplification of all 174 bp alleles (Kamchatka and WE). Allele frequency estimates and forensic genetic parameters were obtained from 115 individuals successfully identified by genotyping 434 hair samples. All markers met the Hardy-Weinberg and linkage equilibrium expectations, and the power of discrimination ranged from 0.667 to 0.962. The total average probability of identity from the 15 STRs was 1.4 $\times 10^{-14}$ ($F_{ST} = 0.05$) while the total average probability of sibling identity was 6.0 $\times 10^{-6}$. Relationship tests revealed several parent-cub and full sibling pairs demonstrating that the marker set would be valuable for the study of family structures. The population data is the first of its kind from the Kamchatka brown bear population. Population pairwise F_{ST}`s revealed moderate genetic differentiation that mirrored the geographic distances to WE populations. The DNA profiling system, providing individual-specific profiles from non-invasive samples, will be useful for future monitoring and conservation purposes

1. Introduction

DNA markers such as microsatellites (STRs) have been applied for over two decades in conservation genetics to study population diversity, impact of genetic drift and level of inbreeding in a variety of species [1,2]. The introduction of STRs in conservation genetics has allowed for use of non-invasive sampling techniques as DNA profiles may be successfully generated from DNA retrieved from materials like hair and faeces [3]. One of the best studied large carnivores in conservation genetics is the brown bear (*Ursus arctos*) [4,5]. The use of such non-invasive genetic sampling techniques has allowed for long-term genetic monitoring of brown bear populations. The Norwegian population has e.g. been monitored more than 15 years by DNA analysis of hair and faeces samples collected annually (see www.rovdata.no). Applying capture-recapture methods important parameters such as abundance, family structure, survival, and migration has been investigated without having to catch or disturb the animals [6–8].

The International Society for Forensic Genetics (ISFG) recommends

* Corresponding author at: Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway.
E-mail address: rune.andreassen@oslomet.no (R. Andreassen).

https://doi.org/10.1016/j.fsiae.2024.100092
Received 21 November 2023; Received in revised form 3 July 2024; Accepted 3 July 2024
Available online 4 July 2024

the application of tetranucleotide STRs in human forensics as well as in wildlife forensic casework as they are less prone to technical artifacts like stutter [9]. Although some tetranucleotide markers have been validated in black bears and brown bears [10,11], the dinucleotide STRs have been utilized in many non-invasive wildlife genetic studies including those of European brown bear [12–16]. Despite tetranucleotide STRs were recommended by ISFG, 13 dinucleotide markers were chosen for validation by Andreassen et al., 2012 [17]. The reason why these STRs were preferred was that they were already applied for monitoring bear populations, and they seemed to perform well in sample materials like faeces and hair due to relatively short amplicon sizes [17–21]. Tetranucleotide STRs, on the other hand, were at the time not used at all for DNA analysis of brown bears. These validations resulted in a marker set combining twelve dinucleotide STRs commonly used for studying brown bears that were among the first wildlife marker sets that were fully validated according to the ISFG recommendations [17]. This forensic profiling system also fulfilled the needs for a marker set for non-invasive capture-recapture methods for conservation genetics as the relatively short amplicon sizes of the dinucleotide markers are beneficial for sample materials like faeces and hair [17–21]. This twelve-marker set, providing individual specific DNA profiles, has been widely applied for forensic purposes [22] as well as for study of abundance, family structure, survival, and migration in North-West European brown bear populations (e.g. [17–21]).

The brown bear populations of Europe vary greatly in size [23], and the species was almost eradicated in several European countries in the 19th and 20th century. This was partly a result of habitat loss and fragmentation, but state-financed extermination campaigns to avoid livestock predation also played a crucial role [24,25]. In recent years, bear management policies and genetic monitoring by use of non-invasive sampling methods have contributed to the recovery of brown bears in Scandinavia [24]. Relatively large populations now exist in the Scandinavian countries (approximately 150 in Norway, 3300 in Sweden and 1800 in Finland) [25–27].

The largest populations of Eurasian brown bear are, however, found in Russia [28]. The Kamchatka brown bears (*Ursus arctos piscator*) have the largest body sizes of the brown bears in Eurasia. In the last 20 years, the population of Kamchatka brown bear has been growing and is estimated at 22,000–24,000 individuals. This subspecies represents the most eastern brown bear population in Russia and therefore is also known as "Far Eastern brown bear". The main threats to the population are trophy hunting and poaching. Although such harvesting is one of the most effective regulators of population size, these activities have been declining [29,30]. The competition on common food resources, in particular salmonid fish, is also a source for human-bear conflicts which have increased in recent years [30]. The Kamchatka brown bears have been studied by observations in the field as well as small-scale marking and satellite tracking (e.g., [28,30,31]), but as far as we know, the Kamchatka brown bear population has not been studied by capture-recapture methods applying autosomal STRs and non-invasive sampling methods. A validated DNA profiling system for Kamchatka brown bears would be an important tool for wildlife forensics. Similar DNA-profiling systems have been applied to solve several cases of illegal hunting and in such cases, provided the prosecution with the conclusive evidence that linked the suspect to the crime [22,32,33]. In one such case example, the individual specific DNA profiles from typing hair and blood stain materials at the kill site matched the DNA-profile from the blood stain on the suspect's firearm as well as being identical to a bear registered in the national bear monitoring database [22]. In cases of fatal bear attacks individual specific DNA-profiling may, on the other hand, be used as a tool to identify harmful individuals [34]. If suitable for use with non-invasive sampling methods and capture-recapture methods, the individual-specific DNA profiles produced would also be valuable for brown bear conservation genetics and population genetics.

The twelve STR markers that were validated for use in the Scandinavian brown bear populations [17] have been successfully applied to study West European (WE) brown bear populations, and large amounts of population specific genetic data already exists (e.g., [17,20,35]). It would be an advantage if the same STR markers could be applied to study the Kamchatka brown bears as this would allow for comparative studies with other Eurasian populations. However, the Kamchatka brown bear populations and the extensively studied WE populations are the brown bear populations that are separated by the largest geographic distances. There could be many reasons why the same STRs/primers would not work well in such a geographically distant brown bear population (e.g., primer site polymorphisms, large differences in allele sizes, low allelic variation in certain markers etc.). Also, population specific allele frequency estimates are needed for implementation of their use in conservation genetics and wildlife forensics. If performing well in the Kamchatka brown bear population, the DNA profiles from the twelve validated STR markers [17] would be expected to provide individual-specific DNA profiles. However, to study family relationships it has been demonstrated that even if the STRs are highly informative it is useful to have an additional number of markers (>15 combined in a DNA profiling system) to allow for estimates of kinship with a satisfying probability [36].

One objective of this study was, therefore, to carry out ISFG recommended validation tests [9] in sample materials (hairs) collected from the Kamchatka brown bear population. The validations included measurements of sensitivity, precision, stutter, and heterozygote balance in five PCR-multiplexes. The multiplexes included the twelve validated STRs [17] and a sex-specific amelogenin marker [37,38]. The STRs were originally cloned and characterized from American black bear and European brown bear [12,15]. Four additional STRs, originally identified in American black bear [15], that have not previously been validated according to ISFG recommendations, were also included. For ease of reading, they are referred to as the new STRs. Furthermore, we aimed to investigate the success rates when applied on DNA-template extracted from hair traps (commonly used in capture-recapture methods), as well as allele frequency distributions, and relevant forensic genetic parameters for the selected 16 STR markers. Finally, inbreeding, population substructure and their genetic distance to WE brown bear populations were investigated in this population for the first time. Altogether, we aimed to provide a validated DNA profiling system for the Kamchatka brown bear population that is useful for forensic purposes, conservation, and population genetic studies as well as studies of family relationships by use of DNA profiles generated from non-invasive sample materials.

2. Materials and methods

2.1. Population sample materials from Kamchatka brown bears

Hair samples from brown bears in the Kamchatka region (n=434) were collected in hair traps placed at 16 different locations around the Azabache Lake over four weeks in July-August (traps were checked every 3–4 day during this period). The area in which the hair traps were located encompassed approximately 200 square km (Fig. 1). The coordinates and the name of the closest river for each of the hair traps are given in Table 1. The samples were stored dry and dark (in paper envelopes) at room temperature until DNA extraction.

2.2. Positive control samples for tests of marker performance

High quality tissue samples from four WE brown bears (*Ursus arctos*) with known genotypes at all twelve validated markers (G1A, G10B, G1D, G10L, MU05, MU09, MU10, MU15, MU23, MU50, MU51 and MU59) [17] were included in the STR analysis pipeline as positive controls for each run. Results from these samples were also used for measuring the performance of the four new STRs validated in this study (G10C, G10J, G10O, and G10X). Template DNA in these controls was extracted from bear tissue samples using the DNeasy Blood & Tissue kit

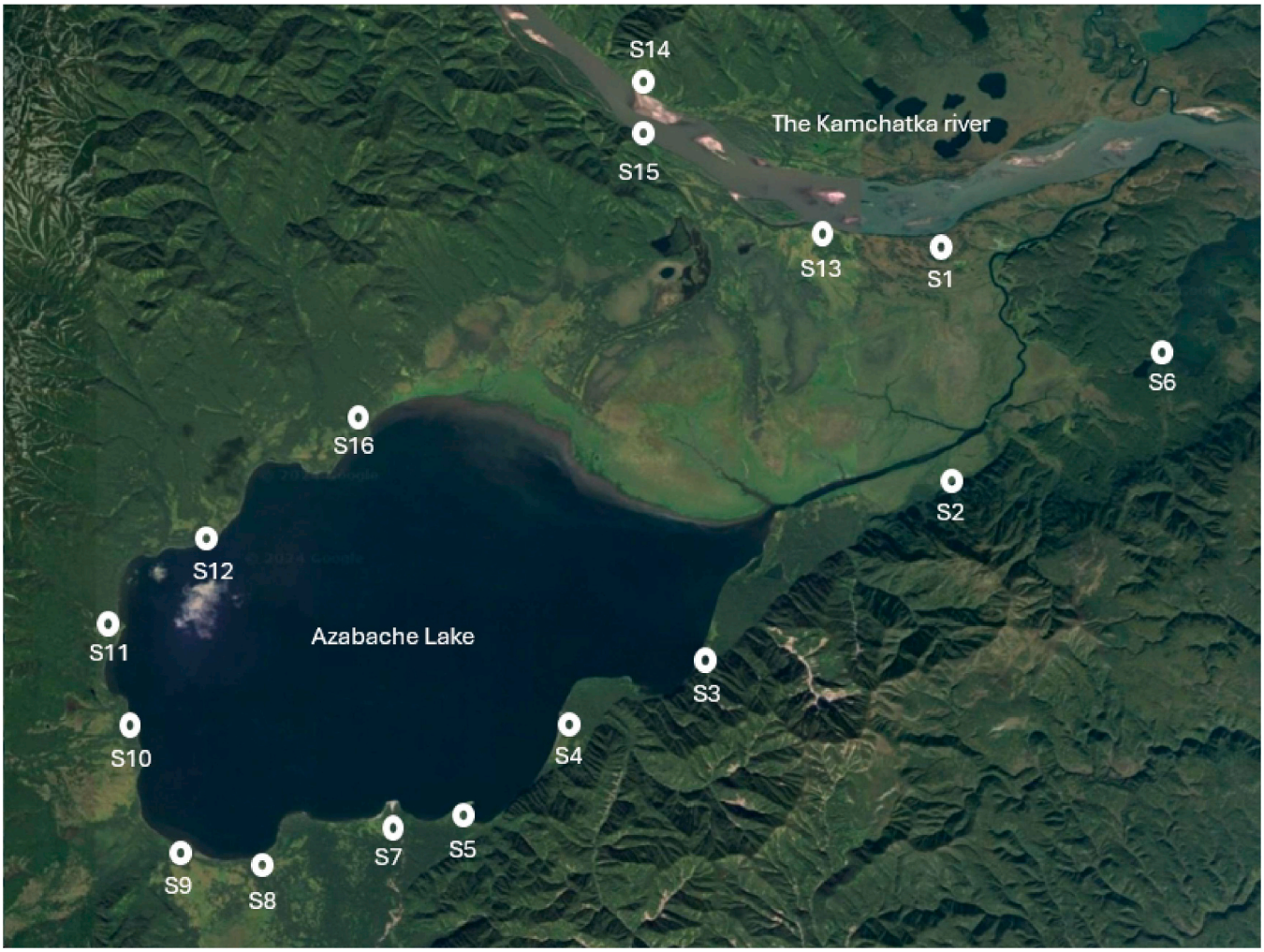

Fig. 1. The geographic area around Azabache Lake (approximately 200 square km) on the Kamchatka Peninsula where 16 hair traps for brown bear were placed. Each hair trap is marked with a white circle and denoted S1 to S16.

Table 1
Location of the 16 Kamchatka hair trap stations.

Station	Coordinates	River
S1	N56 12.687 E161 57.618	Kamchatka
S2	N56 10.213 E161 57.547	Olezkina
S3	N56 08.395 E161 53.224	Geshkin
S4	N56 07.930 E161 50.822	Lotnaya
S5	N56 07.044 E161 49.083	Vydrovaya
S6	N56 11.443 E162 01.330	Kursinka
S7	N56 06.990 E161 47.780	Bushujka
S8	N56 06.516 E161 45.387	Topolinaya
S9	N56 06.674 E161 43.927	Lamutka
S10	N56 07.842 E161 43.254	Kultuchnaya
S11	N56 08.818 E161 43.057	Arishkin
S12	N56 09.686 E161 44.700	Snovidovskiy
S13	N56 12.787 E161 55.234	Kamchatka
S14	N56 14.209 E161 51.963	Kamchatka
S15	N56 13.737 E161 51.988	Kamchatka
S16	N56 10.824 E161 47.388	Ponomarka

(Qiagen, Hilden, Germany) following the manufacturers protocol for tissue samples and concentrations measured with NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific, MA, USA). All control samples originated from legally harvested bears in Norway (T8 (male, 2007), T11 (female, 2009)) and Karelia, Russia (T9 and T10, both males, harvested in 2005). These samples have been used as positive controls in the accredited DNA profiling analysis at NIBIO Svanhovd laboratory (ISO/IEC 17025) that carried out the yearly monitoring of the Norwegian brown bear population until 2021.

2.3. DNA extraction, PCR amplification and STR analysis of sample materials

The root tip of hairs from each of the 434 samples were cut and all hairs from a sample (1–10 hairs) were pooled. DNA was extracted from hair-roots using the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) following the manufacturers protocol (extraction of DNA from tissue). The DNA extract was eluted in 50 μl buffer when there were 4–10 hairs in a sample (n = 240), and in 30 μl when there were 1–3 hairs in a sample (n = 194) to counteract a reduction of DNA concentrations when less materials were extracted.

Eight of the twelve validated STRs [17] were combined in three multiplex assays (MP). MP1 amplified MU09, MU10, MU23, MU59, while MP2 amplified MU05, G10L, MU51, and MP3 amplified MU50 and amelogenin XY (sex-marker, [38]). These multiplex assays have been routinely applied in brown bear STR analyses with template DNA extracted from hair-roots, faeces, and tissue as part of the ongoing annual Norwegian monitoring of brown bears that is part of the Norwegian Large Predator Monitoring Program (see www.rovdata.no). The remaining four validated STRs (G1D, G1A, G10B, MU15) were combined with the four new STRs G10C, G10J, G10O, G10X in two new multiplexes (MP8 amplifying G10B, G10C, G10O, G10X and MP14

amplifying G1D, G1A, G10J, MU15). See Table 2 for an overview of the different STRs and their combinations in the multiplex assays. Primer sequences used for all loci were selected from those used in a nested PCR strategy for amplifying the STRs in brown bear [12], but with slightly modified reverse primers. A complete overview of primer sequences, concentrations, dye labels and modifications [39] for all 16 STR loci and the sex specific marker (amelogenin)[37,38] is given in Supplementary File 1. Dilution of the control extracts with known template concentrations were used as positive controls in all runs (see also Section 2.2). Negative controls (ddH_2O as substitute for template) were included for every seventh sample run in the analysis pipeline.

No quantification of DNA concentrations was carried out for hair sample extracts as concentration measurements would not reveal the template quality (degree of degraded DNA) or the origin of the DNA (bear or bacteria). Instead, the genotyping results of the 434 samples in MP1 were used to qualitatively judge the amount of useful template DNA in each sample. Samples that were negative or of too poor quality were excluded from further analysis.

Multiplex PCRs were performed in 10 μl reaction volumes containing 1 x Multiplex PCR Master Mix (Qiagen, Hilden, Germany), 1 μl primer mix, 0.5 x BSA, 2.95 μl ddH2O and 1 μl template DNA from the eluted hair-sample extracts. Samples with low RFUs or very strong (offscale) results were reamplified with adjusted template volumes (2 μl in cases with weak results), or they were run in single plex assays (see below) while samples with strong results were diluted prior to reamplification. The addition of template DNA was routinely performed by a pipetting robot (epMotion 5070, Eppendorf, Hamburg, Germany), but at low template volumes, manual pipetting was also executed. PCR conditions were 10 minutes at 95 ºC, 35 cycles of 30 seconds at 94 ºC, 30 s at 58 ºC, and 1 min at 72 ºC, and final extension for 45 min at 72 ºC.

Single PCRs were performed for samples where the genotyping of a single STR locus was incomplete when genotyped in the multiplex assay. These PCRs were performed in 10 μl reaction volumes containing 1.5 mM PCR Gold buffer, 200 μM dNTP, 1.5 mM $MgCl_2$, 0.5 μM of each primer, 1 U AmpliTaqGold DNA polymerase (ThermoFisher Scientific, MA, USA), 1 x BSA and 1 μl template DNA. PCR conditions were as before mentioned for all STRs except the single analysis of G10X that used an annealing temperature (Ta) of 54 ºC.

Capillary electrophoresis (CE) with 4 second injection time was carried out on an ABI3730 (ABI, ThermoFisher Scientific, MA, USA). Formamide (10 μl) was added for denaturation of the PCR product (1 μl) prior to CE, and allele sizes were measured using Genescan 500LIZ standard and GeneMapper 4.0 software (ABI, ThermoFisher Scientific, MA, USA).

The analytical peak height thresholds (APHT) of 300 RFUs and homozygous peak height thresholds (HPHTs) of 800 RFU were used when analysing STRs from MP1, MP2 and MP3. This was based on prior in-house validations. An APHT of 300 RFU was also applied when analysing the STRs in MP8 and MP14. The homozygous peak height thresholds (HPHTs) applied for the analysis of the four new STR markers and the other STR markers in MP8 and MP14 were set based on the results from the validation tests of the controls (Section 2.4). The HPHTs were 1800 RFU for G10J and G10B while it was 800 RFU for G10C and G10O as well as the other STRs.

Each sample (n=434) was amplified in three multiplex assays (MP1, MP2 and MP3) followed by fragment analysis in the following manner. Samples which were initially genotyped as homozygous at any of the STR-loci in the first run were confirmed as homozygous by a minimum of two more replicate analyses. Heterozygous genotypes were confirmed by one additional replicate analysis. In practice, if there was a third genotyping of a particular STR locus from a given sample, this was usually carried out in a single PCR assay for the specific marker in question. The final allele designation of all 16 genotypes from a sample included manual inspection of all electropherograms.

Based on the results from Andreassen et al., 2012 [17], a successful genotyping of all eight STR loci in MP1, MP2 and MP3, as well as a confirmed sex, was set as a requirement for accepting the DNA profile as representative of one individual. Any samples with identical DNA profiles when comparing results in these eight markers were assumed to be recaptures of the same individual (individual specific profile). For analysis of the remaining eight STR loci (MP8 and MP14), only one of the representative samples of an individual (N=115) was analysed. Samples with no results in any STR loci were defined as negative for DNA from brown bear. All samples with a satisfactory genotype result in at least one of the initial eight markers analysed were designated as positive for DNA from brown bear. Applying a conservative approach, the samples with genotype results in less than eight markers were designated as brown bears without individual specific DNA profiles and annotated as unknown individuals. A sample with three alleles or more in at least one marker, was discarded as a mixed sample (a hair sample with hairs from more than one individual).

Table 2
The 16 dinucleotide STRs combined into five STR multiplex assays.

Multiplex assay (MP)	Validated STR-markers[a]	New STR-markers[b]
MP1	MU09	
	MU10	
	MU23	
	MU59	
MP2	MU05	
	G10L	
	MU51	
MP3	XY[c]	
	MU50	
MP8	G10B	
		G10C
		G10O
		G10X
MP14	G1D	
	G1A	
		G10J
	MU15	

[a] The dinucleotide STR loci validated in Andreassen et al. 2012 [17]
[b] The new STRs added for validation in this study (in multiplex assays MP8 and MP14)
[c] MU50 is multiplexed with the sex-informarmative amelogenin gene marker denoted XY

2.4. Validation tests of sensitivity, precision, stutter ratio and heterozygote balance in the new STR-loci

Measurements of sensitivity, stutter and heterozygous balance were part of the validation of the four new STRs (G10C, G10J, G10O, G10X) when multiplexed with the other STRs. The performance of the other STRs was also recorded in the same manner. Measurements of sensitivity will reveal at what lower template concentration the STR marker can not reliably amplify the alleles in a genotype (either allele drop-out in a heterozygous genotype or complete drop out of both alleles). It is important that a new marker has a sensitivity that allows for successful genotyping in materials usually investigated. Also, it should be a similarly good sensitivity as that of the other markers if to contribute to improving the power of the DNA profiling system. The measurements of precision will reveal if the CE runs have the power to discriminate between alleles. With a reproducibility with standard deviation (SD) less than 0.3 bp, any dinucleotide alleles (2 bp size difference) will be correctly assigned in 99.7 % of cases if applying a bin range of 3 x SD [40,41]. Stutter ratios and variation in heterozygote balance [17,42,43] are measured for each STR marker to ensure it performs in a manner that allows implementation of locus-specific analytical thresholds to avoid two common types of genotyping errors. One such error is to mistype a heterozygous genotype as a homozygous genotype (drop-out of the allele with the smaller peak-height). This could occur if the interval between the homozygote peak height threshold (HPHT) and the analytical peak height threshold (APHT) does not reflect the

heterozygote balance ratios. The other common type of genotyping error would be to mistype a homozygous genotype as a heterozygous genotype due to a large stutter. This could occur if the values of heterozygote balance ratios and stutter ratios overlap, e.g. if both stutter ratio and the heterozygote balance ratio are close to one.

The measurements were carried out in template dilutions of the control samples. Two of the control samples (T10 and T11) were analysed 16 or more times with three different template inputs: 0.6 ng, 0.2 ng and 0.05 ng as part of the validation of sensitivity. The two other controls (T8 and T11) were analysed with template inputs assumed to provide genotype results (0.3 ng and 0.6 ng) and used for measurement of stutter, heterozygous balance and precision. Between-run precision was measured in the four new STR-loci validated in this study as well as all other STRs by $\geq$20 independent amplifications and subsequent CE runs. The results used to measure heterozygote balance were from T8 (STRs G10X and G10J as well as STRs in MP1, MP2, and MP3) and T11 (STRs in MP8, MP14). Measurements of stutter ratio were also acquired from analysis of these controls. Stutter ratio was calculated by dividing the peak height (RFU) of the stutter peak in position −1R (i.e., one repeat less than the true allele) by the peak height of the true allele [17,44]. Heterozygote balance was calculated by dividing the peak height (RFU) of the short allele by the peak height of the longer allele [17,45].

2.5. Sequencing of tandem repeat arrays and immediate flanking sequences of locus G10X

A suboptimal PCR amplification of one allele (174 bp) was observed in STR locus G10X. The tandem repeat array and the immediate upstream and downstream sequences at locus G10X were analysed by DNA sequencing to reveal whether any polymorphisms in the primer binding sites could explain this observation. The sequence from the original *Ursus americanus* clone (Genbank acc. U22093) was aligned with the orthologous region from *Ursus arctos horribilis* (Genbank acc. NW020656124: 66830130–66830700) to provide DNA sequence information of this particular locus that expanded beyond the primer binding sites in both directions. The aligned sequences were used to pick primers located more distant from the tandem repeat region by use of primer3 software (https://primer3.org/). The new forward primer: 5'*TGTAAAACGACGGCCAGT*GCTAGTGGTACTGGTGATGGT3' and reverse primer: 5'*CAGGAAACAGCTATGACC*CAGCAGTGTCAACCATAGCC3' were extended with M13 universal primer sequences (italics) to allow for sequencing of PCR products with BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, Waltham, MA, USA). PCR products amplified from individuals that were homozygous genotypes 174/174, 180/180 and 182/182 were sequenced, and Sequencher 4.7 software (GeneCodes, Ann Arbor, USA) was used to align and compare the sequences.

2.6. Analysis of data for population genetic parameters, forensic efficiency parameters and kinship

The GDA software v1.1 [46] was used for evaluation of Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) by applying Fisher's exact test and 3200 permutations. The significance levels ($p = 0.05$) were Bonferroni corrected for LD tests (120 tests, $p = 0.0004$) and HWE (16 tests, $p = 0.003$). The LD test setting included the preserving genotypes option in GDA. Estimations of expected heterozygosity (HE) and population structure (F_{IS} and F_{ST}) were also performed with GDA v.1.1. When estimating F_{ST} (both total and pairwise), the data from the eight populations studied in Andreassen et al., 2012 [17] were compared to the data obtained from the Kamchatka population, using the F-statistics analysis (95 % bootstrap confidence interval, 1000 permutations) in the GDA software. In addition, pairwise F_{ST} was estimated among these populations using GenAlEx v.6.501 [47,48]. F_{ST}`s is a widely used estimator applied to investigate the genetic differentiation between populations [49].

PowerStats (the gene count method) was used to obtain the observed allele frequencies and for the calculation of the forensic efficiency parameters [49]. API-Calc v.1.0 [50] was used to estimate average probability of identity (PI) for each marker as well as total average PI (the probability of observing identical profiles in two not closely related individuals in the given population). This statistical program also allows for calculations of average PI based on allele frequencies when accounting for the effects of population substructure (theta), inbreeding (F_{IS}) and close relatedness e.g. full siblings; PI_{sib} (the probability of observing two copies of any profile in dizygotic full siblings). The PI's (PI_{ave} and PI_{sib}) were estimated applying a conservative measure of $F_{ST} = 0.05$ to account for population substructure.

FAMILIAS, which is widely used in forensic genetics, was used to reveal close family relationships (parent-cub, siblings) applying the blind search function with $F_{ST} = 0.05$ to account for population substructure [51,52]. All pairwise combinations among the 115 individuals from the Kamchatka materials were tested to identify parent-cub or full siblings against the alternative hypothesis that they were unrelated and allowing no mutations. The threshold for likelihood ratios (LR) supporting such close family relationships was set to $LR \geq 100$. In addition, parent-cub pairs were tested against the alternative hypothesis that they were full siblings, and finally, the 115 individuals were searched to detect putative father-mother-cub trios.

3. Results

3.1. Measurements of sensitivity, precision, stutter ratios and heterozygote balance

The sensitivity tests showed that all new markers (G10C, G10J, G10O, G10X) were successfully amplified and detected with template inputs of 0.6–0.2 ng. However, when template input was reduced to 0.05 ng there were multiple cases with allele drop-outs in both the four new STR markers and the 12 previously validated STRs. The complete results for sensitivity tests of all 16 STR markers, when combined in the five multiplex assays, are given in Supplementary File 2.

The between-run measurement of precision of the alleles from the four new STR markers (Table 3) revealed standard deviations (SD) of 0.15 bp or less at loci G10C and G10O, while at G10X and G10J the SD measurements of the alleles were $\leq$ 0.20 and 0.24 bp, respectively. The other 12 STRs revealed similar SDs $<$ 0.22 (Supplementary File 3).

The median stutter ratios at the four new markers ranged from 0.18 to 0.56 and given along with their upper 95 percentiles in Table 3. As expected, the longer allele in a heterozygote genotype also showed slightly larger stutter peaks [17,45]. While the largest stutter ratio observed in the new markers was 0.56 at G10C the largest stutter ratio in the 12 other STRs was 0.77 (Supplementary File 3).

The calculation of heterozygote balance ratio provided information about the direction of the imbalance. The heterozygote balance measurements in three of the new STR markers (G10C, G10O and G10J) were in the range 1.3–3.0 indicating that the shorter allele would be the one with the higher peak height. The heterozygote balance in the remaining 12 STR markers were investigated in Andreassen et al., 2012 [17]. Here, they were tested when amplified in different multi-plex assays. All results from these loci showed heterozygote balances above 1 and were comparable to the results from our initial study [17] (Supplementary File 3).

The only exception to heterozygote balance ratios being above 1 was seen in locus G10X. This marker revealed a suboptimal amplification of the 174 bp allele. The control sample with genotype 174/180 showed a heterozygous balance of 0.27 when amplified in the multiplex-assay MP8 (Table 3). This observation was not limited to the Kamchatka materials but seen in all samples tested (controls and other samples from Scandinavia, data not shown). All other allele combinations at G10X did, however, show heterozygous balance ratios above 1.2 (Table 3). Analysis of the 174/180 genotype applying single PCR and reduced

Table 3
Between-run measurements of precision, heterozygote balance and stutter ratio in four new dinucleotide STRs from ≥ 20 runs of control T11 in multi-plex assays MP8 and MP14 and, additionally, T8 in G10X and G10J.

Assay	Locus	Alleles[a]	Mean (bp)[b]	SD (bp)[c]	HB[d]	Stutter ratio[e]
	G10C allele A	155	154.80	0.09		0.42 (0.47)
MP8	G10C allele B	163	163.21	0.11	1.58 (1.20–2.21)	0.56 (0.65)
	G10O allele A	201	201.23	0.15		0.28 (0.32)
MP8	G10O allele B	203	203.26	0.12	1.42 (1.04–1.72)	–
	G10X allele A	180	179.87	0.19		0.36 (0.40)
MP8	G10X allele B	188	188.14	0.17	1.17 (0.91–1.85)	0.54 (0.57)
	G10X allele A	174	173.77	0.20		0.27 (0.34)
MP8	G10X allele B	180	179.95	0.18	0.27 (0.14–0.62)	0.39 (0.45)
	G10J allele A	99	98.96	0.24		0.18 (0.20)
MP14	G10J allele B	101	101.04	0.22	1.26 (1.01–2.84)	–
	G10J allele A	101	101.20	0.19		0.24 (0.25)
MP14	G10J allele B	115	115.54	0.17	3.02 (1.75–5.53)	0.52 (0.56)

[a] Alleles as designated with a size-based nomenclature (basepairs).
[b] Mean value of allele sizes when measured with POP7 on ABI3730.
[c] Standard deviations (SD) from between-run measurements of ≥ 20 runs of a control sample.
[d] Median heterozygote balance with upper and lower 95 percentiles in parenthesis.
[e] Median stutter ratios of alleles with upper 95 percentiles in parenthesis.

stringency in the annealing step (T_a=54 °C) showed a much-improved amplification efficiency of the 174 allele. The results from amplifying the 174/180 genotype in MP8 (T_A=58 °C) or as a single-plex (T_a=54 °C) are shown in Fig. 2.

The cause for the suboptimal amplification of the 174 bp allele was explored by sequencing the tandem repeat region as well as the sequences upstream and downstream of the repeat region including the PCR-primer binding sequences. Samples with three different homozygous genotypes 174/174, 180/180, and 182/182 were successfully sequenced (Supplementary File 4). The likely cause for the suboptimal amplification was a deletion of four base pairs in the sequence flanking the CA-repeats. The deleted bases included the two 3' bases of the reverse primer binding sequence. This led to one mismatch between the primer's 3' end and the primer binding sequence in the G10X 174 bp alleles.

3.2. Kamchatka hair samples: genotyping success rate, individual profiles and mixtures

Sixty-six percent (286 of 434) hair samples from the Kamchatka population were successfully genotyped at eight or more STR markers. The success rate of genotyping was, however, very much influenced by the number of hair roots collected in a sample. The success rate when sample extracts were from 1 to 3 hair roots (30 µl eluate) *versus* 4–10 hair roots (50 µl eluate) is illustrated in Fig. 3. While 79.2 % of the 50 µl eluate DNA samples (190 out of 240) were successfully genotyped in at least eight STRs only 49.5 % of the 30 µl eluate DNA samples (96 out of194) showed the same success rate. In accordance with the conservative thresholds used in genotyping of the Scandinavian brown bear populations [17] only the samples with results from eight or more STRs were denoted as individual-specific profiles, while the ones with a lesser number of successfully typed STRs were denoted as brown bear specific but not individual specific. Pooling results from both sample groups resulted in 286 samples (65.9 %) with individual DNA profiles (8 STRs). As expected, the group with extracts from 4 to 10 hair roots showed a larger number of mixed samples compared to the 1–3 hair extracts (5 % *versus* 0.5 %), while the group with extracts from 1 to 4 hairs showed a larger percentage of negative samples than the 4–10 hair extract group (36 % vs 7 %).

Comparing the 286 samples with individual-specific DNA profiles revealed they were from 115 individuals (59 males and 56 females). The remaining 171 samples were one or more recaptures as they had identical profiles to one of those 115 individuals. An overview of results from all 434 samples collected at the different sample locations is given in Supplementary File 5. The complete genotype results from the 115 individuals, the number of times they were recaptured, and their sampling locations are given in Supplementary File 6.

3.3. Population data and forensic efficiency parameters for the Kamchatka bears

Allele frequencies derived from the 115 individuals genotyped at each of the 16 STR markers are given in Table 4 and Table 5. Several alleles not previously observed in the WE populations [17] were present

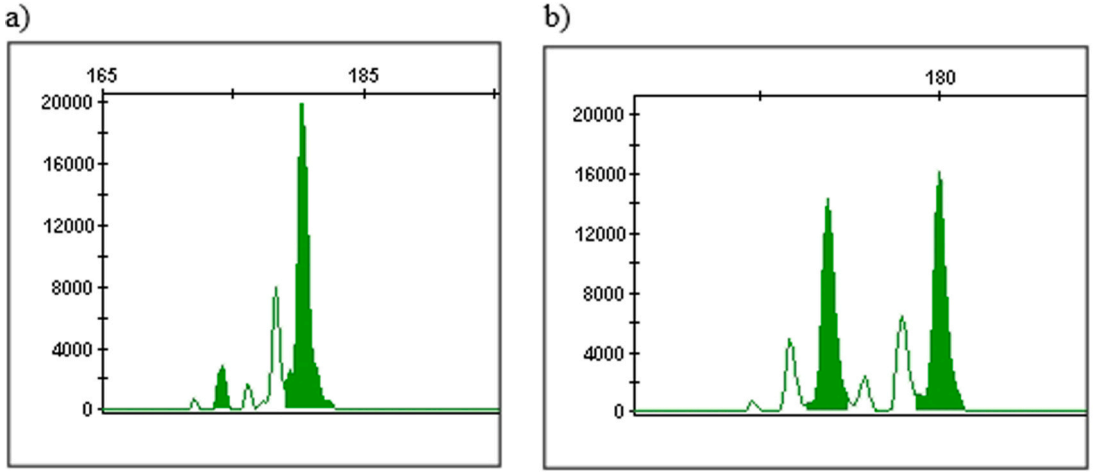

Fig. 2. Electropherograms with results for the control sample with genotype 174/180 at STR locus G10X. The two alleles are marked green. a) Results from a multiplex PCR assay (MP8) with Ta=58 °C. Allele 174 has a much lower peak height than allele 180. The heterozygote balance ratio is 0.14. b) Results from a singleplex PCR assay with Ta=54 °C. Heterozygote balance ratio: 0.89.

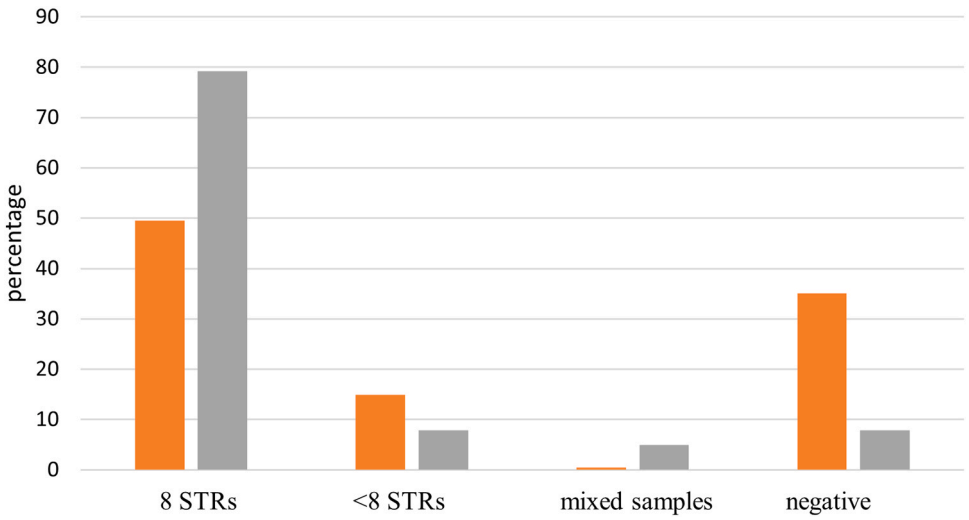

Fig. 3. Genotyping success rate (in percent) in samples with DNA extracted from either 1–3 hair roots (30 µl eluate, orange bars) or 4–10 hair roots (50 µl eluate, grey bars). The results from the two groups are divided into samples with successful STR genotyping in eight STRs (8 STRs), less than eight STRs (< 8 STRs), samples with DNA originating from two or more individuals (mixed samples) or samples with no results (negative).

Table 4
Allele frequencies, H_O and H_E and forensic efficiency measures for eight STR loci.

Allele	G10B	MU09	MU15	MU50	**Allele**	MU05	G1D	MU10	MU51
96	-	-	***0.218***	-	**113**	0.157	-	-	-
98	0.044	-	-	-	**115**	0.217	-	-	-
102	-	0.035	-	-	**117**	0.209	-	-	-
104	-	0.122	-	-	**119**	***0.022***	-	-	-
106	0.259	-	-	-	**121**	-	-	-	-
108	0.044	-	-	-	**123**	0.078	-	-	-
110	-	0.248	0.045	0.126	**125**	0.022	0.023	-	-
112	0.254	0.087	0.105	-	**127**	0.113	0.090	-	-
114	0.110	0.209	0.168	-	**129**	0.078	0.194	-	-
116	0.018	0.004	0.277	-	**131**	0.052	0.248	-	-
118	0.175	0.009	0.155	-	**133**	0.048	0.378	-	-
120	0.083	0.057	0.014	-	**135**	**0.004**	0.045	-	-
122	0.013	0.157	**0.018**	0.135	**137**	-	0.005	-	-
124	-	0.070	-	0.013	**139**	-	0.018	0.013	-
126	-	-	-	0.057	**141**	-	-	0.017	-
128	-	0.004	-	0.187	**142**	-	-	-	0.053
130	-	-	-	0.304	**143**	-	-	0.187	-
132	-	-	-	0.039	**144**	-	-	-	0.132
134	-	-	-	0.126	**145**	-	-	0.235	-
136	-	-	-	0.004	**146**	-	-	-	0.364
138	-	-	-	**0.009**	**147**	-	-	0.530	-
139	-	-	-	-	**148**	-	-	-	0.333
141	-	-	-	-	**149**	-	-	0.013	-
142	-	-	-	-	**150**	-	-	-	0.118
143	-	-	-	-	**151**	-	-	0.004	-
H_O	0.772	0.852	0.736	0.887	**H_O**	0.843	0.793	0.678	0.746
H_E	0.818	0.842	0.814	0.821	**H_E**	0.857	0.747	0.631	0.725
PD	0.941	0.952	0.938	0.933	**PD**	0.956	0.887	0.797	0.855
PE	0.548	0.699	0.487	0.769	**PE**	0.682	0.586	0.395	0.502
PM	0.059	0.048	0.062	0.067	**PM**	0.044	0.113	0.203	0.145
PI_{typ}	2.19	3.38	1.90	4.42	**PI_{typ}**	3.19	2.41	1.55	1.97
N	114	115	109	115	**N**	115	111	115	114

Observed and expected heterozygosity (HO, HE), Power of Discrimination (PD), Power of Exclusion (PE), Matching probability (PM), typical paternity index (PItyp). N=number of Kamchatka brown bears genotyped at each locus.
Allele frequencies that are marked in bold are from alleles not previously observed in the Western European brown bear population [17]. Allele frequencies of alleles that has only been observed once (a single allele) in individuals from Arkhangelsk and Kami Oblast in Russia (NIBIO database, Norway) are marked in bold and cursive.

in rather high frequencies in MU05, MU15, MU50, and MU59 (0.018–0.186, bold font in Tables 4 and 5). The STR markers MU05, MU15, and G10L also revealed alleles that, prior to this study, had only been observed once in individuals from Arkhangelsk and Komi Oblast in Russia (NIBIO genetic brown bear database, Norway). The allele frequencies of these three alleles are marked in bold and cursive font in Tables 4 and 5. Notably, one of these alleles (96) was among the most frequent in locus MU15 in the Kamchatka population. No microvariation (allele size differences of 1 bp) was discovered in any of the markers when analyzing the Kamchatka brown bear population.

The expected and observed heterozygosity frequencies (Ho and He) for all loci are given in Tables 4 and 5. Except for G10O that revealed a relatively low observed heterozygosity (0.402), the heterozygosity frequencies of the other 15 STR markers ranged from 0.626 to 0.887. The

Table 5
Allele frequencies, H_O and H_E, and forensic efficiency measures for eight STR loci.

Allele	G10L	G1A	MU59	MU23	Allele	G10J	G10C	G10X	G10O
168	-	-	-	0.100	**99**	0.305	-	-	-
170	-	-	-	0.200	**101**	0.009	-	-	-
172	-	-	-	0.017	**111**	0.027	-	-	-
174	0.167	-	-	0.448	**113**	0.004	-	-	-
176	-	-	-	0.183	**115**	0.142	-	-	-
177	-	0.308	-	-	**117**	0.274	-	-	-
178	0.018	-	-	0.052	**119**	0.031	-	-	-
179	-	0.018	-	-	**121**	0.142	-	-	-
180	0.110	-	-	-	**123**	0.053	-	-	-
181	-	0.022	-	-	**125**	-	-	-	-
182	0.022	-	-	-	**127**	0.013	-	-	-
183	-	0.161	-	-	**153**	-	0.004	-	-
184	0.202	-	-	-	**155**	-	0.540	-	-
185	-	0.281	-	-	**157**	-	0.263	-	-
186	0.053	-	-	-	**159**	-	0.058	-	-
187	-	0.129	-	-	**161**	-	0.085	-	-
188	0.057	-	-	-	**163**	-	0.009	-	-
189	-	0.071	-	-	**165**	-	0.018	-	-
190	0.053	-	-	-	**167**	-	0.022	-	-
192	0.044	-	-	-	**174**	-	-	0.104	-
193	-	0.009	-	-	**176**	-	-	0.059	-
194	0.180	-	-	-	**178**	-	-	0.041	-
196	***0.004***	-	-	-	**180**	-	-	0.045	-
198	**0.048**	-	-	-	**182**	-	-	0.532	-
202	**0.044**	-	-	-	**184**	-	-	0.153	-
228	-	-	0.004	-	**186**	-	-	0.005	-
232	-	-	0.053	-	**192**	-	-	0.005	-
234	-	-	0.040	-	**194**	-	-	0.023	-
238	-	-	**0.186**	-	**195**	-	-	-	0.045
242	-	-	0.142	-	**196**	-	-	0.036	-
244	-	-	0.119	-	**201**	-	-	-	0.692
246	-	-	0.066	-	**203**	-	-	-	0.214
248	-	-	0.058	-	**205**	-	-	-	-
250	-	-	0.168	-	**207**	-	-	-	0.049
252	-	-	0.164	-	**209**	-	-	-	-
H_O	0.877	0.759	0.841	0.626	**H_O**	0.796	0.652	0.667	0.402
H_E	0.875	0.781	0.868	0.716	**H_E**	0.790	0.625	0.677	0.473
PD	0.962	0.918	0.960	0.878	**PD**	0.917	0.794	0.871	0.667
PE	0.749	0.525	0.677	0.323	**PE**	0.592	0.358	0.379	0.115
PM	0.038	0.082	0.040	0.122	**PM**	0.083	0.206	0.129	0.333
PI_{typ}	4.07	2.07	3.14	1.34	**PI_{typ}**	2.46	1.44	1.50	0.84
N	114	112	113	115	**N**	113	112	111	112

Observed and expected heterozygosity (HO, HE), Power of Discrimination (PD), Power of Exclusion (PE), Matching probability (PM), typical paternity index (PItyp). N=number of Kamchatka brown bears genotyped at each locus. Allele frequencies in bold are from alleles not previously observed in the Western European brown bear population [17]. Frequencies of alleles that has only been observed once (a single allele) in individuals from Arkhangelsk and Kami Oblast in Russia (NIBIO database, Norway) are marked in bold and cursive.

mean Ho and He averaged across all 16 STR loci were identical (0.75). When removing marker G10X the Ho and He were 0.75 and 0.76, respectively. Deviations from Hardy-Weinberg equilibrium (HWE) before adjustment for number of tests ($p < 0.05$) were observed in two out of the 16 STR loci tested (MU23 and G10O). The differences between observed and expected heterozygosity at these two loci were 0.626 and 0.716 (MU23), and 0.402 and 0.473 (G10O). However, after Bonferroni correction of significance levels (16 tests, $p = 0.003$), there were no deviations from HWE in any of the 16 STR markers. The HWE test details for all markers is given in Supplementary File 7.

Tables 4 and 5 report other commonly used forensic efficiency measures including power of discrimination (PD), power of exclusion (PE), matching probability (PM), and typical paternity index (PI_{typ}). Eleven of the STR markers genotyped in the WE population [17] revealed a similar power of discrimination (PD) in the Kamchatka brown bear population with PD's ranging from 0.855 (MU51) to 0.962 (G10L) while MU10 revealed a somewhat lower PD of 0.797. The two new STR markers G10J and G10X showed PDs of 0.917 and 0.871, respectively, while G10C and G10O displayed somewhat less power with PDs of 0.794 and 0.667. G10O showed one allele that was very frequent (allele frequency 0.692) and there were only four different alleles. Together, this led to a low PD in this marker.

Tests for deviation from linkage equilibrium across loci showed linkage disequilibrium (LD) at a significance level of $p < 0.05$ in eight out of 120 pairwise tests applying the "preserve genotypes" option in GDA. After Bonferroni corrections of significance levels (120 tests, $p < 0.0004$) none of the pairwise comparisons remained significant. Results from STR loci showing significant deviations before Bonferroni correction of the significance level are given in Supplementary File 7.

The results from estimations of the inbreeding coefficient (F_{IS}) applying the 16 STR markers and the 115 individuals included in this

Table 6
Measurements of Fst (relative to WE populations) and Fis (within Kamchatka).

Number of loci included	F_{IS} (bootstrap 95 % CI)	F_{ST} (bootstrap 95 % CI)[c]
16 loci[a]	0.011 (-0.017–0.042)	
15 loci (G10X removed)[b]	0.011 (-0.018–0.043)	
12 loci		0.11 (0.09–0.12)

The FIS-value and 95 % bootstrap confidence interval was estimated for all 16 loci included in this study[a] and with G10X removed (15 loci)[b] applying GDA v.1.1

[c]The overall F_{ST}-value was estimated after comparing DNA profiles from the twelve STRs studied in eight populations (combined) reported in Andreassen et al. [17] applying GDA v.1.1

study are shown in Table 6. As locus G10X could need further investigation before being fully included as a validated marker (see Section 3.1), the F_{IS}-value was also estimated with G10X removed. The inbreeding coefficient F_{IS} was 0.011 (with or without G10X), but it was not significant (95 % bootstrap confidence interval from −0.018–0.043, Table 6).

The probability of identity (PI) and the probability of PI_{sib} was estimated at each locus and are given in Supplementary File 7. The power of the DNA profiling system to differentiate between individuals based on their DNA profiles, the total PI_{ave}, is given for different marker combinations in Table 7. The four STR marker combinations shown are: the eight markers in MP1, MP2 and MP3, the same twelve markers as validated in Andreassen et al., 2012 [17], fifteen STR markers (G10X removed) and all sixteen STR markers. Table 7 also shows the total Pi_{sibs} for same STR marker combinations. The magnitude of the total PI_{ave}s, even for the eight-locus combination (6.2×10^{-9}), indicated that these DNA-profile system combinations all resulted in individual-specific DNA-profiles. Full sibling pairs are the family constellation with the highest probability of having identical DNA-profiles (except monozygotic twins). Although the probability for this to happen is much higher than for unrelated individuals, the total PI_{sib} was still relatively low if applying eight STR markers (8.6×10^{-4}). Applying all validated fifteen STR markers to maximize the power of the DNA-profiling system showed a PI_{sib} of 6.0×10^{-6} indicating that siblings could be reliably distinguished.

3.4. Genetic distance measures between Kamchatka and the West-European populations

The eight populations and twelve markers studied in Andreassen et al., 2012 [17] were used when measuring population substructure (F_{ST}) between the Kamchatka and WE populations. The test for population substructure revealed an F_{ST}-value of 0.11 (95 % bootstrap confidence interval 0.09–0.12, Table 6).

The results from estimations of pairwise F_{ST} were all significant ($p < 0.002$), and the genetic distances from the Kamchatka population to the WE populations are shown in Table 8. The first column in the table lists the WE populations, from smallest to largest genetic distance (measured as F_{ST}), to the Kamchatka population ranging from 0.0837 (comparison to Karelia) to 0.1673 (comparison to Vesterbotten). A map (Fig. 4) illustrates the geographic distance between all populations. The genetic distance (pairwise F_{ST}) is also illustrated in a principal coordinates (PCoA) plot in Fig. 5. The PCoA plot showed that the genetically closest populations were Karelia, Kainuu and Pinega (P6, P7 and P8) that together formed a cluster in the PCoA plot (possibly including Finmark). The three south-west populations (P3, P4 and P5) formed another cluster that showed the largest genetic distance to Kamchatka.

Table 7
Results from estimations of the total average probability of identity with API-calc.

	Total average PI[a]	Total average PI_{sib}[b]
8 loci	6.2×10^{-9}	8.6×10^{-4}
12 loci	5.3×10^{-13}	1.6×10^{-5}
15 loci	1.4×10^{-14}	6.0×10^{-6}
16 loci	7.8×10^{-16}	9.8×10^{-7}

[a] The total average probability of identity when applying eight (MP1, MP2, MP3), twelve (previously validated in [17]), fifteen (G10X removed) or all sixteen STRs ($F_{ST} = 0.05$) typed in the Kamchatka brown bear population (N=115).

[b] The total average probability of full sibling identity when applying eight (MP1, MP2, MP3), twelve (previously validated in [17]), fifteen (G10X removed) or all sixteen STRs ($F_{ST} = 0.05$) typed in the Kamchatka brown bear population (N=115).

3.5. Application of the 16 STR markers for close kinship tests

The blind search functions of FAMILIAS [40] was applied to reveal pairs of individuals that were likely parent-cub or full siblings. There were 24 such pairs supported by a likelihood ratio of ≥ 100 (Table 9). Seventeen of these cases indicated parent-cub as the most likely relationship, and the search for trios of father-mother-cub supported there were two such trios among the 115 individuals (1–4, Table 9). Full siblings were the most likely relationship in seven cases. Five of these pairs (21–25, Table 9) showed exclusions of parent-cub relationships. Four of these exclusions were in one marker only, while one of these (25, Table 9) could alternatively be explained as a parent-cub pair where there was a *de novo* one-step mutation in the cub. All but one of the parent-cub pairs consisted of one female, and in most cases, the individuals in the 14 putative mother-cub pairs were sampled at the same hair trap or traps located next to each other (Table 9 and Fig. 1). This additional information agreed very well with the general knowledge and observations in the field; that females with young cubs move together in search of food. The tests demonstrated that the STR marker set can be very useful for identifying close family relationships. We do, however, note that differentiating between parent-cubs and full siblings solely based on genetic evidence can be challenging even with 16 STRs. As an example, individual Kam08 was identified in four parent-cub relationships (5–8, Table 9) which could point towards this individual being the mother of four cubs. This would, however, not agree with Kam013 and Kam015 (9, Table 9) being a parent-cub constellation. Alternatively, they may be full siblings, as the comparison of Kam013 and Kam015 being parent-cub *vs* full siblings showed an LR of only 3. Together with the other blind search results that supported they were cubs of Kam08 (6 and 7, Table 9) it seems equally likely they are full siblings. This demonstrated that even if the blind searches against unrelated results in high LRs supporting that a pair are e.g. parent-cub, an additional comparison between alternative close relationships (e.g. parent-cub *vs* full siblings) may reveal that the genetic data does not convincingly differ between such alternatives.

4. Discussion

4.1. Performance of the 16 STR markers in the Kamchatka population

The validation results of the twelve STR markers extensively used for forensic wild-life purposes, brown bear monitoring and conservation genetics in WE populations [17] showed that all performed similarly well in the Kamchatka population when applied for analysis of non-invasive source materials. Four new STR markers were also included in the validations. They revealed similar sensitivity as the other markers, reliably amplified in 0.2 ng DNA template concentrations. Species specificity was not tested as they are in MPs with other STRs that have been validated as bear specific. We assume that these new STRs, like the other 12 ones, will amplify in other bear species than brown bear. If applied for study of bears in areas with more than one bear species all samples should initially be tested with a species-specific bear test.

The measures of standard deviations indicated a high precision in allele sizing (largest SD less than 0.24 bp) that allowed all alleles to be reliably differentiated using the current DNA extraction, PCR-multiplex and electrophoretic conditions. No stutter ratio larger than 0.56 was observed when testing the new markers indicating that true alleles could be reliably distinguished from stutter peaks. The heterozygous peak height balances also indicated that alleles could be reliably discriminated from stutter peaks (range 1.2–3.0) in heterozygous samples with neighboring alleles. STR marker G10X did, however, show an allele-specific deviation in peak height balance. The suboptimal amplification of this allele (174 bp) was due to a small deletion polymorphism in the sequence targeted by the primer. This seems to be an ancient allele-specific polymorphism as it was seen in both the Kamchatka and the WE

Table 8
Results from estimations of pairwise F_{ST} between Kamchatka and WE populations.

	Kamchatka	Karelia	Kainuu	Pinega	Finnmark	Troms	Trøndelag	Hedmark
Karelia	**0.0837**							
Kainuu	**0.0892**	0.0088						
Pinega	**0.0924**	0.0459	0.0556					
Finnmark	**0.1067**	0.0477	0.0612	0.0720				
Troms	**0.1417**	0.1180	0.1304	0.1311	0.1031			
Trøndelag	**0.1622**	0.0977	0.1204	0.1043	0.0852	0.0990		
Hedmark	**0.1670**	0.1057	0.1267	0.1189	0.1073	0.1566	0.0516	
Vesterbotten	**0.1673**	0.0995	0.1211	0.0982	0.0918	0.1122	0.0131	0.0532

Finnmark, Norway (n = 74); Troms, Norway (n = 34); Trøndelag, Norway (n = 81); Hedmark, Norway (n = 101); Vesterbotten, Sweden (n = 84); Pinega, Russia (n = 26); Kainuu, Finland (n = 44); Karelia, Russia (n = 35); Kamchatka, Russia (n = 115). All estimates of pairwise FST were significant ($p < 0.002$).

Fig. 4. Map of Eurasia including populations from Norway, Sweden, Finland, and Russia. The sample locations of each of the eight brown bear populations (P1-P8) included in Andreassen et.al. [17] are shown to the left (west). P1 = Finnmark, Norway (n = 74), P2 = Troms, Norway (n = 34), P3 = Trøndelag, Norway (n = 81), P4 = Hedmark, Norway (n = 101), P5 = Vesterbotten, Sweden (n = 84), P6 = Pinega, Russia n = 27, P7 = Kainuu, Finland (n = 44), and P8 = Karelia, Russia (n = 35). The sample location of the Kamchatka population (n = 115) is shown to the right (east).

brown bear populations tested. Single PCR amplification of DNA from an individual with the 174/180 genotype applying a less stringent annealing temperature led to a large improvement in heterozygosity balance. However, markers considered for inclusion in a DNA-profiling system for wildlife genetics should be combined in multi-plex assays to provide sufficient genetic data from non-invasive sample extracts that usually have rather low DNA template concentrations. Although being a promising marker, design of new primers followed by validation of performance when used in a multi-plex assay is recommended before being fully included in the DNA profiling system.

4.2. Genotyping success rate, individual profiles and mixtures in hair samples

The average success rate (66 %) when analyzing hair samples was similar to what has been reported for non-invasive samples from WE brown bear populations (65–70 %) [53,54]. Analysing these materials we did, however, note that there was a substantially increased success rate when more hairs were extracted (Fig. 3). This suggests that, if possible, the number of hairs collected and extracted should preferentially be in the range of 4–10. A larger number of hairs included in a sample extraction did result in a small increase (4.5 %) in cases with

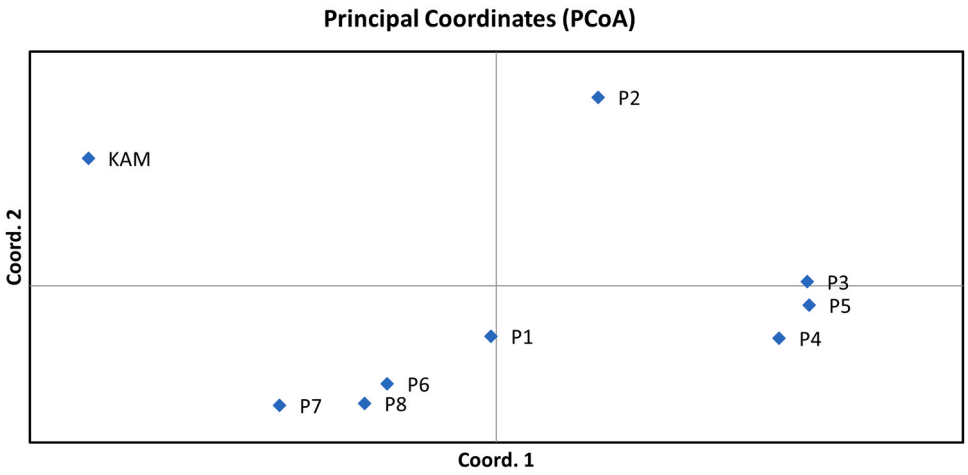

Fig. 5. A principal coordinates plot (PCoA plot) of the genetic distance between Kamchatka (KAM) and the eight West-European brown bear populations sorted based on their pairwise F_{ST}-values (GenAlEx v. 6.501): Kainuu (P7), Karelia (P8), Pinega (P6), Finnmark (P1), Troms (P2), Trøndelag (P3), Vesterbotten (P5) and Hedmark (P4).

Table 9
Results from relationship tests.

No	Location[a]	Individual 1[b]	Individual 2[b]	Relationship	LR	Alt. Relationship[c]	LR alt rel.[c]
1	S4, S3	KamO14, M	KamO25, F	Father-Daughter	1.422	Full Siblings	-
2	S4	KamO24, F	KamO25, F	Mother-Daughter	1.496	Full Siblings	400
3	S5, S3	KamO18, M	KamO107, F	Father-Daughter	29.348	Full Siblings	2.773
4	S5	KamO43, F	KamO107, F	Mother-Daughter	177	Full Siblings	-
5	S2	KamO8, F	KamO5, F	Parent-Cub	604	Full Siblings	-
6	S2, S3	KamO8, F	KamO13, F	Parent-Cub	11.608	Full Siblings	1.14
7	S2, S3	KamO8, F	KamO15, F	Parent-Cub	692	Full Siblings	-
8	S2, S7	KamO8, F	KamO50, M	Parent-Cub	109	Full Siblings	-
9	S3	KamO13, F	KamO15, F	Parent-Cub	4.604	Full Siblings	1.62
10	S12, S3 and S4	KamO84, F	KamO10, M	Parent-Cub	7.086	Full Siblings	287
11	S3	KamO12, F	KamO17, M	Parent-Cub	12.287	Full Siblings	2.547
12	S6	KamO47, F	KamO48, F	Parent-Cub	64.58	Full Siblings	12.572
13	S9, S15	KamO60, F	KamO110, F	Parent-Cub	711	Full Siblings	-
14	S11, S16	KamO80, M	KamO105, M	Parent-Cub	17.495	Full Siblings	11.791
15	S16	KamO102, F	KamO103, M	Parent-Cub	1.916	Full Siblings	1.476
16	S6, S2	KamO47, F	KamO6, F	Parent-Cub	186	Full Siblings	-
17	S15, S1	KamO110, F	KamO111, M	Parent-Cub	124	Full Siblings	-
18	S16	KamO99, F	KamO100, M	Full Siblings	198.13	Parent-Cub	44.844
19	S4, S5	KamO22, M	KamO34, M	Full Siblings	113	Parent-Cub	-
20	S9, S14	KamO60, F	KamO91, F	Full Siblings	3.222	paternity excluded	
21	S16	KamO88, M	KamO101, M	Full Siblings	6.695	paternity excluded	
22	S1, S5	KamO1, F	KamO41, F	Full Siblings	679	paternity excluded	
23	S12, S16	KamO86, F	KamO101, M	Full Siblings	456	paternity excluded	
24	S19, S11	KamO64, F	KamO77, M	Full Siblings	382	paternity excluded	

[a] Location as shown in Fig. 1 and Table 1. Individuals collected at multiple sampling sites have locations separated by a comma.
[b] Individual genotypes given in supplementary file 6. F denotes female, M denotes male.
[c] The alternative close relationship with LR indicated if > 100.

mixtures from two individuals (Fig. 3). The heterozygosity level of the STR markers and the criterion that results from at least eight STR markers is necessary for accepting an individual specific DNA-profile would, however, assure that any sample with DNA from two or more individuals would be spotted as a mixture as it would have three or more alleles in one or more loci. The use of at least eight STR markers as a threshold for generating an individual specific DNA-profile does, thus, ensure that any mixed sample would be identified. However, with the large stutter ratios and heterozygote balance ratios in the dinucleotide STRs validated it would be challenging to interpret the individual DNA-profiles in such mixtures. Instead, these should simply be identified as mixed samples and removed. Nevertheless, the relatively low risk of combining hairs from two individuals when DNA extraction is made from >3 hairs seem to be an acceptable trade-off to the 30 % gain in individual DNA-profiles retrieved when including >3 hairs.

4.3. A DNA profiling system providing individual-specific profiles of Kamchatka brown bears

Allele frequencies and population data were obtained from the 115 individuals collected from a relatively small geographic sampling area (approximately 200 square km, Fig. 1). However, the brown bear population in the Kamchatka area (about 270,000 km^2) is large (>22,000 bears) [28]. The individuals included in this study, thus, represent only a small proportion of the Kamchatka population. However, brown bears may travel long distances to the sampling area at this time of the year as they return annually to feed from the large number of salmon that enter the rivers and lakes to spawn [31]. Therefore, the population sampled, and the number of individuals identified likely provides representative population data from Kamchatka brown bears that is sufficient to provide common forensic power estimates [55,56]. Including more remotely sampled individuals could result in a higher heterozygosity

than observed in our sample. If so, one would achieve an even higher discriminating power (PD) when applying the DNA-profiling system validated here. The level of population substructure (theta) is typically substantially higher in wildlife species than in human populations. Here, we used a conservative estimate ($F_{ST} = 0.05$) as a replacement for an empirically calculated value from a geographically closer population. We found no significant inbreeding, there were no large deviations from HWE, and none of the loci pairwise comparisons revealed LD. Accounting for population substructure, even full siblings could be reliably differentiated (PI_{sib} 9.8×10^{-7}). Given these estimates, we conclude that the validated DNA profiling system has the power to provide individual-specific identification of Kamchatka brown bears. Applied in the courtroom, a match between DNA profiles from trace evidence (e.g. blood stains, hair, skin or bone) and an illegally killed bear would be very strong evidence. Typically, one would report this match as more than a million times more likely that the trace evidence originates from the illegally hunted individual than being from another individual. Whether the court accepts such evidence also depends on proper validations, as conducted in this study, so that the prosecution can confirm that the samples were analyzed according to internationally accepted forensic standards.

4.4. Genetic diversity and genetic distance

The heterozygosity averaged across all loci is a parameter commonly used to describe the genetic diversity in brown bear populations [35]. Our findings from the Kamchatka population revealed little difference in the expected (0.79) and observed heterozygosity (0.78) averaged over the same twelve STR loci as genotyped in the WE brown bear population [17]. The same STR loci were genotyped in the geographically distant WE populations in materials that had approximately the same sample sizes. This allowed for direct comparisons of genetic diversity. This comparison revealed a slightly higher genetic diversity in the Kamchatka population compared to Norwegian and Swedish brown bear populations (average heterozygosity levels: 0.69–0.75, P2 -P5 in [17]). This is not an unexpected observation in a large and dense population of bears as that of Kamchatka. Also, several new private alleles were detected in the Kamchatka population that have never been registered before. This finding is easily explained by the large geographic distance to the WE populations. Another interesting finding was that the 96 bp allele was among one of the most frequent alleles at locus MU15 in Kamchatka while it had only been observed once before in a WE population (a single allele, registered in the NIBIO Svanhovd brown bear genetic database). Alleles with micro size variation (1 bp difference to other alleles) were observed in low frequencies at locus MU23 in some of the WE populations [17]. However, these microvariant alleles were not observed at all in the Kamchatka population. These findings might suggest that the microvariant alleles at MU23 result from relatively recent mutation(s) at this locus in the WE populations, and that the alleles with shared sizes in all populations (including Kamchatka) are the ancestral alleles for this STR marker.

This is the first time the genetic distance (estimated as pairwise F_{ST}`s) of the Kamchatka brown bear population to the eight WE populations [17] was assessed. In general, these genetic distance measurements (Fig. 5) mirrored the geographic distances of the WE populations to Kamchatka.

4.5. Kinship testing applying the 15 validated STR markers

The kinship testing showed that the STR markers can be powerful means for revealing close family relationships. Notably, the use of $F_{ST} = 0.05$ to correct for population substructure had a large impact on the calculated LRs (about a tenfold reduction of LR's). Although it is correct to adjust for population substructure, the LR threshold used here may be too conservative (LR $\geq$ 100). Especially, if a certain family relationship is already supported by non-genetic information (e.g. location and timepoint where/when samples were collected, observations in the field) one could argue that a less strict threshold could be used as the genetic evidence is compensated by the weight of non-genetic information.

Using genetic data as the only source for identifying parent-cub relationships makes it impossible to determine which one is the parent. It is also challenging to properly distinguish between alternative close relationships like parent-cub *vs* full siblings even with results from sixteen STRs. Again, non-genetic information like observations from the field would be important and may help to reliably solve such issues. During the time of the year when the hair samples were collected (summer) a mother is together with her cub(s), but a cub is not likely to be together with its father. Except from the putative father-daughter relationships predicted from the trio-blindsearch, there were five cases with one female and one male individual in the parent-cub pairs (Table 9). Based on the additional non-genetic information one could assume that these cases likely were a mother with a male cub. This exemplifies that the addition of information from field and non-genetic sources could help when aiming to make reliable and detailed kinship assumptions in the population studied.

5. Conclusion

In conclusion, this is the first time non-invasively collected hair samples from brown bears of the Kamchatka population have been analyzed with the autosomal STR markers commonly used in WE populations. The validations of marker performance and the forensic efficiency parameters calculated show that the DNA profiling system provides individual-specific DNA profiles. Furthermore, the individual DNA profiles may also be applied to support or reject close family relationships providing strong evidence that indviduals are closely related based solely on the use of the genetic data. Finally, the DNA profiling system will be valuable for population genetic studies and may be applied for monitoring the Kamchatka brown bear population for conservation purposes.

Ethical statement

No human samples were investigated. No animal experiments were carried out in this study.

CRediT authorship contribution statement

Ida Fløystad: Writing – review & editing, Methodology. **Cornelya F. C. Klütsch:** Writing – review & editing, Visualization, Formal analysis. **Daniel Kling:** Writing – review & editing, Software. **Vladimir Zhakov:** Writing – review & editing, Methodology, Investigation. **Liya Pokrovskaya:** Writing – review & editing, Methodology, Investigation. **Berit K Hansen:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Rune Andreassen:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Formal analysis, Data curation, Conceptualization. **Snorre B. Hagen:** Writing – review & editing, Software, Resources, Funding acquisition, Conceptualization. **Hans G. Eiken:** Writing – review & editing, Visualization, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Siv Grete Aarnes for skilful supervision and technical assistance in the laboratory. We would like to thank Finn-Arne

Haugen for providing Fig. 4. The work is financed in part by research funding from NIBIO.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsiae.2024.100092.

References

[1] N.J. Ouborg, et al., Conservation genetics in transition to conservation genomics, Trends Genet. 26 (4) (2010) 177–187.
[2] Y. Willi, et al., Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species, Proc. Natl. Acad. Sci. USA 119 (1) (2022) e2105076119.
[3] L.P. Waits, D. Paetkau, Noninvasive genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection, J. Wildl. Manag. 69 (4) (2005) 1419–1433.
[4] I. Baciu, A. Fedorca, G. Ionescu, Noninvasive genetics knowledge from the brown bear populations to assist biodiversity conservation, Divers. -Basel 14 (2) (2022) 121.
[5] M.F. Proctor, et al., Review of field methods for monitoring Asian bears, Glob. Ecol. Conserv. 35 (2022) e02080.
[6] P.M. Lukacs, K.P. Burnham, Review of capture-recapture methods applicable to noninvasive genetic sampling, Mol. Ecol. 14 (13) (2005) 3909–3919.
[7] M.K. Schwartz, G. Luikart, R.S. Waples, Genetic monitoring as a promising tool for conservation and management, Trends Ecol. Evol. 22 (1) (2007) 25–33.
[8] A.A. Karamanlidis, A. Kopatz, M.D. Hernando, Dispersal patterns of a recovering brown bear (Ursus arctos) population in a human-dominated landscape, J. Mammal. 102 (2) (2021) 494–503.
[9] A. Linacre, et al., ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations, Forensic Sci. Int. -Genet. 5 (5) (2011) 501–505.
[10] E.P. Meredith, J.K. Adkins, J.A. Rodzen, UrsaPlex: An STR multiplex for forensic identification of North American black bear (Ursus americanus), Forensic Sci. Int. -Genet. 44 (2020) 102161.
[11] O. Kleven, et al., Identification and evaluation of novel di- and tetranucleotide microsatellite markers from the brown bear (Ursus arctos), Conserv. Genet. Resour. 4 (3) (2012) 737–741.
[12] P. Taberlet, et al., Noninvasive genetic tracking of the endangered Pyrenean brown bear population, Mol. Ecol. 6 (9) (1997) 869–876.
[13] K.L. Dicks, et al., Validation studies on dinucleotide STRs for forensic identification of black rhinoceros Diceros bicornis, Forensic Sci. Int. -Genet. 26 (2017) E25–E27.
[14] N. Dawnay, et al., A forensic STR profiling system for the Eurasian badger: A framework for developing profiling systems for wildlife species, Forensic Sci. Int. -Genet. 2 (1) (2008) 47–53.
[15] D. Paetkau, C. Strobeck, Microsatellite Analysis of Genetic-Variation in Black Bear Populations, Mol. Ecol. 3 (5) (1994) 489–495.
[16] L. Waits, et al., Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (Ursus arctos), Mol. Ecol. 9 (4) (2000) 421–431.
[17] R. Andreassen, et al., A forensic DNA profiling system for Northern European brown bears (Ursus arctos), Forensic Sci. Int. -Genet. 6 (6) (2012) 798–809.
[18] Kopatz, A., et al., *Female brown bears in Sør-Varanger, Norway: localities and mother-cub relationships analyzed by genetic methods, in Bioforsk Report*. 2014, NIBIO. p. 25.
[19] Kopatz, A., et al., *Family groups of brown bears in Sør-Varanger*, in *NIBIO reports*. 2017, NIBIO. p. 39.
[20] A. Kopatz, et al., Admixture and Gene Flow from Russia in the Recovering Northern European Brown Bear (Ursus arctos), Plos One 9 (5) (2014) e97558.
[21] J. Schregel, et al., Sex-specific genetic analysis indicates low correlation between demographic and genetic connectivity in the Scandinavian brown bear (Ursus arctos), Plos One 12 (7) (2017) e0180701.
[22] Eiken, H.G., et al., *Population data for 12 STR loci in Northern European brown bear (Ursus arctos) and application of DNA profiles for forensic casework*. Forensic Science International: Genetics Supplement Series., 2009. 2: p. 273-4.
[23] McLellan, B.N., et al., *Ursus arctos (amended version of 2017 assessment). The IUCN Red List of Threatened Species 2017: e.T41688A121229971*. 2017.
[24] J.E. Swenson, et al., The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden, Wildl. Biol. 1 (1) (1995) 11–25.
[25] U. Breitenmoser, Large predators in the Alps: The fall and rise of man's competitors, Biol. Conserv. 83 (3) (1998) 279–289.
[26] J. Kindberg, et al., Estimating population size and trends of the Swedish brown bear Ursus arctos population, Wildl. Biol. 17 (2) (2011) 114–123.
[27] Commission, E. *Conservation status of large carnivores* 2012; Available from: http://ec.europa.eu/environment/nature/conservation/species/carnivores/conservation_status.htm.
[28] I.A. Revenko, Status of Brown Bears in Kamchatka, Russian Far East, Ursus 10 (1998) 11–16.
[29] Pokrovskaya, L., V.V. Zhakov, and I. Ivan Pokrovsky, *Behavioral ecology and genetics of Kamchatka brown bear (Ursus arctos piscator)*. International Bear News Fall 2016, 2016. 25(3): p. 17-19.
[30] S.A. Kolchin, et al., Consequences of a sockeye salmon shortage for the brown bear in the basin of Lake Kurilskoe, Southern Kamchatka, Nat. Conserv. Res. 6 (2021).
[31] I.V. Seryodkin, et al., Home ranges of brown bears on the Kamchatka peninsula and Sakhalin Island, Contemp. Probl. Ecol. 10 (6) (2017) 599–611.
[32] R. Lorenzini, et al., Wildlife molecular forensics: identification of the Sardinian mouflon using STR profiling and the Bayesian assignment test, Forensic Sci. Int Genet 5 (4) (2011) 345–349.
[33] F. Barbanera, et al., Conservation of endemic and threatened wildlife: molecular forensic DNA against poaching of the Cypriot mouflon (Ovis orientalis ophion, Bovidae), Forensic Sci. Int Genet 6 (5) (2012) 671–675.
[34] C. Frosch, et al., Case report of a fatal bear attack documented by forensic wildlife genetics, Forensic Sci. Int Genet 5 (4) (2011) 342–344.
[35] J.E. Swenson, P. Taberlet, E. Bellemain, Genetics and conservation of European brown bears Ursus arctos, Mammal. Rev. 41 (2) (2011) 87–98.
[36] M. Poetsch, et al., The new guidelines for paternity analysis in Germany-how many STR loci are necessary when investigating duo cases? Int. J. Leg. Med. 127 (4) (2013) 731–734.
[37] A. Kopatz, et al., Connectivity and population subdivision at the fringe of a large brown bear (Ursus arctos) population in North Western Europe, Conserv Genet 13 (2012) 681–692.
[38] K. Yamamoto, et al., Sex identification of Japanese black bear, Ursus thibetanus japonicus, by PCR based on amelogenin gene, J. Vet. Med. Sci. 64 (6) (2002) 505–508.
[39] M.J. Brownstein, J.D. Carpten, J.R. Smith, Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping, Biotechniques 20 (6) (1996) 1004.
[40] T.R. Moretti, et al., Validation of STR typing by capillary electrophoresis, J. Forensic Sci. 46 (3) (2001) 661–676.
[41] H. Wenz, et al., High-precision genotyping by denaturing capillary electrophoresis, Genome Res. 8 (1) (1998) 69–80.
[42] P. Gill, R. Sparkes, C. Kimpton, Development of guidelines to designate alleles using an STR multiplex system, Forensic Sci. Int. 89 (3) (1997) 185–197.
[43] D. Shinde, et al., Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites, Nucleic Acids Res. 31 (3) (2003) 974–980.
[44] C. Brookes, et al., Characterising stutter in forensic STR multiplexes, Forensic Sci. Int. -Genet. 6 (1) (2012) 58–63.
[45] J.M. Butler, STR Alleles and Amplification Artifacts. in *Advanced Topics in forensic DNA typing: interpretation*, Elsevier, San Diego, 2015, pp. 47–86.
[46] Lewis, P.O. and D. Zaykin, *Genetic Data Analysis: Computer program for the analysis of allelic data*. 2001.
[47] R. Peakall, P.E. Smouse, GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research, Mol. Ecol. Notes 6 (1) (2006) 288–295.
[48] R. Peakall, P.E. Smouse, GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update, Bioinformatics 28 (19) (2012) 2537–2539.
[49] Promega, *PowerStats*.
[50] K.L. Ayres, A.D.J. Overall, API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives, Mol. Ecol. Notes 4 (2) (2004) 315–318.
[51] T. Egeland, et al., Beyond traditional paternity and identification cases Selecting the most probable pedigree, Forensic Sci. Int. 110 (1) (2000) 47–59.
[52] D. Kling, A.O. Tillmar, T. Egeland, Familias 3-Extensions and new functionality, Forensic Sci. Int. -Genet. 13 (2014) 121–127.
[53] Smith, M.E., et al., *Hair snares applied to detect brown bears in the vicinity of farms in the Pasvik Valley 2008*, in Bioforsk Report 169 2008, NIBIO.
[54] T. Skrbinsek, et al., Highly efficient multiplex PCR of noninvasive DNA does not require pre-amplification, Mol. Ecol. Resour. 10 (3) (2010) 495–501.
[55] R. Chakraborty, Sample-Size Requirements for Addressing the Population Genetic Issues of Forensic Use of DNA Typing, Hum. Biol. 64 (2) (1992) 141–159.
[56] J.M. Butler, STR Population Data Analysis. in *Advanced Topics in forensic DNA typing: interpretation*, Elsevier, San Diego, 2015, pp. 239–279.