Høgskolen i Telemark

År: 2009

Review of microsatellites in ecology and parentage analysis of Norwegian beavers (*Castor fiber*) based on cross-specific microsatellites

Imke Jürgens

[Illustrasjon av bever er kun tilgjengelig i den trykte utgaven]

Master Thesis

Telemark University College

Avdeling for allmennvitenskapelige fag
Hallvard Eikas plass, 3800 Bø

Høgskolen i Telemark

Tittel:	**Review of microsatellites in ecology and parentage analysis of Norwegian beavers (*Castor fiber*) based on cross-specific microsatellites**
Nøkkelord:	*Castor fiber*, microsatellites, cross-specific,hair, parentage analysis, nuclear DNA
Forfatter	Imke Jürgens
Studentnr.:	023452
Fagkode:	4317
Oppgavetype:	Mastergradsoppgave
Studiepoeng:	60
Studium:	Mastergrad i Natur-, helse- og miljøvern
Konfidensiell:	Nei

Avdeling for allmennvitenskapelige fag
Hallvard Eikas plass, 3800 Bø

Preface

This Master's thesis is part of the Master's degree programme at the Department of Environmental and Health Studies at the Telemark University College in Bø, Norway. The work presented in this study has been carried out in the period of 2006 to 2009, and is based on the molecular genetic analyses conducted at the laboratories of the Department of Environmental and Health Studies at the Telemark University College in Bø, Telemark.

I want to thank Professor Elin Kure for inspiring me to conduct studies in molecular genetics and my supervisors Frank Rosell and Mona Sæbø for their great support and technical guidance throughout the process of writing this thesis.

I also want to thank Ruairidh D. Campbell for his advice and humour, Jan Heggenes for technical and personal feedback, Karin Brekke Li for her patience and spirit at the lab in Bø, and my friends and study colleagues for their belief in me.

Seljord, 2 November 2009

Imke Jürgens

Contents

1 Abstract

The Eurasian Beaver (*Castor fiber*) in Scandinavia was once close to extinction due to a bottleneck resulting from vigorous overhunting in the middle of the 19th century. After measures of protection, the population number increased again until it reached a level that brought it into the focus of ecological research and monitoring. Molecular ecology has advanced rapidly over the last decades, and molecule markers like microsatellites have become a powerful tool for such genetic population analyses and monitoring.

In this study a set of nine microsatellite pairs designed for the North American beaver (*Castor canadensis*) is used to analyze parentage in a Eurasian beaver population in Telemark, Norway. To this purpose, DNA was extracted from 148 hair samples. Microsatellite genotypes at five variable loci were obtained from 136 of the 148 extracted DNA samples. Based on previously collected observation data, 50 parent-offspring relationships were constructed from the 136 genotyped individuals. By using the parentage exclusion method, there were found 16 cases of mismatching genotypes between an offspring and a parent or parent pair. These results suggest that extra-pair mating might have occurred in the genotyped beaver population. However, the acceptance of the exclusions proved difficult since 13 out of 16 exclusions were based on homozygous genotypes. The derivation from Hardy Weinberg equilibrium and the low heterozygosity of the microsatellite loci used in this study indicated the presence of null alleles which might have lead to false homozygotes suggesting that further genetic analyses might be necessary for successful parentage analysis.

2 Introduction

The Eurasian beaver (*Castor fiber*) is a semi-aquatic nocturnal rodent living near rivers, creeks, swamps and lakes down from the sea level up to over 100 meters above sea level. Famous for its engineering and ability to shape the landscape according to its demands, the beaver is capable of transforming dry wooden areas into wetlands by constructing dams that create a stable water level. The beaver is said to be monogamous and lives in family groups consisting of the adult parents, yearlings and kids, or as single adult individuals or pairs (Rosell, Pedersen, 1999).

Previously the beaver was found throughout wide parts of Asia and Europe until overhunting and habitat loss almost led to its extinction by the middle of the 19th century, with a remaining 1200 individuals spread in eight small populations in Europe (Nolet, Rosell, 1998). In Norway, a population of c.100 individuals survived the bottleneck in the south-east region of the country, representing the founder population for the current 70 000 animals in Norway today (Parker, Rosell, 2003).

Typically, genetic diversity is lost as a consequence of short periods at small population sizes (Frankham *et al.*, 2002). The Swedish beaver population, consisting of animals reintroduced from Norway, shows indeed a paucity of genetic variability (Ellegren *et al.*, 1993).

In Norway, research on the beaver has been conducted since 1992 comprising management measures and behaviour ecology studies (H. Parker, pers. comm.). Research on behavioural ecology included amongst others observation and mark and recapture. In the course of individual marking, hair samples were taken for further investigation and analyses (F. Rosell pers. comm.). The non-invasive sampling of animal populations has become of prime importance in conservation genetics and behavioural ecology, as only thirty or forty years earlier, more destructive sampling techniques were used where animals often had been killed for the scientific study. In order to avoid such destructive approaches, scientists began to explore the applicability of tissue that could be collected non-destructively (Beebee, Rowe, 2004). With the discovery of the Polymerase Chain Reaction (PCR) (Saiki *et al.*, 1985) it became possible to quantify DNA from even minute amount of sample DNA as it usually is the case of non-invasively collected tissue samples (Taberlet, Luikart, 1999b). Microsatellites are genetic markers consisting of repeated DNA sequences that are highly variable due to frequent mutations between generations (Hancock, 1999). They have been widely used in the analysis of non-invasive genetic samples collected in the field to answer questions about genetic diversity in populations (Ferrando *et al.*, 2007; Waits *et al.*, 2000), population

structure (Natoli *et al.*, 2004), relatedness structure (Burland *et al.*, 2002; Taylor *et al.*, 1997), mating systems (Goossens *et al.*, 1998a), migration patterns (Bergl, Vigilant, 2007), gender determination (Taberlet *et al.*, 1997; Takami *et al.*, 1998), hybridization (Schwartz *et al.*, 2004) and parentage (Constable *et al.*, 2001; Lilland *et al.*, 2001; Schnabel *et al.*, 2000).
The goal of this study was to investigate parentage in the Telemark beaver population by using microsatellites as a genetic marker. A study about mating and kinship by Crawford et al (2008b) revealed the occurrence of extra-pair mating in a North American beaver (*Castor canadensis*) population based on microsatellite analyses. The testing of already existing non-specific microsatellites is a convenient and less expensive and time-consuming way of providing genetic information prior to the development of new primers (Selkoe, Toonen, 2006a). The goal of this study was to use these cross-specific microsatellites the analysis of parentage in the Telemark beaver population. By using a population genetic method, information about the kinship amongst colonies in this population would be obtained as an important supplementation to observation data. Another aspect of interest addressed in this study was to test for the occurrence of extra-pair mating in the beaver by performing parental exclusion based on genetic incompatibility. The applicability of cross-specific microsatellites in the analysis of parentage in the beaver in Norway was also explored. By using hair as DNA source valuable experience was gathered about the suitability of non-invasive sampling methods in the population genetic research of the beaver in Norway.

3 Theory

3.1 Microsatellites

3.1.1 General character of microsatellites

Microsatellites are short, tandemly repeated sequences of DNA (also called short tandem repeats, STRs, (King *et al.*, 2006) or simple sequence repeats, SSRs (Frankham *et al.*, 2002), and are found in the genome of most species studied so far. They are found evenly distributed along the chromosome in a eukaryotic organism, mostly within the non-coding sequences of the DNA (King *et al.*, 2006).

There is not quite consensus about the number of base pairs defining a microsatellite, but the definition most used in the literature reviewed in this study, is that a microsatellite consists of tandemly repeated units of 1-6 base pairs. Another definition by Chambers and MacAvoy (2000) proposes the minimum repeat unit length of two base pairs, represented by the general formula $(N_1N_2N_3\ldots N_x)_n$, with x being the repeat unit size within the range 2-6 and the number of repeat units n having a lower limit of $x*n > 8$ nucleotides. In this study, the minimum number of base pairs defining a microsatellite repeat unit is set to one. Typically, a microsatellite locus varies from 5 and 40 nucleotide repeats in length, although longer sequences are also possible (Selkoe, Toonen, 2006b).

Microsatellites have been found in any eukaryotic and prokaryotic organism so far analyzed. Nevertheless, the distribution of these sequences seems to vary among different taxonomic groups. As for vertebrate organisms, most of the repetitive sequences consists of dinucleotide repeats (34%), followed by mononucleotide and trinucleotide repeats with respectively 20% and 15% of all repeat units in the genome, compare Figure 1 (Toth *et al.*, 2000).

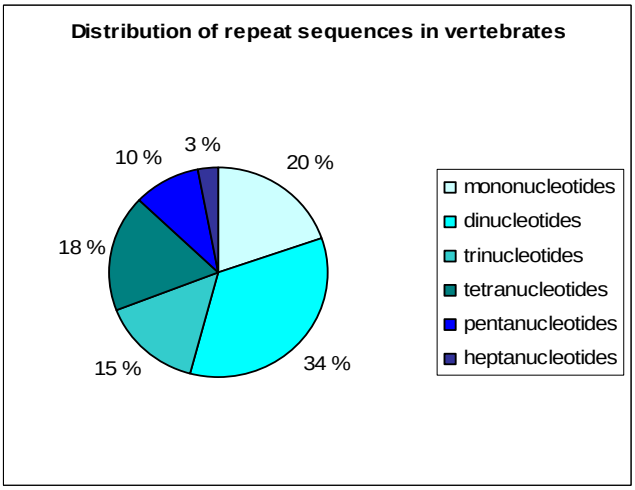

Figure 1 The average distribution of repeat sequences in a group of vertebrate organisms. The percentages are based on numbers from Toth et al. (2000)

The distribution of the repeat types in mammals (Figure 2)is quite similar to the one for vertebrates in general apart from trinucleotide repeats, which only account for an 8% of the total repeat number in the mammal genome, compared to 15% in vertebrates (Toth *et al.*, 2000).

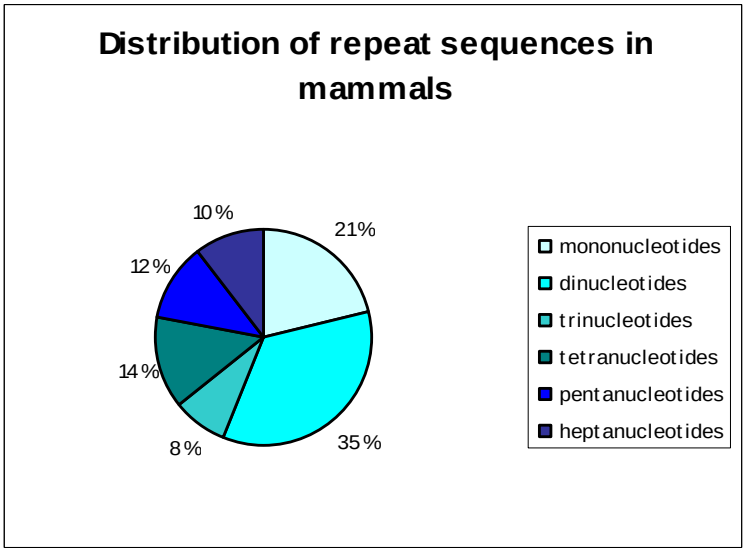

Figure 2 The average percentage distribution of repeat sequences in a group of mammals. Numbers are based on observations from Toth et al. 0(2000)

In rodents, the dinucleotide repeats provide by far the largest part of the repetitive sequences in the genome, consisting of 40%. Mononucleotides however, are far less abundant in the rodent genome compared to vertebrates and mammals with only 13% of all repeat sequences (Figure 3).

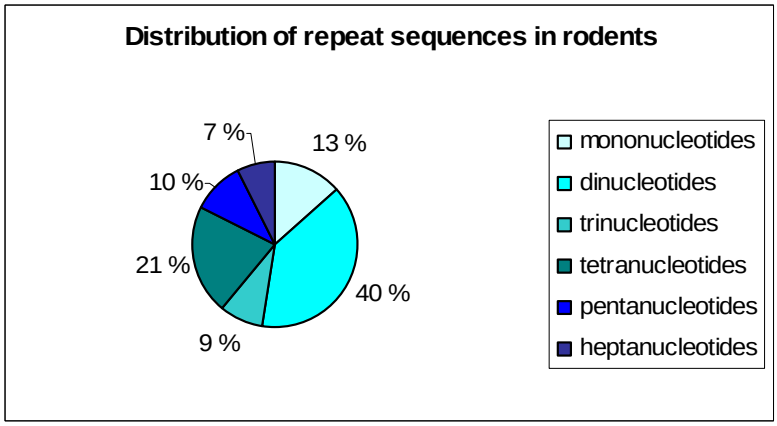

Figure 3 The distribution of repeat sequences in rodents. The percentages are based on numbers from Toth et al. (2000).

From all possible microsatellite types, dinucleotide, trinucleotide and tetranucleotide repeats are the most common choice for molecular genetic studies (Selkoe, Toonen, 2006b).

Microsatellite sequences can be classified into families according to their nucleotide composition:

a) Perfect microsatellites, consisting of a single repeat motif with no interruptions
b) Imperfect microsatellites, in which one ore more repeat unit carry a base pair that does not fit in the repeat structure
c) Interrupted microsatellites, containing a small number of basepairs that do not fit the repeat structure
d) Compound or composite microsatellites, consisting of two ore more adjacent microsatellites with different repeat types (Goldstein, Schlötterer, 1999)

Examples of microsatellites from four different families are shown in Table 1.

Table 1 Different families of microsatellites, after the terminology of Goldstein and Schlötterer (1999)

Microsatellite type	Example
Perfect	CACACACACACACACACACACA
Imperfect	CACACACACA**G**ACACACACACACA
Interrupted	CACACACACACA**GGG**CACACACAC
Compound or composite	CACACACACACA**GATGATGATGAT**

The DNA on either side of the microsatellite is termed the flanking region, and is usually conserved across individuals of the same or different species. It is on these sequences that the

primers will bind to under the polymerase chain reaction (PCR) amplification of the microsatellite locus (Selkoe, Toonen, 2006b).

3.1.2 Microsatellite evolution

The mechanism by which microsatellites expand or contract is predominantly slipped-strand mispairing (Levinson, Gutman, 1987). Strand slippage occurs under DNA synthesis, when one nucleotide strand forms a loop (Figure 4). When the looped out nucleotides are on the newly synthesized strand, an insertion will be the result. Under the next round of replication, the insertion will be copied, and a longer DNA strand will be produced. If the looped-out nucleotides are on the template strand, the newly replicated strand will get shorter, since the looped-out nucleotides are not available for base pairing (Pierce, 2003).

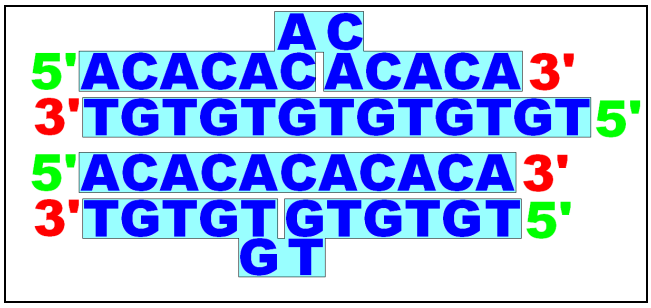

Figure 4 In slipped-strand mispairing, nucleotide loops lead to the insertion of a sequence when the loop is on the newly synthesized strand, while a loop on the template strand leads to the loss of nucleotides.

Recombination might also play a functional role in the altering of length of microsatellites, especially for those sequences that consists of long tandem repeats (Hancock, 1999). Whereas crossing over is a process in which genetic information is exchanged between homologous chromosomes in order to create genetic recombination (King *et al.*, 2006), the unequal crossing over involves crossing over between misaligned chromosome strands and results in one DNA molecule with an insertion and one molecule with a deletion (Pierce, 2003). However, it seems that recombination is not the predominant mechanism that generates microsatellite variability, as strains of *Escherichia coli* without a functional recombination systems have shown a similar microsatellite mutation rate compared to strains where recombination occurred. Also, microsatellite mutations usually consists of gains or losses of single repeat units, whereas mutations due to recombination lead to a wider range of novel mutants (Levinson, Gutman, 1987). In eukaryotic organisms, errors due to the process of slipped-strand mispairing are usually higher in repetitive sequences like microsatellites

compared to non-repetitive DNA, and the error rate increases with the growing length of the repetitive sequence (McCulloch, Kunkel, 2008). Not only the very length of the repetitive DNA sequence, but also the number of repeat units in a microsatellite is related to the rate at which strand-slippage mutations occur (Eisen, 1999) and a larger number of repeat units will increase the rate of microsatellite mutations (Wierdl *et al.*, 1997).

Despite of high mutation rates in repetitive sequences, most of the errors that occur under DNA synthesis are immediately corrected by DNA repair systems. While the DNA mismatch repair system removes loops formed under strand-slippage (Pierce, 2003), the exonucleolytic proofreading process is less effective in repetitive DNA and may explain the higher rate of mutation rates in these sequences (Eisen, 1999), which may be $10^{-2} – 10^{-6}$ events per locus per generation compared to the more stabile DNA sequences in coding DNA (Li *et al.*, 2002). Due of the instability that is created by the occurrence of these mutations and the partly correction of theses errors, microsatellites can become multi-allelic in a population and bi-allelic in an individual. Microsatellites are inherited co-dominantly from parent to offspring following Mendelian inheritance, and when analysed, they can show both homozygous and heterozygous genotypes in an individual. Even the use of only a small number of microsatellite loci can reveal genotypic differences in a large number of individuals in a population, at least when the microsatellites are highly variable. In a population genetic analysis using even a relatively small microsatellites each individual will show a unique multilocus genotype when the microsatellites are highly variable (Wan *et al.*, 2004).

Usually, the mutation rate of a particular microsatellite locus is not known and can only be assumed based as an approximate average (Balloux, Lugon-Moulin, 2002). Generally, the mutation rate is quite variable. Pure or compound microsatellites have a higher mutation rate than interrupted microsatellites (Petes *et al.*, 1997). The mutation rate also differs both between larger taxonomic groups, between closely related species (Crawford, Cuthbertson, 1996) and even between loci and alleles in the same species (Ellegren, 2000). It has also been shown that for humans, male individuals possess a higher microsatellite mutation rate than females do (Xu *et al.*, 2000).

Two theoretical models of microsatellite mutation

In the traditional stepwise mutation model (SSM), a mutation step consists of the addition or removal of one microsatellite repeat unit at a fixed rate (Ellegren, 2004). In this model, originally proposed as early as in 1973 by Ohta and Kimura (Ohta, Kimura, 1973), it is assumed that the number of repeat units increases or decreases independently of the length of

the microsatellite, with the smallest number of repeat units being one (Calabrese, Sainudiin, 2005). However, several studies have shown that this is not always the case. Mutations in microsatellites do not exclusively consist of the increase or decrease of one single repeat unit, but occur as well in steps of two, three, four or five dinucleotides. It has also been observed that there are more frequent contraction mutations with increasing microsatellite length, while the rate for expansion mutations stay constant (Xu *et al.*, 2000).

In the second main microsatellite mutation model, the infinite allele model (IAM), it is assumed that every mutation event creates a new allele whose size is independent from the original allele. This model is quite often used as a default model in population genetics analyses, since it is more simple and general (Selkoe, Toonen, 2006b). This model is called the infinite alleles model, and has been proposed by Kimura and Crow (1964).

In addition to these two main models, there are several propositions of microsatellite mutation models that try to take into account the various complications and derivations from the more general models as does the two phase model from DiRienzo (1994).

However, choosing the appropriate mutation model will be most important on large-scale phylogenetic applications and cross-specific comparison (Chambers, MacAvoy, 2000), whereas some computer programs used in likelihood-based parentage analyses allow the user to specify parameter according to the most appropriate mutation model (Jones, Ardren, 2003). Generally, the best microsatellite models will assume that long microsatellites mutate more often that short ones and are more likely to contract than to expand and that short microsatellites show a bias towards expansions (Calabrese, Sainudiin, 2005).

3.2 Microsatellite analysis

A general population genetic study based on the analysis of microsatellites will generally involve the following steps:

1. Finding and choosing appropriate microsatellites
2. Sample collection and preservation
3. DNA extraction
4. DNA amplification
5. Estimation of genotyping errors
6. Microsatellite data analysis and statistics

3.2.1 Finding and choosing appropriate microsatellites

The fact that microsatellites show a high level of polymorphism and many alleles per locus in a wide variety of species makes them a versatile tool for answering population genetic

questions (Frankham *et al.*, 2002). When considering the use of microsatellite markers for a certain population genetic study it may be worth the effort to search literature for any existing microsatellite marker for the target species and any closely related species (Queller *et al.*, 1993; Selkoe, Toonen, 2006b). At present, most microsatellite markers developed are being reported as "Primer notes" in the journal Molecular Ecology Resources (formerly Molecular Ecology Notes), which purpose is the dissemination of technical advances in fields of molecular ecology. In addition to report on new molecular marker development it comprises also new computer programs and methodological innovations which are published online as computer or technical notes. Molecular Ecology Resources provides an online searchable database containing an archive of all published notes (http://www.blackwell-synergy.com/loi/MEN). When searching for DNA sequences from certain species, the extensive online database from the United States' National Center for Biotechnological Information (http://www.ncbi.nlm.nih.gov/) offers large information on molecular biology. The web site provides among others the database GenBank®, which is the U.S. National Institutes of Health (NIH) genetic sequence database offering an annotated collection of all publicly available DNA. Being a part of the International Nucleotide Sequence Database Collaboration, GenBank is exchanging data daily with the DNA DataBank of Japan (DDBJ) and the European Molecular Biology Laboratory (EMBL), thus providing continually updated information on molecular biology (Benson *et al.*, 2008).

Since the DNA surrounding a microsatellite locus, the flanking region, is generally conserved across individuals of the same species, it might be possible that primers developed for a certain species might amplify DNA from other species in the same genus or even family (Selkoe, Toonen, 2006b). The potential for successful cross-specific primer transfer appears to be highest in species with long generation time, mixed or outcrossing breeding systems and where genome size in the target species is small compared to the source species. As for mammals, 50% of the microsatellites amplified by cross-specific primers showed polymorphism, with the highest probability of obtaining polymorphic markers when transferring primers between genera and the lowest across different families (Barbara *et al.*, 2007). Despite the limited potential of cross-specific transfer of microsatellites, existing microsatellites are often tested before species-specific primers are developed (Goossens *et al.*, 1998a). Unsuitable markers are usually discarded from further use (Carpenter *et al.*, 2005), or cross-specific microsatellites are tested and used in addition to markers developed for the focal species (Kruckenhauser, Pinsker, 2004). In some population genetic studies, exclusively

cross-specific microsatellites show sufficient polymorphism, especially when the target and non-target species are closely related (Burton, 2002). In other cases, microsatellites are developed intentionally for the species of interest when there are no amplifying cross-specific microsatellites available (Crawford *et al.*, 2008a) or when specifically designed primers are intended to be used further in other closely related species (Gondek *et al.*, 2006).

There are several basic assumptions behind the application of microsatellite data for population genetic purposes. Selkoe and Toonen (2006b) present a detailed guide for the evaluation of microsatellite loci for inclusion in a population genetic study. They propose a preliminary screening of the loci on a subset of samples and to test if assumptions on the microsatellite loci are met. A suggestion of quality control checklist adopted from Selkoe and Toonen and slightly modified is given below (Table 2).

Table 2 Quality control checklist adopted and slightly modified from Selkoe and Toonen (2006b)

Assumptions
1. Accurate genotyping
2. Complete allele amplification
3. Linkage equilibrium
4. Neutrality
5. Mendelian inheritance
6. Unambiguous allele identification

1. Accurate genotyping

In order to detect and identify genotyping errors in a microsatellite study, 10% of the samples run in the study should be genotyped repeatedly, and both the locus-specific and the overall genotyping error be calculated and reported (Hoffman, Amos, 2005).

2. Complete allele amplification

In order to ensure that all alleles are amplified throughout the study, a positive control should be run with every PCR batch (Selkoe, Toonen, 2006b). Alleles that do not amplify in a heterozygous individual are usually described as an allelic dropout (ADO) and the rate of ADO calculated and reported (Broquet, Petit, 2004). An allele that consistently fails to amplify a PCR product at a detectable level is called null allele (Dakin, Avise, 2004). The presence of null alleles might be detected with the test for the Hardy-Weinberg equilibrium

(HWE) in which observed genotypes frequencies are compared with the expected frequencies (Selkoe, Toonen, 2006b). The Hardy Weinberg law states that in an infinitely large, interbreeding population in which mating is random and there is no mutation, selection or migration, allele frequencies and genotype frequencies remain unchanged from generation to generation (Beebee, Rowe, 2004; Frankham *et al.*, 2002; King *et al.*, 2006). In the test on the HWE the observed allele frequencies are compared with the expected allele frequencies and the deviations from the equilibrium be reported. A commonly detected deviation of the HWE is the excess of homozygotes. The reason for this deficit in heterozygotes may lie in the biological properties of the population analysed, for example in the case of strong inbreeding. Another reason might be the occurrence of null alleles, which are alleles that fail to amplify in the PCR due to non ideal PCR conditions or the presence of mutations in the primer binding regions so that primer binding (hybridization) is inhibited. In order to detect null alleles, it should be determined whether the unamplified locus in the individuals remains unamplified, while the other loci in the individual produce alleles normally. When the locus after repeated DNA extraction and amplification still fails to produce any alleles, it is likely that the individual is homozygote for a null allele. For population genetic analyses that require high accuracy in genotyping, such as parentage analysis, even rare null alleles can confound results and any loci with strong evidence of null alleles should be excluded (Selkoe, Toonen, 2006b). When alleles in a heterozygote are very different in size, PCR can be more efficient in the replication of the shorter allele, and the longer allele appears too faint to be detected in the genotype scoring process. In this case it might be useful to re-amplify the individuals that are homozygous for small alleles and increase the DNA concentration in the DNA sequencer run (Wattier *et al.*, 1998).

3. Gametic disequilibrium (Linkage disequilibrium)

When to microsatellite loci are found very close together on a chromosome, they might not assort independently, but be transmitted to their offspring as a pair. There are several computer programs designed for microsatellite data analysis that also test and detect gametic disequilibrium by searching for correlations between alleles at different loci (Selkoe, Toonen, 2006b). Some population genetic analyses like parentage testing do not require the test for linkage equilibrium prior to analysis, and due to the sensitivity of most suitable computer programs for linkage testing, even a highly significant linkage disequilibrium between a pair of loci may be rather weak (Kalinowski *et al.*, 2007).

5. Selective neutrality

In order to use molecular variation in the identification of individuals, populations or species, it is essential to have genetic markers that are neutral (Hedrick, 1996). To test microsatellite loci for selective neutrality several tests are proposed by Selkoe and Toonen (2006b). Manual comparison of allele frequencies work as a test for selective neutrality (Lewontin, Krakauer, 1973) as well as computer programs that include linkage tests like GENEPOP (Rousset, 2007).

Mendelian inheritance

Another important requirement for population genetic studies is that the molecular markers follow the rules of Mendelian inheritance (Jarne, Lagoda, 1996; Selkoe, Toonen, 2006b). Although performing defined crosses and genotyping a large number of offspring in a parentage analysis may be challenging and impractical, the evaluation of inheritance in a population genetic study should be conducted and loci that show more than two alleles per diploid individual discarded (Selkoe, Toonen, 2006b).

6. Unambiguous allele identification

When two alleles have the same size and sequence but not descent from the same ancestral allele, they show homoplasy. In large-scale phylogenetic analyses, homoplasy might lead to the underestimating of the actual population divergence, while it might not be of much concern in population genetic studies (Jarne, Lagoda, 1996).

Numbers of microsatellite loci

It is generally agreed in the field of molecular ecology that in most cases, the more loci included in the study, the more reliable the data results will be (Selkoe, Toonen, 2006b). However, including inappropriate loci that don't meet the quality requirements suggested above could lower both the precision and accuracy of genetic estimates. There is a clear trade-off between the adjustment of loci number and the resulting change in accuracy and precision of the resulting data set, which makes it quite challenging to choose a certain number of loci. Computer programs like CERVUS v.3.03 can simulate a parentage analysis based on given parameters and the results of the simulation be used to determine the appropriate number of loci (Kalinowski *et al.*, 2007). Generally, the selection of microsatellites according to their displayed level of polymorphism will depend on the question of interest (Selkoe, Toonen, 2006b). In the case of paternity assignment based on likelihood, the number of microsatellite loci required to exclude all non-parental males for 99% of mother-offspring pairs will

increase with the number of non-parental males that have to be excluded. At the same time, the number of loci needed will decrease with increasing heterozygosity of the loci (Queller *et al.*, 1993).

3.2.2 Sample collection and preservation

Generally, in studies that ultimately require statistical analysis the sample size should be as large as reasonably possible. Practically, this means a minimum number of 10 individuals per population, although around 20 would be preferable. However, statistical power is usually more strongly affected by the number of markers and their polymorphism than by sample size, and the choice of sample size and marker should be adapted and optimized according to the type of analysis (Beebee, Rowe, 2004). In contrast to more traditional sampling methods where specimen were collected, "sacrificed" and prepared for documentation and collection, many recent survey methods rely rather on the observation of the animals and the collection of tissue samples for further genetic studies. In order to avoid capturing and restraining of animals, a variety of non-invasive methods to collect DNA samples has been developed over the last years (Morin, Woodruff, 1996). Hairs have been non-invasively collected and successfully used in the course of population genetic studies from the northern hairy-nosed wombat (*Lasiorhinus kreftii*) by suspending tape close to the burrows (Sloane *et al.*, 2000), from the Eurasian badger (*Meles meles*) by setting out barbed-wire enclosures around bait stations (Frantz, 2004), from the American marten (*Mares americana*) by using glue patches attached to trees (Foran *et al.*, 1997) and from the Eurasian otter (*Lutra lutra*) by using a tunnel-like construction with Velcro® patch on the inside (Anderson, 2006). As another alternative for non-invasive genetic sampling, shed hairs were collected in the field from Gombe chimpanzees (*Pan troglodytes schweinfurthi*) from individual sleeping nests (Constable *et al.*, 2001) and from wolves (*Canis lupus*) by following wolf tracks in the winter (Scandura *et al.*, 2006). Hairs can also be directly plucked from animals when using live traps for capture (Goossens *et al.*, 1998b), although this is no longer defined as non-invasive but rather as non-destructive sampling since the animals are both disturbed and caught under sample collection (Taberlet *et al.*, 1999a). Both shed and directly plucked hair can provide a low but sufficient amount and an adequate quality of DNA for genetic typing (Gagneux *et al.*, 1997). However, when the DNA quantities in those samples become very small, there is an increased for genotyping errors, and some extra guidelines should be followed in order to obtain reliable genotyping (Taberlet, Luikart, 1999b).

When the genetic material for a study is totally based on non-invasively collected DNA samples, new sample collection methods should be tested and evaluated prior to the main study. For example, the number of hair traps can be adjusted in order to collect the necessary number of hair samples without delaying the course of the actual study (Frantz *et al.*, 2004). A useful approach for evaluating the quality of non-invasively collected hair samples is to conduct DNA extraction on samples from different sources like blood or faeces (Scandura *et al.*, 2006). Quite frequently, the collection of non-invasive DNA samples is preceded by regular observations of the animals of interest in order to reduced the probability of misidentification (Constable *et al.*, 2001; Kholodova *et al.*, 2000). By collecting more than one sample from each individual it is possible to overcome variations in sample quality (Vigilant, 2002). In order to avoid contamination of the collected DNA samples with cross-specific DNA, the samples should not get in contact with human skin, especially when some of the genetic methods includes ingredients that are not specific for the target species (Taberlet *et al.*, 1997). It is especially important to avoid human contact with the DNA containing root-end of the hairs (Morin *et al.*, 1994), which can be obtained by simply using forceps when handling the hairs (Kholodova *et al.*, 2000; Takami *et al.*, 1998). Usually, DNA from hair samples is successfully extracted when the hair samples are stored in dry conditions like in paper envelopes or vials including a desiccant (Anderson, 2006; Morin *et al.*, 1994; Sloane *et al.*, 2000; Taberlet *et al.*, 1997), but storage of hair samples in a Ziplock bag in a freezer at -20^{o}C has proved to be a good alternative with even higher amplification success compared to dry storage conditions success (Roon *et al.*, 2003). Suboptimal storage conditions, like keeping hair samples in airtight tubes can lead to decreased PCR amplification success (Galan *et al.*, 2003). The DNA in hair is subjected to degradation over time, and consequently, storage conditions should be controlled regularly, and the DNA be extracted within six month after collection (Roon *et al.*, 2003).

3.2.3 DNA extraction

Many population studies of DNA are based on the extraction of total cellular DNA with the following selection of subsets of the genome by the PCR. The isolation of most of the DNA present in a sample is important for studies where much of the native DNA in the samples may be degraded, and will provide much information due to the availability of total cellular DNA compared to what is obtained by extraction from only organellar DNA. Fortunately, the PCR is so sensitive that it only requires little initial DNA template (Milligan, 1998) defined

segments of DNA in a sample can be amplified to microgram quantities from as little as a single template molecule (Hoelzel, Green, 1998).

The array of available DNA extraction techniques is extensive since the chemical composition of tissue derived from different sources might be quite variable. There are a variety of protocols for DNA isolation available, differing in both the time required for DNA extraction, and the quality of the isolated DNA. As long-lasting incubations for each sample might appear time-consuming and thus ineffective for larger sample sizes, practically it might be possible to perform other tasks under sample incubation, and often, the extra time invested in a more complete DNA purification might make the subsequent analysis easier due to higher DNA quality (Milligan, 1998).

Pre-examination/Pre-screening

Quite often in genetic studies using hair as DNA source, the suitable hairs are chosen after a closer examination. Microscopy seems suitable in order to choose suitable hairs according to the presence of dry cells (Scandura *et al.*, 2006; Taberlet *et al.*, 1997).

When using non-invasive methods for collecting hair in the field, hair samples are usually identified based on hair morphology (Mowat, Strobeck, 2000) and the presence of tracks around collection sites before the DNA extraction in order to make sure that they come from the species of interest (Sloane *et al.*, 2000).

When hair is plucked directly from the animals, a binocular microscope is useful to select hairs for extraction with an intact follicle (Anderson, 2006). If hairs are not checked or examined before DNA extraction, not all hairs might contain follicle with cells, resulting in a comparatively low percentage of successfully isolated DNA (Kholodova *et al.*, 2000).

For the analysis of nuclear DNA, the root of the hair(s) is the essential part of the DNA extraction sample, while parts of the hair shaft might be included as well. The total length of the hair shaft in the sample may not influence the success of DNA extraction when the objective of the extraction is genomic DNA (Kholodova *et al.*, 2000; Sloane *et al.*, 2000; Taberlet *et al.*, 1997). However, the number of hairs used for DNA extraction might influence the genotyping error rate thus care must be taken when deciding how many hairs to use in order to extract DNA. Genotyping errors can become substantial when using a single plucked hair for DNA extraction due to the resulting low DNA quantity. When including several hairs in a sample, the genotyping error can be lowered considerably, depending on the number of hairs (Goossens *et al.*, 1998b). However, in the case where the source or identity of a hair sample is uncertain, it is more favourable to extract DNA from a single hair at a time (Sloane

et al., 2000; Taberlet *et al.*, 1997) and if single hair provide a more even quantity of DNA opposed to the use of a larger numbers of hairs (Walker *et al.*, 2006). In order to make the genotyping of single hairs more reliable the multi-tubes approach suggested by Goossens et al. (1998b) may be considered. It is possible to compare the relative quality and quantity of DNA extracted from hair by comparing the result to DNA extracted from other types of tissue (Anderson, 2006). If the DNA extraction did not result in any detectable DNA, it might simply be repeated (Vigilant, 2002).

Negative control/Contamination control

It is essential to avoid the contamination of DNA samples with non-target DNA. Some technical precautions proposed by Taberlet et al. (1999a) are shown in Table 3.

Table 3: Some technical precautions in order to avoid contamination with non-target DNA, adopted and modified from Taberlet et al. (1999a).

Methods for avoiding contamination
• physical separation of the laboratories designed for DNA extraction and PCR experiments • avoidance of concentrated DNA extracts in the DNA extraction room • use of pipettes with aerosol-resistant pipette tips and • monitoring of reagents for DNA contamination by running negative PCR controls

The contamination of a DNA samples with non-focal DNA in the course of DNA extraction might also be avoided by alcohol-flaming forceps and tweezers (Scandura *et al.*, 2006) or the use of a sterile scalpel (Anderson, 2006) before handling a new sample.

In order to detect whether contamination with exogenous DNA has occurred during DNA extraction, tubes without DNA can be treated identically with the ordinary samples through the extraction procedure as an extraction negative control (Sloane *et al.*, 2000; Taberlet *et al.*, 1997).

Sufficient DNA quantity and quality

Sometimes, although following an appropriate protocol for DNA extraction, the quantity of DNA obtained from the extraction may not be sufficient to be used in the purpose of a certain study (Taberlet *et al.*, 1997).

Usually, the probability of genotyping errors like non-amplifications and allelic dropout will increase with decreasing concentration of DNA in a sample (Fernando *et al.*, 2003).
It has been shown that there is a greater risk of genotyping errors due to incomplete allele amplification when the DNA concentration falls under a certain concentration in the template (Gagneux *et al.*, 1997). In order to maximize DNA yield from single hairs, DNA extractions can be carried out in the field on the day of hair collection (Sloane *et al.*, 2000).

The determination of the DNA quantity obtained from extracted hair samples could principally allow the evaluation of extracts by DNA content, increase the reliability of results and speed up the genotyping process by detecting samples not suitable for further amplification. Unfortunately, most conventional fluorimetric assays of DNA content lack sufficient sensitivity to determine the quantity of samples with very low DNA concentrations, and neither is it possible to distinguish between the target DNA and non-focal DNA in the assay. A more useful method to evaluate the suitability of the DNA samples would be to measure the amount of amplifiable DNA of interest present in a sample extract (Morin *et al.*, 2001). An alternative to the determination of DNA quantities in extraction samples is to test the PCR products amplified by conserved primers (Frantz *et al.*, 2004).

3.2.4 DNA amplification and sequencing

The amplification of microsatellite DNA implies the use of the polymerase chain reaction (PCR). With this technique, defined segments of DNA can be amplified to microgram quantities from as little as one single DNA template molecule, which makes it a suitable tool for the analyses of even minutes DNA amounts (Taberlet, Luikart, 1999b).
The procedure is a simple and fast way to generate ample DNA material for further genetic analysis. The combination of PCR amplification and DNA sequencing has been shown to be a powerful tool for population genetic studies (Hoelzel, Green, 1998).

Multiplex PCR vs. singleplex PCR

Running multiplex PCR including groups of primers is a feasible choice in order to establish effective in the laboratory when the sample size of the genetic study is large. Although it might not be possible to include all primers in one multiplex PCR without redesigning some of the primers, combining two or three loci is usually a realistic alternative which makes the PCR highly elaborate (Schlötterer, 1998).

Multi-tubes approach vs. single-tube approach

In the microsatellite amplification of DNA samples from hair the amount of DNA available for sequencing can be very low and is often in the pictogram range (Taberlet *et al.*, 1996). In this case, the probability increases that some of the alleles are not detected or misinterpreted as background contamination. One way of minimizing this possibility is the multi-tubes approach, where the sample DNA is divided among several tubes and then amplified, typed separately and the obtained genotypes compared to each other. This approach may provide more reliable results than the single-tube procedure, although this depends on how small the DNA amounts are in the samples (Navidi *et al.*, 1992). Although quite low amounts of DNA in the sample will be further diluted in this approach with the increasing probability of genotyping errors (Paetkau, 2003) and a huge number of PCRs must be performed (Vigilant, 2002), it represents a feasible method of reducing genotyping, also when followed with modifications (Ferrando *et al.*, 2007).

Internal control/positive control

One requirement for the successful amplification of microsatellite in a PCR is that there are no mutations in the flanking regions of the microsatellites. If mutations occur in the primer binding regions, some samples will have only one allele amplified, or even fail to amplify at all (Paetkau, Strobeck, 1994). In order to check for the presence of amplifiable DNA in the sample, and as a control for the PCR, a conserved primer can be included (Ferrando *et al.*, 2007).

Negative control/contamination control/check

In order to avoid contamination of samples with foreign DNA in the course of PCR experiments, the precaution guidelines for DNA extraction (Taberlet *et al.*, 1996) summarized in Table 3 apply also for the performance of the PCR amplification (Scandura *et al.*, 2006). Additional measures can be taken like the ultraviolet sterilization of laboratory equipment prior to setting up a PCR reaction, and the use of a 90% bleach solution to clean working surfaces (Constable *et al.*, 2001). A negative control run simultaneously with the DNA samples is a simple way of monitoring contamination with foreign DNA and can contain all the ingredients necessary for PCR but no the template DNA (Ferrando *et al.*, 2007; Scandura, 2005) or only water (Romain-Bondi *et al.*, 2003).

Amplification check before sequencing

The outcome of the DNA amplification can be verified by running a certain amount of each PCR product on a gel for electrophoresis. A DNA ladder might be included in each run, as a reference for fragment size and intensity (Scandura *et al.*, 2006).

Quality assessment

Scandura et al (2006) developed a quality control to described and define the quality of single-locus microsatellite amplification, based on the attribution of a quality score (Q-score) to every single-locus genotype. In the first quality check, the band intensity, sharpness and absence of non-specific products were evaluated and given as the PCR score, while the second check comprised evaluated the peak shape and height together with the presence and conformation of shadow peaks of the obtained microsatellite profile (SEQ score). Results from these evaluations are made comparable by assigning to the scores.

Amplification success

When the PCR amplification success is low, PCR conditions should be modified in order to obtain a sufficient number of genotypes. When amount of extracted DNA is expected to be low, the number of PCR cycles could be increased in order to get a PCR product that can be detected and analyzed. However, this might not be necessary if the detection methods are sensitive enough to detect and analyze even small amounts of DNA (Taberlet, Luikart, 1999b).

PCR product sizing/sequencing

A traditional method to determine the size of the microsatellites amplified in the PCR is the polyacrylamide gel electrophoresis based on either denaturing or native gels (Schlötterer, 1998). Despite of this being a convenient method for PCR product detection, the sizes of microsatellite bands can usually only be roughly determined and additional bands like stutter bands might also lead to misinterpretation of band patterns. A more appropriate system for accurate sizing of microsatellites is the capillary electrophoresis (Schlötterer, 1998). Automated fluorescence technologies based capillary electrophoresis are widely used for detection and size determination of PCR products, and make it possible to run multiplex PCRs by using primer pairs that are differently labelled, either with fluorescent dye (David, Menotti-Raymond, 1998) or radioactively (Blouin *et al.*, 1996; Semple *et al.*, 2001).

It might not even be necessary to conduct gel electrophoresis on all PCR products, but rather rely on an automatic DNA sequencer (Bergl, Vigilant, 2007; Carpenter *et al.*, 2005; Kruckenhauser, Pinsker, 2004; Scandura, 2005; Schnabel *et al.*, 2000; Walker *et al.*, 2006) and detected and size alleles using the software connected to the sequencing system (Bergl, Vigilant, 2007; Carpenter *et al.*, 2005; Scandura, 2005; Schnabel *et al.*, 2000).

However, especially in DNA extractions based on non-invasively collected DNA samples like hair and faeces it might be useful to verify the outcome of the PCR preliminary to further genetic analysis (Constable *et al.*, 2001; Lorenzini *et al.*, 2004; Scandura *et al.*, 2006) as recommended by Chambers and MacAvoy (2000). A standard PCR buffer contains all the necessary ingredients in certain concentrations, although it might be necessary to adjust the concentration of some of the ingredients in order to increase the efficiency of the PCR (Dean, Milligan, 1998). By monitoring PCR amplification results the requirements for high specificity, high signal output and purity might be met. Whether finally to rely merely on automated systems or use combine the two electrophoresis approaches remains with the investigator in a study (Chambers, MacAvoy, 2000).

3.2.5 Estimation of genotyping errors

It should be taken into account that genotyping errors do happen (Paetkau, 2003). Genotyping errors are usually defined as the differences observed between two or more molecular genotypes which were obtained independently from the same sample. These errors occur due to the imperfectness of molecular assays and manual sample handling, and can be created during every step of the genotyping process (sampling, DNA extraction, molecular analysis, genotype scoring and data analysis) and through a variety of factors like technical artefacts, human causes or simply by chance (Bonin *et al.*, 2004).

There are numerous causes for the presence of genotyping errors. In the case of low DNA quantity and/or quality, errors may occur during PCR due to non-amplification of certain alleles (Hoffman, Amos, 2005). This is the case for the allelic dropout (ADO), when under these conditions one allele in a heterozygous individual is not detected at all after PCR amplification (Taberlet *et al.*, 1996) and the individual is identified as a false homozygote (Goossens *et al.*, 1998b). It seems that the probability of ADO under PCR increases with the decreasing amount of DNA in a PCR tube. The DNA quantities that are obtained from non-invasively collected samples like hair and faeces are usually quite low, and therefore the probability of ADO is high. In order to obtain a reliable genotyping with a confidence of 99%,

taking into account the stochastic sampling of template DNA, the possibility of generating false alleles and the risk of contamination, the multiple tubes approach is recommended for samples with very low DNA quantity (Taberlet *et al.*, 1996).

Another source for genotyping errors is the presence of null alleles (Dakin, Avise, 2004). Whereas ADO occurs stochastically, a null allele might be prevented from being amplified during PCR due to mutations in the flanking regions of the microsatellites (Dakin, Avise, 2004) that inhibit or event prevent primer binding (Callen et al. 1993) or short allele dominance, where larger alleles remain undetected when amplified together with shorter alleles (Wattier et al. 1998).

Another type of genotyping error is false alleles which are generated from strand slippage artefacts in the first cycles of the PCR, amplify together with the genuine alleles and often become visible as shadow bands on a gel (Taberlet *et al.*, 1996). Usually, these additional bands are smaller than the original allele (Fernando *et al.*, 2003) and differ from the expected allele size most likely with only few repeat units (Schlötterer, 1998). The rates of strand slippage under PCR are related to repeat unit length of the microsatellite, as there are more stutter bands for dinucleotide repeats than for trinucleotide or tetranucleotide repeats (Schlötterer, Tautz, 1992). Other additional bands that are longer than the expected allele can form as the result from the terminal transferase activity of *Taq* polymerase adding an A nucleotide to the PCR product (Schlötterer, 1998).

Other amplified bands with unexpected sizes can be generated due to the contamination of the sample with foreign DNA or from unspecific amplification. In contrast to false alleles, these contamination events can be detected by using negative controls during PCR (Broquet, Petit, 2004). If the amount of the contaminating DNA amount is much smaller than the target DNA, the band signal will be much weaker compared to the strongest signal of an expected allele, and should thus be ignored in the interpretation (Navidi *et al.*, 1992).

Another cause for genotyping errors is the phenomenon of electrophoresis artefacts which have been observed to arise when high concentrations of PCR products are electrophoresed in an ABI 377 automated sequencer (Fernando *et al.*, 2001).

Scoring errors can occur during automated sequencing due to the subjectivity during the band interpretation process when there are more than one experimenter that evaluates whether an amplified DNA fragment with low intensity should be excluded or not (Bonin *et al.*, 2004) or to the misinterpretation of allele peaks by sequencer software (Ginot *et al.*, 1996).

Consequences of genotyping errors for population genetic studies

The genotyping errors described above usually produce a pseudo-homozygosity or heterozygote deficiency, both in the case of ADO and null-alleles (Chakraborty) and the phenomenon of short allele dominance (Wattier). The consequence of these types of genotyping error in a parentage analysis will be an increased possibility of excluding true parents or offspring on the base of their apparent homozygous state (Dakin, Avise, 2004). Detailed recommendations for tracking and the assessment of genotyping errors are given in Bonin et al. (2004).

Since genotyping errors in a population genetic study seem to be unavoidable (Paetkau, 2003), and the impact of even a low rate of genotyping errors on a population analysis is considerable (Hoffman, Amos, 2005), it is recommended that the genotyping error rate within each study is quantified and reported (Selkoe, Toonen, 2006b). Generally, there is no concordant way of expressing the genotyping error rate throughout genetic studies. In order to make it possible to compare error rates between studies, it can be reported as both the number of errors per allele and per reaction, and summarize these for each locus individually and across all loci. In order to include typographical and scoring errors, error rate per reaction can also be calculated as the number of incorrect genotypes divided by the total number of reactions used for comparison, and the error rate can be calculated as the number of incorrect alleles divided by the total number of alleles (Hoffman, Amos, 2005).

3.2.6 Microsatellite data analysis and statistics

Microsatellite data can provide a wealth of genetic information. Usually, most analytical methods for retrieving information from this genetic data are based on the analysis of allele frequency data. A variety of computer programs is available in order to perform statistical analysis that obtain information on relatedness and parentage, population dispersal and migration, inbreeding and population size (Luikart, England, 1999). Allele frequency data obtained by a computer program like CERVUS v.3.0.3. usually comprises statistic parameters for each locus and summary statistics across all loci (Kalinowski *et al.*, 2007).

Parentage analysis methods

Information on genetic parentage is not only useful in order to obtain information about the kinship structure in a population, but also essential in the study of several other aspects of population genetics, like the impact of inbreeding, the verification of pedigrees and the

determination of the effective population size (Frankham *et al.*, 2002) and mating systems (Parker *et al.*, 1996).

In principal, microsatellites are used as Mendelian genetic markers which estimate genetic paternity or maternity by excluding adults as parents whose genotypes are incompatible with those of the juveniles under consideration. Statistical exclusion probabilities are derived from the variability of the markers and the available information about the biological nature of the actual parentage problem (Avise, 2004).

Exclusion method

The classical method of parentage determination in a population is the exclusion method. This approach is based on the Mendelian rules of inheritance, and uses incompatibilities between parents and offspring to reject particular parent-offspring hypotheses. An individual is excluded as a candidate parent if it possesses an allele that is incompatible with the alleles of the putative offspring. Under strict exclusion conditions, a single allele mismatch will exclude a candidate parent, and the single remaining non-excluded candidate parent be assigned to the offspring. In some parentage analysis computer programs the number of mismatches necessary for exclusion can be specified in order to make the method more robust. This is particularly useful when the pool of candidate parents and the number of loci is large since additional loci or additional individuals increases the likelihood that a dataset will contain genotyping errors like null alleles and mutations (Jones, Ardren, 2003). Exclusion is a simple method for testing potential parent-offspring relationship, while assignment of an offspring to a parent usually only is possible by using likelihood-based methods when multiple parents are excluded (Marshall *et al.*, 1998).

Categorical allocation

The categorical allocation assigns progeny to non-excluded parents based on the likelihood scores derived from their genotypes. In this method, the likelihood of one or a pair of individuals being the parent(s) of a given offspring is divided by the likelihood that these individuals are unrelated. The logarithm of this likelihood ratio is determined, and from all the candidate parents, offspring are assigned to the pair with the highest LOD score. When multiple parent-offspring relationships obtain the highest LOD score, parentage remains ambiguous (Jones, Ardren, 2003).

Fractional allocation

The fractional allocation method calculates the likelihoods for putative parent-progeny relationships the same way as the categorical allocation method. Here, a fraction between 0 and 1 of each offspring is assigned to all non-excluded parents. The fraction of an offspring assigned to a particular candidate parent is proportional to its likelihood of parenting the offspring compared to all other non-excluded candidate parents (Jones, Ardren, 2003). In contrast to the categorical method, paternity will be assigned for all progeny, even though some progeny will not be assigned to a single farther (Devlin *et al.*, 1988).

Genotypic reconstruction

In this method, the multilocus genotypes of parents and offspring are used to reconstruct the genotypes of unknown parents that contributes gametes to a progeny array for which one is known a priori (Jones, 2001). The reconstructed genotypes are compared to the genotypes of a pool of candidate parents or to one another. This method is quite computationally intensive and time-consuming (Jones, Ardren, 2003).

Choosing appropriate parentage analysis method

The choice of the appropriate parentage analysis method will largely depend on the types of samples than can be collected in the study. The optimal set-up is a large number of offspring collected from known mated pairs and adults. The prospects for complete assignment of parentage are also quite good when samples of offspring can be collected in family groups with their mothers, and when a complete sample of adult males from the population can be obtained. With fewer data available, the probability of correct parentage assignment decreases. In order to compensate for a less ideal sample construction, molecule markers with a greater resolving power can be used. If only parts of the candidate parents can be sampled, it is still important to know the total number of putative parents in the population since that number is a relevant parameter in parentage assignment (Jones, Ardren, 2003).

A large number of candidate males will lead to a lower percentage of resolved paternity tests compared to a group with fewer candidates. In order to resolve more paternity tests, more males can be sampled, which would lead to an increase in the assignment of paternity (Marshall *et al.*, 1998). Still, the number of maximum number of individuals that can be resolved depends on the number of loci (Weller *et al.*, 2006) and the locus heterozygosity (Paetkau, 2003). In order to choose an appropriate parentage analysis method according to the particular sample set-up, Jones and Ardren (2003) give some suggestions about preferred

parentage analysis techniques dependent on the available genotypes and a priori information about kinship relations.

4 Material and methods

4.1 Study population and data collection

The study population consists of several beaver colonies in Telemark County in Southern Norway. From this population, hair samples have been collected since 1999 in the course of the ongoing research at the Telemark University College in Bø. (F. Rosell, pers. comm.). In this study, DNA was extracted from 148 hair samples representing beaver individuals that previously had been captured, marked or identified and released.

The hair samples consisted mostly of guard hair directly plucked from the animals' coat and some samples containing hair from the undercoat. The hair samples were stored in paper envelopes at room temperature, and the storage time varied from two months to eight years until extraction.

4.2 DNA extraction

In order to avoid contamination, DNA extractions were carried out in a separate laboratory dedicated to this purpose and spatially separated from PCR products (and aerosol-resistant pipette tips were used). From the hair samples, guard hairs were preferably chosen for DNA extraction, and if not available, up to five hairs from the undercoat were used. Each hair was checked for the presence of hair follicle either by eye-sight or by using a binocular. Generally, DNA was extracted from two to five hairs per sample by cutting each hair 0.5cm above the hair root. If no follicle was visible, the whole length of the hairs were used, cut in segments of 0.5cm. The DNA extraction method followed the user-developed protocol for purification of total DNA from hair using the DNeasy®Blood&Tissue Kit (QIAGEN®, 2006) with minor modifications.

4.3 Microsatellites

Nine primer pairs developed by Crawford et al. (2008a) for the North American beaver (*Castor canadensis*) were used for genotyping.

A conserved primer pair (BK) designed by Kocher et al. (Kocher *et al.*, 1989) was used in order to check for the presence of DNA in the extraction samples and as a positive control in some of the PCR runs.

Table 4 Primer sequences used in the study, all designed by Crawford et al. (2008a) for the North American Beaver (*Castor canadensis*) and their characteristics in the target species.

Locus	Primer sequence (5'- 3')	Repeat motif	T_a	Allele size	H_o	H_E	GenBank Accession no.
Cca4	GATTTCAGACACAGCCACCA AGTGATGGGATTGAACTCCAG	$AC_{(17)}$	61	352-364	0.7	0.772	EF524501
Cca5	TGCTTTCATCTGCTCTATGAAAAT CAGTGATGAAGGGAAGAGGAA	$CT_{(21)}$	61	157-185	0.32	0.621	EF524502
Cca8	GGGCTCAGAGGAAAAAGGAG GATCAGGCAAAAGGCTGGTA	$GATA_{(12)}$	61	356-426	0.8	0.837	EF524503
Cca9	TCTTTCTTGTTGGTCCTGGAA TGGGAGAGTGGTTGCCTATC	$TG_{(21)}$	60	136-156	0.77	0.753	EF524504
Cca10	TTTTGTTGGGAAATATGCTGTT TGCAGAACAAAGAAAATATTGAAAG	$TC_{(19)}$	60	120-154	0.83	0.862	EF524505
Cca13	CCCTAGACTTTGATTATACGG AGGTTGCCTAGAGAGAGGTGTG	$GT_{(11)}GT_{(7)}$	60	277-295	0.45	0.481	EF524506
Cca15	TCTGCCTTATGTGATGGTCAA CTCAAAGCACACAGGTCAGC	$AG_{(6)}AG_{(7)}$	59	177-185	0.65	0.583	EF524508
Cca18	CTGCTGTGGGATCTTGGATT TGGTATGTGCTACACAGAAAACAA	$CT_{(10)}$	59	205-220	0.5	0.513	EF5245010
Cca19	TTGAGGTCAACCTGTGGCTA TTAGACATGCACCGCCATAC	$TG_{(12)}AG_{(10)}$	59	220-266	0.87	0.815	EF5245011

4.4 DNA amplification

The PCR amplification was carried out in two steps, PCR I and PCR II.

4.4.1 PCR I

The first PCR was used as a test for the presence of amplifiable DNA in the extraction samples, carried out as a singleplex PCR on all DNA samples using primer pairs Cca5, Cca13, Cca18 and the universal primer pair. The PCR products were visualized on a polyacrylamide gel after native gel electrophoresis. Volumes, concentrations and PCR conditions for the electrophoresis are given in Table 5.

Table 5 Concentrations and volumes for the initial PCR

	Concentration	Volume	Final concentration
dH_2O		10.3 µL	
10xPCR buffer	10x	2.5 µL	1x
$MgCl_2$	25mM	4.0 µL	2mM
dNTP	2mM each	2.5 µL	0.2mM
Forward primer	10 pmol/µL	0.75 µL	0.3 pmol/µL
Reverse primer	10 pmol/µL	0.75 µL	0.3 pmol/µL
AmpliTaqGold	5U/µL	0.2 µL	0.04U
DNA template	unknown	4.0 µL	
Final volume		25 µL	

In the PCR runs, samples were heated to 95 ^{o}C for five minutes, followed by 30 cycles of 95 ^{o}C for 60 s, 60 s at the primer-specific annealing temperature, and 72 ^{o}C for 90 s. The final extension temperature at 72^{o}C was followed by a final extension for 5 min. at 72^{o}C on an Eppendorf Mastercycler®gradient (Eppendorf). In addition to the DNA samples negative controls were included in each PCR run, containing all the PCR ingredients but no DNA template. All samples were amplified using all primers, and a universal primer was run on all samples as a positive control.

4.4.2 Gel electrophoresis

Following the first PCR amplification step, the PCR products from all samples were electrophoresed on a 6% polyacrylamide gel for 30 min at 150V using a Mini-PROTEAN®3 Cell (Bio-Rad). The ØX174 RF DNA *Hae* III size marker (ABgene) was included on each gel for fragment size determination and as a positive control for the electrophoresis. Samples containing DNA extraction ingredients but no DNA were run as a negative control on each gel. The gels were stained with ethidium bromide and then imaged and digitized on a Syngene gel documentation system and fragment sizes were analyzed using the GeneSnap software (Syngene). The presence of amplified DNA segments were checked, and the position and intensity of the bands evaluated. Samples that showed strong bands after amplification with Cca5, Cca13 and Cca18 were further tested with the primer pairs Cca9, Cca10, Cca15 and Cca19, while samples that did not amplify a detectable PCR product were re-amplified with the universal primer pair BK. Samples that consistently did not produce any visible DNA fragment were considered to contain too low DNA quantities to be detected by manual gel electrophoresis.

4.4.3 PCR II and microsatellite genotyping

In the second PCR step, multiplex PCRs were performed in 50µL reactions volumes which included Multiplex PCR Master Mix, Q-solution and distilled water from the QIAGEN® Multiplex PCR Kit (Quiagen) and primers labelled with different colours. Volumes and concentrations are given in Table 6. All 148 DNA samples were included in this second step, and all PCR runs contained negative control samples containing all PCR ingredients and water but no DNA template. Six markers (Cca4, Cca5, Cca8, Cca13 and Cca18, primer mix 1) and three markers (Cca9, CCa10 and Cca15, primer mix 2) were run together as multiplex PCRs. Concentrations and volumes for the multiplex PCR runs are given in Table 6.

Table 6 Concentrations and volumes for the multiplex PCR.

	Concentration	Volume	Final concentration
RNA-free H_2O		11.0 µL	
Multiplex PCR master mix	10x	25.0 µL	1x
Primer mix 1 or 2	2µM pr. primer	5.0 µL	0.2 µM
Q-solution	5x	4.0 µL	0.5x
DNA template	(Unknown)	4.0 µL	
Final volume		50.0 µL	

The multiplex PCR programme started with a denaturation step at 95°C for 15 min, followed by 40 cycles consisting of 0.5min at 95°C, 1.5min at annealing temperature (58 °C) and 1min at 72°C. A final extension step at 72°C for 10min was added, and samples kept at 4°C. From the PCR products, 1µL was loaded into a well of a 96-well tray, previously filled with 12µL of formamide and 0.5µL of the GeneScan™-500 ROX™ size-standard (Applied Biosystems). The diluted samples were heated at 95°C and placed on ice before genotyping.
The diluted samples were analyzed and sized by capillary electrophoresis on the ABI PRISM® 3130x/ Genetic Analyzer. The allele lengths were determined by using the GeneMapper software (Applied Biosystems). The success rate of PCR amplification at each locus was calculated by dividing the number of positively amplified PCR products by the total number of PCR reactions at that locus, using the results obtained after sequencing.

4.5 Data analysis

4.5.1 Statistical methods and genotyping errors

Allele frequencies and descriptive statistics for each locus (observed and expected heterozygosity, PIC, exclusion probabilities, deviations from Hardy Weinberg equilibrium and null allele frequencies) were calculated using the computer program CERVUS v.3.0 (Kalinowski *et al.*, 2007).
In this study, genotyping errors were defined as the observed differences between two or more molecular genotypes that are obtained independently from the same sample.

Amongst the methods for genotyping error detection proposed by Hoffman and Amos (2005) two methods were found suitable and used for estimating error rates in this study. These two methods for genotype error estimation included the repeat-genotyping of a randomly selected subset of samples and the comparison the genotypes of deliberately resampled individuals. From all 148 samples, 15 samples (10%) were randomly chosen for repeat-genotyping together with 13 other deliberately plucked samples. The total of 28 samples were genotyped

in a multiplex PCR containing Cca4, Cca8, Cca13 and Cca18.The genotyping results from the two reactions for each of the 28 samples were compared to each other and the number of non-consistent and missing genotypes determined. It was also recorded whether a genotype at a locus was missing after only one or both reactions. All the numbers are given on the genotype level.

The ADO and FA rates were calculated according to Bonin (2004) at the allelic level as the ratio between the observed number of different alleles and the total number of comparisons. The occurrence of ADOs was reported in the case when an allele in a heterozygous individual failed to amplify, thus leading to a false homozygote. The rate of ADO was calculated as the number of positive amplifications involving the loss of one allele divided by the number of all positive amplifications of individuals determined as heterozygotes according to Broquet and Petit (2004).

False alleles were reported and the rate of false alleles calculated by dividing the number of false alleles by the number of cases in which a false allele could be detected, following the suggestions by Creel et al. (2003).

4.5.2 False exclusion probabilities

In order to assess the probability of falsely excluding a parent-offspring dyad based on false homozygotes, the probabilities of false exclusions were calculated for all loci used in the exclusion cases. The calculations were based on the equation (Equation 1) proposed by Dakin and Avise (2004), assuming that an offspring might receive a null allele from one parent and any visible allele from the other parent, resulting in the possibility of falsely excluding either the false homozygous offspring or parent.

$$P = \sum_{i=1}^{k} p_i p_k (1 - p_i)$$

Equation 1 Equation for calculating the probability of false exclusion of a true parent (Dakin, Avise, 2004) for a population with k-1 visible alleles with population frequencies p_i (i = 1 to k-1) and a null allele with frequency p_k.

Observational data on strongly assumed true parent-offspring relationships were used as a comparison to exclusions based on genotypes. In the case of homozygote-homozygote mismatches between known parent-offspring dyads at a locus with a significant deviation from HWE, the possibility of false exclusions due to non-amplifying alleles was considered.

4.5.3 Parentage testing and exclusions

Based on observational data collected throughout the ongoing research at the Telemark University College since 1997, a total of 50 parent-offspring relationships were constructed. These relationships included assumed kinship between father or mother and their putative offspring as well as between offspring and parent pairs.

In order to investigate parentage, the offspring genotypes were compared to the genotypes of the putative parent by using the classical exclusion method. Offspring were required to share one allele at each locus with the mother and the second allele with the father. If there was any allele at a locus in the offspring that did not match any allele at that locus in the potential parent, the parent-offspring relationship was considered to be not true and defined an exclusion case. However, genotype mismatches between maternal and offspring genotypes were interpreted as errors resulting from the genotyping process as in Marshal et al (1998), and no exclusions of a mother and her offspring were accepted as probable in this study.

5 Results

5.1 DNA extraction

DNA was successfully extracted from 136 out of 148 hair samples. From the 12 samples that did not amplify any detectable PCR product, three samples consisted of hair shafts of the beaver's undercoat without visible follicles while one sample contained shafts without visible cells from guard hairs.

5.2 DNA amplification and gel electrophoresis

The DNA amplification success after the initial PCR I was evaluated as the percentage of samples that showed visible bands after gel electrophoresis with the respective primers. Most of the samples (82 to 85%) that had not shown visible bands after amplification with primers Cca5, Cca13 and Cca18 produced visible fragments when amplified with the universal primer. All percentages and counts of the gel electrophoresis products are given in Table 7.

Table 7 Percentages of samples that showed visible bands in gel electrophoresis after the initial PCR I round, numbers are given in brackets.

locus	samples	locus	samples
Cca5	70% (107/153)	BK	82% (27/33)
Cca13	55% (81/148)	BK	85% (29/34)
Cca18	69% (102/147)	BK	85% (29/34)
Cca9	100% (15/15)		
Cca10	100% (15/15)		
Cca15	100% (14/15)		

5.3 PCR II and microsatellite genotyping

Of the 148 hair samples, 136 gave rise to a genotype profile after multiplex PCR amplification, with twelve samples not containing any amplifiable DNA. These samples were excluded from further analysis and calculations.

From the nine primer pairs used in the amplification, five pairs (Cca4, Cca8, Cca13, Cca18 and Cca19) successfully amplified polymorphic microsatellites, while the remaining showed to be monomorphic. The monomorphic microsatellites Cca5, Cca9, Cca10 and Cca15 were excluded from further analysis due to their lack of polymorphism.

The average PCR success rate for primer pairs Cca4, Cca8, Cca13, Cca18 and Cca19 was 77.3%, varying from 69.9 to 92.6%.

5.4 Data analysis

The average number of alleles was 3.2 with two loci being bi-allelic and three loci showing three, four and five alleles (Table 8). Tests for deviation from HWE revealed three deviations, with one of them being significant, at locus CCa4. At the bi-allelic loci Cca18 and Cca19 the test for HWE-deviation was not done due to the extremely low frequencies of the one of the alleles at both loci. Expected heterozygosities at the polymorphic loci Cca4, Cca8 and Cca19 ranged from 0.449 to 0.505, while loci Cca18 and Cca19 showed very low heterozygosities with 0.017 and 0.010 (Table 9).

Table 8 Alleles, allele frequencies and heterozygosities

Locus	Number of alleles	Allele size	Count	Heterozygotes	Homozygotes	Frequency
Cca4	4	380	1	1	0	0.0045
		386	127	37	45	0.5670
		388	94	34	30	0.4196
		390	2	2	0	0.0089
Cca8	3	382	60	34	13	0.3158
		390	128	36	46	0.6737
		394	2	2	0	0.0105
Cca13	5	267	91	49	21	0.3611
		269	161	49	56	0.6389
		270	1	1	0	0.004
		271	1	1	0	0.004
		290	1	1	0	0.004
Cca18	2	216	228	2	113	0.9913
		218	2	2	0	0.0087
Cca19	2	261	3	1	1	0.0144
		263	205	1	102	0.9856

Table 9 Characteristics of the five polymorphic loci used in the study. N: number of samples typed at the locus, H_O: observed heterozygosity, H_E: expected heterozygosity, PIC: polymorphic information content, **: significant at the 1%-level, NS: not significant, NA: not available

Locus	Alleles	N	PCR success rate	H_O	H_E	PIC	Null allele frequency estimate	Deviation from HWE
Cca4	4	112	82.4%	0.330	0.505	0.389	0.2075	**
Cca8	3	95	69.9%	0.379	0.449	0.356	0.0798	NS
Cca13	2	126	92.6%	0.389	0.463	0.355	0.0853	NS
Cca18	2	115	84.6%	0.017	0.017	0.017	0.0007	NA
Cca19	2	104	76.5%	0.010	0.029	0.028	0.2843	NA
overall			77.3%					

5.4.1 Genotyping errors

Of all 28 samples that were chosen for repeat-genotyping, there were found three cases of non-consistent genotypes (an error rate of 0.1 on the genotype level). In two cases, allelic drop-outs led to false homozygotes, while in the other case a false allele might have amplified

in either genotype. The rate of ADO on the allelic level was calculated as 0.08, while the detection of one possible false allele led to a false allele rate on the allelic level of 0.04. Both the allelic drop-out and the false allele were found at locus Cca13.
Of a total of 40 comparisons between maternal and offspring genotypes, there are nine cases of homozygous mother-offspring genotype mismatches (22.5%).

5.4.2 False exclusion probabilities

Based on the allele frequencies and null allele estimates, false exclusion probabilities were calculated for locus Cca4, Cca8, Cca13 and Cca19 (Table 10). Due to the negative null estimate frequency at locus Cca18 it was not possible to calculate the false exclusion probability for this locus.

Table 10 False exclusion probabilities for locus Cca4, Cca8, Cca13 and Cca19, based on equation 1.

Locus	Estimated null allele frequency	False exclusion probability
Cca4	0.2075	0.1644
Cca8	0.0798	0.0743
Cca13	0.0853	0.0780
Cca19	0.2843	0.2034

5.4.3 Parentage testing and exclusions

Based on available observational data, 50 putative parent-offspring relationships were constructed. Of these relationships, 34 comparisons between offspring and both their parents (parent pairs) were possible. In 16 cases, the genotype of only either parent was determined, resulting in 9 mother–offspring and 7 father–offspring comparisons. In these relationships there were found 16 cases of non-compatible genotypes between the offspring and either mother or father or both parents (Table 11). Of these incompatible genotypes, 10 homozygous mismatches (case 1, 3, 4, 5, 6, 7, 8, 9, 10 and 14) were found between a mother and her offspring and not accepted as probable. Seven homozygous mismatches between an offspring and the putative father were found (case 2, 3, 4, 5, 13, 14 and 16). Three times (case 11, 12 and 15), it was possible to exclude the parents as a pair from their putative offspring, although either genotype of the parent was compatible with the offspring.
The mismatches consisted of homozygotes apart from case 11 (at Cca8 and Cca13), case 12 (at Cca13) and case 15 (at Cca13). Neither genotype mismatches were exclusively between heterozygous loci; at least one of the mismatching genotypes in a comparison was a homozygote. In two cases, the mismatching genotypes between two individuals were based on more than one locus (case 11 and 13).

Table 11 Cases of parentage exclusions in the Eurasian beaver. Offspring genotypes mismatching the respective genotypes of one or both of the putative parent(s) are underlined. Single excluding alleles are shown in bold.

Case no.		ID	Cca4	Cca8	Cca13	Cca18	Cca19	Excluded
	father	Jørn	386/388	390/390	267/269	216/216	263/263	
	mother	Hanne	386/386	382/382	267/269	216/216		
1	offspring	Bram	386/386	390/390	267/267	216/216	263/263	mother
2	offspring	Bruno	386/386	382/382	267/269	216/216	263/263	father
	father	Ola By	-	382/390	269/269	216/216	263/263	
	mother	Christina	388/388	-	269/269	216/216	263/263	
3	offspring	#193	388/388	-	267/267	216/216	-	mother; father
	father	Helgenen	386/388	390/390	267/269	216/216	263/263	
	mother	Unni	388/388	390/390	267/269	216/216	263/263	
4	offspring	Hege	388/388	382/382	269/269	216/216	263/263	mother; father
5	offspring	Bård	388/388	382/382	269/269	216/216	263/263	mother; father
	father	Carl	386/388	382/390	269/269	216/216	263/263	
	mother	Stina	386/386	382/390		216/216	263/263	
6	offspring	Lasse	388/388		269/269	216/216	263/263	mother
7	offspring	Kyrgyz boy	388/388	382/390	269/26/	216/216	263/263	mother
	father	Harald	386/386	382/390	267/269	216/216		
	mother	Sonja	388/388	-	-	-	263/263	
8	offspring	Ari	386/386	-	267/269	216/216	263/263	mother
9	offspring	Mett-Marit	386/386	382/382	267/269	216/216	263/263	mother
	mother	Fatima	386/386	390/390	267/269	216/216	263/263	
10	offspring	Marcus	388/388	382/390	269/269	216/216	263/263	mother
	father	Horst	386/386	390/390	269/269	216/216	-	
	mother	Tanja	386/386	390/390	269/269	216/216	263/263	
11	offspring	Hildegunn	386/386	**382**/390	**267**/269	216/216	263/263	parent pair
12	offspring	Rocky II	386/386	390/390	**267**/269	216/216	263/263	parent pair
	father	Erlend #157	386/386	382/382	267/267	216/216	261/261	
	mother	Lisbeth	388/388	382/390	269/269	216/216	263/263	
13	offspring	Magnhild	386/388	390/390	369/269	-	263/263	father
	father	Tommy	388/388	-	267/267	-	-	
	mother	Olive	388/388	390/390	267/269	216/216	263/263	
14	offspring	Karin	386/386	382/390	267/269	216/216	263/263	mother; father
	father	Jon	386/386	-	269/269	216/216	263/263	
	mother	Trude	386/386	382/390	269/269	216/216/	263/263	
15	offspring	Villy	386/386	382/390	**267**/269	216/216	263/263	parent pair
	father	Kjartan	386/386	382/382	267/267	216/216	263/263	
16	offspring	Matthias	-	-	269/269	216/216	-	father

6 Discussion

6.1 DNA extraction

In this study, it was possible to extract amplifiable DNA from 91% (136 out of 148) of the hair samples. Four hairs samples consisted of hair shafts without a visible hair root, which might have provided too low DNA quantities to be extracted with the method that was applied. In the remaining eight samples, DNA extraction might have failed due to other circumstances. Remnants of PCR-inhibiting ingredients in the samples used during DNA extraction might have caused a failure in the PCR amplification. A procedure for DNA extraction where the hairs are digested in a PCR-compatible buffer would allow the processing of the hair without the chance of loss of sample material (Vigilant, 1999).

In this study, the available hair samples had been collected over a period of several years. Due to DNA degradation process that usually occurs over a certain period of time if storage conditions are not optimal, a certain amount of DNA in some of the hair samples might have become denaturated in a way that made the DNA extraction impossible. In order to increase the amount of extractable DNA in the samples, they should have been extracted within the shortest time possible after the sample collection.

6.2 DNA amplification and gel electrophoresis

The presence of extracted DNA in the samples was checked by using traditional polyacrylamide gel electrophoresis. Strong band signals on the gels were interpreted as the presence of a rather large quantity of PCR products in the samples. Since accurate measurement of DNA quantity in the extracted DNA samples was not possible, DNA differences in DNA quantities in the samples were compared only in relation to each other. Strong signals were interpreted as a high amount of PCR products in the PCR samples, which could either been due to an initial high amount of DNA present in the tissue, or the more successful PCR reactions compared to samples with weaker signals.

Due to the low resolution of the gel, it was not possible to determine whether the amplification products in a sample consisted of one or two alleles at a locus, which made the use of a fluorescence-based PCR technique necessary.

Although it was not possible to determine whether an amplification product in a sample consisted of a single or two different alleles, the gel electrophoresis served as a method to reveal the presence of DNA in the sample, and was used as a cheap and simple method for detecting lack or presence of extracted DNA in the samples primarily to the genotyping process.

6.3 PCR II and microsatellite genotyping

In the multiplex PCR process, several primers were successfully combined to amplify DNA and gave rise to genetic profiles.
Compared to singleplex PCR the multiplex PCR showed to be less time-consuming during both DNA amplification and interpretation of the PCR signal output.

Useful microsatellite markers should amplify in the majority of individuals in the target species and show sufficient polymorphism (Selkoe, Toonen, 2006b). In this study, both the amplification success rates and the marker polymorphisms were rather low. Since the microsatellites originally were designed for the North American beaver, it was likely to obtain a less complete genotyping profile when using the primers on the European beaver, as cross-specific primers tend to have a lower amplification success (Barbara *et al.*, 2007). One possible reason for this is the occurrence of mutations in the primer-binding sites in the non-target species (Selkoe, Toonen, 2006b) which prohibited the DNA amplification. These mutations might have developed by chance in the European beaver and not in the North American beaver population. As opposed to the extraction of DNA from blood, the DNA in hair is more exposed to degradation over time, and if the sample storage of hair samples becomes too long, there will be a significant decrease in microsatellite amplification success across time (Roon *et al.*, 2003). In this study, hair sample storage time varied widely across all samples. In order to compare the effect of storage time on the quality of DNA in the hair samples and the DNA amplification rate some further analyses will be necessary.

6.4 Data analysis

From the nine microsatellites tested in the study, five loci showed some degree of polymorphism. Loci Cca5, Cca9, CCa10 and Cca15 amplified only a single allele. When using cross-specific genetic markers, one might expect a lower degree of polymorphism in the study species compared to the species those markers were designed for. As for mammals, the mean percentage of polymorphic markers can be expected to be >40% when transferring markers between species within the same family and up to 70% when transferring between species within genus (Barbara *et al.*, 2007). With five of nine microsatellites (55.55%) designed in the North American beaver being polymorphic in the Eurasian beaver, the transferability of the markers lie in the expected range. However, two of the polymorphic loci (Cca18 and Cca19) displayed even lower heterozygosities than expected at 0.017 and 0.029. The least frequent alleles of the two alleles were amplified in two different individuals.

Practically, the applicability of these two markers in terms of population analyses is quite low. Although there is a probability of using these alleles in the traditional parentage exclusion tests, they do not display the preferred qualities of suitable microsatellite markers. Observed heterozygosities at loci Cca4, Cac8 and CCa13 ranged from 0.33 to 0.38, which is assumed to be rather low compared to an ideal genetic microsatellite marker exhibiting high heterozygosities ranging from 0.6 to 0.8 (Taberlet, Luikart, 1999b). As a consequence, the statistical power of these markers will be low and a larger number of polymorphic microsatellite loci would be necessary in order to conduct population genetic analyses that are based on marker statistics like allele frequencies and heterozygosities. However, in other studies, like parentage analyses based on the traditional parentage exclusion method, a few loci with a moderate polymorphism might be sufficient (Wan *et al.*, 2004). Having gone through a bottleneck no longer than a hundred years ago, the Eurasian beaver population in Norway shows a low degree of genetic variation (Ellegren *et al.*, 1993) which might also be expressed in low levels of microsatellite genetic diversity (Frankham *et al.*, 2002). However, levels of genetic diversity for microsatellites are not straightforward comparable between studies, as microsatellite mutation rates differ between species and repeat types. Furthermore, primers designed for one species can lead to lower levels of variation when used in other species (Ellegren *et al.*, 1997; Gunn *et al.*, 2005).

6.4.1 Genotyping errors

In earlier population ecology studies based on microsatellite genotyping, the error rate is still not reported, although strongly recommended. Due to this lack of comparable genotyping error rate data, in this study, general guidelines concerning the importance of genotyping error assessment are given rather than a comparison of error rate numbers.

The genotyping error rate in this study is 0.1 on the genotype level. For studies of relatedness like paternity exclusions, Taberlet and Luikart (1999b) suggest that an error rate of less than 1% may be required. In order to reduce the amount of ADOs, there are several possible strategies. Since the proportion of ADOs decreases with the increasing amount of template DNA in the samples (Morin *et al.*, 2001), it would be useful to obtain the largest possible amount of DNA under the DNA extraction process. Measures to achieve this are described earlier in this study (chapter 3.2.2 and 3.2.3) and include recommended procedures for the sample collection and preservation and DNA extraction. However, it has to be taken into consideration that increasing the amount of extract not only may improve PCR amplification process, but simultaneously increase the amount of PCR inhibitors (Morin *et al.*, 2001). To

ensure that sufficient DNA for reliable genotyping is provides after the extraction, the multiple tubes-approach (Goossens *et al.*, 1998b) is recommended to make the genotyping results more reliable.

In this study, the error rate caused by ADO is 8% at the allelic level comprising only two incidences of ADO in the process of repeated genotyping. This number might appear rather low, but the total number of alleles analyzed in this study is equally low. When increasing the number of alleles, as it might happen when increasing the number of genotyped individuals, the ADO rate will simultaneously increase. As it has been shown that the ADO rate increases with growing PCR product size (Sefc *et al.*, 2003), avoiding markers yielding larger product sizes is suggested as well as the possibility of redesigning the primers to amplify smaller fragments (Hoffman, Amos, 2005).

It was possible to detect the occurrence of false alleles in one case, resulting in a FA rate of 0.04 on the allelic level in this study. False alleles are produced under the PCR process due to strand slippage, and thus the use of markers consisting of tri- and tetranucleotides instead of dinucleotide repeats represents one means of slippage prevention (Schlötterer and Tautz, 1992 get ref).

Other suitable means of reducing genotyping errors are suggested by Ewen (2000) and comprises the amelioration of laboratory routines. Simple suggestions are double typing the sample sheets and checking the label concordance both before and after the PCR run and by running simple macros written in Microsoft Excel where incorrect sheets can easily be replaced. In this study, the use of a simple formula at the end of each row in a Microsoft Excel worksheet has proven useful as a simple error check.

Another source of genotyping errors is connected to the interpretation of band patterns on the capillary electrophoresis output. The conversion of the microsatellite pattern into genotypes is often complicated by both stutter bands and additional bands from non-template DNA. These complications arise especially in the analysis of dinucleotide markers when two alleles in a heterozygous individuals differ by only a single repeat unit (Johansson *et al.*, 2003) and may lead to confusion between homozygote and adjacent allele heterozygote genotypes. In this case, it is possible to distinguish between homozygotes and adjacent allele heterozygotes by comparing them with the pattern of known single alleles (Hoffman, Amos, 2005).

Currently, special algorithms are being developed in order to automate the band interpretation process (Johansson *et al.*, 2003). An alternative to automated signal interpretation is to choose microsatellites based on longer repeat units.

An easy and most widely accessible form of error-checking is the comparison of mother and offspring genotype mismatches with known field data. Usually, these comparisons detect only a subset of errors (Hoffman, Amos, 2005). In this study, several mismatches between a mother and her offspring were found (10 out of 40 comparisons between offspring and mother genotypes). This indicates a higher presence of undetected false homozygotes. One alternative, less probable explanation for this finding is that these offspring – mother relationships indeed are not true but based on falsely interpreted or recorded field observation. Taking into consideration the general consensus in current literature on population genetic analyses based on microsatellite data, in this study it was chosen to follow the advice of not accepting the exclusions of previously assumed mother – offspring relationships. Rather, theses results are used as a means of error checking as suggested by Hoffman and Amos (2005) and Gagneux et al (1997).

Finally, although the error rate calculated in this study was not directly comparable to error rates from similar studies due to the lack of standard calculation methods, the current emphasis in published reviews seems to lie on the importance of reporting the error rate in itself. Reporting the error rate should be regarded as a means to evaluate data quality rather than discrediting the final results and conclusions from the data. It has even been suggested that stating the genotyping error rate in population genetic studies should become a convention similar to the *P*-value in statistical tests (Bonin *et al.*, 2004).

6.4.2 False exclusion probabilities

The calculation of false exclusion probabilities was used as a way of assessing the single-locus and overall probability of falsely excluding a true parent. In this study, the calculated false exclusion probabilities are quite high for locus Cca4 and Cca19. Exclusions based on these loci will imply a certain risk of falsely excluding the actual parent, 16.44% and 20.34%, respectively. Those numbers were taken into account when interpreting the results gained from the parentage exclusions.

6.4.3 Parentage testing and exclusions

Although the PCR run with the microsatellites Cca4, Cca8, Cca13, Cca18 and Cca19 could not provide complete genotypes for all individuals genotyped, it was possible to test parentage in some parent-offspring relationships. The relationships between parental and offspring individuals were constructed based on the frequencies of observations of the individuals in the same or nearby locations, which suggested possible kinship. Relationships were tested pair-wise between father or mother and the possible offspring and between parents and offspring.

In the case of compatible genotypes, the relationships were accepted as possibly true. Mismatching genotypes led to the assumption that the relationship was not true, despite earlier observations.

Likelihood based parentage analysis methods that assign parents to their offspring usually assume parameters that were not given in this study. Consequently, in the case of parent – offspring exclusions, no alternative parent was considered and tested. Especially the incomplete sampling of candidate parents in this study prevented the assignment of parents to offspring since the exclusion of all but one parent candidate here was not possible. In order to assign paternity or maternity to an offspring, the genotypes of all candidate parents should be available and not only a fraction like it was the case in this study. One other assumption that was not met was that all microsatellite loci be in Hardy Weinberg equilibrium. Three of the five microsatellites used in this study did deviate from the HWE and thus were not appropriate for likelihood based parentage assignment methods. Also, locus Cca4 showed a null allele frequency of >0.2, which suggests removing the locus from parentage analysis that are based on exclusion probabilities.

The traditional exclusion method which is more of a binary character rather than based on likelihood was found to be the most feasible parentage method in order to obtain information from the genotyped samples in this study. In all but three cases, the exclusions are based exclusively on homozygous genotypes in both offspring and parent(s). In case 11, there are found two alleles in the offspring's genotype at two loci that mismatch the genotypes of the homozygous parent pair. In case 12 and 15, the heterozygous genotypes of the offspring include a single mismatching allele which excludes them genetically from their parents (as a pair). Theoretically, a single allele mismatch could be treated as a conclusive evidence for exclusion of a parent from parentage. Practically however, this mismatch could have resulted from either the genuine non-relationship between the individuals or from genotyping errors. As the occurrence of genotyping errors is accepted and taken into account in this study, recommendations are followed concerning parent-offspring mismatches based on homozygote-homozygote genotypes. Where observational data strongly suggested that the genetically mismatching parent-offspring relationship was true, the exclusions were not accepted, as in the case of the genetically mismatching mother-offspring genotypes. The homozygote excess in the genotype frequencies suggests the possible occurrence of null alleles, which might have led to masked heterozygotes and false exclusions. In order to avoid these false exclusions, it was chosen to rely on data obtained by long-term observations rather than genotypes. One option of verifying the genetic data would be the repeated extraction and

amplification of DNA from these samples in order to obtain more reliable genotypes. However, in the case of null alleles even frequently repeated PCR runs might not lead to a more complete amplification as mutations in the flanking regions prevents primer binding. As an alternative, primers might be redesigned in order to avoid the mutated sequence(s).

In the course of this study, genetic research on the North American beaver, previously associated with monogamous mating systems, showed that extra-pair mating within or between colonies had occurred (Crawford *et al.*, 2008b). Thus, it is possible that some of the parent-offspring exclusions resulting from this study actually are true, and should not, as has been done here, be rejected and used as a means of detecting genotyping errors. However, the quality and applicability of the genetic tools used in this study are not as high as those tools used in the study on the North American beaver. It remains to be seen what results could be achieved with more appropriate genetic tools on the beaver population in Telemark. Microsatellites displaying a higher variability, a larger percentage of extractable DNA from hair samples and a more complete sampling of beaver colonies could provide results more suitable to supply the ongoing population ecology research on this field with information about the genetic composition of the beaver population in Telemark.

References

ABgene *Thermo Fisher Scientific Inc.*

Anderson (2006) Non-invasive genetic sampling of the eurasian otter (LUTRA LUTRA) using hairs. *Hystrix Italian Journal of Mammalogy* **17**, 65-77.

Applied Biosystems *California, USA.*

Avise JC (2004) *Molecular Markers, Natural History, and Evolution*, 2 edn. Sinauer Associates; 2 edition

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.

Barbara T, Palma-Silva C, Paggi GM, *et al.* (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Molecular Ecology* **16**, 3759-3767.

Beebee TJC, Rowe G (2004) *An introduction to molecular ecology* Oxford University Press.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2008) GenBank. *Nucl. Acids Res.* **36**, D25-30.

Bergl RA, Vigilant L (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology* **16**, 501-516.

Bio-Rad *Bio-Rad Laboratories, Inc, California, USA.*

Blouin MS, Parsons M, Lacaille V, Lotz S (1996) Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* **5**, 393-401.

Bonin A, Bellemain E, Bronken Eidesen P, *et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261-3273.

Broquet T, Petit E (2004) Quantifying genotyping errors in noninvasive population genetics. *Molecular Ecology* **13**, 3601-3608.

Burland MT, Bennett N, C. , Jarvis JUM, Faulkes CG (2002) Eusociality in African mole-rats: new insights from patterns of genetic relatedness in the Damaraland mole-rat (*Cryptomys damarensis*). *Proc. R. Soc.Lond. B:* **269**, 1025-1030.

Burton C (2002) Microsatellite analysis of multiple paternity and male reproductive success in the promiscuous snowshoe hare. *Canadian Journal of Zoology* **80**, 1948-1956.

Calabrese P, Sainudiin R (2005) Models of Microsatellite Evolution. *Statistical Methods in Molecular Evolution*, 289-306.

Carpenter P, Pope LC, Greig C, *et al.* (2005) Mating system in the Eurasian badger, Meles meles, in a high density population. *Molecular Ecology* **14**, 273-284.

Chambers GK, MacAvoy ES (2000) Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology -- Part B: Biochemistry and Molecular Biology* **126**, 455-476.

Constable JL, Ashley MV, Goodall J, Pusey AE (2001) Noninvasive paternity assignment in Gombe chimpanzees. *Molecular Ecology* **10**, 1279-1300.

Crawford AM, Cuthbertson RP (1996) Mutations in sheep microsatellites. *Genome Res.* **6**, 876-879.

Crawford J, Liu Z, Nelson C, Bloomquist C (2008a) Isolation and characterization of microsatellite loci in the beaver (*Castor canadensis*). *Molecular Ecology Resources* **8**, 616-618.

Crawford JC, Liu Z, Nelson TA, Nielsen CK, Bloomquist CK (2008b) Microsatellite analysis of mating and kinship in beavers (Castor canadensis). *Journal of Mammalogy* **89**, 575-581.

Creel S, Spong G, L SJ (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology* **12**, 2003-2009.

Dakin E, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* **93**, 504-509.

David VA, Menotti-Raymond M (1998) Automated DNA detection with fluorescent-based technologies. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR). Oxford University Press Inc., New York.

Dean M, Milligan B (1998) Detection of genetic variation by DNA conformational and denaturing gradients methods. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR). Oxford University Press Inc., New York.

Devlin B, Roeder K, Ellstrand NC (1988) Fractional paternity assignment: theoretical development and comparison to other methods. *Theoretical and Applied Genetics* **76**, 369-380.

DiRienzo AD, Peterson AC, Garza JC, *et al.* (1994) Mutational Processes of Simple-Sequence Repeat Loci in Human Populations. *Proceedings of the National Academy of Sciences* **91**, 3166-3170.

Eisen JA (1999) Mechanistic basis for microsatellite instability. In: *Microsatellites - Evolution and Applications* (eds. Goldstein DB, Schlötterer C), pp. 34-48. Oxford University Press, Oxford.

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* **16**, 551-558.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* **5**, 435-445.

Ellegren H, Hartman G, Johansson M, Andersson L (1993) Major Histocompatibility Complex Monomorphism and Low Levels of DNA Fingerprinting Variability in a Reintroduced and Rapidly Expanding Population of Beavers. *Proceedings of the National Academy of Sciences* **90**, 8150-8153.

Ellegren H, Moore S, Robinson N, *et al.* (1997) Microsatellite evolution--a reciprocal study of repeat lengths at homologous loci in cattle and sheep. *Mol Biol Evol* **14**, 854-860.

Eppendorf *Biochip Systems GmbH, Hamburg, Germany*.

Ewen KR, Bahlo M, Treloar SA, *et al.* (2000) Identification and Analysis of Error Types in High-Throughput Genotyping. *American journal of human genetics* **67**, 727-736.

Fernando P, Evans BJ, Morales JC, Melnick DJ (2001) Electrophoresis artefacts-a previously unrecognized cause of error in microsatellite analysis. *Molecular Ecology Notes* **1**, 325-328.

Fernando P, Vidya TNC, Rajapakse C, Dangolla A, Melnick DJ (2003) Reliable Noninvasive Genotyping: Fantasy or Reality? *J Hered* **94**, 115-123.

Ferrando A, Lecis R, Domingo-Roura X, Ponsà M (2007) Genetic diversity and individual identification of reintroduced otters (Lutra lutra) in north-eastern Spain by DNA genotyping of spraints. *Conservation Genetics* **online (1)**, 1-11.

Foran DR, Minta SC, Heinemeyer KS (1997) DNA-based analysis of hair to identify species and indivuals for population research and monitoring. *Wildlife Society Bulletin* **25**, 840-847.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to conservation genetics* Cambridge University Press, Cambridge.

Frantz AC (2004) Non-invasive genetic typing in the study of badger (*Meles meles*) Ecoloy. *PhD thesis, University of Sussex*.

Frantz AC, Schaul M, Pope LC, *et al.* (2004) Estimating population size by genotyping remotely plucked hair: the Eurasian badger. *Journal of Applied Ecology* **41**, 985-995.

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* **6**, 861-868.

Galan M, Baltzinger C, Hewison MAJ, Cosson J-F (2003) Deer species identification using DNA extracted from hair samples and the polymerase chain reaction (PCR)method. *?*

Ginot F, Bordelais I, Nguyen S, Gyapay G (1996) Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. *Nucl. Acids Res.* **24**, 540-541.

Goldstein DB, Schlötterer C (1999) *Microsatellites. Evolutions and Applications* Oxford University Press Inc., New York.

Gondek A, Verduijn M, Wolff K (2006) Polymorphic microsatellite markers for endangered spotted suslik, Spermophilus suslicus. *Molecular Ecology Notes* **6**, 359-361.

Goossens P, Graziani L, Waits LP, *et al.* (1998a) Extra-pair paternity in the monogamous Alpine marmot revelaed by DNA microsatellite analysis. *Behaviour Ecol. Sociobiol.*, 281-288.

Goossens P, Waits LP, Taberlet P (1998b) Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology* **7**, 1237-1241.

Gunn MR, Dawson DA, Leviston A, *et al.* (2005) Isolation of 18 polymorphic microsatellite loci from the North American red squirrel, Tamiasciurus hudsonicus (Sciuridae, Rodentia), and their cross-utility in other species. *Molecular Ecology Notes* **5**, 650-653.

Hancock JM (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites - Evolution and Applications* (eds. Goldstein DB, Schlötterer C), pp. 1-9. Oxford University Press, Oxford.

Hedrick PW (1996) Conservation genetics and molecular techniques: a perspective. In: *Molecular genetic approaches in conservation* (eds. Smith T, Wayne RK), pp. 459-477. Oxford University Press, New York.

Hoelzel AR, Green A (1998) PCR Protocols and population analysis by direct DNA sequencing and PCR-based DNA fingerprinting. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR). Oxford University Press Inc., New York.

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**, 599-612.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* **11**, 424-429.

Johansson Å, Karlsson P, Gyllensten U (2003) A novel method for automatic genotyping of microsatellite markers based on parametric pattern recognition. *Human Genetics* **113**, 316-324.

Jones AG (2001) GERUD1.0: a computer program for the reconstruction of parental genotypes from progeny arrays using multilocus DNA data. *Molecular Ecology Notes* **1**, 215-218.

Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. *Molecular Ecology* **12**, 2511-2523.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099-1106.

Kholodova MV, Easton E, Milner-Gulland EJ (2000) The Use of Fur Collected in the Field for Studying the Genetic Diversity of Wild Ungulates. *Biology Bulletin* **27**, 585-590.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725-738.

King RC, Stansfiedl RD, Mulligan PK (2006) *A Dictionary of Genetics*, 7 edn. Oxford University Press Inc, USA, New York.

Kocher TD, Thomas WK, Meyer A, *et al.* (1989) Dynamics of Mitochondrial DNA Evolution in Animals: Amplification and Sequencing with Conserved Primers. *Proc. Nati. Acad. Sci.* **86**, 6196-6200.

Kruckenhauser L, Pinsker W (2004) Microsatellite variation in autochthonous and introduced populations of the Alpine marmot (*Marmota marmota*) along a European west-east transect. *Journal of Zoological Systematics and Evolutionary Research* **42**, 19-26.

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* **4**, 203-221.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175-195.

Li Y-C, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* **11**, 2453-2465.

Lilland B-G, Bensch S, von Schantz T (2001) Parentage determination in kin-structured populations: microsatellite analsyses in the Siberian jay (*Perisoseus infaustus*) during a 25-year population study. *Avian Science* **1**, 3 - 14.

Lorenzini R, Posillico M, Lovari S, Petrella A (2004) Non-invasive genotyping of the endangered Apennine brown bear: A case study not to let one's hair down. *Animal Conservation* **7**, 199-209.

Luikart G, England PR (1999) Statistical analysis of microsatellite DNA data. *Trends in Ecology & Evolution* **14**, 253-256.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655.

McCulloch SD, Kunkel TA (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* **18**, 148-161.

Milligan BG (1998) Total DNA isolation. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR), pp. 29-64. Oxford University Press Inc., New York.

Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology*, 1835-1844.

Morin PA, Wallis J, Moore JJ, Woodruff DS (1994) Paternity exclusion on a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular Ecology* **3**, 469-478.

Morin PA, Woodruff DS (1996) Noninvasive genotyping for vertebrate conservation. In: *Molecular genetic approaches in conservation* (eds. Smith T, Wayne RK), p. 298313. Oxford University Press, New York.

Mowat G, Strobeck C (2000) Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. *Journal of Wildlife Management*, 183-193.

Natoli A, Peddemors VM, Rus Hoelzel A (2004) Population structure and speciation in the genus Tursiops based on microsatellite and mitochondrial DNA analyses. *Journal of Evolutionary Biology* **17**, 363-375.

Navidi W, Arnheim N, Waterman MS (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR-statistical considerations. *American journal of human genetics* **50**, 347-359.

Nolet BA, Rosell F (1998) Comeback of the beaver Castor fiber: an overview of old and new conservation problems. *Biological Conservation* **83**, 165-173.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetical Research* **22**, 201-204.

Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology* **12**, 1375–1387.

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* **3**, 489-495.

Parker H, Rosell F (2003) Beaver management in Norway: a model for continental Europe? *Lutra* **46**, 223-234.

Parker PG, Waite TA, Peare T (1996) Paternity studies in animal populations. In: *Molecular Genetic Approaches in Conservation* (eds. Smith TB, Wayne RK), pp. 413-423. Oxford University Press, New York Oxford.

Petes TD, Greenwell PW, Dominska M (1997) Stabilization of Microsatellite Sequences by Variant Repeats in the Yeast Saccharomyces cerevisiae. *Genetics* **146**, 491-498.

Pierce BA (2003) *Genetics: A conceptual approach*, 1 edn. W.H. Freeman and Company, New York.

QIAGEN® (2006) User-Developed Protocol: Purification of total DNA from nails, hair, or feathers using the DNeasy® Blood & Tissue Kit. 1-3.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology & Evolution* **8**, 285-288.

Quiagen *Quiagen GmbH, Hilden, Germany*.

Romain-Bondi KA, B.Wielgus R, Waits L, *et al.* (2003) Density and population size stimates for North Cascade grizzly bears using DNA hair-sampling techniques. *Biological Conservation* **117**, 417-428.

Roon DA, Waits LP, Kendall KC (2003) A quantitative evaluation of two methods for preserving hair samples. *Molecular Ecology Notes* **3**, 163-166.

Rosell F, Pedersen KV (1999) *Bever* Landbruksforlaget.

Rousset F (2007) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. . *Mol. Ecol. Notes in press.*

Saiki RK, Gelfand D, Stoffel S, *et al.* (1985) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Scandura, Capitani, Iacolina (2006) An empirical approach for reliable microsatellite genotyping of wolf DNA from multiple noninvasive sources. *Conservation Genetics* **7**, 813-823.

Scandura M (2005) Individual sexing and genotyping from blood spots on the snow: A reliable source of DNA for non-invasive genetic surveys. *Conservation Genetics* **6**, 871-874.

Schlötterer C (1998) Microsatellites. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR), pp. 237-261. Oxford University Press Inc., New York.

Schlötterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. *Nucl. Acids Res.* **20**, 211-215.

Schnabel RD, Ward TJ, N DJ (2000) Validation og 15 microsatelittes for parentage testing in North American bison, Bison bison and domestic cattle. *Animal Genetics* **31**, 360-366.

Schwartz MK, Pilgrim KL, McKelvey KS, *et al.* (2004) Hybridization between Canada lynx and bobcats: Genetic results and management implications. *Conservation Genetics* **5**, 349-355.

Sefc MK, Payne RB, Sorensen MD (2003) MIcrosatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. *The Auk* **120**, 1-8.

Selkoe KA, Toonen RJ (2006a) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Selkoe KA, Toonen RJ (2006b) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Semple K, Wayne RK, Gibson RM (2001) Microsatellite analysis of female mating behaviour in lek-breeding sage grouse. *Molecular Ecology* **10**, 2043-2048.

Sloane MA, Sunnucks P, Alpers D, Beheregaray LB, Taylor AC (2000) Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censusing method. *Molecular Ecology* **9**, 1233-1240.

Syngene *Synoptics Ltd, Cambridge, UK*.

Taberlet P, Camarra JJ, Griffin S, *et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology* **6**, 869-876.

Taberlet P, Griffin S, Goossens P, *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleid Acids Research* **24**, 3189-3194.

Taberlet P, Luikart G (1999b) Non-invasive genetic sampling and invidiual identification. *Biological Journal of the Linnean Society* **68**, 41-55.

Taberlet P, Waits LP, Luikart G (1999a) Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution* **14**, 323-327.

Takami K, YOSHIDA M, YOSHIDA Y, KOJIMA Y (1998) Sex Determination in Giant Anteater (*Myrmecophaga tridactyla*) using Hair Roots by Polymerase Chain Reaction Amplification. *J. Reprod. Dev.* **44,**, 73-78.

Taylor AC, Horsup A, Johnson CN, Sunnucks P, Sherwin B (1997) Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an

endangered marsupial, the northern hairy-nosed wombat Lasiorhinus krefftii. *Molecular Ecology* **6**, 9-19.

Toth G, Gaspari Z, Jurka J (2000) Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Res.* **10**, 967-981.

Vigilant (2002) Technical Challenges in the Microsatellite Genotyping of a Wild Chimpanzee Population Using Feces. *Evolutionary Anthropology* **Suppl 1**, 162-165.

Vigilant L (1999) An Evaluation of Techniques for the Extraction and Amplification of DNA from Naturally Shed Hairs. *Biol. Chem.* **380**, 1329-1331.

Waits L, Taberlet P, Swenson JE, Sandegren F, Franzen R (2000) Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (Ursus arctos). *Molecular Ecology* **9**, 421-431.

Walker FM, Sunnucks, taylor (2006) Genotyping of "captured" hairs reveals burrow-use and ranging behavior of southern hairy-nosed wombats. *Journal of Mammalogy* **87**.

Wan Q-H, Wu H, Fujihara T, Fang S-G (2004) Which genetic marker for which conservation genetics issue? *Electrophoresis* **25**, 2165-2176.

Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998) Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in Gracilaria gracilis (Rhodophyta). *Molecular Ecology* **7**, 1569-1573.

Weller JI, Seroussi E, Ron M (2006) Estimation of the number of genetic markers required for individual animal identification accounting for genotyping errors. *Animal Genetics* **37**, 387-389.

Wierdl M, Dominska M, Petes TD (1997) Microsatellite Instability in Yeast: Dependence on the Length of the Microsatellite. *Genetics* **146**, 769-779.

Xu X, Peng M, Fang Z, Xu X (2000) The direction of microsatellite mutations is dependent upon allele length. *Nat Genet* **24**, 396-399.