From Faeces to Ecology and Behaviour – Non-Invasive Microsatellite Genotyping as a Means to Study Wild Otters (*Lutra lutra*)

Dissertation

To Fulfil the

Requirements for the Degree of

„doctor rerum naturalium“ (Dr. rer. nat.)

Submitted to the Council of the Faculty

of Biology and Pharmacy

of the Friedrich Schiller University Jena

by Dipl. Biologist

Simone M. Lampa

born on 24.03.1980 in Hünfeld, Germany

Supervised by:

Prof. Dr. Klaus Henle, Helmholtz-Centre for Environmental Research – UFZ, Leipzig, Germany

Prof. Dr. Bernd Gruber, University of Canberra, Australia

Prof. Dr. Stefan Halle, University of Jena, Germany

Reviewer:

Prof. Dr. Günter Köhler, University of Jena, Germany

Prof. Dr. Klaus Henle, Helmholtz-Centre for Environmental Research – UFZ, Leipzig, Germany

Prof. Dr. Niko Balkenhol, University of Göttingen, Germany

Submitted to the University of Jena: 24.10.2014

Date of public presentation and thesis defence: 23.04.2015

Table of Content

Chapter One

Introduction 1

1.1 General Introduction 2

1.2 Aim and Structure of the Dissertation 4

1.3 The Eurasian Otter 5

1.4 The Study Area 8

1.5 General Methodology 10

1.6 Overview of Manuscripts 12

Chapter Two

An Optimisation Approach to Increase DNA Amplification Success of Otter Faeces 15

Chapter Three

How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation – A Review of Available Methods Illustrated by a Case Study 26

Chapter Four

Non-Invasive Genetic Mark-Recapture as a Means to Study Population Sizes and Marking Behaviour of the Elusive Eurasian Otter (*Lutra lutra*) 49

4.1 Abstract 50

4.2 Introduction 50

4.3 Methods 52

4.4 Results 57

4.5 Discussion 62

4.6 Acknowledgements 68

Chapter Five

Non-Invasive Genetic Mark-Recapture as a Means to Study Population Dynamic and Spatial Use of Eurasian Otters (*Lutra lutra*) in a fish pond landscape 69

5.1 Abstract 70

5.2 Introduction 70

5.3 Methods 72

5.4 Results 77

5.5 Discussion 82

5.6 Acknowledgements 86

Chapter Six

Discussion 87

6.1 Overview of Research Outline and Main Results 88

6.2 Key Findings 89

6.3 Limitations and Methodological Constraints 93

6.4 Suggestions for Further Research 95

6.5 Suggestion for Otter Conservation 96

Summary 98

Zusammenfassung 100

Acknowledgements 102

References 104

Ehrenwörtliche Erklärung 111

Appendix to Chapter Three 112

Chapter 1

Introduction

1.1 General Introduction

The Eurasian otter (*Lutra lutra*, Linnaeus, 1758) is a member of the *Mustelidae* or weasel family of mammals. As a semi-aquatic carnivore they inhabit coastal and freshwater habitats including all kinds of running and standing waters as well as wetlands (Kruuk 2006). Their main prey is fish, which brings them into direct competition with humans who use fish either as food (e.g. aquaculture) or for recreation (e.g. angling) (Santos-Reis et al. 2013). Also their fur was highly coveted, leading to a heavy hunting pressure especially in the late 19th/early 20th century. Furthermore, river regulation and canalisation, destruction of riparian vegetation, intensified agricultural land use accompanied by a loss of structural diversity, as well as water pollution, draining of wetlands, and a decrease in prey species resulted in a massive population decline all over their distribution range, but especially in Europe (Kruuk 2006; Ruiz-Olmo et al. 2008). Here, the otter vanished in many parts of Middle Europe such as Western Germany, the Netherlands, Belgium, Luxembourg, Eastern France, Switzerland, parts of Austria, or Central Italy (Ruiz-Olmo et al. 2008). As a consequence, the otter received protection statuses from the Bern Convention (1979 – strictly protected), the Habitats Directive (1992 – Annexes II and IV), the convention on international trade in endangered species (CITES; 1977 – Appendix I), and the world conservation union (IUCN; 2000 – vulnerable; 2004 – near threatened). This protection might have been one of several reasons why otters started to recolonise former haunts in Europe within the last decades (Ruiz-Olmo et al. 2008). In Germany, the remaining populations of Eastern Saxony, Brandenburg, and Mecklenburg-Western Pomerania started to rise and to expand towards west (Reuther 2004), despite the increasing road-kill risk that is the major threat nowadays (Hauer et al. 2002a). But our knowledge about otters is still very limited. For example, Kruuk (2006) stated that "we still know little about actual numbers over larger areas, about population sizes, and about changes in areas where previous estimates have been made." Also little is known about the process and speed of the recolonisation in Germany, about migration routes, or about the population dynamic and population numbers of the source populations in Eastern Germany. But a solid basis for conservation management also in the face of environmental changes requires information on population dynamic and actual numbers to e.g. understand the current spread.

However, studying otters is rather challenging since they are difficult to observe directly because of their mostly nocturnal activity (Ruiz-Olmo et al. 2008) and they are also difficult to live-trap (Kruuk 2006). So far, there are only few studies using telemetry (Durbin 1998; Ó Néill et al. 2008; Quaglietta et al. 2012), since there are several problems associated with this technique. First, the animal has to be live-trapped; this requires usually an official permit and veterinary assistance. Then, the telemetry device has to be affixed to the otter, either externally or by abdominal implantation that involves a consequential risk for the animal. Standard collars are also risky and not recommended for otters (Kruuk 2006). But Quaglietta et al. (2012) just recently reported of a new approach that use harnesses with a GPS-GPRS device. However, the obtained information is often only received by a few individuals (e.g. five otters by Durbin (1998)) and is usually about spatial use or spatial organisation

(Durbin 1998; Ó Néill et al. 2009; Quaglietta et al. 2014). Direct observations were used in several studies of Hans Kruuk and colleagues to gain information on otter numbers, individual ranges or spatial organisation (Kruuk & Moorhouse 1991), on recruitment (Kruuk et al. 1991), or on scent marking behaviour (Kruuk 1992). However, Kruuk (2006) conceded that it required several years to be able to distinguish individuals by ear-tags or characteristic throat patches to receive such information. Furthermore, direct observations were only possible because otters at the coast of Shetland were diurnal (Kruuk 2006). To imply population sizes, Kruuk et al. (1989) also counted otter holts along the coast of Shetland. Here, holts were easy to find, but in freshwater areas this method was unsuitable (Kruuk 2006). Another approach with which one can obtain ecological information (e.g. body conditions, reproductive performance) is to collect carcasses. This method was applied several times (e.g. Kruuk & Conroy 1991; Ansorge et al. 1997; Ruiz-Olmo et al. 1998; Elmeros & Madsen 1999; Hauer et al. 2002b, a), but requires an elaborate system to find and collect dead animals, to store them and is usually done over a long period to receive reliable information. Moreover, some information drawn from the data (e.g. sex ratio, age pattern) can be biased if the probability to die is not equally distributed among age and gender.

Instead of directly observing or handling the animal, there is also an indirect way by searching for their tracks or faeces. Especially the latter are easy to find as they are usually placed on conspicuous points throughout an otter's home range and were often used for monitoring purposes (Mason & Macdonald 1987) or for diet analysis (e.g. Almeida et al. 2012). But inferring from the number of found faecal samples on the number of animals was as often criticised (Kruuk et al. 1986; Chanin 2003) and although faeces indicate that otters are present, it is not valid vice versa (Kruuk et al. 1986). However, since each faecal sample contain sloughed gut cells from the originator, DNA techniques such as microsatellite genotyping (Bruford & Wayne 1993) can be used to assign an individual genetic fingerprint to each sample (Kohn & Wayne 1997). Microsatellites – also known as simple sequence repeats (SSRs) (Tautz 1989) or short tandem repeats (STRs) (Edwards et al. 1991) – consist of tandemly repeated sequences of 2–6 base pairs (bp) up to a total length of < 1000 bp. They are spread throughout the genome, very common, and highly polymorphic due to their high mutation rate (10^{-2}–10^{-6}) (Hancock 1999). That makes them to excellent markers for individual identification (Bruford & Wayne 1993). This so-called non-invasive genetic sampling was first used on wild animals in the 1990ies (Höss et al. 1992; Taberlet & Bouvet 1992). If individuals are genetically tagged, repeated sampling enables to track them in time and space, producing capture-recapture histories that can be applied to respective capture-mark-recapture (CMR) models. Non-invasive genetic sampling and the combination with CMR models can provide diverse information about population size and dynamic (e.g. survival, migration, growth rate, fecundity), behavioural biology, home range and territory size, genetic variation, phylogeography, relatedness, gene flow, as well as diet and diseases (Queller et al. 1993; Kohn & Wayne 1997; Taberlet et al. 1999; Lukacs & Burnham 2005b). For otters microsatellite genotyping was only available since Dallas and Piertney (1998) designed primers for 13 polymorphic

microsatellites, that were first applied on wild-collected faecal samples by Dallas et al. (2003). In this work, the authors applied the method on carcasses and on wild-collected faecal samples and demonstrated that both sample groups generated similar estimates of population genetic composition and sex ratio, suggesting that faeces can be used to derive such information. Also for population size it was shown that microsatellite genotyping using faecal samples produced reliable estimates compared to classical field methods either by simulation (Petit & Valière 2006) or on wild mammal populations (Solberg et al. 2006; Guschanski et al. 2009; Marucco et al. 2012), also for otters (Arrendal et al. 2007). In those studies, non-invasive genetic methods often revealed to be even cheaper and more accurate than classical field methods.

The drawback of non-invasive genetic genotyping is the low quality of the samples involving low success rates, the problem of genotyping errors, and contamination susceptibility. Genotyping errors occur when the observed genotype is not corresponding to the true genotype of an individual (Pompanon et al. 2005). The lower the quality of a sample, the higher is the genotyping error rate. However, a bunch of methods were developed within the last years that either increase success rates and hence the probability of a correct genotype (Piggott et al. 2004; Beja-Pereira et al. 2009), minimises genotyping errors (Taberlet et al. 1996; Frantz et al. 2003), detect and quantify them (Miller et al. 2002; McKelvey & Schwartz 2005), or incorporate them into subsequent statistical analysis (Lukacs & Burnham 2005a; Wang 2007; Wright et al. 2009).

1.2 Aim and Structure of the Dissertation

With this thesis I aim to contribute to the research on the threatened Eurasian otter that is of high conservation concern, because of its earlier massive decline and range contractions, its important role as a top predator in its ecosystem (Ripple et al. 2014), and because of its current increase and expansion in Germany that evokes conflicts with humans living on aquaculture. To understand the population dynamic and hence the current spread and to be able to manage it, we require knowledge that is either unavailable, not well understood, or has to be checked whether it also applies to fish pond systems, the main otter habitat in the Upper Lusatia that is the main source population for the recent expansion in Saxony, Germany. For this purpose, I decided to use non-invasive genetic CMR methods to gain information about actual population sizes, population dynamic parameters, marking behaviour, and spatial use of one source population in Eastern Germany. To make this method more efficient, especially for otter scats, I first optimised the required genetic methods to receive high success rates and low genotyping errors rates (Chapter two). Following this, I applied non-invasive genetic CMR to the first sampling year (2006) to demonstrate pitfalls and risks of this method and present a road map in which I offer solutions for the outlined problems (Chapter three). This road map is not only valid for otters, but is written as a general guideline when using non-invasively collected samples with low quality (see also Appendix). Finally, I used this road map to obtain reliable estimates and information on population size, sex ratio, and marking behaviour (Chapter four), as well as survival, temporary

migration, dispersal, and spatial use (Chapter five) over six sampling years (2006–2012). Chapter six completes the thesis by first giving an overview of the conducted research, followed by a synthesis of the four key findings and a discussion about limitations and methodological constraints. The chapter closes with suggestions for further research and for otter conservation.

1.3 The Eurasian Otter

The Eurasian otter is widely distributed covering parts of Europe, Asia, and Northern Africa (Ruiz-Olmo et al. 2008). In Europe the principal occurrences are divided into a "western distribution area" with Portugal, most parts of Spain, Western France, Britain, and Ireland and an "eastern distribution area" including Eastern Germany, Poland, Czech Republic, Slovenia, the Balkans (Southeast Europe), the Baltic states, Finland, and parts of Sweden, Norway, and Denmark (Fig. 1.1). In Germany, the eastern states Saxony, Brandenburg, and Mecklenburg-Western Pomerania are nowadays nearly nationwide inhabited. Otters can also be found again in Saxony-Anhalt, Thuringia, Schleswig-Holstein, Lower Saxony, and Bavaria (BfN 2012) (Fig. 1.2).

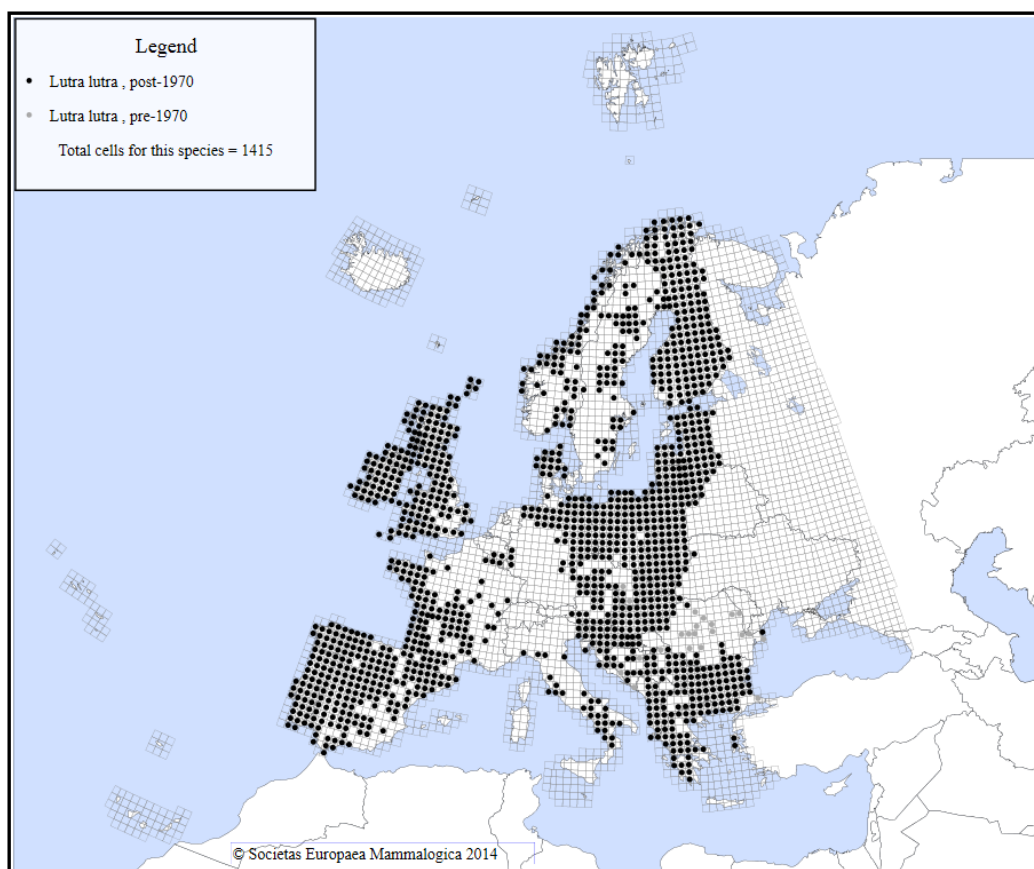

Figure 1.1 Distribution of the Eurasian otter in Europe. Black dots indicate confirmed otter presence after 1970, grey dots before 1970. Source: European Mammal Society (http://www.european-mammals.org/php/showmap.php?latname=Lutra+lutra&latname2=; [accessed: 22.04.2014]).

The otter is a medium-sized carnivore with an average body weight of 10 kg for males and 7 kg for females and a total length (including the tail) of about 1.2 m for a large male and 1 m for females (Kruuk 2006). They are semi-aquatic inhabiting all kind of water bodies such as lakes, ponds, rivers,

streams, marshes, swamps and coastal areas. Their diet mainly consists of fish representing up to 80%, but they also feed on aquatic insects, crustaceans, amphibians, reptiles, birds, or small mammals, depending on the region, habitat, and the season (Clavero et al. 2003; Almeida et al. 2012). It is mostly crepuscular and nocturnal, but Kruuk (2006) found diurnal active otters in coastal habitats and supposed the activity pattern to be reversely bound to the activity of their prey. The otter lives predominantly solitary. It occupies a home range with a core area that usually does not overlap with core areas of other adult otters. Although the home range itself can overlap between females and between opposite sexes (Erlinge 1968; Kruuk 2006; Quaglietta et al. 2014). Many studies found home ranges of males to be larger than those of females depending on the reproduction status (Sjöåsen 1997), residency (Kruuk & Moorhouse 1991), and the age (Arrendal 2007). On lakes and streams, female home ranges ranged between 1–12 km in length and male home ranges between 10–21 km (Erlinge 1967).

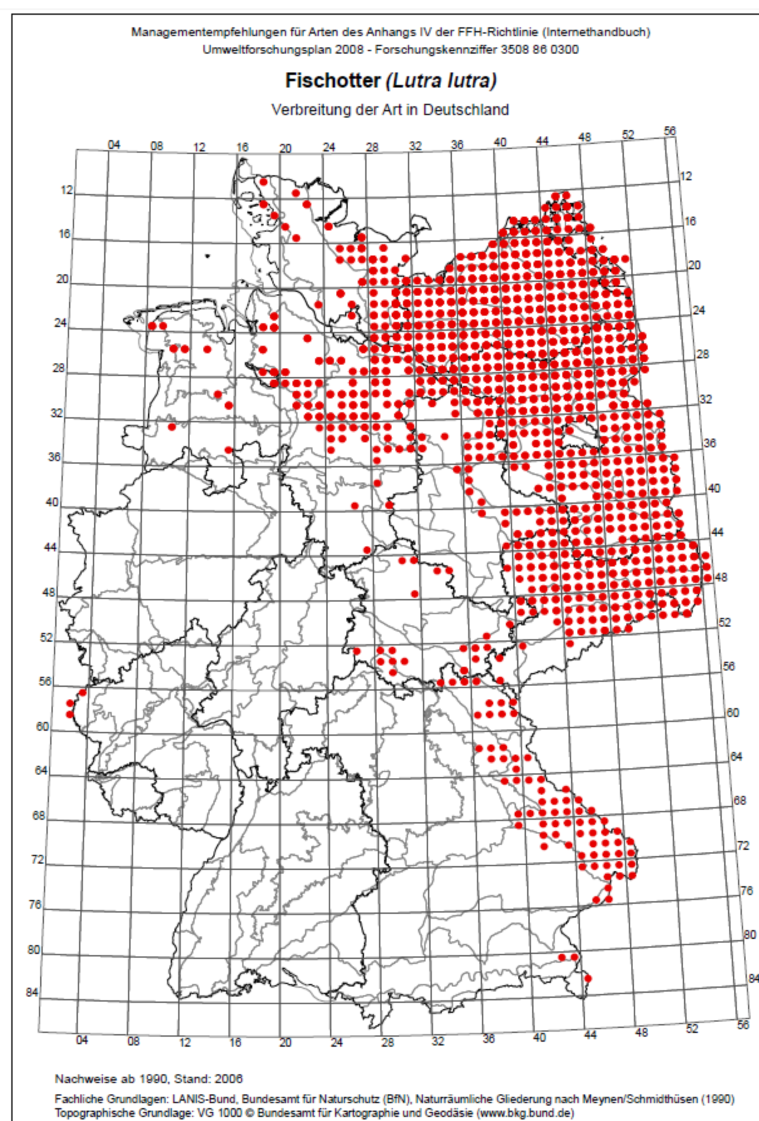

Figure 1.2 Distribution of the Eurasian otter in Germany as of 2006. Red dots illustrate occupied areas. Source: Federal Agency for Nature Conservation (BfN) (http://www.ffh-anhang4.bfn.de/fileadmin/AN4/documents/mammalia/Lutra_lutra_Verbr.pdf#page=2; [accessed: 22.04.2014]).

Travel distances per night in winter amounted to 3–4 km by females with cubs and 9–10 km by male otters (Erlinge 1967). Longest travel distances that were so far measured over several nights are about 68 km (Jenkins 1980) or even 84 km (Durbin 1998). For sleeping and resting they use either dens below ground level or between tree roots or piles of rocks, but they also use thick vegetation like reeds as resting sites (Kruuk 2006).

Otters communicate via olfaction by scent marking. Their scats, so called spraints, are one source of scent marking. It consists mainly of food remains, sometimes with secretions of the two anal scent glands. But there is also a jelly-like substance produced in the intestine that can be deposited with and without the glandular secretion (Trowbridge 1983; Kruuk 2006; Kean et al. 2011). Spraints and anal jellies are often placed on prominent locations (Mason & Macdonald 1987), such as rocks, trees, on scratch piles, or under bridges. Since their digestion is comparably fast with average minimum passage rates of 3 h 15 min (Jurisch & Geidezis 1997), one otter can mark up to 30 spraints per day (Kruuk 1992, 2006). There are several assumptions regarding the communicative function of sprainting such as territory defence (Gosling 1982), resource utilization (Kruuk 1992, 2006), or communicating sexual status or for mate attraction (Kean et al. 2011; Remonti et al. 2011).

Sexual maturity is attained between 18 and 24 months (Hauer et al. 2002b; Ruiz-Olmo et al. 2008). Since otter females are continuously polyoestrous (Mason & Macdonald 1986), mating and hence breeding can occur at all times of the year, with seasonal peaks (Sidorovich 1991; Beja 1996; Elmeros & Madsen 1999) or with evident seasonality (Kruuk et al. 1987) in some regions. Hauer et al. (2002b) for example found a seasonal birth peak in summer for otters living in Eastern Germany. The gestation period lasts between 61–74 days (Kruuk 2006). Litter sizes at birth can reach up to five cubs (Hauer et al. 2002b), but averages to minima of 1.7 in coastal habitats and 2.9 in freshwater systems (Beja 1996). For Eastern Germany a mean litter size at birth of 2.4 is reported (Hauer et al. 2002b). The cubs are reared by the female and start to be independent with 9 to 13 months (Kruuk et al. 1991; Hauer et al. 2002b).

Within the first year of life, especially in the first few months, the mortality seems to be moderately high with up to 42% dead yearlings (Ansorge et al. 1997; Ruiz-Olmo et al. 1998; Kruuk 2006). Although the mortality in subsequent years is lower, Kruuk (2006) reported a linear increase in probability of death with age and most studies reported short life expectancies and populations that mainly consist of young otters (Kruuk & Conroy 1991; Ruiz-Olmo et al. 1998; Bjorklund & Arrendal 2008). The oldest found otter in the wild was about 16 years old (Ansorge et al. 1997; Gorman et al. 1998). Major threats are on one hand human interventions in the aquatic systems, such as damming or canalisation of rivers, removal of bank side vegetation, draining of wetlands, or pollution that either directly reduces reproduction rate or life expectancy or indirectly the food resources (Kruuk 2006; Ruiz-Olmo et al. 2008). On the other hand humans decimate otters directly through road-traffic that killed up to 87% of dead found individuals (Zinke 1991, 2000) or through fish-traps or even (often illegal) hunting (Sidorovich 1991; Hauer et al. 2002a). Besides the danger that emerges from humans,

there are occasionally observations of bite wounds mostly from dogs or attacks by raptors (Kruuk & Conroy 1991; Sidorovich 1991; Hauer et al. 2002a).

The otter is a specialist in its habitat and there are no sympatric native species in Europe that live in exactly the same ecological niche. However, since the beginning of the 20th century the American mink (*Neovison vison*, Schreber 1777) started to spread in the wild, first in Northern Europe (e.g. Sweden: 1920s; Finland 1930s), later also in Middle (e.g. Germany: 1950s; Czech Republic: 1960s), and Southern Europe (e.g. Spain: 1970s; Italy: 1980s) (Bonesi & Palazon 2007). The mink was reported to be a competitor to Eurasian otters (Bonesi et al. 2006; McDonald et al. 2007; Melero et al. 2012). However, many studies found that minks are dominated by otters, with the latter being able to reduce mink densities or to slow down their colonisation (e.g. Bonesi & Macdonald 2004b; Bonesi et al. 2006; Bonesi & Palazon 2007). But there are also researchers that propose probable co-existence between both (Bonesi & Macdonald 2004a; Harrington et al. 2009), with minks changing their diet to more terrestrial prey (Bonesi et al. 2004) and/or catching smaller fishes than otters (Bueno 1996).

1.4 The Study Area

All otter faecal samples required for this thesis were collected in a study area located in a region called Upper Lusatian heath and pond landscape in Eastern Saxony, Germany. The landscape in this region is a patchwork of ponds, creeks, moor, cropland, pasture, and forests, but also small settlements and abandoned opencast mines that are mostly flooded nowadays. Although the dominant land use is agriculture and forest, the region is characterised by about 5000 ha ponds that are used for fish farming (Myšiak et al. 2013). Together with Lower Lusatia, it is one of the biggest continuous pond region of Central Europe (Schwerdtner & Gruber 2007). Already in the 13th century the people started to construct fish ponds on the moor-, marsh-, and swampland (Myšiak et al. 2013). This was possible because of the numerous creeks and rivers out of which ditches were created that supply the ponds with water and connect them with each other (Böhnert et al. 1996). The ponds are on average about 1 m deep and often clustered to pond areas. Main stocked fish species are carps (*Cyprinus carpio*), making up 89% of the fish production (Myšiak et al. 2013), but also tench (*Tinca tinca*), rainbow trout (*Oncorhynchus mykiss*), pike (*Esox lucius*), wels catfish (*Silurus glanis*), and perches (*Percidae*). In autumn, most ponds are drained. Saleable fish is harvested and sold; younger fish is inserted into the few smaller and deeper ponds for wintering. In spring, the fish is usually reinserted into the larger summer ponds.

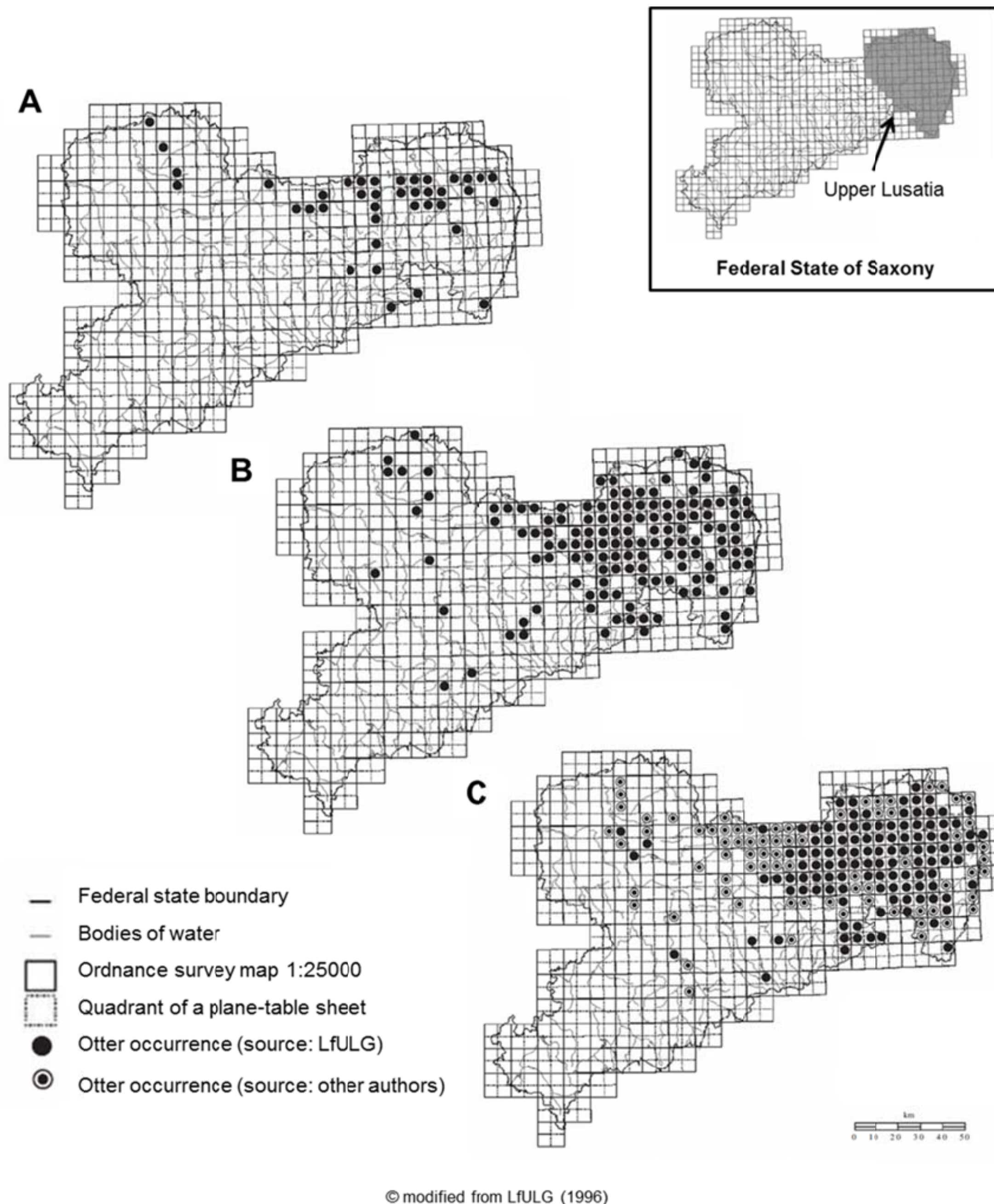

Figure 1.3 Evidences of otter occurrences in the Federal State of Saxony, Germany, during the period (A) 1950–1969, (B) 1970–1989, (C) 1994–1995. Maps are derived from "Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG – Saxon Federal State Office for Environment, Agriculture and Geology) (Ed.) 1996: Artenschutzprogramm Fischotter in Sachsen – Materialien zu Naturschutz und Landschaftspflege. Radebeul".

Besides the commercially function, the ponds play a very important role as secondary habitat for several endangered species including the Eurasian otter. Upper Lusatia is assumed to host one of the biggest and most viable populations in Central Europe (Ansorge et al. 1996, 1997; Klenke 1996). Here, the otter never got extinct during its depression. Generally, landscapes dominated by fish ponds are important habitats for otters and functioned as haven during their massive decline (Kranz, 2000). As can be seen in the chronological sequence of otter distributions in Saxony (Fig. 1.3), the population in Upper Lusatia is a source population for the re-expansion in Saxony. Hence, this population is

together with the ones in Brandenburg and Mecklenburg-Western Pomerania of prime importance for the otter conservation in Germany, even more because otters in fish pond systems are not well-received by fish farmers (Klenke et al., 2013; Kranz, 2000). However, fish pond systems were so far underrepresented in studies that tried to gain information about the species.

The chosen study area is located on the western margin of the UNESCO Biosphere Reserve "Upper Lusatian Heath and Pond Landscape" ("Oberlausitzer Heide- und Teichlandschaft"), between the villages "Königswartha" (51°19′ N, 14°20′ E) and "Groß Särchen" (51°22′ N, 14°19′ E). The study area includes seven pond areas, each comprising 8–13 ponds of varying size (0.36–39.6 ha), and one single pond (7.6 ha) (Fig. 1.4). In total, the study area included 64 ponds with an overall water surface of 505 ha. All ponds are connected by a complex system of ditches and streams and are framed by naturally vegetated embankments, partly used as agricultural roads. Islands, extensive reed belts, and heavily vegetated peninsulas induce a heterogeneous structure. The pond areas (PA) are surrounded by pastureland, cropland, forest, and small villages. One pond area is disconnected by a railway line, two are separated by roads out of which one is the federal road "B 96".

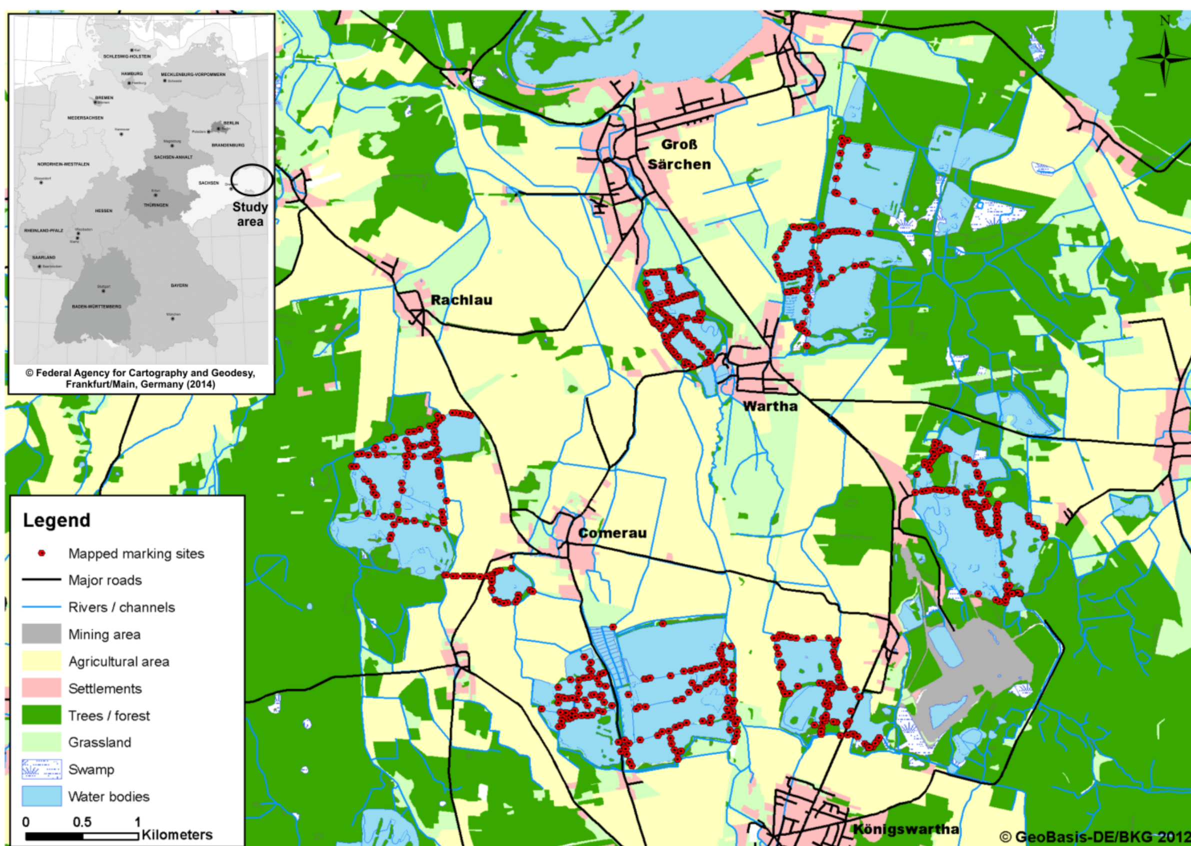

Figure 1.4 Geographical location of the study area and land use types. Blue areas are freshwater bodies (ponds or rivers/channels). Red dots depict mapped marking sites over all six years.

1.5 General Methodology

Since otters use their spraints for intraspecific communication, they tend to mark on frequently visited conspicuous terrestrial sites at specific locations throughout their home range (e.g. rocks, trunks, under bridges, at junctions of water channels, on runways). These markings sites are used by all members of

the population, regardless of their sex, reproductive status, or age (Kruuk 2006) and they can be easily detected by collectors. Each active otter marking site along the pond banks of all ponds filled with water and close-by ditches or streams were first mapped in early March 2006. They were tagged in a map and described in detail to facilitate a recover of each site later on. Although each pond in a pond area was rounded (if pond banks were not fully overgrown and hence impassable), most marking sites were located on banks between ponds rather than on the edge (Fig. 1.4). With the help of several field helpers, I conducted my first sampling period end of March 2006 from 26th–31st. The first day was used to train all field helpers how to search for faeces and to get acquainted with their pond area and tagged marking sites. Detected spraint samples at this first day were not collected but marked with materials of the surrounding to facilitate recognition of fresh spraints the next day. On the following five consecutive days, all detected fresh scat and anal jelly samples were collected from the tagged marking sites and from sites not previously detected. We collected mainly in the morning on days without rain or frost with two collectors for each pond area. This sampling regime was five times repeated in the years 2007–2012 (Tab. 1.1). Most marking sites were in use over several years or even over the entire six years. Due to the seasonality of the water regime, the study area size differed for each sampling year (Tab. 1.1).

Table 1.1 Overview of the six sampling years with sampling time, size of the water area, and number of collected samples.

Sampling year	Sampling time												Area of water in ha	Number of collected samples
	J	F	M	A	M	J	J	A	S	O	N	D		
2006													261	356
2007													399	282
2008													449	198
2009														
2010													294	381
2011													366	461
2012													360	454

The external layer of each detected spraint sample – containing sloughed gut cells – was wiped off with a commercially available cotton swab, placed in a separate sterile 10 ml cryovial (Biozym Scientific, Hessisch Oldendorf, Germany), and either DNA-extracted on day of collection (year 2006) or stored at –80°C in 1.8 ml ASL buffer (Qiagen, Hilden, Germany) (years 2007–2012). The DNA of all samples was purified (DNA extraction) using the QIAamp® DNA Stool Mini Kit (Qiagen). For microsatellite genotyping, I chose seven polymorphic microsatellite markers: Lut435, Lut457, Lut604, Lut615, Lut701, Lut733, and Lut914 (Dallas & Piertney 1998; Dallas et al. 2000, 2002). For the sex identification, I used two markers located on two genes linked to the Y chromosome: the SRY gene (sex-determining region Y) with the marker Lut-SRY (Dallas et al., 2000) and the DBY gene (DEAD box on the Y) with the marker DBY7Ggu (Hedmark et al., 2004). Each DNA-extracted sample was then amplified for the chosen markers using the polymerase chain reaction (PCR) (Saiki et al. 1988).

Since the forward primers (DNA sequence serving as starting point for PCR) were labelled with a fluorescent dye, the PCR-products can be separated by length and visualised (due to the dye) in a DNA sequencer and the corresponding software. Both allele sizes – the length of the microsatellite at each of the homologous chromosomes in base pairs – present the “genotype” of this individual at the specific microsatellite. The derived individual multi-locus genotypes were then used for further genetic and statistical analyses.
For chapter two, I used samples that were either collected opportunistically in one pond area within the above described study area (experiments on preservation and extraction) or from the first systematic sampling year in 2006 (experiments on sample types and PCR protocols). Chapter three makes use of the collected samples in 2006, whereas chapters four and five present results of analyses using all collected samples over the entire six sampling years.

1.6 Overview of Manuscripts

Chapters two to five present the results of the research conducted and are written as scientific manuscripts. Out of these, chapters two and three are published in international peer-reviewed journals. Chapters four and five are submitted manuscripts. In the following, I give a brief overview of the four scientific manuscripts:

Chapter two

Title: An optimisation approach to increase DNA amplification success of otter faeces
Authors: Simone Lampa, Bernd Gruber, Klaus Henle and Marion Hoehn
Summary: This manuscript presents results of comparative experiments on the amplification success rate of different otter sample types, of different storage times, of two DNA extraction methods and of three PCR protocols. Our results suggested that anal jelly samples are of highest amplification success and that storage without a reagent at –20°C decrease the amplification success rate with increasing storage time. Furthermore, we could demonstrate that the more expensive and time-consuming Qiagen kit extraction produced significantly higher success rates compared to the cheap and quick Chelex® 100 extraction method. Finally, we presented a two-step multiplex PCR protocol that significantly increased success rates and decreased genotyping error rates compared to the original PCR conditions for the employed markers described in Dallas et al. (1999).
Author contributions: S.L. organised and carried out the field work, conducted the laboratory work, analysed the data statistically, and wrote the manuscript (contribution: 80%). B.G. helped in sample collection, supervised the statistical analyses, and revised the manuscript (contribution: 10%). K.H. provided the facilities and revised the manuscript (contribution: 4%). M.H. supervised parts of the lab work and revised the manuscript (contribution 6%). The idea and the design of the experiment was a collaborative effort of S.L. and B.G.

Current status: This chapter has been published as "Lampa S, Gruber B, Henle K & Hoehn M (2008) An optimisation approach to increase DNA amplification success of otter faeces. Conservation Genetics 9(1), 201-210. DOI 10.1007/s10592-007-9328-9"

Chapter three

Title: How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation – A review of available methods illustrated by a case study

Authors: Simone Lampa, Klaus Henle, Reinhard Klenke, Marion Hoehn and Bernd Gruber

Summary: In this manuscript, we reviewed the literature and the pros and cons of each step required for non-invasive genetic mark-recapture (CMR) analyses: sampling design; sampling, preservation, and extraction methods; microsatellite genotyping; population size estimation models. The review is strengthened by a case study on otters with which we tested several methods for their appropriateness to accommodate for genotyping errors. As a result, we offer a step-by-step protocol for non-invasive genetic CMR studies that target to reliably estimate population sizes in the presence of high genotyping error rates. This step-by-step protocol is also summarised in a table that can be found as supplemental material in the online version of this article at the publisher's website and is attached in the appendix of the dissertation.

Author contributions: S.L. organised and conducted the field work, analysed the samples in the laboratory, analysed the samples statistically, reviewed the literature, as well as conceptualised and wrote the manuscript (contribution: 83%). K.H. provided the facilities, revised the manuscript and gave helpful comments that improved the manuscript in the process of re-submission (contribution: 7%). R.K. helped in sample collection and revised the manuscript (contribution: 2%). M.H. helped in mapping of marking sites and with shortening the manuscript (contributions: 2%). B.G. helped mapping marking sites, collecting samples and in finding ideas for the outline of the manuscript, supervised parts of the statistics, and revised the manuscript (contribution: 6%).

Current status: This chapter has been published as "Lampa S, Henle K, Klenke K, Hoehn M & Gruber B. (2013) How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation – A Review of Available Methods Illustrated by a Case Study. Journal of Wildlife Management 77(8), 1490–1511. DOI 10.1002/jwmg.604"

Chapter four

Title: Non-invasive genetic mark-recapture as a means to study population sizes and marking behaviour of the elusive Eurasian otter (*Lutra lutra*)

Authors: Simone Lampa, Jean-Baptiste Mihoub, Reinhard Klenke, Bernd Gruber and Klaus Henle

Summary: In this manuscript, we used 2132 otter faeces collected over a period of six years (2006–2012) to study the marking behaviour and to estimate population sizes and sex ratios employing

misidentification closed population CMR models. We further tested whether faecal sample densities can be used to infer on otters abundances.

Author contributions: S.L. organised and conducted the field work, analysed collected samples in the laboratory and statistically, reviewed the literature, as well as conceptualised and wrote the manuscript (contribution: 85%). J-B.M gave support in statistical analyses and reviewed the manuscript (contribution: 3%). R.K. helped with sample collection (contribution: 3%). B.G. helped mapping marking sites and collecting samples, contributed to the basic research idea, and reviewed the manuscript (contribution: 3%). K.H. helped with sample collection, provided the facilities, and revised the manuscript (contribution: 6%).

Current status: This chapter has been submitted to PLOS ONE on 1st of October 2014

Chapter five

Title: Non-invasive genetic mark-recapture as a means to study population dynamic and spatial use of Eurasian otters (*Lutra lutra*) in a fish pond landscape

Authors: Simone Lampa, Jean-Baptiste Mihoub, Reinhard Klenke, Bernd Gruber and Klaus Henle

Summary: This manuscript builds upon the data and results of chapter four. Using the samples collected from 2006–2012, we estimated apparent survival and temporary migration employing misidentification robust design models and tested for sex-biased dispersal. Additionally, we estimated activity range indices, tested for differences in sex, and for patterns in activity range overlaps between individuals.

Author contributions: S.L. organised and conducted the field work, analysed collected samples in the laboratory and statistically, reviewed the literature, as well as conceptualised and wrote the manuscript (contribution: 85%). J-B.M gave support in statistical analyses and reviewed the manuscript (contribution: 3%). R.K. helped with sample collection (contribution: 3%). B.G. helped mapping marking sites and collecting samples, contributed to the basic research idea, and reviewed the manuscript (contribution: 3%). K.H. helped with sample collection, provided the facilities, and revised the manuscript (contribution: 6%).

Current status: This chapter has been submitted to PLOS ONE on 1st of October 2014

____________________	____________________
Signature Simone Lampa	Signature Prof. Dr. Stefan Halle

Chapter 2

An Optimisation Approach to Increase DNA Amplification Success of Otter Faeces

Simone Lampa

Bernd Gruber

Klaus Henle

Marion Hoehn

Conservation Genetics 9(1), 201–210 (2008)

Conserv Genet (2008) 9:201–210
DOI 10.1007/s10592-007-9328-9

RESEARCH ARTICLE

An optimisation approach to increase DNA amplification success of otter faeces

Simone Lampa · Bernd Gruber · Klaus Henle · Marion Hoehn

Received: 13 December 2006 / Accepted: 29 March 2007 / Published online: 25 April 2007
© Springer Science+Business Media B.V. 2007

Abstract Faeces have proved to be a suitable non-invasive DNA source for microsatellite analysis in wildlife research. For the success of such studies it is essential to obtain the highest possible PCR amplification success rate. These rates are still relatively low in most carnivorous species, especially in the otter (*Lutra lutra*). We therefore optimised the entire microsatellite genotyping process by combining our findings with results from previous studies to gain a high rate of reliable genotypes. We investigated the influence of otter faecal quality in relation to the quantity of slimy secretions and three levels of storage periods at –20°C on amplification success. Further, we tested the cost-effective and time-saving Chelex extraction method against the profitable QIAamp® DNA Stool Kit (Qiagen), and compared three PCR methods - a standard single-step PCR protocol, a single-locus two-step PCR procedure and a multiplex two-step PCR procedure - regarding success rate and genotyping errors. The highest amplification success rate (median: 94%; mean: 78%) was achieved using faecal samples consisting only of jelly extracted with the QIAamp® DNA Stool Mini Kit (Qiagen) immediately after collection and amplified following the time and cost efficient multiplex two-step PCR protocol. The two-step procedure, also referred to as pre-amplification approach, turned out to be the main improvement as it increases amplification success about 11% and reduces genotyping errors about 53%, most notably allelic dropouts.

S. Lampa (✉) · K. Henle · M. Hoehn
Department of Conservation Biology,
UFZ—Helmholtz Centre for Environmental Research,
Permoserstr. 15, Leipzig 04318, Germany
e-mail: simone.lampa@ufz.de

B. Gruber
Department of Computational Landscape Ecology,
UFZ—Helmholtz Centre for Environmental Research,
Permoserstr. 15, Leipzig 04318, Germany

Keywords Faecal DNA · *Lutra lutra* · Microsatellites · Non-invasive samples · Pre-amplification

Introduction

Microsatellite genotyping of non-invasive DNA sources like faeces is a novel and increasingly applied approach to analyse the genetic structure of species. So far it is the only available technique to study population structure, population size, genetic diversity, and relatedness of elusive animals (Amos and Pemberton 1992; Bruford and Wayne 1993; Queller et al. 1993; Kohn and Wayne 1997; Reed et al. 1997) such as otters (*Lutra lutra*).

However, faecal samples typically contain low quantities of target DNA in a bacterial-enriched environment that includes PCR-inhibitors (Sidransky et al. 1992; Tschirch 1995; Murphy et al. 2000) and is likely exposed to hydrolytic, oxidative, and enzymatic degradation (Kohn et al. 1995; Frantzen et al. 1998; Idaghdour et al. 2003). Thus, the success of a microsatellite analysis is significantly influenced by the age of the scat (Jansman et al. 2001; Dallas et al. 2003) and the exposure to weather conditions (Farrell et al. 2000; Murphy et al. 2000). It has been demonstrated that diet also affects the amplification success rate strongly being high for herbivorous species (Flagstad et al. 1999; Banks et al. 2002), intermediate for omnivorous species (Gerloff et al. 1995; Goossens et al. 2000; Frantz et al. 2003), and usually rather low in studies with carnivores (Reed et al. 1997; Kohn et al. 1999; Piggott and Taylor 2003). The first microsatellite studies analysing otter faeces from wild populations obtained amplification success rates

of only 20% (Coxon et al. 1999; Dallas et al. 2003), which is close to the lower end even for carnivorous species. Such a low amplification success reduces considerably the suitability of faecal samples for genetic studies in this species. Therefore, it is of paramount importance to optimise the genetic techniques to obtain an adequate amplification success rate. The first attempt to optimise microsatellite analysis for the otter was presented recently (Hájková et al. 2006). Hájková et al. (2006) investigated the impact of collection temperature and sample type on amplification success, tested three buffers and ethanol for their efficiency in preservation and compared the extraction results of two similar stool kits (Qiagen, Invitek). In our approach we tried to systemise the optimisation by breaking down the whole procedure of a microsatellite analysis into its relevant parts. Four successive steps, each depending on the previous one, are crucial for microsatellite analyses: (1) the collection of samples, (2) the subsequent storage method, (3) the extraction of DNA, and (4) the amplification of DNA using PCR (polymerase chain reaction). In this study, we concentrated our optimisation effort on each of the four steps.

Hájková et al. (2006) found a pronounced effect of sample quality. However, even within the same species sample quality can vary depending on study area, diet, time of year, and also on the microsatellite loci and amplification protocol used. Therefore, we tested the influence of faecal quality on the amplification success as well. Hájková et al. (2006) reported that storage time had no effect on DNA amplification success. This is in stark contrast to a number of studies (Frantzen et al. 1998; Murphy et al. 2002) and needs further investigation. Hence, we studied the impact of storage periods on the success rate using three different levels of storage time at –20°C (one day/one week/two weeks). Although it has been shown that for the extraction of faecal DNA stool kits (e.g. Qiagen) produce high success rates (e.g. Goossens et al. 2000; Roeder et al. 2004; Hájková et al. 2006), they are also the most expensive. We examined whether the cheap and time-saving Chelex method can achieve comparable results and could be used instead. Finally, to include recent advances in PCR techniques, we optimised the PCR conditions of single-locus and multiplex PCR and investigated whether a two-step amplification approach (similar to Piggott et al. 2004; Hedmark and Ellegren 2006) yielded higher amplification success rates and lower genotyping errors than standard PCR.

Materials and methods

Sample collection

Spraint and anal jelly samples were collected from a wild otter population in Upper Lusatia, Saxony, Eastern Germany. Collections were made mainly during the morning hours on days without rain or frost. In all trials only freshly deposited faeces from the previous night were used. The external layer of the spraint which contains sloughed gut cells was wiped off with a commercially available cotton swab. Each cotton swab sample was stored in a separate sterile 500 ml tube. A pilot study demonstrated that this technique decreases the risk of sampling prey hard parts, such as bones or fish scales, while increasing the proportion of sloughed gut cells sampled. It is also manageable in the field, and has a reduced risk of contamination while maintaining a sufficient amplification success rate relative to other methods.

To test for the effect of the sample quality on the PCR amplification, 20 samples were classified into three types of faeces according to the quantity of slimy secretions: spraint (consisting of prey remains and almost no mucus), spraint plus mucus (consisting of prey remains and a layer of mucus), and jelly (gelatinous secretion of anal scent glands without prey remains). Supplementary data, such as weather conditions and collection time, were recorded to test these factors for correlation with DNA amplification success. Faecal samples were extracted with the most reliable extraction method (see section ‘‘Comparison of extractions’’), PCR amplified at six loci using several PCR protocols, and separated in an ABI PRISM® 3100 Genetic Analyser (see section ‘‘PCR amplification’’).

Storage time

The effect of storage time on genotyping success was tested using 15 faecal samples. Three swabs were taken from each faecal sample at three different positions of the surface using a separate cotton swab each time. All swabs were frozen at –20°C in a 500 ml tube within 10 h of sampling. Amplification success could thus be tested for each scat for each of the three storage times: (1) one day, (2) one week, and (3) two weeks. To avoid any methodological bias the order of the subsamples was randomised before proceeding with the following steps. DNA from all 45 subsamples was extracted with the most reliable extraction method (see section ‘‘Comparison of extractions’’). Eight microsatellite loci with fragment length 127–211 bp (Lut 435, 457, 604, 615, 701, 717, 733, 832; Dallas and Piertney 1998; Dallas et al. 2002) were amplified following the single-locus two-step PCR protocol (see section ‘‘DNA amplification’’; annealing temperature for Lut 717: 57°C; Lut 832: 55°C), and separated by electrophoresis on 1.75% agarose gels in TBE buffer (68.5 mM Tris HCl, 89 mM boric acid, 2.5 mM EDTA). DNA was ethidium bromide stained and visualised using BIO-RAD Gel Doc 1000.

DNA extraction

DNA extractions were carried out in a separate laboratory that was free of concentrated otter DNA or PCR products. Aerosol resistant pipette tips were used in all working steps. Negative controls were included in each extraction to monitor contamination.

Two different extraction methods were tested: The Chelex® 100 method is a very fast, simple, and cost-effective technique that has been used in previous studies to isolate DNA from hair (Walsh et al. 1991; Vigilant 1999; Frantz et al. 2004) and faeces (Paxinos et al. 1997; Reed et al. 1997; Palomares et al. 2002; Berry and Sarre, unpublished). In contrast, the QIAamp® DNA Stool Mini Kit (Qiagen) is more time-consuming and costly but produces relatively high-quality template DNA. The Qiagen kit is based on the GuSCN/silica method (Frantz et al. 2003) and has been applied to faecal DNA extractions several times (Goossens et al. 2000; Morin et al. 2001; Frantz et al. 2003; Nsubuga et al. 2004; Roeder et al. 2004).

The Chelex extraction protocol involved an initial wash with 0.75 ml PBS puffer (pH 7.4), which was added to the cotton swab in the 500 ml tube and homogenised by vortexing. 500 µl of the supernatant was transferred to a fresh tube containing 500 µl 10% H_2O-Chelex-solution and 4 µl of proteinase K (10 mg/ml) was added. Samples were then vortexed before incubation over night at 55°C with rotation. The following day samples were briefly vortexed, then boiled for 20 min followed by a 5 min centrifugation at 16,000*g*. The supernatant was removed into a new tube and centrifuged again for 5 min at 16,000*g* before the supernatant was transferred into a fresh tube and stored at –20°C.

The Qiagen kit extraction was carried out according to the manufacturer's instructions except for the initial steps. Here the cotton swab was suspended in 1.7 ml of ASL buffer (warmed to 70°C) in the 500 ml tube and vortexed for 20 s. After 2 min of incubation at room temperature the extraction was performed as from step 4 of the manufacturer's instructions.

DNA was extracted on the day of collection from 47 faecal samples, which were each wiped at two different positions with separate cotton swabs to allow a comparison of both extraction methods for the same scat. Only samples for which at least three microsatellites could be successfully amplified were included in the final comparison, this was achieved for 24 samples. Amplification was carried out with six microsatellites (Lut 435, 457, 604, 615, 701, 733; Dallas and Piertney 1998; Dallas et al. 2002) following the single-locus two-step PCR protocol (see section ''DNA amplification''). PCR products were initially screened by agarose gel electrophoresis and only successfully amplified samples were genotyped on ABI PRISM® 3100 Genetic Analyser (Applied Biosystems).

PCR amplification

PCRs were prepared using aerosol resistant pipette tips in a DNA UV-cleaner box and all reactions included a PCR negative control.

In a pilot study 12 microsatellites designed by Dallas and Piertney (1998) and Dallas et al. (2002) (Lut 435, 457, 604, 615, 701, 715, 717, 733, 782, 818, 832, 833) were tested for their amplification success rate, allelic richness and heterozygosity in otter faecal DNA originating from Saxony. Three of the 12 markers had a very low amplification success rate (Lut 782: 14,9%; Lut 818: 12,8%; Lut 833: 10,6%). Moreover, Lut 782 turned out to be monomorphic.

Six of the remaining nine microsatellites had the same optimal annealing temperature (58°C) and were hence suitable for multiplex PCR. Conditions were optimised for single and multiplex PCR of these six loci (combinations: Lut 435, 604, 701; Lut 457, 615, 733) based on the original PCR conditions described in Dallas et al. (1999). The major difference between the original standard PCR protocol and the optimised single-locus and multiplex protocols described here is that two consecutive PCR reactions were carried out, with PCR product from the first amplification being used as the template for a second PCR reaction. During the single-locus two-step procedure only one locus was amplified per reaction, whereas the multiplex two-step approach contained the primers for three markers in both (first and second) PCR steps. Furthermore, compared to the original PCR protocol the optimised protocols included higher (Taq DNA Polymerase, primers) and lower (MgCl) concentrations of PCR reagents, longer reaction times during amplification (30 s vs. 15 s), the use of locus-specific annealing temperatures and a greater extent of cycles (first PCR: 45, second PCR: 40 vs. 35). To test the effect of performing two consecutive PCRs, the product from the first amplification was also genotyped for all samples identified as positive using the single-locus PCR protocol.

All three protocols (original, single-locus, multiplex) were performed in 25 µl volumes containing 3 µl of DNA extract (6 µl of PCR product for the second PCR). The final reaction concentrations for both single-locus and multiplex PCRs consisted of 1 × reaction buffer (Taq PCR Core Kit, Qiagen) with 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.6 µM of each primer (0.4 µM for the second PCR) and 0.5 units Taq Polymerase (Taq PCR Core Kit). Whereas the original protocol uses a touch-down profile, we amplified at the locus-specific annealing temperature. The PCR profile was: initial 2 min 15 s at 90°C, and then cycles of 30 s at 90°C, 30 s at 58°C and 30 s at 72°C. The first PCR was replicated for 45 cycles and the second PCR for 40 cycles. Amplification ended with a final extension at 72°C for 1 min. Reactions were carried out in a BIOMETRA T3

Thermocycler. Forward primers were end-labelled at the 5′-end with a fluorescent dye and the pigtail 'GTTGCTT' was added to the 5′-end of reverse primers to enhance the 3′ adenosine overhang. This avoids typing error due to variability in non-templated nucleotide addition at the 3′-end of PCR products (Brownstein et al. 1996).

DNA from 20 faecal samples was extracted immediately after collection using the most reliable extraction method (see section ''Comparison of extractions'') and stored at –20°C. For each of these 20 samples, three PCRs were performed with the six loci, giving a total of 60 PCRs per amplification protocol tested. PCR products were separated and visualised in an ABI PRISM® 3100 Genetic Analyser (Applied Biosystems) and analysed using ABI Prism® GeneMapper Software V.3.0.

In all comparisons PCR amplification success rate was used as an indicator for the quality of the particular method and was estimated as the median percentage of successfully amplified products either over all samples per locus (see sections ''Effect of storage time'', ''Comparison of extraction'', ''Comparison of amplification protocols'') or over all loci per sample (see sections ''Sample collection'', ''Comparison of amplification protocols''). Due to sample sizes between 15 and 24, non-parametric tests (Kruskal–Wallis test, multiple comparisons and Wilcoxon test) were used to test for significant differences between methods ($P = 0.05$ after Bonferroni adjustments for multiple testing, thus in three tests (see sections ''Sample collection'', ''Effect of storage time'', ''Comparison of amplification protocols'') the level of significance is $P = 0.0167$).

In addition, the three PCR protocols (original, single-locus, multiplex) were evaluated by comparing the rate of false alleles and allelic dropout. Due to three replications of each sample with each PCR protocol (nine PCRs per locus and sample in total) the criteria established by Frantz et al. (2003) could be used to obtain reliable genotypes for each successfully amplified sample. Therefore, genotyping errors were ascertained by comparing scored genotypes with the reliable genotype. False alleles can occur in all positive samples, whereas allelic dropout can only be detected in positive heterozygous samples (Creel et al. 2003; Broquet and Petit 2004). Hence, both rates were estimated considering only such genotypes.

Results

Sample collection

There was no correlation between collection time (within the first 20 h after defecation) and amplification success rate in any of the three trials (storage time, extraction, amplification) (Kruskal–Wallis test: H $(2, n = 59) = 2.82$, $P = 0.24$). Due to constant weather conditions during all sample collections, predictions about weather effects on PCR amplification could not be made. Rather, the quantity of slimy secretions influenced amplification success significantly in all three PCR protocols (Kruskal–Wallis test: original: H $(2, n = 20) = 9.92$, $P = 0.0070$; single-locus: H $(2, n = 20) = 9.42$, $P = 0.0090$; multiplex: H $(2, n = 20) = 11.13$, $P = 0.0038$). PCR amplification success was low for spraint and spraint plus mucus samples but high for samples that consisted only of jelly (Table 1). Pairwise comparisons using multiple comparisons showed a significant increase in amplification success rate for jelly samples compared with spraint samples. Compared with spraint plus mucus samples only the multiplex protocol achieved a significantly increased success rate for jelly samples (Table 1).

Effect of storage time

On average, DNA was obtained from 80% of the samples ($n = 15$) extracted after one day, while only 63% and 60% of the same faecal samples ($n = 15$ each) that were extracted after one or two weeks amplified successfully (represented as bands in agarose gel, Fig. 1). Thus, by increasing the storage time we observed a significant decrease in amplification success rate, whereby the highest decline was detected after one week (Wilcoxon tests: 1 day–1 week: $Z = 2.366$, $P = 0.018$; 1 day–2 weeks: $Z = 2.521$, $P = 0.012$; 1 week–2 weeks: $Z = 1.400$, $P = 0.161$).

Comparison of extractions

As shown in Fig. 2, the two extraction methods differ significantly in amplification success rate (Wilcoxon test: $Z = 2.201$, $P = 0.028$). Also, differences between extraction methods in successful PCRs were highly significant for each microsatellite loci (Fisher test: $P < 0.001$, $df = 1$). All six loci amplified in 17 out of 24 (70.8%) kit-extracted samples, whereas none of the Chelex-extracted samples yielded positive amplifications at all six loci simultaneously. Although positive extracts purified with the Chelex method often showed higher signal intensity as the same kit-extracted sample, some negative Chelex-extracted faeces did not even display primer dimers, which indicate the presence of PCR inhibitors (Kohn et al. 1995; Reed et al. 1997; Vigilant 1999; Palomares et al. 2002). Potential PCR inhibition, caused by components in faecal extracts, were tested in supplementary assays by amplifying tissue DNA (already successfully amplified in former PCRs) in which kit- or Chelex-extracted faecal DNA were added. Negative amplifications were only noted from

Table 1 Median values and statistical significance of the amplification success rate corresponding to quantity of slimy secretions (spraint; spraint plus mucus; jelly)

	Median of amplification success rate			Multiple comparisons		
	Spraint (1)	Spraint plus mucus (2)	Jelly (3)	1–2	1–3	2–3
Original	0.00 (n = 5)	0.06 (n = 6)	0.76 (n = 9)	P = 1	P = 0.016	P = 0.076
Single-locus	0.22 (n = 5)	0.28 (n = 6)	1.00 (n = 9)	P = 1	P = 0.014	P = 0.15
Multiplex	0.17 (n = 5)	0.19 (n = 6)	0.94 (n = 9)	P = 1	P = 0.019	P = 0.018

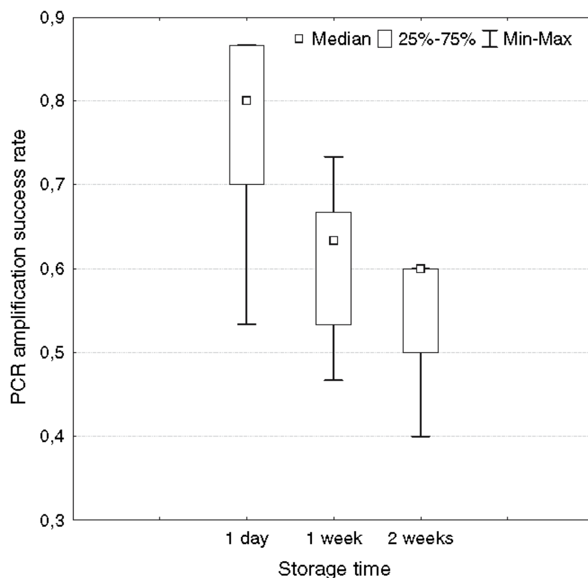

Fig. 1 Comparison of amplification success rate of three different storage times at –20°C. Results are from 45 subsamples of 15 faecal samples that were extracted using the Qiagen kit one day, one week, or two weeks after storing at –20°C. Each sample from each treatment was amplified once with eight microsatellite loci. The percentage of successful amplification (detected as bands in an agarose gel) for each locus in each storage time was calculated by pooling the data across the 15 samples. Boxes represent the interquartile range with the median square, while error bars signify outliers

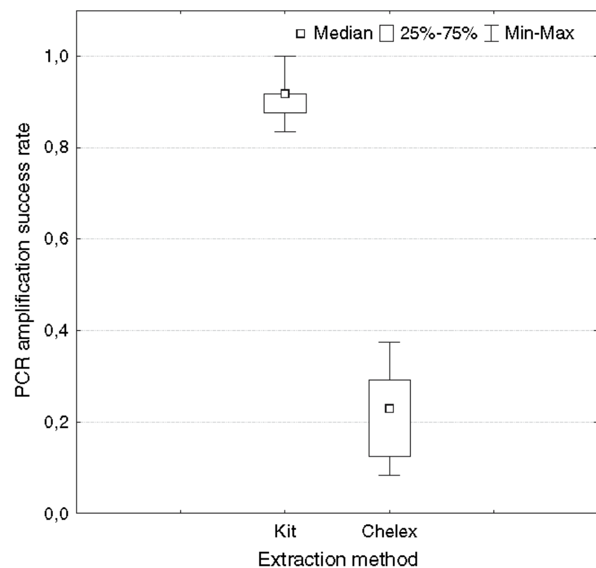

Fig. 2 Median of the amplification success rate of two extraction methods tested with 24 subsamples respectively. One amplification with six microsatellite loci was carried out for each sample and each extraction method. The amplification success rate (represented as bands in an agarose gel and in an ABI PRISM® 3100 Genetic Analyser) was obtained by averaging the positive samples over six loci. Variance across loci is illustrated by a box plot chart. Boxes represent the interquartile range with the median square, while error bars signify outliers

mixtures of tissue and Chelex-extracted faecal DNA, whereas all mixtures containing tissue and kit-extracted faecal DNA resulted in positive PCRs.

Comparison of amplification protocols

Positive amplification products occurred in 27% of faecal samples amplified using the original protocol compared with 51% using the optimised single-locus protocol, and 47% using the multiplex protocol (Fig. 3). Thus, the optimised single-locus two-step PCR protocol showed a near significant improvement in amplification success compared to the original protocol (Wilcoxon tests: Single-locus–Original: $Z = 2.201$, $P = 0.028$; Multiplex–Original: $Z = 2.201$, $P = 0.028$). Between the single-locus and the multiplex protocol no significant difference was found (Wilcoxon test: $Z = 0.944$, $P = 0.345$). Additionally, the percentage of samples, in which a genotype could be obtained at least once for all six loci simultaneously, declined from 40% using the multiplex protocol to 35% using the single-locus protocol to finally 25% using the original protocol, thus highlighting the improved PCR conditions (Fig. 4). A similar pattern was noted when the number of positive replicates over all samples and microsatellite loci was considered (Fig. 4).

When considering genotyping errors, false alleles differed between 0% following the original or the optimised single-locus protocol and 0.02% using the multiplex protocol. Allelic dropout varied from 13 to 56% in the original protocol (median 27%), from 16 to 48% in the single-locus protocol (median 38%), and from 24 to 42% in the multiplex protocol (median 29%). In all positive amplification

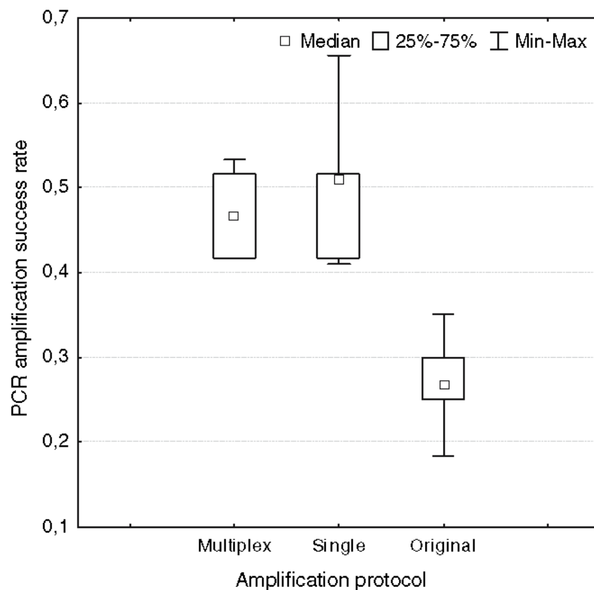

Fig. 3 Amplification success rate of three PCR protocols obtained by amplifying 20 faecal samples in triple replication with six microsatellite loci. The percentages of successful amplification per treatment were calculated as the median of all positive samples across each loci. Boxes represent the interquartile range with the median square, while error bars signify outliers

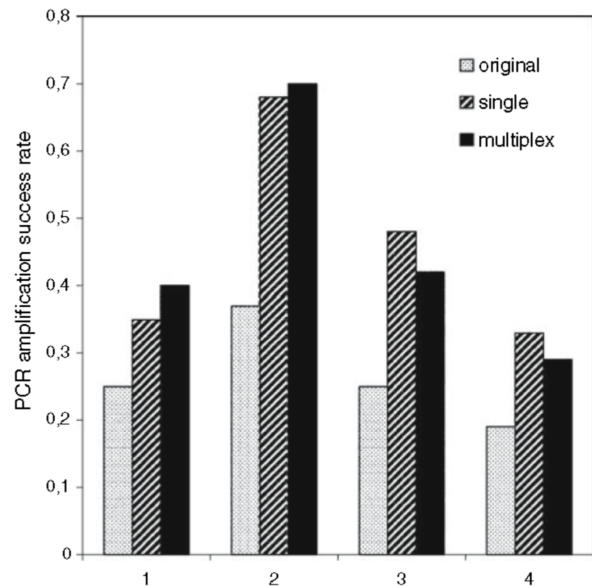

Fig. 4 Amplification success rate of three PCR protocols (original, single-locus, multiplex) regarding several criteria: (1) samples with at least one positive amplification in all six loci, (2) at least one positively amplified replicate over all samples and loci, (3) at least two positively amplified replicates over all samples and loci, (4) three positively amplified replicates over all samples and loci. Data obtained by amplifying 20 faecal samples in triplicate with six microsatellite loci

products amplified according to the original protocol, the signal intensity was conspicuously reduced compared to those of the optimised protocol, making precise allele detection much more complicated. Furthermore, separation and visualisation by the less sensitive electrophoresis in agarose gels frequently failed due to low peak size (often < 50 units).

Of each successfully amplified sample using the single-locus two-step protocol the first PCR product was also genotyped. Compared to the first PCR, the second amplification showed an about 11% increase in success rate that ranged from 6 to 28% across samples. Moreover, the appearance of genotyping errors (mainly allelic dropouts, 46 of 47) decreased about 53% after performing the second amplification ranging from 14% to 100% between loci. From all samples that generated a genotyping error during the first PCR only 47% repeated this error after the second amplification. Only once a sample showed artefacts after performing the second amplification, although the first PCR amplified the reliable genotype.

Discussion

In order to maximise the amplification success of DNA from otter faeces, we investigated the effect of four factors: faecal quality, storage period at –20°C, extraction methods, and PCR conditions. All four factors had a strong influence on the amplification success rate.

The first crucial step of using faeces for microsatellite analysis is the sample collection which is related to faecal quality. The freshness of faeces is a determining factor for PCR amplification success (Jansman et al. 2001; Dallas et al. 2003), hence only freshly deposited scats from the previous night (up to 20 h after defecation) were collected. However, within these approximately 20 h after defecation, we could not detect a decline in amplification success rate. Comparable results were obtained by Coxon et al. (1999) and Hájková et al. (2006), who also failed to detect significant differences within the first hours after defecation. In addition, the quantity of slimy secretions is an important influencing factor. We found that samples consisting only of jelly showed a very high PCR success rate, whereas PCR success rate was low for samples of spraints with or without mucus. This result confirms previous studies that deal with the amplification success of faecal DNA from otters (Coxon et al. 1999; Hájková et al. 2006) and seals (Reed et al. 1997). Anal jelly may contain less PCR inhibitors, bacteria, and enzymes than spraint that is composed mainly of prey remains. One reason, why spraint plus mucus samples achieved comparable low amplification success rates as spraint samples, may be that the

mucilage layer on spraint cannot be equated with the secretion of anal scent glands called jelly. In the large intestine mucus, secreted from goblet cells, is used as a lubricant for faeces which must pass over membranes (Liebich 1999; Welsch 2006). Whereas the secretion of anal scent sacs either act as visual and olfactory stimuli used in the social organisation of the population (Gorman et al. 1978; Macdonald and Mason 1987) or is of gastric origin being produced when otters have not eaten for 18–24 h (Conroy and French 1991; Carss and Parkinson 1996). Hájková et al. (2006) chose the classification (i) spraint, (ii) spraint with jelly, (iii) jelly and equated consequently the thin layer of mucus on spraint with the anal jelly. In contrast, we distinguish in our study between these two types of slimy secretion, which is supported by the high difference in amplification success between spraint plus mucus and jelly samples. Also, climatic conditions can have a high impact on the amplification rate of faeces (Farrell et al. 2000; Murphy et al. 2000). Humidity, for instance, might provide a better microclimate for bacteria and enzymes, while longer periods of rainy weather might wash away cells from the surface of the scat. Reduced amplification success rate in wet periods were detected for faeces from carnivores in western Venezuela (Farrell et al. 2000). During our pilot study we were able to observe the same findings. Hence, we only sampled faeces during periods of dry diurnal weather (i.e. without rain or frost). Finally, the collection method itself might interfere with the amplification success. It has been demonstrated that homogenisation of faecal samples yields reduced PCR success rates compared to surface wash or homogenisation of surface scrape (Piggott and Taylor 2003). In light of this and the results of our pilot study we used commercially available cotton swabs for the sample collection.

Storing faecal samples in a freezer at –20°C is a frequently practised method that should provide protection against further degradation (Tikel et al. 1996; Reed et al. 1997; Frantzen et al. 1998; Wasser et al. 1997; Ernest et al. 2000; Frantz et al. 2003; Piggott and Taylor 2003). Therefore, we considered this storage method to be suitable in combination with our collection technique as reagents or buffer solutions may remove cells from the cotton swab. However, we could show that amplification success rate decreased drastically with increased storage time (i.e. 20% within two weeks). As a result, we recommend that DNA extraction should be performed immediately after collection. Degradation during the thawing process is an unlikely reason for this effect since all subsamples were treated in the same manner. In contrast to our study, Hájková et al. (2006) recently reported that PCR success rate did not decline in otter faecal samples that were preserved up to 234 days in a freezer at –20°C, stored in 96% ethanol or buffers of kits. Our results demonstrated that the decline of amplification success rate was highest after one week of storage. Afterwards the decline was rather low. A possible explanation for the results of Hájková et al. (2006) may be that they extracted most of the samples after the first week of storing. However, contrasting results may also have occurred because samples were stored in a buffer solution which may preserve faecal samples over an intermediate period of time.

In addition to collection and preservation, the extraction of samples is the third factor strongly influencing quality and quantity of template DNA (Wasser et al. 1997; Flagstad et al. 1999; Banks et al. 2002). The efficiency of DNA purification can vary greatly among species and even among individuals (Taberlet and Luitkart 1999; Gossens et al. 2000; Piggott and Taylor 2003). Due to low DNA concentration in faeces (Gerloff et al. 1995; Murphy et al. 2000) and a high proportion of PCR inhibitors, nucleases, bacteria, and enzymes (Deuter et al. 1995; Kohn et al. 1995; Reed et al. 1997; Frantzen et al. 1998) a rapid and easy to handle (Reed et al. 1997) species specific protocol should be used to isolate a maximum of DNA while removing PCR inhibitors simultaneously. Therefore, two extraction methods were tested. The crucial difference between the quick, cheap, and simple Chelex method and the more time-consuming QIAamp® DNA Stool Mini Kit (Qiagen) is the thorough purification of extracts with the Qiagen kit, whereas Chelex, as an alkaline chelating resin, removes only polyvalent metal ions (Walsh et al. 1991; Reed et al. 1997). In consideration of the 69% higher amplification success rate obtained by extracting samples with the Qiagen kit, we suspect that the washing and purification steps of the Qiagen kit remove PCR inhibitors to the greatest possible extent. The higher amplification success justifies the time-consuming washing steps. In our study the Chelex extractions must still have contained a large number of PCR inhibitors, since some Chelex-extracted samples not only failed to amplify, but also showed no sign of primer dimers, which indicate the presence of PCR inhibitors (Kohn et al. 1995; Reed et al. 1997; Vigilant 1999; Palomares et al. 2002). Moreover, control amplifications that contained tissue extracts and Chelex-purified faecal samples also failed to amplify. This indicates that the addition of Chelex extracts to successfully amplifying tissue samples can inhibit their amplification. Other studies on otter faeces compared the Qiagen kit with the similar Invitek kit whereby the Invitek kit yielded a higher amplification success rate (Hájková et al. 2006).

PCR conditions need to be adjusted according to the quality and quantity of the DNA faecal extracts. For these purposes we optimised the amplification protocol designed for microsatellite loci by Dallas et al. (1999). The optimised single-locus two-step PCR protocol achieved a 24%

(10% regarding positive PCRs at six loci) increase in amplification success compared to the standard PCR conditions outlined by Dallas et al. (1999). Major enhancements are the performance of two consecutive PCR reactions with the first PCR product being the template for the second PCR reaction, the use of locus-specific annealing temperatures, a greater extent of cycles, and modifications in reagent concentrations. Based on the improved protocol for single-locus PCR conditions a more time and cost efficient multiplex protocol was developed, despite previous opinions that multiplexing primers of faecal DNA is difficult to achieve (Ernest et al. 2000). By using less DNA extract the multiplex protocol allows to analyse more loci and to perform the necessary PCR repetitions. Genotyping errors (i.e. false alleles and allelic dropout) occurred at a similar rate in all three PCR protocols and can only be overcome by a large number of replicated amplifications (Taberlet et al. 1996; Kohn et al. 1999; Ernest et al. 2000; Frantz et al. 2003; Broquet and Petit 2004). Due to the low signal intensity, the detection of alleles was hampered when using the original protocol instead of the optimised protocols. As we could demonstrate the use of a two-step procedure, also referred as to pre-amplification approach, offers an increase in quality and quantity of the template DNA. This corroborates the results of recent studies (Piggott et al. 2004; Hedmark and Ellegren 2006) and highlights again the advantages of the pre-amplification approach. Piggott et al. (2004) reported an improvement in amplification success rate and genotyping error rate, whereas Hedmark and Ellegren (2006) found that allelic dropouts generated during the first PCR step were repeated to a high extent in the second amplification. Our results are in line with the ones of Piggott et al. (2004), as we observed that only 47% of allelic dropouts from the first PCR appeared also in the second amplification, while 53% of the samples generating a genotyping error during the first step (46 allelic dropouts vs. 1 false allele) amplified the reliable genotype after the second PCR. Moreover, we were able to optimise the pre-amplification approach further and obtained a still more time and cost effective protocol. Instead of performing an initial multiplex PCR with all six primers followed by a second separate amplification for each marker, we amplified three markers at once in both consecutive amplifications. Thus, the procedure outlined by Piggott et al. (2004) need seven PCR reactions per sample to amplify six loci, using our approach it demands merely four amplification steps per sample for six markers.

For comparability purposes, the amplification success rate is often calculated over all samples and loci in the literature (e.g. Hájková et al. 2006). We provide an additional amplification success rate that is based on the amplification of all six microsatellites simultaneously (Fig. 4). The restricted estimate of amplification success offers more information about the value of primers when dealing with questions such as population size estimates or parentage analyses.

In summary, DNA of otter faeces is only available in low concentrations (Gerloff et al. 1995; Murphy et al. 2000), is exposed to degradation (Frantzen et al. 1998), contains a large number of PCR inhibitors (Deuter et al. 1995; Kohn et al. 1995), and is thus vulnerable to genotyping errors (Goossens et al. 2000; Broquet and Petit 2004). This may explain why previous PCR amplification attempts have met limited success (Coxon et al. 1999; Dallas et al. 2003). However, these difficulties can be overcome with a suitable preservation technique that avoids further degradation and an extraction method that removes PCR inhibitors to a large extent. Efficient amplification conditions can compensate for low DNA concentrations while replicated PCRs can remediate genotyping errors. We could assert that the highest amplification success rate could be achieved by an extraction of jelly samples with the QIAamp® DNA Stool Mini Kit followed immediately after collection and amplified using the optimised multiplex PCR protocol. With this combination we obtained a median amplification success of 94% (mean: 78%) compared to the 20% of the original methods described by Dallas et al. (2003). To apply this method to future studies it is important to know whether the quantity of samples will be sufficient if only jelly samples can be collected. However, even if all types of faecal samples are collected and analysed the amplification success is still 47% (mean: 47%) over all samples per loci and 40% over all locus per samples that obtained a genotype at least once for all six loci. With such success rates, genetic analyses of otter faeces can provide a powerful way to study otter populations.

Acknowledgements We would like to thank R. Klenke for helpful discussions and pertinent knowledge about the otter biology. A. Vallentin and U. Hempel are thanked for collecting faecal samples. Thanks to E. Küster and the Department of Bioanalytical Ecotoxicology and also to W. Durka and the Department of Community Ecology for giving access to the laboratories. We are grateful to the members of the Institute of Applied Ecology, University of Canberra for proof-reading the manuscript and to the anonymous reviewers who have greatly contributed to improve a first version of this paper. We kindly acknowledge private landowners for their permission to access the ponds.

References

Amos B, Pemberton J (1992) DNA fingerprinting in non-human populations. Curr Opin Genet Dev 2:857–860

Banks SC, Piggott MP, Hansen BD, Robinson NA, Taylor AC (2002) Wombat coprogenetics enumerating a common wombat population by microsatellite analysis of faecal DNA. Aust J Zool 50:193–204

Broquet T, Petit E (2004) Quantifying genotyping errors in noninvasive population genetics. Mol Ecol 13:3601–3608

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1006

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939–943

Carss DN, Parkinson SG (1996) Errors associated with otter *Lutra lutra* faecal analysis. 1. Assessing general diet from spraints. J Zool 238:301–317

Conroy JWH, French DD (1991) Seasonal patterns in the sprainting behaviour of otters (*Lutra lutra* L.) in Shetland. In: Reuther C, Röchert R (eds) Proceedings of 5th International Otter Colloquium, Hankensbüttel, 1989

Coxon K, Chanin P, Dallas J, Sykes T (1999) The use of DNA fingerprinting to study the population dynamics of otters (*Lutra lutra*) in southern Britain a feasibility study. Environment Agency R&D Project W1-025

Creel S, Spong G, Sands JL, Rotella J, Zeigle J, Joe L, Murphy KM, Smith D (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. Mol Ecol 12:2003–2009

Dallas JF, Piertney SB (1998) Microsatellite primers for the Eurasian otter. Mol Ecol 7:1247–1263

Dallas JF, Bacon PJ, Carss DN, Conroy JWH, Green R, Jefferies DJ, Kruuk H, Marshall F, Piertney SB, Racey PA (1999) Genetic diversity in the Eurasian otter*Lutra lutra* in Scotland Evidence from microsatellite polymorphism. Biol J Linn Soc 68:73–86

Dallas JF, Marshall F, Piertney SB, Bacon PJ, Racey PA (2002) Spatially restricted gene flow and reduced microsatellite polymorphism in the Eurasion otter *Lutra lutra* in Britain. Conserv Genet 3:15–29

Dallas JF, Coxon KE, Sykes T, Chanin PRF, Marshall F, Carss DN, Bacon PJ, Piertney SB, Racey PA (2003) Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. Mol Ecol 12:275–282

Deuter R, Pietsch S, Hertel S, Müller O (1995) A method for preparation of fecal DNA suitable for PCR. Nucleic Acids Res 23:3800–3801

Ernest HB, Penedo MCT, May BP, Syvanen M, Boyce WM (2000) Molecular tracking of mountain lions in the Yosemite Valley region in California genetic analysis using microsatellites and faecal DNA. Mol Ecol 9:433–441

Farrell LE, Roman J, Sunquist ME (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. Mol Ecol 9:1583–1590

Flagstad O, Roed K, Stacy JE, Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. Mol Ecol 8:879–883

Frantz AC, Pope LC, Carpenter PJ, Roper TJ, Wilson GJ, Delahay RJ, Burke T (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Mol Ecol 12:1649–1661

Frantz AC, Schaul M, Pope LC, Fack F, Schley L, Muller CP, Roper TJ (2004) Estimating population size by genotyping remotely plucked hair the Eurasian badger. J Appl Ecol 41:985–995

Frantzen MAJ, Silk JB, Fergusen JWH, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. Mol Ecol 7:1423–1428

Gerloff U, Schlötterer C, Rassmann K, Rambold I, Hohmann G, Fruth B, Tautz D (1995) Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (Pan paniscus). Mol Ecol 4:515–518

Goossens B, Chikhi L, Utami SS, de Ruiter J, Bruford MW (2000) A multi-samples multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. Conserv Genet 1:157–162

Gorman ML, Jenkins D, Harper RJ (1978) The anal scent sacs of the otter (*Lutra lutra*). J Zool 186:463–474

Hájková P, Zemanová B, Bryja J, Hájek B, Roche K, Tkadlec E, Zima J (2006) Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Mol Ecol Notes 6:559–562

Hedmark E, Ellegren H (2006) A test of the multiplex pre-amplification approach in microsatellite genotyping of wolverine faecal DNA. Conserv Genet 7:289–293

Idaghdour Y, Broderick D, Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustards. Conser Genet 4:789–792

Jansman HAH, Chanin PRF, Dallas JF (2001) Monitoring otter populations by DNA typing of spraints. IUCN Otter Spec Group Bull 18:12–19

Kohn M, Knauer F, Stoffella A, Schröder W, Pääbo S (1995) Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. Mol Ecol 4:95–103

Kohn MH, Wayne RK (1997) Facts from feces revisited. Trends Ecol Evol 12:223–227

Kohn MH, York EC, Kamradt DA, Haught G, Sauvajot RM, Wayne RK (1999) Estimating population size by genotyping faeces. Proc R Soc Lond B 266:657–663

Liebich HG (1999) Funktionelle Histologie der Haussäugetiere: Lehrbuch und Farbatlas für Studium und Praxis, 3rd edn. Schattauer, Stuttgart

Macdonald SM, Mason CF (1987) Seasonal marking in an otter population. Acta Theriol 32:449–462

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 versus*). Mol Ecol 10:1835–1844

Murphy MA, Waits LP, Kendall KC (2000) Quantitative evaluation of fecal drying methods for brown bear DNA analysis. Wildlife Soc B 28:951–957

Murphy MA, Waits LP, Kendall KC, Wasser SK, Higbee JA, Bogden R (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. Conserv Genet 3:435–440

Nsubuga AM, Robbins MM, Roeder AD, Morin PA, Boesch C, Vigilant L (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Mol Ecol 13:2089–2094

Palomares F, Godoy A, Piriz A, O'Brien J, Johnson WE (2002) Faecal genetic analysis to determine the presence and distribution of elusive carnivores design and feasibility for the Iberian lynx. Mol Ecol 11:2171–2182

Paxinos E, Mcintosh C, Ralls K, Fleischer R (1997) A noninvasive method for distinguishing among canid species amplification and enzyme restriction of DNA from dung. Mol Ecol 6:483–486

Piggott MP, Taylor AC (2003) Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced species. Aust J Zool 51:341–355

Piggott MP, Bellemain E, Taberlet P, Taylor AC (2004) A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. Conserv Genet 5:417–420

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. Trends Ecol Evol 8:285–288

Reed JZ, Tollit DJ, Thompson PM, Amos W (1997) Molecular scatology the use of molecular genetic analysis to assign species sex and individual identity to seal faeces. Mol Ecol 6:225–234

Roeder AD, Archer FI, Poinar HN, Morin PA (2004) A novel method for collection and preservation of faeces for genetic studies. Mol Ecol Notes 4:761–764

Sidransky D (1992) Identification of rare mutations in the stool of patients with curable colorectable tumors. Science 256:102–105

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 24:3189–3194

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. Biol J Linn Soc 68:41–55

Tikel D, Blair D, Marsh HD (1996) Marine mammal faeces as a source of DNA. Mol Ecol 5:456–457

Tschirch W (1995) Neue biomedizinische Methoden für die Feldforschung - dargestellt am Beispiel des Fischotters (*Lutra lutra*). In: Stubbe M, Stubbe A, Heidecke D (eds) Methoden der feldökologischen Säugetierforschung (Methods in mammalian field ecology), vol 1. Martin-Luther-Universität, Halle/Saale, pp 203–210

Vigilant L (1999) An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. Biol Chem 380:1329–1331

Walsh PS, Metzger DA, Higuchi R (1991) Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing form forensic material. Biotechniques 10:506–513

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of ursids. Mol Ecol 6:1091–1097

Welsch U (2006) Lehrbuch Histologie, 2nd edn. Urban & Fischer, München

Chapter 3

How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation – A Review of Available Methods Illustrated by a Case Study

Simone Lampa

Klaus Henle

Reinhard Klenke

Marion Hoehn

Bernd Gruber

The Journal of Wildlife Management 77(8), 1490–1511 (2013)

The Journal of Wildlife Management 77(8):1490–1511; 2013; DOI: 10.1002/jwmg.604

Review

How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation—A Review of Available Methods Illustrated by a Case Study

SIMONE LAMPA,[1] *UFZ—Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstrasse 15, 04318, Leipzig, Germany; and Institute of Ecology, Friedrich Schiller University Jena, Dornburger Strasse 159, 07743, Jena, Germany*

KLAUS HENLE, *UFZ—Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstrasse 15, 04318, Leipzig, Germany*

REINHARD KLENKE, *UFZ—Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstrasse 15, 04318, Leipzig, Germany*

MARION HOEHN,[2] *UFZ—Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstrasse 15, 04318, Leipzig, Germany*

BERND GRUBER, *Institute for Applied Ecology and Collaborative Research Network for Murray-Darling Basin Futures, University of Canberra, ACT 2601, Canberra, Australia*

ABSTRACT The main goal of non-invasive genetic capture-mark-recapture (CMR) analysis is to gain an unbiased and reliable population size estimate of species that cannot be sampled directly. The method has become an important and widely used tool to research and manage wildlife populations. However, researchers have to struggle with low amplification success rates and genotyping errors, which substantially bias subsequent analysis. To receive reliable results and to minimize the time and costs required for non-invasive microsatellite genotyping, one must carefully choose a species-specific sampling design, methods that maximize the amount of template DNA, and methods that could overcome genotyping errors, especially when using low-quality samples. This article reviews the literature and the pros and cons of the main methods used along the process described above. The review is strengthened by a case study on Eurasian otters (*Lutra lutra*) using feces; we tested several methods for their appropriateness to accommodate for genotyping errors. Based on this method testing, we demonstrated that high genotyping error rates are the key problem in this process leading to a severely flawed dataset if no consensus genotype is formed. However, even if generating consensus genotypes minimizes errors dramatically, we show that it may not achieve a definite eradication of all errors, which results in overestimated population sizes if conventional estimators are used. In conjunction with these findings, we offer a step-by-step protocol for non-invasive genetic CMR studies to achieve a reliable estimate of population sizes in the presence of high genotyping error rates. © 2013 The Wildlife Society.

KEY WORDS capture-mark-recapture (CMR), consensus genotypes, Eurasian otter (*Lutra lutra*), fecal DNA, microsatellites, screening approach.

The protection and management of cryptic, elusive, and vulnerable species often require reliable information about population size and distribution, relatedness, sex ratios, dispersal distances, and genetic diversity. Non-invasive genetic sampling (NGS) is becoming increasingly popular to obtain such parameters that were previously difficult to assess. Thus far, feces (Table 1), hair (Table 1), urine (e.g., Hausknecht et al. 2007), saliva (e.g., Sundqvist et al. 2008), eggshells (e.g., Martín-Gálvez et al. 2011), sloughed skin (Palsboll et al. 1997), and feathers (e.g., Johansson et al. 2012) have been used successfully as non-invasive DNA sources to generate multilocus genotypes for individual identification on a range of species (for a review: Waits and Paetkau 2005, Beja-Pereira et al. 2009). If samples are collected at several points in time, capture-mark-recapture models (CMR) can be used to obtain demographic parameters such as survival, migration, fecundity, and population growth or size (Lukacs and Burnham 2005*a*). Each sample is genotyped at multiple molecular loci, such as microsatellites or nucleotide polymorphisms (SNP). Matching genotypes are deemed to derive from the same individual and classified as recaptures. Non-matching genotypes indicate a newly captured animal. Hence, for each sampling

Received: 11 September 2012; Accepted: 7 June 2013
Published: 8 October 2013

[1]*E-mail: simone.lampa@ufz.de*
[2]*Present address: Commonwealth Scientific and Industrial Research Organisation, ACT 2601, Canberra, Australia*

Table 1. Characteristics of studies from 2002 to present that estimated population size using microsatellite genotyping of non-invasive samples of wild mammal populations assuming closed populations. We located studies using a search in ISI Web of Knowledge using the keywords microsatellite genotyping capture mark recapture, non-invasive genetic CMR, and non-invasive population size estimation, and limited search results to mammal studies only. We included a maximum of 4 studies per species. We only included 1 study from identical author groups using the same methods but for different species. Parameters related to the population size estimation are species, DNA source (F = feces; H = hair), probability of identity (PI = expected PI; PI_{sib} = PI for siblings; PI_{obs} = observed PI), heterozygosity (het; He = expected heterozygosity; Ho = observed heterozygosity), number of loci used for population size estimation (+s = plus sex marker), sampling period (SP) in days, genotyping error rate (GER; AD = allelic dropout rate; FA = false allele rate; PCR = polymerase chain reaction), method to overcome genotyping errors (GEs), genotyping success rate (GSR), and method for population size estimation (PSE). Missing information is marked by a dash (–).

Study	Species	DNA	PI	Het	Loci	SP	GER (%)	Method to overcome GEs[a]	GSR (%)	Method for PSE
Arandjelovic et al. (2011)	*Pan troglodytes troglodytes* (central chimpanzee)	F	$PI_{sib} \leq 1 \times 10^{-3}$	—	8	182–365	AD = 16 FA = 2[b] (over all PCRs)	• Pre-screening with qPCR and sex marker and discard of LQS • Multiplex pre-amplification • Similar assumptions to cMTA • 3 PCRs and discard of LQS • 1–7 repetitions for missing loci • 1 MM, 2 MM, 3 MM check	For 8 loci: 40.3	CAPWIRE (TIRM)
Arrendal et al. (2007)	*Lutra lutra* (Eurasian otter)	F	$PI_{sib} = 5.5 \times 10^{-3}$	—	8 +s	17	AD = 7.7	• A variation of MTA similar to cMTA	30.7	CAPWIRE (TIRM)
Banks et al. (2002)	*Vombatus ursinus* (wombat)	F	For 6 loci: $PI = 7.7 \times 10^{-7}$ $PI_{sib} = 6.6 \times 10^{-3}$	For 6 loci: He = 0.72	5 +s	12	—	• 3 PCRs	82.8	CAPTURE ($M_{h\text{-jackknife}}$)
Bellemain et al. (2005)	*Ursus arctos* (brown bear)	F	$PI = 1.4 \times 10^{-6}$ $PI_{sib} = 4.5 \times 10^{-3}$	Ho = 0.70	6 +s	77–91	GER = 2 (for 5% of samples)	• Multiplex pre-amplification of Piggott et al. (2004) • 4 PCRs	68.2–79.4	Hyperbolic accumulation curve (Kohn et al. 1999); exponential accumulation curve (Eggert et al. 2003); closed capture models in MARK; Lincoln–Peterson estimator
Brinkman et al. (2010*a*, 2011)	*Odocoileus hemionus sitkensis* (Sitka black-tailed deer)	F	$PI = 3 \times 10^{-4}$ $PI_{sib} = 2.1 \times 10^{-2}$	—	7	3 × 62 within 3 years	GER = 10.2 (for 31% of samples)	• MTA	51	Huggins closed capture models in MARK (per year)
Chang et al. (2012)	*Rhino-pithecus roxellana* (Sichuan snub-nosed monkeys)	F	$PI = 2.7 \times 10^{-11}$ $PI_{sib} = 2.1 \times 10^{-5}$	He = 0.59 Ho = 0.59	16	~365	—	• Pre-screening with qPCR and discard of LQS • Multiplex pre-amplification • Similar assumptions to cMTA • 3 PCRs • Use of DROPOUT • Use of MICRO-CHECKER	58.2	CAPWIRE, CAPTURE ($M_{h\text{-Chao}}$)

(*Continued*)

Table 1. (*Continued*)

Study	Species	DNA	PI	Het	Loci	SP	GER (%)	Method to overcome GEs[a]	GSR (%)	Method for PSE
Cullingham et al. (2010)	*Vulpes velox* (swift fox)	F	$PI = 2.6 \times 10^{-9}$ $PI_{sib} = 3.8 \times 10^{-4}$	He = 0.74 Ho = 0.76	5–9	153[c]	GER (per locus) = 0–4.3 (for 26.3% of the samples)	• Pre-screening with qPCR and 1 marker and discard of LQS • 2 PCRs for 26.3% of the samples • Use of DROPOUT	For 9 loci: 63 (after pre-screening)	CAPWIRE
Dreher et al. (2007)	*Ursus americanus* (black bear)	H	$PI = 2.7 \times 10^{-6}$ $PI_{sib} = 8.7 \times 10^{-3}$	He = 0.76 Ho = 0.75	5	35	GER = 4.4 (for 9.4% of the samples)	• Quality-control protocol (Paetkau 2003)	60.6	Closed capture models in MARK; misidentification model in MARK
Ebert et al. (2012)	*Sus scrofa* (wild boar)	F	$PI = 4.7 \times 10^{-8}$ $PI_{sib} = 1.3 \times 10^{-3}$	He = 0.79 Ho = 0.83	6 +s	14	AD = 4.72[b] FA = 0.002[b]	• Pre-screening with qRT-PCR and discard of LQS Assumptions of cMTA 3–6 PCRs and discard of LQS	90.3 (47.4 of all samples)	Closed capture models in MARK
Eggert et al. (2003)	*Loxodonta cyclotis* (African forest elephant)	F	For 7 loci: $PI = 6 \times 10^{-5}$ $PI_{sib} = 1.8 \times 10^{-3}$	For 7 loci: He = 0.55 Ho = 0.51	5–7 +s	15	GER = 15.7 (no information about calculation)	Pre-screening with mtDNA and discard of LQS • Assumptions of cMTA • For ambiguous samples MTA	For ≥5 loci: 60–72	CAPTURE ($M_{h\text{-jackknife}}$); exponential accumulation curve; hyperbolic accumulation curve
Flagstad et al. (2004)	*Gulo gulo* (wolverine)	F	$PI = 1.2 \times 10^{-5}$ $PI_{sib} = 7.8 \times 10^{-3}$	He = 0.53–0.55 Ho = 0.54–0.58	9 +s	91	AD = 9.8[b] FA < 0.5	Pre-screening with 1 marker and discard of LQS • Assumptions of cMTA • 2–6 PCRs	77	CAPTURE
Frantz et al. (2003)	*Meles meles* (Eurasian badger)	F	For 5 loci: $PI_{obs} = 1 \times 10^{-3}$ $PI_{sib} \approx 6 \times 10^{-3}$	—	7	10	AD = 27[b] FA = 8	• cMTA	74	CAPTURE
Guertin et al. (2012)	*Lontra canadensis* (river otter)	F	$PI = 2.4 \times 10^{-6}$ $PI_{sib} = 5 \times 10^{-3}$	He = 0.57 Ho = 0.52	8	6–8	AD = 12[b] FA = 2[b]	• Pre-screening with 3 markers and discard of LQS • cMTA • Check for non-random geospatial distribution of samples (Smith et al. 2006)	For ≥ 7 loci: 12.3	Robust design in MARK; CAPWIRE
Hajkova et al. (2009)	*Lutra lutra* (Eurasian otter)	F	Mean for 10 loci: $PI = 8.9 \times 10^{-6}$ $PI_{sib} = 4.4 \times 10^{-3}$ Mean for 6 loci: $PI = 8.2 \times 10^{-4}$ $PI_{sib} = 3.5 \times 10^{-2}$	For 10 loci: He = 0.53 Ho = 0.5	6–10 +s	30–120	AD = 18[b] FA = 2.9[b]	• MTA	54.8–63.1	CAPWIRE
Harris et al. (2010)	*Ovis ammon* (argali)	F	For 10 loci: $PI = 4.5 \times 10^{-11}$ $PI_{sib} = 1.1 \times 10^{-4}$ For 6 loci: $PI = 1.1 \times 10^{-6}$ $PI_{sib} = 4.4 \times 10^{-3}$	He = 0.74	6–10 +s	381	AD = 5 FA = 1	• Pre-screening with 1 marker and discard of LQS • 3–4 PCRs for selected samples • DCH test using DROPOUT	For ≥ 6 loci: 59	Closed capture models in MARK; CAPWIRE

(*Continued*)

Table 1. (*Continued*)

Study	Species	DNA	PI	Het	Loci	SP	GER (%)	Method to overcome GEs[a]	GSR (%)	Method for PSE
Mondol et al. (2009)	*Panthera tigris* (tiger)	F	For 10 loci: $PI_{sib} = 5 \times 10^{-4}$ For 7 loci: $PI_{sib} = 7 \times 10^{-4}$	For 10 loci: He = 0.77 Ho = 0.59	5–10	42	AD = 0.67 FA = 0 (over all PCRs)	• 4 PCRs	For 10 loci: 20.7 For ≥ 5 loci: 65.5	CAPTURE
Mowat and Paetkau (2002)	*Martes americana* (American marten)	H	For 6 loci: $PI_{sib} = 8.1 \times 10^{-3}$	For 6 loci: Ho = 0.69	4–6	14	—	• 1–3 PCRs and discard of LQS • 1 MM check	For 6 loci: 64.4 For ≥ 4 loci: 77.2	CAPTURE
Piggott et al. (2006)	*Petrogale penicillata* (brush-tailed rock-wallaby)	F	For 6 loci: $PI = 2.2 \times 10^{-6}$ $PI_{sib} = 7.2 \times 10^{-3}$	—	7 +s	8	AD = 2.77 FA = 1.48 (over all PCRs)	• Assumptions of cMTA • Maximum of 6 PCRs • 1 MM and 2 MM check	70–81	CAPTURE; hyperbolic accumulation curve (Kohn et al. 1999)
Poole et al. (2011)	*Oreamnos americanus* (mountain goat)	F/H	$PI = 1.8 \times 10^{-6}$ $PI_{sib} = 3.2 \times 10^{-3}$	Ho = 0.66	7 +s	43 (57 including scat age)	—	• Quality-control protocol (Paetkau 2003) • 4 PCRs for intermediate quality samples and discard of LQS • 1 MM and 2 MM check	35 (53 of analyzed samples)	Closed capture models in MARK; CAPTURE
Prigioni et al. (2006)	*Lutra lutra* (Eurasian otter)	F	—	—	10	166	—	• MTA	41.2	Hyperbolic accumulation curve (Kohn et al. 1999); exponential accumulation curve (Eggert et al. 2003)
Puechmaille and Petit (2007)	*Rhinolophus hipposideros* (lesser horseshoe bat)	F	For 8 loci: $PI = 1 \times 10^{-8}$ -8.2×10^{-8} $PI_{sib} = 7.4 \times 10^{-4}$ -1.2×10^{-3}	For 8 loci: He = 0.70–0.72 Ho = 0.70–0.77	6-8	1	AD = 7.65[b] FA = 2.72[b]	• Pre-screening with mtDNA and discard of LQS • Assumptions of cMTA and discard of LQS • 1 MM and 2 MM check	For ≥ 6 loci: 91.1	CAPWIRE; sequential Bayesian method (Petit and Valière 2006)
Robinson et al. (2009)	*Ursus americanus* (black bear)	H	For 7 loci: $PI = 1.2 \times 10^{-8}$ $PI_{sib} = 1 \times 10^{-3}$ For 5 loci: $PI = 1.4 \times 10^{-6}$ $PI_{sib} = 7 \times 10^{-3}$	—	5–7	11	GER = 1.9 (no information about calculation)	• 1 PCRs each and discard of LQS • 1–2 repetitions • 1 MM and 2 MM check	For 7 loci: 29.6 For ≥ 5 loci: 55.4	CAPWIRE; CAPTURE; closed capture models in MARK
Ruell et al. (2009)	*Lynx rufus* (bobcat)	F	$PI_{obs} = 0$ $PI_{sib} = 2 \times 10^{-2}$	—	4 +s	16	GER = 0.0004 (expected)	• Pre-screening with mtDNA and discard of LQS • 3–6 PCRs following MTA • 1 MM and 2 MM check	88–90 (after pre-screening)	CAPTURE ($M_{h\text{-jackknife}}$; $M_{h\text{-Chao}}$); CAPWIRE
Ruibal et al. (2009)	*Dasyurus maculatus* (spotted-tailed quoll)	F	$PI = 3.2 \times 10^{-6}$ $PI_{sib} = 3 \times 10^{-3}$	—	7–10	56	AD = 5.4[b] FA = 0.6[b] (calculated for the first 2 PCRs)	• cMTA • Comparison with reference profiles from tissue samples • Use of DROPOUT for unknown genotypes	For 10 loci: 39 For ≥ 8 loci: 57	Closed capture models in MARK; CAPWIRE; misidentification model in MARK
Scheppers et al. (2007)	*Meles meles* (Eurasian badger)	H	$PI_{sib} = 6.6 \times 10^{-3}$	—	7	28	No error was detected in repetitions	• Repetition of unique genotypes • 1 MM and 2 MM check	94.6	CAPWIRE

(*Continued*)

Table 1. (*Continued*)

Study	Species	DNA	PI	Het	Loci	SP	GER (%)	Method to overcome GEs[a]	GSR (%)	Method for PSE
Stenglein et al. (2010)	*Canis lupus* (gray wolf)	F/H	For 9 loci: $PI_{sib} = 3.8 \times 10^{-4}$	—	9 +s	2 × 67	For scat: AD = 13[b] FA = 3[b] For hair: AD = 20[b] FA = 5[b]	• Pre-screening with mtDNA and discard of LQS • Assumptions of cMTA • 2 PCRs and discard of LQS • 1–3 repetitions for missing loci and discard of LQS • Use of RELIOTYPE for unique genotypes	51–52	CAPWIRE; Bayesian method (Petit and Valière 2006)
Solberg et al. (2006)	*Ursus arctos* (brown bear)	F	For 7 loci: $PI = 1.4 \times 10^{-6}$ $PI_{sib} = 4.5 \times 10^{-3}$	Ho = 0.70	6 +s	77–91	GER < 2 (calculated for 5% of PCRs)	• Multiplex pre-amplification of Piggott et al. (2004) • 4 PCRs	70–80	Closed capture models in MARK
Tredick et al. (2007)	*Ursus americanus* (black bear)	H	Exclusion of genotypes with $PI_{sib} \geq 0.01$	—	6	56	—	• Quality-control protocol (Paetkau 2003)	85–87.8	CAPTURE
Van Manen et al. (2012) (only post-construction phase)	*Ursus americanus* (black bear)	H	$PI = 1.1–5.8 \times 10^{-9}$ $PI_{sib} = 3.0–4.6 \times 10^{-4}$	—	10 +s	49	—	• Pre-screening with 7 markers and discard of LQS • 1 repetition for missing loci and discard if not successful • 1 MM and 2 MM check	51–65	Closed capture models in MARK
Williams et al. (2009)	*Martes americana*/ *M. pennanti* (American marten/ fisher)	H	$PI = 3.5/3.4 \times 10^{-6}$ $PI_{sib} = 5.6/5.5 \times 10^{-3}$	He = 0.72/0.71	6 +s	35	GER = 4.3/5.8 (obtained by 1 MM and 2 MM check)	• Quality-control protocol (Paetkau 2003) with 2 initial PCRs	—	CAPWIRE; CAPTURE; misidentification model in MARK
Wilson et al. (2003)	*Meles meles* (Eurasian badger)	F	$PI_{sib} = 0.01$ only cited from Frantz et al. (2003)	—	7	10	—	• Assumptions of cMTA • 3–4 PCRs	74	Hyperbolic accumulation curve (Kohn et al. 1999)
Zhan et al. (2006, 2009)	*Ailuropoda melanoleuca* (giant panda)	F	—	Ho = 0.63	9 +s	335	—	• MTA	—	CAPWIRE

[a]LQS = low-quality samples, cMTA = comparative multiple-tubes appraoch of Frantz et al. (2003), MTA = multiple-tubes appraoch of Taberlet et al. (1996), MM = genotypes that mismatched at 1 (1 MM), 2 (2 MM), or 3 alleles (3 MM), DCH = difference in capture history test.
[b]AD and FA rate were calculated following Broquet and Petit (2004) and are comparable with the rates obtained in our study.
[c]Calculated for all samples of 2 sampling periods with 61 days each.

occasion, the respective individual is determined to be newly captured, recaptured, or not captured, resulting in individual capture histories that are used for CMR analysis. The combination of non-invasive microsatellite genotyping and CMR models, referred to as non-invasive genetic CMR, have been used increasingly to study and monitor populations, predominantly to estimate population size, of several species (Table 1).

Despite the great potential of non-invasive genetic CMR, many difficulties must be overcome in practice, such as low success rates or genotyping errors and hence, potentially biased population size estimates. Three consecutive steps are involved in genetic CMR approaches to estimate population size, each entailing specific challenges (Fig. 1). The first step is to obtain target DNA from field collected samples. Here, the difficulty is to find a sampling design that is appropriate for the target species (e.g., ensuring sufficient numbers of collected samples) and the assumptions of the selected CMR model (e.g., population closure, equal capture probabilities among individuals; Williams et al. 2002). As target DNA in non-invasive samples, especially in feces, is often degraded, only available in low quantities, and contaminated by polymerase chain reaction (PCR) inhibitors (Kohn et al. 1995, Creel et al. 2003, Idaghdour et al. 2003), researchers are also challenged to find suitable collection, preservation, and extraction methods that inhibit further DNA degradation and yield a sufficient amount of high quality target DNA. The second step is the microsatellite genotyping of the DNA samples. Because of low DNA quantity and quality in non-invasive samples, researchers have to cope with low success rates, genotyping errors, and contamination susceptibility; with low numbers of template molecules, contaminant molecules will likely be amplified (Adams et al. 2003). All 3 factors can substantially hamper the subsequent analysis (Creel et al. 2003, McKelvey and Schwartz 2004, Hoffman and Amos 2005, Schwartz et al. 2006). At the third step, the population size needs to be estimated based on the genotyped samples. Here, the challenge is to decide if the data still correspond to the assumptions of the CMR model selected in the first step (e.g., closed vs. open population, equal vs. varying capture probability, no-error vs. error-incorporating models) and to fit this or the adjusted model to the data.

As difficulties can arise at each step and affect the outcome and suitability of methods in the next one, the process is not straightforward. A number of comprehensive review articles address data collection, preservation, extraction, amplification, genotyping errors, and/or data analysis (Lukacs and Burnham 2005*a*, Pompanon et al. 2005, Waits and Paetkau 2005, Beja-Pereira et al. 2009, Marucco et al. 2011). But, a thorough and systematic evaluation of advantages and disadvantages of existing methods and a concrete road map for method selection that covers each step of the process from data collection to analysis is still lacking. We provide a rigorous review that tackles difficulties along each of the steps from field sampling to data analysis and offer a tested step-by-step protocol for conducting more reliable population size estimations using non-invasive genetic methods. To further strengthen our discussion and to offer on-the-ground guidelines, we complement the review by testing some of the methods in a non-invasive genetic CMR study on a wild otter (*Lutra lutra*) population using fecal DNA. These data are particularly relevant for this review because otter feces are notorious for low DNA quality and quantity, with high error rates and low genotyping success (Prigioni et al. 2006, Arrendal et al. 2007, Ferrando et al. 2008, Lampa et al. 2008, Lanszki et al. 2008). Using the employed example, we illustrate that some currently favored methods may not be suitable when facing low genotyping success and very high genotyping errors. Thereby, we offer tools to optimize the allocation of time and money and to improve the reliability of the results for science and application.

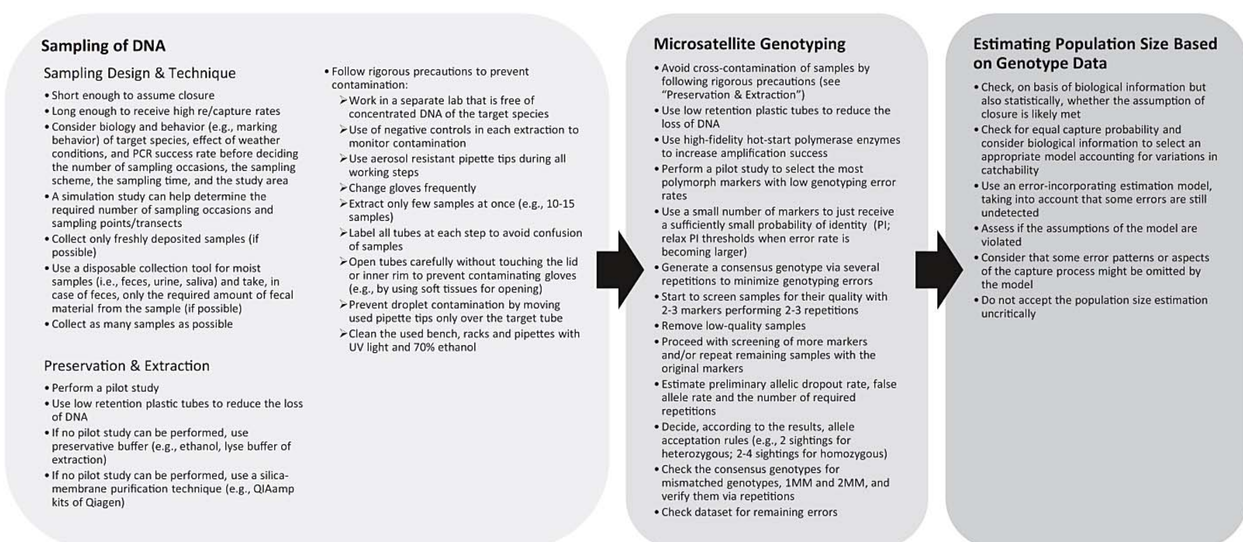

Figure 1. Flow chart and summary of points that should be considered if non-invasive samples with low genotyping success rate and high genotyping error rate are used for genetic capture-mark-recapture analysis.

The article is divided into 3 sections: 1) sampling of DNA, 2) microsatellite genotyping, and 3) estimating population size. Each section starts with a description of the basic requirements, followed by a review of the most commonly applied approaches and tools, and a discussion of their benefits and weaknesses. Finally, we offer recommendations and conclude with the case study to illustrate several approaches and to further offer a practical workflow and a tree for decision support. Supplementarily, we provide a simple step-by-step guide (Table S1, available online at www.onlinelibrary.wiley.com) that summarizes all relevant tasks and their potential solutions with pros and cons.

Because feces is the most commonly used non-invasive material that is easy to find in the wild and provides more information than other samples (e.g., diet, hormones, parasites; Beja-Pereira et al. 2009) and because of the employed case study, the review focuses on fecal samples. However, as the problem of genotyping errors in CMR analysis arises with all kinds of non-invasive samples, we believe that the conclusions and the protocol are applicable for any non-invasive genetic CMR study dealing with high genotyping error rates. We also note that this review does not attempt to cover all available software for analyses, because some were designed for relatively specific purposes and are not widely applicable.

FIRST STEP: SAMPLING OF DNA

Sampling Design

The sampling design is a crucial factor determining the quality of population size estimation. The most precise and unbiased estimates will be received when the sampled population is—geographically and demographically—closed (i.e., no birth, death, or migration), as this requires fewer estimated parameters and allows more flexible assumptions. The assumption of closed populations is most common in genetic based population size estimation (Table 1)—but note that several CMR models for open populations also allow estimation of population size (Pollock et al. 1990, Williams et al. 2002). Geographic closure (no movement on or off the study area) and demographic closure (no birth or death) may be achieved when the study area is either isolated or large enough so that movements across the borders of the study area by a few individuals are likely to be negligible and when the sampling period is kept as short as possible (e.g., several days; Pollock et al. 1990). To determine how many sampling occasions are necessary for a reliable size estimate, we recommend conducting a simulation in advance (e.g., in Program MARK [White and Burnham 1999]). This requires a worst-case estimate of 4 parameters: the minimum number of samples per day and animal, the maximum number of animals in the study area, the minimum number of samples found, and the minimum genotyping success rate (see the Case Study Section). Simulating and comparing different numbers of sampling occasions (e.g., 3–12) can then be used to select the minimum number of required sampling occasions. Fewer occasions require more restrictive model assumptions that are often difficult to meet (Mowat and Strobeck 2000, see Third Step Section). The interval between each sampling occasion should be long enough to ensure that the behavior of the focal species is not altered extremely by the frequent disturbance and that new samples can be deposited by all individuals. Each individual must have a reasonable chance of being sampled (no individual heterogeneity) at each sampling occasion (no time variation) and that already sampled individuals do not react to the sampling (no behavioral response). Although equal catchability is difficult to achieve in the field and variations can be incorporated to a certain degree into closed CMR models, they should be minimized to the maximum extent as extreme variations may lead to non-estimable parameters, decreased estimate precision, or even inconclusive or over- or underestimated size estimates (Williams et al. 2002). This requires an objective sampling scheme that is determined by the target species, their biology, and the sampled population (Williams et al. 2002, Garton et al. 2012). Depending on this, researchers have to decide what kind of non-invasive material to collect, where to collect those samples, when to conduct the study, and how many samples are required.

The right choice of the DNA source, whether to sample actively (e.g., hair snares around bait stations) or passively (e.g., feces, feathers, eggs, urine), increases the capture probability. Collecting not only 1 sample type of a species but several, like feathers and feces for birds (Jacob et al. 2010), can further increase capture probability. This approach could also be applied to individuals with different attributes (e.g., sex, age) to decrease unequal capture probability (e.g., urine and saliva for subadults but hairs and feces for adults; Inoue et al. 2007). For sparsely distributed mammals with large territories (e.g., bears), researchers often used active hair sampling around bait stations to increase capture rates, minimize variations in capture probability, meet the closure assumption, and receive hairs with many large root bulbs (Beja-Pereira et al. 2009, Marucco et al. 2011). However, researchers might have difficulty preventing possible behavioral responses, avoiding potential cross-contamination between species, or ensuring that all individuals are equally attracted to baits (Boulanger et al. 2006, Beja-Pereira et al. 2009).

If feces were used, most studies collected the samples using transects (e.g., Kohn et al. 1999, Banks et al. 2002, Ruell et al. 2009, Brinkman et al. 2010*a*, Jacob et al. 2010), following trails (e.g., Eggert et al. 2003, Flagstad et al. 2004, Mondol et al. 2009, Cullingham et al. 2010), or using sampling points such as latrine (e.g., Frantz et al. 2003, Wilson et al. 2003), marking (e.g., Ruibal et al. 2009), feeding (e.g., Flagstad et al. 2004), or resting sites (e.g., Piggott et al. 2006, Puechmaille and Petit 2007). Alternative DNA sources were either collected around bait stations (hairs) or passively in nests (hairs, feathers, eggs), near bedding sites (hairs), along trails (hairs, urine), in tracks (hairs), or in bite wounds (saliva; see Beja-Pereira et al. 2009 for a review). Regardless which method is used, all sampling sites should be visited at each sampling occasion to reduce the risk that field personal preferentially search at previously successful locations, which would introduce bias due to

individual heterogeneity and/or a behavioral response. The distribution of sampling points or transects within the study area and the time spend at each site has to ensure that each member of the population has an equal chance to be collected and that several individuals are sampled more than once (see Garton et al. 2012 for a review). For example, choosing only latrines as sampling points for feces in species where dominant animals defecate more often or exclusively on prominent sites, investigators might miss transients, floaters, and subordinates.

The time period of collection should also be well chosen, as variations in animal behavior due to differences in sex, age, social and reproductive status, seasons, or weather conditions may cause individual or temporal heterogeneity in capture probability (e.g., different behavior of males and females during breeding season; Petit and Valière 2006, Ruibal et al. 2009). Seasonal weather conditions can also reduce the DNA quantity or quality of samples (e.g., DNA washed off by rain or DNA degradation by UV light; Murphy et al. 2007, Ruibal et al. 2009, Brinkman et al. 2010*b*).

The best way to minimize chances for variations in capture probability is to get capture probability as high as possible (Lukacs and Burnham 2005*a*). This can be done by collecting large numbers of samples at each sampling occasion, especially when genotyping success rate is low. Previous literature recommends collecting as many samples as can be found (Marucco et al. 2011) or at least 2.5–3 times as many samples than assumed individuals (Solberg et al. 2006). The samples being analyzed can then be chosen randomly or by using a spatial pattern as criterion for prioritization (Adams et al. 2003, Marucco et al. 2011). Remaining samples can be analyzed subsequently if more samples are required for certain sampling sessions or if samples failed to amplify. Thereby, researchers should consider that genotyping success can vary within the same type of non-invasive samples (for instance, the amount of slimy secretions on feces can increase genotyping success rate; Hajkova et al. 2006, Lampa et al. 2008). Furthermore, only fresh samples should be taken, as samples deposited before the first sampling occasion may violate the closure assumption. For feces, fresh samples also result in greater amplification success, as DNA is comparatively less degraded (Santini et al. 2007). Thus, before the first sampling and after each sampling, researchers must remove old scats or mark them if removal may perturb the species social behavior, as the case for otters (see the Case Study Section). This will also avoid double sampling or contamination among individuals.

Sampling, Preservation, and Extraction Methods

Especially for moist samples, the sampling technique can be critical for the success of a study. Saliva is usually taken with swabs (Blejwas et al. 2006, Inoue et al. 2007, Sundqvist et al. 2008, Sastre et al. 2009). Urine is sampled either using disposable plastic tools (Hayakawa and Takenaka 1999, Inoue et al. 2007) or as a frozen snow–urine mixture (Hayakawa and Takenaka 1999, Inoue et al. 2007). Feces are either sampled entirely (Kohn et al. 1999, Solberg et al. 2006), parts are cut off (Bellemain et al. 2005, Hajkova et al. 2009), or only the surface is taken (Frantz et al. 2003, Wilson et al. 2003). For entire scats or cut-off parts, surface wash seemed to increase amplification success and to decrease genotyping errors compared to homogenization (Flagstad et al. 1999, Palomares et al. 2002, Piggott and Taylor 2003). However, sampling entire scats may alter the marking behavior of the target species if feces are used for intraspecific communication (e.g., individuals may defecate more frequently subsequent to the collection of their scats). An alternative could be to scrape off the surface of the scat with a disposable collection tool like toothpicks or cotton swabs to avoid cross-contamination. This technique also decreases the risk of sampling prey hard parts, while increasing the proportion of sloughed gut cells (Lampa et al. 2008). To decide the best sampling technique, preservation method, extraction method, and sampling season, we advise a pilot study.

Although hairs or feathers are usually stored either frozen or at room temperature in an envelope with silica gel (see Beja-Pereira et al. 2009 for a review), the preservation of moist samples (i.e., urine, saliva, feces) is more difficult, since DNA degradation caused by bacteria, enzymes, oxidation, or hydrolysis has to be inhibited. The most favorable storing approach should be easy to handle given the collection method (e.g., cotton swab vs. entire scats), have no negative effect on the employed extraction technique, and simplify the upcoming extraction. Several approaches are available for the preservation of DNA all aiming to deactivate nucleases by removal of water or cations or by storing at low temperatures (see Beja-Pereira et al. 2009 for a review). Regarding feces, most researchers stored samples in ethanol, dried them, or froze them; nevertheless, the use of buffers like DETs (Frantzen et al. 1998, Frantz et al. 2003), RNA later solution (Nsubuga et al. 2004, Beja-Pereira et al. 2009, Vigilant and Guschanski 2009), or the lyses buffer of the employed extraction method (Hajkova et al. 2006, Santini et al. 2007; see the Case Study Section) were also very efficient. Especially, the latter is feasible for all sampling techniques and eases the extraction substantially, since the first extraction step is already done. Furthermore, sloughed gut cells in the feces that are dissolved in the buffer will be targeted to the extraction instead of poured away as with ethanol, DETs, or RNA later solution. Most studies comparing preservation methods (Frantzen et al. 1998, Murphy et al. 2002, Frantz et al. 2003, Hajkova et al. 2006, Santini et al. 2007) indicated that storage in preservative buffers is preferable to merely freezing or drying. In our experience, storage in the first buffer of the employed extraction kit was not only highly practicable and timesaving, but also increased extraction success rate. This method could also be used for other non-invasive samples like hair, feathers, urine, or saliva. However, a pilot study might often be required to decide the best-suited method (Bhagavatula and Singh 2006, Beja-Pereira et al. 2009).

During extraction, PCR inhibitors, bacteria, and enzymes should be removed (Kohn et al. 1995, Reed et al. 1997, Frantzen et al. 1998), and the greatest possible amount of DNA should be isolated. At the same time, the extraction

technique should be quick, cheap, and easy to handle (Reed et al. 1997). Beja-Pereira et al. (2009) provided a helpful review of existing methods for several DNA sources. These can be divided into 3 main groups: 1) phenol–chloroform extraction, 2) resin-based extraction (e.g., Chelex®, Sigma-Aldrich, Seelze, Germany), and 3) silica-based extraction, often in combination with GuSCN buffers, the latter being the most frequently used method for fecal DNA. Almost all studies in our review (Table 1) used either the QIAamp® DNA Stool Mini Kit (Qiagen, Hilden, Germany) or other silica-based spin column kits. Many studies showed that this technique is most efficient in dealing with PCR inhibitors and yields sufficient amounts of high quality target DNA for feces (Idaghdour et al. 2003, Piggott and Taylor 2003, Bhagavatula and Singh 2006, Lampa et al. 2008) and other non-invasive samples (Beja-Pereira et al. 2009). During extraction, contamination is the most critical factor and must be rigorously prevented (see Fig. 1). To reduce the loss of DNA, Beja-Pereira et al. (2009) recommended using low retention plastic tubes.

SECOND STEP: MICROSATELLITE GENOTYPING

Errors in genotyping and insufficient discriminatory power of genotypes may result in overestimation or underestimation of population size. Overestimation arises from ghost individuals that are produced if samples originating from the same individual do not show identical genotypes, because of genotyping errors. This creates non-existent individuals. They are recorded as new captures in CMR models, increase the number of unique individuals, and decrease recapture rates; hence, they lead to an overestimation of population size (Creel et al. 2003, McKelvey and Schwartz 2004). Underestimation will occur if samples from 1 individual are considered to belong to another, already existing individual, because of genotyping errors, or if 2 different individuals share the same genotype (shadow effect) because of a lack of power in distinction or because they are identical twins. This results in fewer unique individuals, raised recapture rates, and thus an underestimated population size. Therefore, genotyping errors, as well as the shadow effect, can cause severe problems in genetic CMR and have to be tackled. In the following section, we first address the shadow effect and subsequently genotyping errors, reviewing the causes for both and how they can be reduced.

Shadow Effect

A shadow effect is introduced into CMR studies if animals that have not been captured before are believed to be recaptures because their genotype is an indistinguishable shadow of previously captured animals (Mills et al. 2000, Waits et al. 2001). The proportion that 2 different individuals will share the same genotype at multiple loci in a given population is defined as the probability of identity (PI). It is a measure for the discriminatory power of the employed marker system and should guide the decision for the number of employed loci. With an increasing number of markers, the PI will decrease and so will the bias induced by the shadow effect. As an initial guideline for adjusting the number of loci, previous studies have recommended using a theoretical unbiased PI between 1×10^{-3} (Mills et al. 2000, Waits et al. 2001) and 1×10^{-6} (McKelvey and Schwartz 2004) or an observed PI (PI_{obs}) of 1×10^{-2} to 1×10^{-4} (Waits et al. 2001). Since the risk for a shadow effect will increase with increasing size of the target population, increasing kinship, decreasing heterozygosity of markers, and decreasing number of individuals sampled, the PI threshold should be chosen more conservatively in these cases. For instance, Paetkau (2003) suggested for small populations ($n < 100$), to use a 6-locus or 5-locus system when heterozygosity (He) is reaching a level of 0.69 or 0.78, respectively, but to increase these minimums to 0.75 or 0.83, respectively, for large populations ($200 < n > 400$). For populations with many related individuals, Waits et al. (2001) recommended the use of the theoretical PI for siblings (PI_{sib}) as a conservative upper bound, albeit Rew et al. (2011) revealed PI_{sib} as an overly conservative measurement.

Although increasing the number of markers will reduce the shadow effect, it will also increase the costs and the probability of genotyping errors. Because genotyping errors occur at the level of the locus, the number of errors will be the per-locus error rate × the number of loci × the number of samples (McKelvey and Schwartz 2004). For instance, when genotyping 7–10 loci, microsatellite genotyping errors can inflate mark–recapture population estimates up to 200% (Waits and Leberg 2000) or even up to 550% if 13 loci are analyzed (Creel et al. 2003). Hence, the number of typed loci should be chosen as a compromise between minimizing genotyping errors and minimizing the probability of identity (Mills et al. 2000, Waits and Leberg 2000, Waits et al. 2001, Creel et al. 2003, Pompanon et al. 2005). Because the bias in population size estimation caused by the shadow effect is small compared to the bias caused by even low levels of genotyping errors (Mills et al. 2000, Waits and Leberg 2000, Ruell et al. 2009), researchers should relax all threshold values and use few but highly variable loci (Lukacs and Burnham 2005*a*, Knapp et al. 2009). Using fewer loci (<10) also has the advantage of saving time, money, and extract. Waits and Leberg (2000), for example, recommended using 4–5 loci when the population in question is small ($n \leq 50$) and error rates are high (0.05 per locus). Previous studies that estimated population sizes applied 4–16 marker systems with heterozygosities above 0.5 and with a PI ranging from 8.2×10^{-4} to 2.7×10^{-11} (Table 1).

Genotyping Errors

Genotyping errors have been found in essentially all studies based on non-invasive samples (Creel et al. 2003, Bonin et al. 2004, Broquet and Petit 2004, Pompanon et al. 2005) with rates usually ranging between 0.7% and 35% (Table 1). Even modest error rates can substantially bias estimates of population size (Creel et al. 2003, McKelvey and Schwartz 2004, Hoffman and Amos 2005, Schwartz et al. 2006). Despite several studies pointing out concern about genotyping errors (Taberlet et al. 1996, Miller

et al. 2002, Creel et al. 2003, Bonin et al. 2004, McKelvey and Schwartz 2004), only 26% of studies between 2009 and 2010 reported genotyping error rates (Guichoux et al. 2011). Although the awareness increased since then, studies are still published that do not report an error rate (Table 1: 30% in 2010–2012).

The causes for genotyping errors can derive from a range of sources associated with molecular analysis (Pompanon et al. 2005). Most influential for CMR studies are human failures (e.g., mishandling data, misscoring of alleles, data entry errors, cross-contamination of samples), biochemical artifacts (e.g., slippage of Taq Polymerase, lack of specificity), and low sample quality (Broquet and Petit 2004, Pompanon et al. 2005). The consequences are either allelic dropout (AD; failure of 1 allele to amplify in a heterozygous locus) or mistaken and false alleles (FA; misprinting; Pompanon et al. 2005). Several authors have developed methods to reduce the risk of errors. These methods can be grouped into 3 strategies: 1) minimizing genotyping errors, 2) detecting and quantifying errors, and 3) integrating errors into statistical analysis (see the Third Step Section).

Minimizing genotyping errors.—One possible method to reduce genotyping errors is to improve the amplification by using engineered polymerases (e.g., high-fidelity and hot-start technique; Beja-Pereira et al. 2009) or choosing the right markers, as AD rates are greater with increasing allele size and in tetranucleotides (Buchan et al. 2005, Broquet et al. 2007). Another possibility is to produce a consensus genotype of each sample through several PCR replications per sample and locus. The number of replications required depends on the assumptions made. Taberlet et al. (1996) devised the multiple-tubes approach (MTA), which includes at least 7 allele sightings for a homozygous locus before a genotype is accepted and at least 2 allele sightings for a heterozygous locus. The MTA became a standard method to deal with genotyping errors, but many authors considered this technique as too conservative (Bayes et al. 2000, Ernest et al. 2000, Banks et al. 2002, Miller et al. 2002, Palomares et al. 2002, Frantz et al. 2003). Frantz et al. (2003) presented a modification of the MTA, the comparative multiple-tubes approach (cMTA), which has less replication steps per locus: 2 allele sightings for heterozygous, 3 for homozygous. This practice was applied in many studies (e.g., Bayes et al. 2000, Bhagavatula and Singh 2006, Piggott et al. 2006, Arrendal et al. 2007), and also variations of it with only 2 (Palomares et al. 2002, Adams and Waits 2007) or 4 allele sightings for a homozygous locus (Bellemain et al. 2005, Ferrando et al. 2008). Although amplifying samples several times is expensive and may be problematic with little amount of extract (Paetkau 2003, Schwartz et al. 2006), we could demonstrate that samples with high error rates generate severely flawed genotypes without repetitions (see the Case Study Section).

A possibility to save extract and to reduce the number of repetitions, while generating a consensus genotype, is multiplex pre-amplification (Bellemain and Taberlet 2004, Piggott et al. 2004). Here, an initial multiplex PCR run with all employed primers is performed. The extract of the first run is used in a second PCR with each primer individually (Piggott et al. 2004) or as a multiplex with 2 or 3 markers (Bellemain and Taberlet 2004). This procedure decreased genotyping errors and increased the amplification success greatly for samples with low quality and quantity of template DNA (Bellemain and Taberlet 2004, Piggott et al. 2004, Lampa et al. 2008, Arandjelovic et al. 2009). Despite the usefulness of the multiplex pre-amplification (e.g., for difficulties at certain markers), it should not be applied without a pilot study.

Nonetheless, low-quality samples with little prospect to generate a reliable genotype will still be present. To minimize costs and efforts, one could discard these samples, for instance by using quantitative PCR (Morin et al. 2001) or mitochondrial DNA analysis (Kohn et al. 1999) prior to microsatellite genotyping. But, this requires additional efforts to the microsatellite analysis and quantitative PCR, although becoming increasingly common, is still expensive (Schwartz et al. 2006). If species identification can be done with the employed marker system, a more straightforward way is to remove low-quality samples during the amplification process based on low amplification success. Paetkau (2003) suggested removing samples that did not produce a genotype and repeating only those samples that are likely to contain errors. He analyzed each sample once at all markers, excluded samples with less than 3 high-confident genotypes, and re-amplified missing data, low confident results, or similar genotypes that mismatched at 1 (1 MM) or 2 alleles (2 MM). After 2 amplifications per locus, Paetkau (2003) discarded the samples with less than 4 reliable genotypes. So far, this quality control approach was only used in hair-based studies (Table 1) for which it might be appropriate, but it seems inappropriate when PCR failing and the genotyping error rate is rather high (see the Case Study Section). Cullingham et al. (2010) selected a small and a large polymorphic marker and interpreted a failed amplification in the larger locus as a sign of degradation and removed these samples.

The question arises where to set a reasonable threshold before excluding samples from the dataset. We advise that researchers should conduct a pilot study. With high genotyping error rates, low amplification success, or low number of samples, the threshold must be more relaxed; otherwise, too many samples that would have generated a genotype are discarded. By combining the advantages of the cMTA (Frantz et al. 2003)—the low number of required repetitions—and the idea to screen the dataset for low-quality samples (Paetkau 2003), we developed a screening approach for samples with very high genotyping error rates and low genotyping success rates. This screening approach consists of 5 amplification steps; afterward, low-quality samples are removed according to certain thresholds (outlined in Fig. 2). Step 1 consists of 3 PCRs for all samples using 3 multiplexed markers to screen the dataset for non-target species (if possible) and for samples with little prospect to produce a complete multilocus genotype. As genotyping error rates will be usually unknown in that early stage, the thresholds to drop low-quality samples should be rather

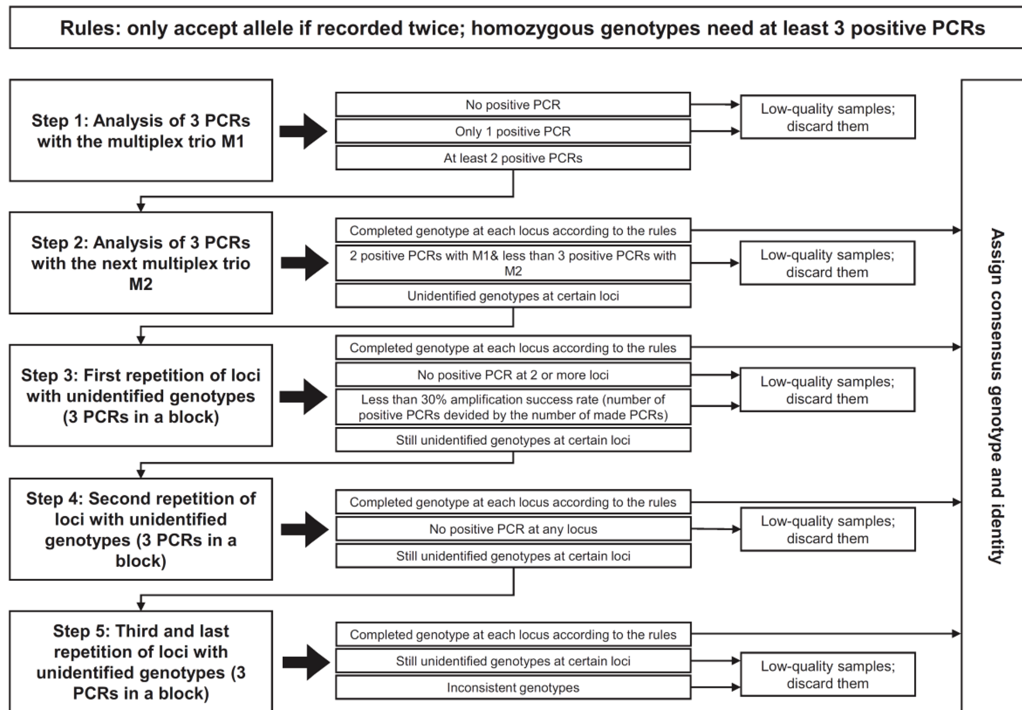

Figure 2. Flow chart of the screening approach we used following the assumptions of Frantz et al. (2003). We scored amplifications as positive if a polymerase chain reaction (PCR) product of the expected size was present, even if the genotype may not have been correct. We calculated the success rate for a sample by dividing the number of positive PCRs at each locus by the total numbers of PCRs made.

relaxed (Fig. 2) In the second step, remaining samples are amplified 3 times with 3 additional multiplexed markers, dropping again poor quality samples during that process. The rules when to accept a genotype highly depends on the genotyping error rate, but in general, heterozygous alleles should be accepted when recorded at least twice, whereas homozygous genotypes could be accepted after 2–4 sightings. All probes with still missing genotypes are re-amplified at the specific loci or removed according to the set thresholds (Fig. 2: step 3–step 5). Compared to the MTA (Taberlet et al. 1996), cMTA (Frantz et al. 2003), and the quality control of Paetkau (2003), we required less repetitions but gained an increased genotyping success rate and more reliable genotypes (see the Case Study Section). Though removing low-quality samples may lead to heterogeneity in catchability if individuals differ in the probability of producing low-quality samples (Creel et al. 2003, Lukacs and Burnham 2005*a*), these samples will very likely not produce any genotype or only a severely flawed one (see the Case Study Section).

The latest developments in high-throughput next-generation sequencing technologies open up new possibilities (Beja-Pereira et al. 2009, Li et al. 2010, Guichoux et al. 2011). This technology allows a considerably faster, cheaper, and stringent selection of the best microsatellite markers with lower error susceptibility or will even allow genotyping hundreds of samples at thousands of loci via sequencing (Guichoux et al. 2011). Furthermore, the potential to sequence a single molecule without preceding amplification might reduce the problems of low amplification success and high genotyping error rates.

Detection and quantification of genotyping errors.—Although approaches to minimize genotyping errors are important, they do not demonstrate that the dataset is error free (Schwartz et al. 2006). Therefore, researchers should screen the dataset for possible errors, quantify them, and direct repetitions only to error-prone samples or loci to further eliminate genotyping errors. Several methods have been developed for this purpose.

If consensus genotypes are derived by multiple repetitions, a genotyping error rate can be calculated by comparing the scored genotypes of each PCR with the consensus genotype. Samples or loci with high error rates can be repeated, ensuring the genotypes, or rejected. Broquet and Petit (2004) criticized that many different calculation methods are employed in the literature for the assessment of AD or FA, leading sometimes to substantial underestimation and often to incomparability between studies (Table 1). They proposed to consider all positive PCRs for the calculation of FA, but only heterozygous genotypes for AD, as these are the genotypes where the respective error can occur. To automate this calculation, Valière (2002) designed the software package GIMLET. Disadvantages are that the program cannot handle varying numbers of repetitions per loci within a sample.

Instead of calculating a genotyping error rate, Miquel et al. (2006) suggested to compute a standardized quality index by assigning 0 to falsely scored genotypes, compared to the consensus genotype, and 1 to correct genotypes. The sum of the scores divided by the number of repeats amounts to the quality index. This, however, drops important information about the type and number of errors within a sample or locus.

Thus far, the majority of studies used genotyping error rates to assess the quality of their analysis.

Johnson and Haydon (2007) devised a maximum-likelihood estimation (MLE) of AD and FA, implemented in Program PEDANT (http://www.stats.gla.ac.uk/~paulj/pedant.html), where only 5–10% of the samples require duplicated genotypes. As only a part of the dataset is included in the analysis, the calculation is completed per locus but not per sample.

Even though genotyping error rates indicate problematic samples or loci, they do not determine if the dataset is error free. Various methods have been developed that can be differentiated into those that attempt to quantify the genotype reliability and those methods that test whether errors are still contained in the sample. The former method is applied in a maximum-likelihood estimator (MLE) and the corresponding program RELIOTYPE (Miller et al. 2002), which estimates the average number of reactions required per sample and marker to identify a correct genotype based on the information of at least 2 positive PCRs per locus. The disadvantage is that the method assumes the dataset to be free of FA or contaminant alleles and accounts only for evenly distributed AD across loci and among alleles. Both assumptions were disproved in many studies (Bjorklund 2005, Buchan et al. 2005, Pompanon et al. 2005, DeWoody et al. 2006). The MLE (Miller et al. 2002) is performed during or after genotyping; however, Valière et al. (2002) devised a simulation tool called GEMINI that determines the number of required repetitions for a correct genotype in advance based on known error rates and heterozygosity. This may not be possible in many studies where such information is not available prior to the analysis. Furthermore, resulting recommendations of required repetitions are not specified for each locus or sample.

Van Oosterhout et al. (2004) created a software tool, MICRO-CHECKER, which tests for null alleles, short allele dominance, or stutter peaks. The method can be applied to genotypes of the first positive PCR. However, the software tool can be cumbersome for datasets containing many loci or populations (DeWoody et al. 2006) and as analyses are directed to loci, it does not indicate which samples might harbor errors. Therefore, almost each sample must be repeated for the loci in question, if the loci are not excluded, and many correct genotypes will be unnecessarily repeated or rejected. Another drawback is that the method is only sensitive to systematic errors; random errors will not be detected (see the Case Study Section).

McKelvey and Schwartz (2004, 2005) developed 2 tests implemented in the software DROPOUT to detect extant errors in the dataset. The bimodal test (EB) indicates which samples differ in 1–3 loci (1 MM–3 MM) and are hence likely to contain errors. The difference in capture history test (DCH) determines the loci producing the most errors. The tests can be used at each state of the analysis to minimize repetitions (Schwartz et al. 2006) and to check whether the dataset is error free. The 2 tests assume equal capture probability among individuals (Lukacs and Burnham 2005*a*), which is often difficult to meet in the field. If the tests are used for genotypes of the first PCR, a marker system of at least 8 loci is recommended and the per-locus error rate should not exceed 0.25 to not bias the DCH test. If employing 6–7 markers, they recommended forming a consensus genotype first and to carefully interpret the test results as some individuals may differ in only 1 or 2 loci. Still, an error-free and sufficiently tagged dataset should have rather low numbers of 1 MM or 2 MM (following Waits and Paetkau [2005] no more than 2 1 MM pairs and 9 2 MM pairs) and the probability of creating new individuals by expanding or changing the tag composition should become infinitesimal (Mills et al. 2000, McKelvey and Schwartz 2005).

Creel et al. (2003) stated that no method might completely eliminate genotyping errors. They proposed producing consensus genotypes and analyzing them with a matching approach to reveal if 2 mismatched samples still belong to the same individual and should be treated as recaptures in a CMR estimate. They calibrated the required threshold with a dataset of known individuals, but without this information, implementing this approach is difficult. A similar possibility is to consider samples having 1 MM or 2 MM to derive from the same individual or to remove those samples (Bellemain et al. 2005). Although these corrections can serve as conservative lower bound, they should only be used with caution, as it may result in underestimated population sizes (see the Case Study Section).

In summary, with our example dataset we could demonstrate that high genotyping error rates lead to untrustworthy genotypes after the first positive PCR and minimizing errors via repeated amplifications to generate a consensus genotype is crucial. This should be conducted following a screening approach (Fig. 2; see the Minimizing Genotyping Errors Section under the Second Step Section) to remove non-target and low-quality samples. The derived genotypes of the first 2 screening steps can be used to estimate a preliminary AD and FA rate using PEDANT (Johnson and Haydon 2007) and the number of required repetitions by employing the MLE or equation 6 of Miller et al. (2002) with locus-specific parameters (see the Case Study Section). According to these results the assumptions for allele acceptance (e.g., Frantz et al. 2003) as well as the number of required positive PCRs can be determined and the thresholds for dropping poor samples can be refined. After generating the final consensus genotype through the required repeats, researchers should check whether all errors are removed from the dataset by using tests like those of McKelvey and Schwartz (2004) and implementing a final 1 MM and 2 MM pair check. Additionally, the program MICRO-CHECKER (Van Oosterhout et al. 2004) could help to identify systematic errors.

THIRD STEP: ESTIMATING POPULATION SIZE BASED ON GENOTYPE DATA

Although population size can be estimated with models for open populations, the estimation is more precise and convergence of models is more likely when the population

is closed (i.e., no births, deaths, or migrations) between capture occasions, as this requires fewer parameters to be estimated, hence less data, and allows relaxing the equal catchability assumption in various ways (Otis et al. 1978, Pollock et al. 1990). Certain tests can be used to check whether the assumption of closure is met (CLOSETEST [Stanley and Richard 2005], Pradel's [1996] recruitment model [Boulanger et al. 2002]), but these should be used with caution, as variations in capture probability can cause an incorrect rejection of closure (Otis et al. 1978), just as the presence of ghost individuals creates artificial transients in the dataset. Therefore, Otis et al. (1978) suggested to assess closure based on biological information and to make sure that the study design meets the closure assumption (e.g., by avoiding seasons of high migration, high mortality). If the number of samples or individuals is very small, if samples are too dispersed across a large study area, or if no definite sampling occasions can be determined, then closed population models will not produce meaningful results. In those cases, some authors used accumulation curves to estimate population sizes (Table 1). These methods assume randomly distributed individuals in space, equal catchability, and are only reliable if practically all individuals have been sampled (Eggert et al. 2003, Bellemain et al. 2005). Therefore, Lukacs and Burnham (2005*a*) strongly advised against their use.

Most closed CMR models use multiple session sampling, where only 1 detection per animal per sampling occasion is considered, although some estimators can handle multiple detections of an individual per sampling occasion (Miller et al. 2005, Yoshizaki et al. 2011) or even single-session sampling (Miller et al. 2005, Wright et al. 2009). The latter may reduce the risk of violating the closure assumption (Petit and Valière 2006) and may be beneficial in terms of time and costs (Luikart et al. 2010). But it allows less flexibility in modeling capture processes and is difficult to apply to non-invasive sampling, as spatial autocorrelation of sequential samples will likely be occur (Luikart et al. 2010).

Because of differences in sex, age, size, or reproductive status, capture probabilities in wild populations often vary temporally, among individuals, as a response to previous capture experience, or combinations thereof (Otis et al. 1978, Pollock et al. 1990). Hence, the challenge is to find the appropriate model for the data. One can test for variations in capture probability (e.g., CAPTURE [White et al. 1982]; CAPWIRE [Miller et al. 2005]; MARK [White and Burnham 1999]; Puechmaille and Petit [2007]), but they can be confounded with other factors. For example, ghost individuals created by genotyping errors can induce or increase individual heterogeneity, which cannot be differentiated from true individual heterogeneity. The shadow effect also creates heterogeneity in apparent capture probabilities (Mills et al. 2000). Thus, for the final model selection one should consider biological information of the target species, the employed sampling design, and which management decisions will be based on the model (Pollock et al. 1990, Manly et al. 2005). For example, when using bait stations for hair sampling, a model accounting for behavioral response could be appropriate (Dreher et al. 2007), as could a model accounting for individual heterogeneity because the attraction may vary with sex or age (Mowat and Strobeck 2000). Individual heterogeneity may also be present using passively collected samples if the study area is comparably small and females have smaller home ranges than males and are hence more often collected (Ruell et al. 2009). Also, when low-quality fecal samples are removed, individual heterogeneity may be induced if the probability of producing those samples is not equally distributed among individuals (Lukacs and Burnham 2005*a*).

Conventional Estimation Models

Conventional estimation models do not integrate genotyping errors. They have been used in a number of DNA-based studies (Table 1). The most commonly used approaches are the estimators in Program CAPWIRE (Miller et al. 2005), CAPTURE (White et al. 1982), and MARK (White and Burnham 1999). Program CAPWIRE (Miller et al. 2005), recently also available as an R-package (Pennell et al. 2013), was designed for small populations and explicitly for NGS allowing researchers to incorporate multiple captures of an individual within a sampling occasion. CAPWIRE includes 1 model for even capture probability (ECM) and 1 for individual capture heterogeneity (TIRM). A likelihood-ratio test (LRT) is proposed to select between the 2 models, but Puechmaille and Petit (2007) argued that it is likely to miss some types of capture heterogeneity. For overdispersed data containing individuals that are captured far too frequently than CAPWIRE predicts, the R-package offers the partitioning method PART for model TIRM (Pennell et al. 2013).

Program CAPTURE, integrated in software MARK (White and Burnham 1999), contains models to estimate population size using capture probabilities that are constant (M_0) or vary individually (M_h), temporally (M_t), or as behavioral responses to previous captures (M_b), as well as combinations thereof except M_{tbh} (Otis et al. 1978). A criterion is offered for selecting the most appropriate model, but it should be used with caution, as it has low power, especially for small populations, for data with low capture probabilities (Pollock et al. 1990) or with genotyping errors (Roon et al. 2005). Via simulation studies, Roon et al. (2005) showed that the model selection most frequently chose M_0 for error-free data, but M_h or other alternatives for data with errors. However, M_h heavily overestimates population sizes if the per-locus error rate is above 0.01 (Roon et al. 2005); we also showed this in our case study.

Program MARK (White and Burnham 1999) offers further models that account for variations in capture probability: the finite mixture model of Pledger (2000) and the conditional likelihood models of Huggins (1989). MARK uses corrected Akaike's Information Criterion (AIC_c) to select among alternative models within each of these model systems. However, it may fail if the assessed models suffer structural deficits (Burnham and Anderson 1998).

Numerous alternatives to the above-mentioned approaches exist; Gazey and Staley (1986) presented a Bayesian

estimator for single-session sampling that can deal with small sample sizes and are thus potentially useful for rare species. However, the model assumes equal catchability (Petit and Valière 2006) and estimates may strongly depend on the selected priors, especially with low capture probabilities (Smith 1991). More flexible conventional model systems that have, to our knowledge, not been used yet for non-invasive genetic data are the sample coverage estimators (Lee and Chao 1994) and the estimating function (Chao et al. 2001) for which free software exists (Chao and Yang 2003). Both approaches allow varying capture probabilities and provide a measure to quantify the degree of heterogeneity but require large datasets. The sample coverage estimator tends to the true population size if capture probabilities follow a gamma distribution. The estimating function offers an estimator for the most general model M_{tbh} (Chao and Huggins 2005).

Several authors using genetic data and conventional estimators found greater population sizes than counted animals (Harris et al. 2010), biologically plausible estimates (see the Case Study Section), or estimates suggested by other non-genetic methods (Zhan et al. 2006, Hajkova et al. 2009, Uno et al. 2012). This could be caused, among other reasons, by undetected ghost individuals. If non-invasive samples with high genotyping error and PCR failing rates are used for genetic CMR, no lab protocol or subsequent error check may completely eliminate genotyping errors with certainty (Creel et al. 2003, Marucco et al. 2011, see the Case Study Section). Since even small amount of undetected errors result in significant overestimation of population size using conventional estimators (Waits and Leberg 2000, Creel et al. 2003, Roon et al. 2005, Lukacs and Burnham 2005*b*, Link et al. 2010, Yoshizaki et al. 2011), models that account for genotyping errors may be preferable.

Models Accounting for Genotyping Errors

Although several approaches are available that account for genotyping errors, only a minor number of studies applied them so far (Table 1). Two approaches, Stevick et al. (2001) and Knapp et al. (2009; GUAVA), assume equal capture probabilities among individuals and are based on 2-sampling estimators, thus they can only be used under restrictive conditions. Furthermore, the method of Stevick et al. (2001) requires a known false negative rate, whereas GUAVA needs information about genotyping error rate, allele frequencies, and Hardy–Weinberg equilibrium. Therefore, Knapp et al. (2009) advised to perform the first population size estimate with established approaches and to apply GUAVA only for subsequent estimations. If capture probabilities differ among individuals—a common situation in natural populations—the employed estimators tend to underestimate population sizes (Pollock et al. 1990), as we also demonstrated in our case study. Thus far, an application of the GUAVA method could not be found in the literature and it should be applied with caution unless underestimation does not result in increased management risks.

A more general approach was developed by Lukacs and Burnham (2005*b*; L&B estimator). The L&B estimator, implemented in Program MARK (White and Burnham 1999), adds to each available model in MARK a parameter α, the probability of a correct classification. A high value of alpha indicates a low probability of still extant genotyping errors. The parameter estimates perform well when α approaches 1, except for very low capture probabilities ($P \leq 0.1$). But, when error rates exceed 5%, the variance may become large, which makes the population size estimate uninformative. The model quality indicator AIC_c can help to determine whether an L&B or a conventional model is better supported by the data. A potential problem of the L&B estimator is that misidentification is difficult to separate from heterogeneities in capture probabilities as both processes can lead to similar capture histories. Several authors criticized the approach for relying on questionable assumptions (Wright et al. 2009, Link et al. 2010, Yoshizaki et al. 2011; Table S1, available online at www.onlinelibrary.wiley.com). All 3 critics designed new estimators that incorporate genotyping uncertainty into mark-recapture models.

Yoshizaki et al. (2011) suggested several approaches that are extensions of the CAPTURE model M_t, including the parameter alpha (the probability of a correct classification). Hence, none of the developed estimators account for individual heterogeneity or behavioral response. Although 3 estimators performed similarly well for high alpha values (0.9–1) and moderately high capture probabilities, they do not recommend employing any misidentification model when capture probabilities is rather low ($p = 0.1$).

Link et al. (2010) used a Bayesian method that assumes, like the L&B estimator, that the same misidentification cannot occur twice (i.e., that ghost individuals are not resighted) though they admitted that this assumption may not be reasonable for many datasets. The estimator is again an extension of the CAPTURE model M_t including the parameter alpha. So far, the estimator does not include other variations in catchability.

Wright et al. (2009) also designed a Bayesian approach that does not require that genotyping errors result in a new genotype. It accounts for an equal AD rate among individuals and alleles but omits other genotyping errors. The model requires at least 2 positive PCRs per locus, a single-sampling session, and an equal capture probability. However, only recently Wright et al. (2012) presented an extension of the model allowing for multiple sampling sessions, other types of errors, and for heterogeneity in capture probability.

To our knowledge, none of these new models have been applied to natural populations, except for the authors' examples, or were evaluated by others for their usefulness, validity, and accurateness. So far, no software is available to implement these methods, although J. Wright offers R scripts upon request (Department of Mathematics and Statistics, University of Otago, New Zealand).

With respect to other studies conducting non-invasive genetic CMR analyses, thus far only a minor number apply a misidentification model (Table 1). Either they assume their error rate to be negligible (Robinson et al. 2009) or they consider it as inappropriate because estimations are

considerably less than the ones calculated with conventional estimators (Williams et al. 2009). Even so, when genotyping errors are present, the conventional estimators should be used with caution. Since most error-incorporating models result in reliable estimations when genotyping error rate is close to 0, we recommend preferring one of these models. If variations in capture probability are high, models that do not account for it have substantial bias (Otis et al. 1978, Pollock et al. 1990). Thus, if biology or statistical tests indicate such variations, currently only the L&B models and the extensions of Wright et al. (2012) allow for all (L&B) or most (Wright) variations in capture probability, although one has to be aware of their limitations. If only time variability exists, then the estimators of Yoshizaki et al. (2011) and Link et al. (2010) may be suitable alternatives by the time a script or software is available. In any case, population size estimates based on non-invasive genetic CMR should never be accepted uncritically.

CASE STUDY

Study Area

We carried out our study in the Upper Lusatian heath and pond landscape in the eastern part of Saxony, Germany. The Upper Lusatia is characterized by about 1,000 ponds, mostly used for fish farming and hosts one of the biggest and most viable Eurasian otter (*Lutra lutra*) populations in Central Europe (Hauer et al. 2002*a*). The area under investigation is located between Königswartha and Groß Särchen (51°20′N, 14°18′E) and includes 6 pond areas plus 1 single pond with a total water surface of 486 ha. During the sample collection in March 2006, most larger summer ponds were dried out and mainly the deeper winter ponds contained fishes resulting in a reduced water surface area (260.53 ha).

Sampling of DNA

Although otters reproduce throughout the entire year, we conducted the sampling in an off-peak season (March; Hauer et al. 2002*b*) and during a period when otters converge on few stocked winter ponds, which may increase the capture and recapture rates. To determine how many sampling occasions were required, we performed a simulation of capture histories in advance employing Program MARK (White and Burnham 1999). Using worst-case assumptions (capture rate = 0.25; genotyping success rate = 0.1; population size = 40), the simulation suggested at least 5 sampling occasions for reliable estimates. Since otters defecate easily 30 scats per day (Kruuk 2006), we treated each of 5 consecutive days as a sampling occasion. Such a short sampling period ensures the comparability of environmental conditions (i.e., weather, habitat structures, food availability) between the sampling occasions and that relevant recruitment or mortality will likely not occur (Kruuk 2006).

Because both sexes of otters, regardless of the reproductive status, and all ages use marking sites for their feces and defecate in similar rates (Kruuk 2006), all individuals will likely have the same chance of being captured if those marking sites are used as sampling points. Along all pond banks, we mapped each active otter marking site and tagged each detected scat with material of the surrounding area (sticks, reed, stones, etc.) to facilitate recognition of fresh scats the next day. On the following 5 consecutive days, we collected all detected fresh scats and anal jelly samples from the tagged marking sites and from sites not previously detected. We collected mainly in the morning on days without rain or frost with 2 collectors for each pond area. We wiped off the external layer of the scat with a cotton swab and stored it in a separate sterile 10-ml cryovial. A pilot study demonstrated that the cotton swab technique decreases the amount of prey hard parts but increases the proportion of sloughed gut cells. Furthermore, this technique decreases the risk of cross-contamination and alteration of marking behavior; otter scats are used for intraspecific communication (Kruuk 2006). Although we extracted the samples on the day of collection, this was logistically difficult. As other preservation methods failed (ethanol, −20° C, silica, 2-step method [Nsubuga et al. 2004]) for our samples (Lampa et al. 2008), we conducted an experiment to test the preservation in the first buffer of the Qiagen extraction kit (lyse buffer ASL) at −80° C. We compared the amplification success rate (SR) between samples collected without a reagent and extracted immediately after collection ($n = 36$; SR = 40.2%), samples collected in ASL-buffer and extracted at the day of collection ($n = 20$; SR = 41.9%), and samples collected in ASL buffer and stored at −80° C for a year ($n = 34$; SR = 53.8%). Although we found no significant difference between the 3 methods (analysis of deviance: $P = 0.428$, $\chi^2_2 = 1.696$), it was slightly more successful to store the samples in ASL buffer and −80° C, which we recommend as it is highly practicable. We extracted all 250 samples using the QIAamp® DNA Stool Mini Kit (Qiagen), which revealed a 69% higher amplification success rate than the resin-based Chelex® extraction method (Lampa et al. 2008).

Microsatellite Genotyping

Minimizing genotyping errors.—From a pilot study, we knew that otter fecal samples from our study area have fairly high genotyping error rates and low genotyping success rates (Lampa et al. 2008). Hence, generating consensus genotypes seemed to be required. Here, we illustrate our screening approach and compare the results with those obtained by 3 widely used approaches—the MTA (Taberlet et al. 1996), cMTA (Frantz et al. 2003), and the quality control approach of Paetkau (2003; Table 2).

To increase genotyping success and to decrease genotyping errors, we applied a variation of the multiplex pre-amplification (Lampa et al. 2008). We performed amplifications following several contamination preventions (see Fig. 1). We separated and visualized the PCR products of the 6 employed microsatellites (multiplex combinations: M1: Lut 457, 615, 733; M2: Lut 435, 604, 701; Dallas and Piertney 1998, Dallas et al. 2002) in an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany) and analyzed the products using ABI PRISM® GeneMapper™ Software V.3.7 (Applied Biosystems). We used M1 as the first multiplex trio in step 1 with $n = 250$

Table 2. Comparison of the number of required positive polymerase chain reactions (PCRs) per locus, the number of created erroneous genotypes compared to the consensus genotype, and genotyping success rate using our screening approach, the multiple-tubes approach (MTA; Taberlet et al. 1996), the comparative multiple-tubes approach (cMTA; Frantz et al. 2003), and the quality control approach of Paetkau (2003). Data are from 115 genotyped Eurasian otter fecal samples collected in Saxony, Germany in March 2006.

Approach	Positive PCRs per locus	Created erroneous genotypes	Genotyping success rate (%)
Screening approach	3.8	0	46
MTA	5.4	0	46
cMTA	4.2	0	42.4
Quality control	1.8	31	≤30

samples and M2 as the second multiplex trio in step 2 (Fig. 2).

We performed all PCRs in a block of 3 amplifications because with high error and PCR failing rates, a blockwise amplification is far less time consuming than a stepwise repetition. For the amplifications of steps 3–5 (Fig. 2), we multiplexed loci if possible and repeated amplifications up to a maximum of 12 positive PCRs. We regarded amplifications as positive if a PCR product of the expected size was present, even if the genotype may not have been correct. To test whether the discarded samples were low-quality samples and, hence, to test whether we chose appropriate thresholds, we processed 33% of rejected samples through the subsequent steps. Out of 135 rejected samples, 6 (4.4%) could have been successfully genotyped. Our procedure thus reduced genotyping success rate by 2.4%. However, these 6 samples had a failing rate of 63% and hence required on average 13 repetitions per locus to obtain a reliable genotype. Moreover, they showed a genotyping error rate of 80.2%, but contained no further information for the dataset because we had found other genotyped samples of the same individual at the same sampling occasion. Thus, we regard our screening approach as appropriate. Only samples with unique genotypes showed mismatches at 1 (1 MM) or 2 alleles (2 MM) relative to multiply found genotypes. We re-amplified the mismatched loci 3 times to ensure that this was not caused by genotyping errors.

Out of 250 samples, we discarded 135 samples (54%) and 115 samples (46%) showed a reliable genotype, referred to as genotyping success rate. These samples revealed 22 different genotypes. The mean amplification success rate was 75.6% regarding 115 genotyped samples or 41.1% regarding all 250 samples. Since the mean failed amplification rate (no DNA signal) was 33.4%, we required on average 6.4 PCRs (range 3–24) per sample and locus to determine a reliable genotype or 3.8 PCRs if only positive PCRs are counted.

The mean expected (He) and mean observed heterozygosity (Ho) reached values of He = 0.54 and Ho = 0.70. The unbiased probability of identity (PI) over all loci was 1.36×10^{-4} and the PI_{sib} was 2.45×10^{-2}. The actual observed PI was 1.3×10^{-3}. Since the studied population was fairly small (<50) and the genotyping error rate high, we were more concerned about producing ghost individuals than failing to discriminate between individuals. Therefore, we accepted comparably high PIs that were at the lower boundary of literature recommendations.

Compared to our screening approach, the MTA would have required 40.5% more positive PCRs but would have led to the same results (Table 2). The cMTA would have resulted in slightly more repetitions and a slightly diminished genotyping success rate. The main difference between the screening approach and the cMTA is that no samples are eliminated in the latter. Although this may work well if success rate is high, low-quality samples such as otter scats should be screened for error-prone samples with little prospect to produce a complete multilocus genotype to save money and time. Compared to the MTA and cMTA, the screening approach seemed to minimize the number of required iterations and maximize the genotyping success rate. The quality protocol of Paetkau (2003) would have resulted in a severe incorrect dataset with many ghost individuals and a substantial decrease in genotyping success rate as samples would be discarded although they could reveal a reliable genotype if repeated more often. With such a low genotyping success rate, estimation of population size is less accurate. However, Paetkau (2003) optimized his protocol for hairs as an origin for DNA with a high success rate (83%). He only had to reject few samples (13%) because of low quality.

Detection and quantification of genotyping errors.—Methods to detect and quantify genotyping errors can be used after the first positive PCR to decrease the number of PCR repetitions but also to test the reliability of consensus genotypes. To test the suitability and efficiency of several methods, we first employed the MLEs of Johnson and Haydon (2007) and Miller et al. (2002), as well as the programs MICRO-CHECKER (Van Oosterhout et al. 2004) and DROPOUT (McKelvey and Schwartz 2005) using the genotypes after the first or second positive PCR. We compared these results with the results of the method of Broquet and Petit (2004) and the programs MICRO-CHECKER and DROPOUT using consensus genotypes.

The MLE for AD and FA of Johnson and Haydon (2007) using all samples with 2 positive PCR products per locus ($n = 130$) revealed a mean AD of 41.5% and a mean FA of 3.3%. When averaging 10 runs of 10% of the samples, we obtained a mean AD of 43.3% (range 35–59%) and a mean FA of 3.4% (range 0–8%). For a genotype reliability of at least 98%, the MLE of Miller et al. (2002) estimated on average 3.9 required replications per sample and locus. Since the assumptions of even dropout rates and heterozygosities across loci were violated for our dataset, we re-analyzed the data with locus-specific heterozygosities and dropout rates (equation 6 of Miller et al. 2002), which resulted in the same estimation.

For the genotypes after the first positive PCR, Program MICRO-CHECKER (Van Oosterhout et al. 2004) did not find any markers with short allele dominance (large AD), but found 3 with scoring errors due to stutter peaks and possibly 5 with null alleles. Unfortunately, the program does not

determine which samples contain errors. Hence, each sample has to be repeated with the marker in question, which applied to all but 1, but in this marker, we found 27 wrong genotypes when compared with the consensus genotype. The EB test in DROPOUT (McKelvey and Schwartz 2005) revealed that all samples except 3 would need a repetition. However, the 3 seemingly correctly genotyped samples contained 2–5 genotyping errors and hence would have been incorrectly accepted. The DCH test in DROPOUT showed that locus Lut733 added significantly more errors to the dataset than the remaining loci and should therefore be excluded. Running all loci through the L_{base+1} position produced 59 new individuals. Thus, all these applied methods showed that the genotypes obtained after the first or second positive PCR still contained many genotyping errors and were not trustworthy; we still needed to form a consensus genotype. After the first positive PCR, 70% of the samples showed wrong genotypes at 1–5 microsatellites compared to the consensus genotype and 53% were wrong at 1–4 markers after 2 positive PCRs.

When all scored genotypes obtained during the screening approach were compared with the consensus genotype ($n = 115$), we obtained an overall genotyping error rate of 44.9% with a mean FA of 1.8% (range: 0.3–4.9%) and a mean AD of 43.1% (range: 15.9–54.5%), following the calculation of Broquet and Petit (2004). These rates correspond to the estimates of the MLE of Johnson and Haydon (2007) and are rather high compared to other genetic CMR studies (Table 1). This might be explained by the typically low quality of otter samples (Hansen et al. 2008) but also by the conservative way of computing that is not representative for all genotypes (Broquet and Petit 2004). However, with such a high error rate, repeating amplifications is necessary. The screening approach required on average 3.8 positive replicates per sample and locus, as was predicted by the MLE of Miller et al. (2002).

For the consensus genotypes, Program MICRO-CHECKER revealed no errors, whereas the 2 DROPOUT tests indicated errors were still present. We presumed that an error-free dataset would produce no new false individuals in the DCH test and at best no 1 MM pairs or at least substantially less than 2 MM pairs. The EB test resulted in 11 1 MM, 6 2 MM, and 5 3 MM pairs. The DCH test revealed that no locus had to be dropped because of significantly higher error rates, but that 3 markers are error prone (Lut435, Lut601, Lut733). The 1 MM and 2 MM pairs emerged only in these 3 error-prone markers and in the 6 unique genotypes. We therefore re-ran the tests with an altered dataset where these unique genotypes were 1) removed from the dataset (dataset a), 2) grouped with a similar group adopting their genotype (meaning a heterozygous genotype could become homozygous; dataset b), or 3) grouped with a similar group assuming that heterozygous genotypes are more reliable than homozygous (because AD is more likely than FA; dataset c).

The reasons for this approach were 1) genotyping errors lead most likely to 1 MM and 2 MM; 2) only unique genotypes and their similar groups showed 1 MM and 2 MM; 3) similar groups consisted mostly of 3–4 samples, which may indicate that these individuals were difficult to amplify and tended to genotyping errors; 4) groups of identical genotypes not showing 1 MM or 2 MM are more likely to be correct or at least reproducible (Paetkau 2003, Lukacs and Burnham 2005*b*); 5) all identical genotypes were sampled at least 3 times and it is unlikely that so many individuals were captured only once but none twice; and 6) if the unique genotypes would be true individuals, at least some should deviate by more than 1 or 2 loci.

One unique genotype showed 5–6 mismatches to spatially proximate groups, but 2 mismatches to a group spatially far away. Such a long distance movement could not be observed for any other individual; thus, an allocation would be ambiguous. Therefore, we decided to keep this sample as a single sample without assigning it to a similar group.

Datasets a and b revealed 4 1 MM and 7 2 MM and 2 error-prone loci (Lut601, Lut733), though not significant. Only dataset c produced no new false individuals and no samples showed 1 MM, thus indicating an error-free dataset.

Estimating Population Size Based on Genotype Data

Using the unmodified and the 3 modified datasets, we compared the 2 most employed conventional estimation models, CAPWIRE and CAPTURE, with 3 error-incorporating methods, GUAVA, L&B estimator, and the Wright model (Wright et al. 2009; Table 3). We treated each of the 5 consecutive sampling days as a capture session. Based on the biology of otters (Kruuk 2006), relevant recruitment or mortality unlikely occurred during such a short period, hence the assumption of closure is reasonable. Pradel's (1996) model employed following Lukacs (2010) also supported this assumption for all datasets (data not shown). Regarding variations in capture probability, the respective test statistics in the employed programs supported M_0, M_b, and/or M_h (Table 3). Individual heterogeneity (M_h) could have been induced by removing low-quality samples and behavioral response (M_b) could have been caused by sampled individuals who changed their marking intensity because of the frequent disturbance. Combined models (M_{bh}, M_{tb}, M_{th}) and model M_t were statistically not supported and would have required more data. The L&B model M_h generated inconclusive results for the unmodified dataset as misidentification and heterogeneity were confounded, a common problem when using limited numbers of sampling occasions and when many individuals were found only once (e.g., because of ghost individuals).

The results for the unmodified dataset differed widely among all models (Table 3). The estimated population size was less than the number of different genotypes in the dataset for all error-incorporating estimators, but equal or greater for the conventional estimators. Using CAPTURE, the $M_{h\text{-jackknife}}$ model was selected (CAPTURE offers 2 M_h estimators—$M_{h\text{-jackknife}}$ and $M_{h\text{-Chao}}$) and estimated 1 otter per 8.5 ha water, which is incredibly high (Kruuk 2006) and unlikely to be true. $M_{h\text{-jackknife}}$ is known to highly overestimate if the data still contain errors (Roon

Table 3. Comparison of population estimates ($\hat{N} \pm$ SE) and 95% confidence intervals (CI) using 2 conventional estimators: the R-package CAPWIRE (* = best fit), CAPTURE (* = chosen as appropriate model), and 3 error-incorporating models, GUAVA (using Bailey's binomial estimator compiling all sampling occasions into 2 [day 1–5] or using only the first 2 sampling occasions [day 1–2]), Lukacs and Burnham (L&B) estimator (* = best corrected Akaike's Information Criterion [AIC_c], α = probability that an individual is correctly genotyped and classified), and the Wright model. Data are from 115 genotyped Eurasian otter fecal samples collected in Saxony, Germany in March 2006 that were compiled into 4 different datasets: an unmodified dataset and 3 modified datasets: unique genotypes are removed (dataset a); unique genotypes are added to similar groups adopting their genotype (dataset b); and unique genotypes are added to similar groups assuming the heterozygous genotype to be true (dataset c). For each dataset, we report the number of individual genotypes for the respective scenario (n).

Approach	Model	Unmodified ($n = 22$)	Dataset a ($n = 17$)	Dataset b ($n = 17$)	Dataset c ($n = 17$)
CAPWIRE	TIRM	$\hat{N} = 23$, CI = 22–25	$\hat{N} = 17$, CI = 17–18	$^*\hat{N} = 17$, CI = 17–21	$^*\hat{N} = 17$, CI = 17–21
	PART-TIRM	$^*\hat{N} = 24$, CI = 22–27	$^*\hat{N} = 17$, CI = 17–18	$\hat{N} = 17$, CI = 17–17	$\hat{N} = 17$, CI = 17–17
CAPTURE	M_0	$\hat{N} = 22 \pm 0.95$, CI = 22–28	$^*\hat{N} = 17 \pm 0.47$, CI = 17–17	$^*\hat{N} = 17 \pm 0.44$, CI = 17–17	$^*\hat{N} = 17 \pm 0.44$, CI = 17–17
	M_h	$^*\hat{N} = 36 \pm 6.45$, CI = 29–55	$\hat{N} = 19 \pm 2.16$, CI = 18–27	$\hat{N} = 19 \pm 2.46$, CI = 18–29	$\hat{N} = 19 \pm 2.46$, CI = 18–29
GUAVA	Day 1–5	$\hat{N} = 14.3 \pm 1.12$, CI = 12–16	$\hat{N} = 14.1 \pm 0.96$, CI = 12–16	$\hat{N} = 13.9 \pm 1.05$, CI = 12–16	$\hat{N} = 13.9 \pm 1.1$, CI = 12–16
GUAVA	Day 1–2	$\hat{N} = 11.2 \pm 0.92$, CI = 9–13	$\hat{N} = 10.8 \pm 0.77$, CI = 9–12	$\hat{N} = 10.8 \pm 0.82$, CI = 9–12	$\hat{N} = 11.1 \pm 0.86$, CI = 9–13
L&B estimator	M_0	$^*\hat{N} = 17.6 \pm 2.7$, CI = 13–24, $AIC_c = 58.1$, $\alpha = 80.2\%$	$^*\hat{N} = 15.6 \pm 1.5$, CI = 13–19, $AIC_c = 53.0$, $\alpha = 91.9\%$	$^*\hat{N} = 15.6 \pm 1.5$, CI = 13–19, $AIC_c = 51.9$, $\alpha = 91.5\%$	$^*\hat{N} = 15.6 \pm 1.5$, CI = 13–19, $AIC_c = 51.9$, $\alpha = 91.5\%$
	M_b	$^*\hat{N} = 17.2 \pm 2.9$, CI = 12–24, $AIC_c = 58.4$, $\alpha = 76.4\%$	$\hat{N} = 15.5 \pm 1.5$, CI = 13–19, $AIC_c = 55.0$, $\alpha = 91.2\%$	$\hat{N} = 15.5 \pm 1.5$, CI = 13–19, $AIC_c = 54.1$, $\alpha = 90.9\%$	$\hat{N} = 15.5 \pm 1.5$, CI = 13–19, $AIC_c = 54.1$, $\alpha = 90.9\%$
	M_h	Confounded parameters, inconclusive results	$\hat{N} = 15.8 \pm 1.5$, CI = 13–19, $AIC_c = 57.1$, $\alpha = 92.7\%$	$\hat{N} = 15.6 \pm 1.5$, CI = 13–19, $AIC_c = 56.3$, $\alpha = 91.9\%$	$\hat{N} = 15.6 \pm 1.5$, CI = 13–19, $AIC_c = 56.3$, $\alpha = 91.9\%$
Wright model	Bayesian estimator	$\hat{N} = 20$, CI = 19–22			

et al. 2005). The 2 tests in DROPOUT (McKelvey and Schwartz 2004) and the estimates of all error-incorporating models suggested that the unmodified dataset still harbored genotyping errors.

Regarding the modified datasets, CAPWIRE (using the R-package) and CAPTURE obtained population size estimates identical to the number of different genotypes, whereas the L&B estimator and the GUAVA method resulted in lower estimates. No estimates were possible for the modified datasets with the Wright model. The GUAVA method seemed to overcompensate for genotyping errors and likely underestimated population sizes, as it would require that identical multilocus errors were replicated in other samples from the same individual and that the genotype of a group of samples was incorrect and not reproducible; both assumptions are rather unlikely (Paetkau 2003, Lukacs and Burnham 2005*b*). The low estimates may have been because of its inability to account for variations in catchability. To test the reliability of the modified datasets, we genotyped 2 additional markers (SRY, Lut914; Dallas et al. 2000) and realized that 1 affiliations was not reasonable, as it showed other genotypes for both markers than the seemingly matching group. Hence, we would have falsely removed a possibly true individual or falsely allocated it to another individual. Thus, correcting a seemingly erroneous dataset can result in an underestimated population size even for conventional estimators.

Interestingly, the L&B estimate of M_0 for the unmodified dataset resulted in the same number as different genotypes in the modified datasets, when the incorrect affiliation was also considered (18 individuals). In a comparison, the 2 best L&B models M_0 and M_b, selected by AIC_c, were 2.2 (M_0) or 5 times (M_b) more likely than the respective conventional model without misidentification. The estimate obtained by the Wright model, kindly computed by J. Wright, suggested 3 more individuals than the averaged L&B models, which could be reasonable, as some individuals might have not been collected during our sampling period. But, it could also be a consequence of not considering FA. Because of the consequences for the protected and threatened otter and because of the high genotyping error rate, we would rather accept a smaller population size. In conclusion, we suggest using the consensus genotypes (here named unmodified dataset) and an error-incorporating method, such as the L&B estimator, or an alternative estimator, such as the Wright model.

MANAGEMENT IMPLICATIONS

Non-invasive genetic CMR is being used increasingly to estimate population sizes and other population parameters. It is an important tool for the management and conservation of natural populations (Schwartz et al. 2007). However, wildlife managers have to be aware of the numerous difficulties in each of the 3 steps involved in genetic CMR analysis (Fig. 1). If these are ignored, especially the presence of genotyping errors, the estimated parameters will be severely biased (e.g., heavily overestimated population sizes;

Waits and Leberg 2000, Pompanon et al. 2005), which can have a devastating effect on the management of wild populations, particularly when the target species is endangered or evokes conflicts with humans. Hence, wildlife managers and conservationists should be interested in a close cooperation between field helpers, lab personnel, ecologists, and statisticians and should employ an a priori developed species-specific sampling design, a rigorous lab procedure that intends to minimize and detect genotyping errors, and an error accounting estimation model (Table S1, available online at www.onlinelibrary.wiley.com).

ACKNOWLEDGMENTS

We would like to thank H. Bayerl for his great assistance in collecting samples and helping to extract the enormous number of samples. We thank A. Tschierschke, U. Hempel, C. Helm, C. Meichsner, J. Kalbitz, Z. Liu, S. Geißler, and J. Kaulfuß for collecting fecal samples. Thanks to E. Küster and the Department of Bioanalytical Ecotoxicology and to W. Durka and the Department of Community Ecology for giving access to the laboratories and for helping with several challenges. We are grateful to J. Wright who spent her time to compile and run our data with their model. We kindly acknowledge private landowners for their permission to access the ponds. This work was supported by the Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE). S. Lampa was funded by the German Science Foundation (DFG), grant GR 1960/2-1.

LITERATURE CITED

Adams, J. R., B. T. Kelly, and L. P. Waits. 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). Molecular Ecology 12:2175–2186.

Adams, J. R., and L. P. Waits. 2007. An efficient method for screening faecal DNA genotypes and detecting new individuals and hybrids in the red wolf (*Canis rufus*) experimental population area. Conservation Genetics 8:123–131.

Arandjelovic, M., K. Guschanski, G. Schubert, T. R. Harris, O. Thalmann, H. Siedel, and L. Vigilant. 2009. Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. Molecular Ecology Resources 9:28–36.

Arandjelovic, M., J. Head, L. I. Rabanal, G. Schubert, E. Mettke, C. Boesch, M. M. Robbins, and L. Vigilant. 2011. Non-invasive genetic monitoring of wild central chimpanzees. PLoS ONE 6(3):e14761.

Arrendal, J., C. Vila, and M. Bjorklund. 2007. Reliability of noninvasive genetic census of otters compared to field censuses. Conservation Genetics 8:1097–1107.

Banks, S. C., M. P. Piggott, B. D. Hansen, N. A. Robinson, and A. C. Taylor. 2002. Wombat coprogenetics: enumerating a common wombat population by microsatellite analysis of faecal DNA. Australian Journal of Zoology 50:193–204.

Bayes, M. K., K. L. Smith, S. C. Alberts, J. Altmann, and M. W. Bruford. 2000. Testing the reliability of microsatellite typing from faecel DNA in the savannah baboon. Conservation Genetics 1:173–176.

Beja-Pereira, A., R. Oliveira, P. C. Alves, M. K. Schwartz, and G. Luikart. 2009. Advancing ecological understandings through technological transformations in noninvasive genetics. Molecular Ecology Resources 9:1279–1301.

Bellemain, E., J. E. Swenson, O. Tallmon, S. Brunberg, and P. Taberlet. 2005. Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for brown bears. Conservation Biology 19:150–161.

Bellemain, E., and P. Taberlet. 2004. Improved noninvasive genotyping method: application to brown bear (*Ursus Arctos*) faeces. Molecular Ecology 4:519–522.

Bhagavatula, J., and L. Singh. 2006. Genotyping faecal samples of Bengal tiger (*Panthera tigris tigris*) for population estimation: a pilot study. BioMed Central Genetics 7:48. http://www.biomedcentral.com/1471-2156/7/48

Bjorklund, M. 2005. A method for adjusting allele frequencies in the case of microsatellite allele drop-out. Mol Ecol Notes 5:676–679.

Blejwas, K. M., C. L. Williams, G. T. Shin, D. R. McCullough, and M. M. Jaeger. 2006. Salivary DNA evidence convicts breeding male coyotes of killing sheep. Journal of Wildlife Management 70:1087–1093.

Bonin, A., E. Bellemain, P. B. Eidesen, F. Pompanon, C. Brochmann, and P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology 13:3261–3273.

Boulanger, J., M. Proctor, S. Himmer, G. Stenhouse, D. Paetkau, and J. Cranston. 2006. An empirical test of DNA mark-recapture sampling strategies for grizzly bears. Ursus 17:149–158.

Boulanger, J., G. C. White, B. N. McLellan, J. Woods, M. Proctor, and S. Himmer. 2002. A meta-analysis of grizzly bear DNA mark-recapture projects in British Columbia, Canada. Ursus 13:137–152.

Brinkman, T. J., D. K. Person, F. S. Chapin, W. Smith, and K. J. Hundertmark. 2011. Estimating abundance of Sitka black-tailed deer using DNA from fecal pellets. Journal of Wildlife Management 75:232–242.

Brinkman, T. J., D. K. Person, M. K. Schwartz, K. L. Pilgrim, K. E. Colson, and K. J. Hundertmark. 2010*a*. Individual identification of Sitka black-tailed deer (*Odocoileus hemionus sitkensis*) using DNA from fecal pellets. Conservation Genetics Resources 2:115–118.

Brinkman, T. J., M. K. Schwartz, D. K. Person, K. L. Pilgrim, and K. J. Hundertmark. 2010*b*. Effects of time and rainfall on PCR success using DNA extracted from deer fecal pellets. Conservation Genetics 11:1547–1552.

Broquet, T., N. Ménard, and E. Petit. 2007. Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conservation Genetics 8:249–260.

Broquet, T., and E. Petit. 2004. Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology 13:3601–3608.

Buchan, J. C., E. A. Archie, R. C. Van Horn, C. J. Moss, and S. C. Alberts. 2005. Locus effects and sources of error in noninvasive genotyping. Molecular Ecology Notes 5:680–683.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information theoretic approach. Springer-Verlag, New York, New York, USA.

Chang, Z. F., Z. J. Liu, J. Y. Yang, M. Li, and L. Vigilant. 2012. Noninvasive genetic assessment of the population trend and sex ratio of the Shennongjia population of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*). Chinese Science Bulletin 57:1135–1141.

Chao, A., and R. M. Huggins. 2005. Modern closed-population capture-recapture models. Pages 58–87 *in* S. C. Amstrup, T. L. McDonald, and B. F. J. Manly, editors. Handbook of capture-recapture analysis. Princeton University Press, Princeton, New Jersey, USA.

Chao, A., and H.-C. Yang. 2003. Program Care-2 (for capture-recapture part 2). Program and user's guide. <http://chao.stat.nthu.edu.tw/>. Accessed 08 May 2013.

Chao, A., P. S. F. Yip, S. M. Lee, and W. T. Chu. 2001. Population size estimation based on estimating functions for closed capture-recapture models. Journal of Statistical Planning and Inference 92:213–232.

Creel, S., G. Spong, J. L. Sands, J. Rotella, J. Zeigle, L. Joe, K. M. Murphy, and D. Smith. 2003. Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. Molecular Ecology 12:2003–2009.

Cullingham, C. I., M. Curteanu, M. C. Ball, and M. Manseau. 2010. Feasibility and recommendations for swift fox fecal DNA profiling. Journal of Wildlife Management 74:849–859.

Dallas, J. F., D. N. Carss, F. Marshall, K.-P. Koepfli, H. Kruuk, S. B. Piertney, and P. J. Bacon. 2000. Sex identification of the Eurasian otter (*Lutra lutra*) by PCR typing of spraints. Conservation Genetics 1:181–183.

Dallas, J. F., F. Marshall, S. B. Piertney, P. J. Bacon, and P. A. Racey. 2002. Spatially restricted gene flow and reduced microsatellite polymorphism in

the Eurasion otter (*Lutra lutra*) in Britain. Conservation Genetics 3:15–29.

Dallas, J. F., and S. B. Piertney. 1998. Microsatellite primers for the Eurasian otter. Molecular Ecology-Primer Notes 7:1247–1263.

DeWoody, J., J. D. Nason, and V. D. Hipkins. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes 6:951–957.

Dreher, B. P., S. R. Winterstein, K. T. Scribner, P. M. Lukacs, D. R. Etter, G. J. M. Rosa, V. A. Lopez, S. Libants, and K. B. Filcek. 2007. Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bear. Journal of Wildlife Management 71:2684–2693.

Ebert, C., F. Knauer, B. Spielberger, B. Thiele, and U. Hohmann. 2012. Estimating wild boar (*Sus scrofa*) population size using faecal DNA and capture-recapture modelling. Wildlife Biology 18:142–152.

Eggert, L. S., J. A. Eggert, and D. S. Woodruff. 2003. Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. Molecular Ecology 12:1389–1402.

Ernest, H. B., M. C. T. Penedo, B. P. May, M. Syvanen, and W. M. Boyce. 2000. Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. Molecular Ecology 9:433–441.

Ferrando, A., R. Lecis, X. Domingo-Roura, and M. Ponsa. 2008. Genetic diversity and individual identification of reintroduced otters (*Lutra lutra*) in north-eastern Spain by DNA genotyping of spraints. Conservation Genetics 9:129–139.

Flagstad, O., E. Hedmark, A. Landa, H. Broseth, J. Persson, R. Andersen, P. Segerstrom, and H. Ellegren. 2004. Colonization history and noninvasive monitoring of a reestablished wolverine population. Conservation Biology 18:676–688.

Flagstad, O., K. Roed, J. E. Stacy, and K. S. Jakobsen. 1999. Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. Molecular Ecology 8:879–883.

Frantz, A. C., L. C. Pope, P. J. Carpenter, T. J. Roper, G. J. Wilson, R. J. Delahay, and T. Burke. 2003. Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology 12:1649–1661.

Frantzen, M. A. J., J. B. Silk, J. W. H. Fergusen, R. K. Wayne, and M. H. Kohn. 1998. Empirical evaluation of preservation methods for faecal DNA. Molecular Ecology 7:1423–1428.

Garton, E. O., J. S. Horne, J. L. Aycrigg, and J. T. Ratti. 2012. Research and experimental design. Pages 1–40 *in* N. J. Silvy, editor. The wildlife techniques manual. Johns Hopkins University Press, Baltimore, Maryland, USA.

Gazey, W. J., and M. J. Staley. 1986. Population estimation from mark-recapture experiments using a sequential Bayes algorithm. Ecology 67:941–951.

Guertin, D. A., M. Ben-David, A. S. Harestad, and J. E. Elliott. 2012. Fecal genotyping reveals demographic variation in river otters inhabiting a contaminated environment. Journal of Wildlife Management 76:1540–1550.

Guichoux, E. L. Lagache, S. Wagner, P. Chaumeil, P. Léger, O Lepais, C. Lepoittevin, T. Malausa, E. Revardel, F. Salin, and R. J. Petit. 2011. Current trends in microsatellite genotyping. Molecular Ecology Resources 11:591–611.

Hajkova, P., B. Zemanova, J. Bryja, B. Hajek, K. Roche, E. Tkadlec, and J. Zima. 2006. Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Molecular Ecology Notes 6:559–562.

Hajkova, P., B. Zemanova, K. Roche, and B. Hajek. 2009. An evaluation of field and noninvasive genetic methods for estimating Eurasian otter population size. Conservation Genetics 10:1667–1681.

Hansen, H., M. Ben-David, and D. B. McDonald. 2008. Effects of genotyping protocols on success and errors in identifying individual river otters (*Lontra canadensis*) from their faeces. Molecular Ecology Resources 8:282–289.

Harris, R. B., J. Winnie, S. J. Amish, A. Beja-Pereira, R. Godinho, V. Costa, and G. Luikart. 2010. Argali abundance in the Afghan Pamir using capture-recapture modeling from fecal DNA. Journal of Wildlife Management 74:668–677.

Hauer, S., H. Ansorge, and O. Zinke. 2002*a*. Mortality patterns of otters (*Lutra lutra*) from Eastern Germany. Journal of Zoology 256:361–368.

Hauer, S., H. Ansorge, and O. Zinke. 2002*b*. Reproductive performance of otters (*Lutra lutra*, Linnaeus, 1758) in Eastern Germany: low reproduction in a long-term strategy. Biological Journal of the Linnean Society 77:329–340.

Hausknecht, R., R. Gula, B. Pirga, and R. Kuehn. 2007. Urine—a source for noninvasive genetic monitoring in wildlife. Molecular Ecology Notes 7:208–212.

Hayakawa, S., and O. Takenaka. 1999. Urine as another potential source for template DNA in polymerase chain reaction (PCR). American Journal of Primatology 48:299–304.

Hoffman, J. I., and W. Amos. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14:599–612.

Huggins, R. M. 1989. On the statistical analysis of capture experiments. Biometrika 76:133–140.

Idaghdour, Y., D. Broderick, and A. Korrida. 2003. Faeces as a source of DNA for molecular studies in a threatened population of great bustards. Conservation Genetics 4:789–792.

Inoue, E., M. Inoue-Murayama, O. Takenaka, and T. Nishida. 2007. Wild chimpanzee infant urine and saliva sampled noninvasively usable for DNA analyses. Primates 48:156–159.

Jacob, G., R. Debrunner, F. Gugerli, B. Schmid, and K. Bollmann. 2010. Field surveys of capercaillie (*Tetrao urogallus*) in the Swiss Alps underestimated local abundance of the species as revealed by genetic analyses of non-invasive samples. Conservation Genetics 11:33–44.

Johansson, M. P., B. J. McMahon, J. Hoglund, and G. Segelbacher. 2012. Amplification success of multilocus genotypes from feathers found in the field compared with feathers obtained from shot birds. Ibis 154: 15–20.

Johnson, P. C. D., and D. T. Haydon. 2007. Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics 175:827–842.

Knapp, S. M., B. A. Craig, and L. P. Waits. 2009. Incorporating genotyping error into non-invasive DNA-based mark-recapture population estimates. Journal of Wildlife Management 73:598–604.

Kohn, M., F. Knauer, A. Stoffella, W. Schröder, and S. Pääbo. 1995. Conservation genetics of the European brown bear—a study using excremental PCR of nuclear and mitochondrial sequences. Molecular Ecology 4:95–103.

Kohn, M. H., E. C. York, D. A. Kamradt, G. Haught, R. M. Sauvajot, and R. K. Wayne. 1999. Estimating population size by genotyping faeces. Royal Society 266:657–663.

Kruuk, H. 2006. Otters—ecology, behaviour and conservation. Oxford University Press, Oxford, United Kingdom.

Lampa, S., B. Gruber, K. Henle, and M. Hoehn. 2008. An optimisation approach to increase DNA amplification success of otter faeces. Conservation Genetics 9:201–210.

Lanszki, J., A. Hidas, K. Szentes, T. Revay, I. Lehoczky, and S. Weiss. 2008. Relative spraint density and genetic structure of otter (*Lutra lutra*) along the Drava river in Hungary. Mammalian Biology 73:40–47.

Lee, S. M., and A. Chao. 1994. Estimating population-size via sample coverage for closed capture-recapture models. Biometrics 50:88–97.

Li, R. W. Fan, G. Tian, H. Zhu, L. He, J. Cai, Q. Huang, Q. Cai, B. Li, Y. Bai, Z. Zhang, Y. Zhang, W. Wang, J. Li, F. Wei, H. Li, M. Jian, J. Li, Z. Zhang, R. Nielsen, D. Li, W. Gu, Z. Yang, Z. Xuan, O. A. Ryder, F. C.-C. Leung, Y. Zhou, J. Cao, X. Sun, Y. Fu, X. Fang, X Guo, B. Wang, R. Hou, F. Shen, B. Mu, P. Ni, R. Lin, W. Qian, G. Wang, C. Yu, W. Nie, J Wang, Z. Wu, H. Liang, J. Min, Q. Wu, S. Cheng, J. Ruan, M. Wang, Z. Shi, M. Wen, B. Liu, X. Ren, H. Zheng, D. Dong, K. Cook, G. Shan, H. Zhang, C. Kosiol, X. Xie, Z. Lu, H Zheng, Y. Li, C. C. Steiner, T.T.-Y. Lam, S. Lin, Q. Zhang, G. Li, J. Tian, T. Gong, H. Liu, D. Zhang, L. Fang, C. Ye, J. Zhang, W. Hu, A. Xu, Y. Ren, G. Zhang, M. W. Bruford, Q. Li, L. Ma, Y. Guo, N. An, Y. Hu, Y. Zheng, Y. Shi, Z. Li, Q. Liu, Y. Chen, J. Zhao, N. Qu, S. Zhao, F. Tian, X. Wang, H. Wang, L. Xu, X. Liu, T. Vinar, Y. Wang, T. -W. Lam, S.-M. Yiu, S. Liu, H Zhang, D. Li, Y. Huang, X. Wang, G. Yang, Z. Jiang, J. Wang, N. Qin, L. Li, J. Li, L. Bolund, K. Kristiansen, G. K. -S. Wong, M. Olson, X. Zhang, S. Li, H. Yang, J. Wang, and J. Wang. 2010. The sequence and de novo assembly of the giant panda genome. Nature 463:311–317.

Link, W. A., J. Yoshizaki, L. L. Bailey, and K. H. Pollock. 2010. Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics 66:178–185.

Luikart, G., N. Ryman, D. A. Tallmon, M. K. Schwartz, and F. W. Allendorf. 2010. Estimation of census and effective population sizes: the

increasing usefulness of DNA-based approaches. Conservation Genetics 11:355–373.

Lukacs, P. M. 2010. Closed population capture-recapture models. Page 509 *in* E. Cooch and G. C. White, editors. Program MARK—a gentle introduction. http://www.phidot.org/software/mark/docs/book/. Accessed 25 Apr 2013.

Lukacs, P. M., and K. P. Burnham. 2005*a*. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14:3909–3919.

Lukacs, P. M., and K. P. Burnham. 2005*b*. Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. Journal of Wildlife Management 69:396–403.

Manly, B. F. J., S. C. Amstrup, and T. L. McDonald. 2005. Capture-recapture methods in practice. Princeton University Press, Princeton, New Jersey, USA.

Martín-Gálvez, D., J. M. Peralta-Sanchez, D. A. Dawson, A. M. Martín-Platero, M. Martínez-Bueno, T. Burke, and J. J. Soler. 2011. DNA sampling from eggshell swabbing is widely applicable in wild bird populations as demonstrated in 23 species. Molecular Ecology Resources 11:481–493.

Marucco, F., L. Boitani, D. H. Pletscher, and M. K. Schwartz. 2011. Bridging the gaps between non-invasive genetic sampling and population parameter estimation. European Journal of Wildlife Research 57:1–13.

McKelvey, K. S., and M. K. Schwartz. 2004. Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. Journal of Wildlife Management 68:439–448.

McKelvey, K. S., and M. K. Schwartz. 2005. Dropout: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. Molecular Ecology Notes 5:716–718.

Miller, C. R., P. Joyce, and L. P. Waits. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357–366.

Miller, C. R., P. Joyce, and L. P. Waits. 2005. A new method for estimating the size of small populations from genetic mark-recapture data. Molecular Ecology 14:1991–2005.

Mills, L. S., J. J. Citta, K. P. Lair, M. K. Schwartz, and D. A. Tallmon. 2000. Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. Ecological Applications 10:283–294.

Miquel, C., E. Bellemain, C. Poillot, J. Bessière, A. Durand, and P. Taberlet. 2006. Quality indexes to assess the reliability of genotypes in studies using noninvasive sampling and multiple-tube approach. Molecular Ecology Notes 6:985–988.

Mondol, S., K. U. Karanth, N. S. Kumar, A. M. Gopalaswamy, A. Andheria, and U. Ramakrishnan. 2009. Evaluation of non-invasive genetic sampling methods for estimating tiger population size. Biological Conservation 142:2350–2360.

Morin, P. A., K. E. Chambers, C. Boesch, and L. Vigilant. 2001. Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes versus*). Molecular Ecology 10:1835–1844.

Mowat, G., and D. Paetkau. 2002. Estimating marten (*Martes americana*) population size using hair capture and genetic tagging. Wildlife Biology 8:201–209.

Mowat, G., and C. Strobeck. 2000. Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. Journal of Wildlife Management 64:183–193.

Murphy, M. A., K. C. Kendall, A. Robinson, and L. P. Waits. 2007. The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. Conservation Genetics 8:1219–1224.

Murphy, M. A., L. P. Waits, K. C. Kendall, S. K. Wasser, J. A. Higbee, and R. Bogden. 2002. An evaluation of long-term preservation methods for brown bear (*Ursus Arctos*) faecal DNA samples. Conservation Genetics 3:435–440.

Nsubuga, A. M., M. M. Robbins, A. D. Roeder, P. A. Morin, C. Boesch, and L. Vigilant. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology 13:2089–2094.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical-inference from capture data on closed animal populations. Wildlife Monographs 62:7–135.

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

Palomares, F., A. Godoy, A. Piriz, J. O'Brien, and W. E. Johnson. 2002. Faecal genetic analysis to determine the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. Molecular Ecology 11:2171–2182.

Palsboll, P. J. J. Allen, M. Berube, P. J. Clapham, T. P. Feddersen, P. S. Hammond, R. R. Hudson, H. Jorgensen, S. Katona, A. H. Larsen, F. Larsen, J. Lien, D. K. Mattila, J. Sigurjonsson, R. Sears, T. Smith, R. Sponer, P. Stevick, and N. Oien. 1997. Genetic tagging of humpback whales. Nature 388:767–769.

Pennell, M. W., C. R. Stansbury, L. P. Waits, and C. R. Miller. 2013. Capwire: a R package for estimating population census size from non-invasive genetic sampling. Molecular Ecology Resources 13:154–157.

Petit, E., and N. Valière. 2006. Estimating population size with noninvasive capture-mark-recapture data. Conservation Biology 20:1062–1073.

Piggott, M. P., S. C. Banks, N. Stone, C. Banffy, and A. C. Taylor. 2006. Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. Molecular Ecology 15:81–91.

Piggott, M. P., E. Bellemain, P. Taberlet, and A. C. Taylor. 2004. A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. Conservation Genetics 5:417–420.

Piggott, M. P., and A. C. Taylor 2003. Extensive evaluation of faecel preservation and DNA extraction methods in Australian native and introduced species. Australian Journal of Zoology 51:341–355.

Pledger, S. 2000. Unified maximum likelihood estimates for closed capture–recapture models using mixtures. Biometrics 56:434–442.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs 107: 1–97.

Pompanon, F., A. Bonin, E. Bellemain, and P. Taberlet. 2005. Genotyping errors: causes, consequences and solutions. Nature Reviews Genetics 6:847–859.

Poole, K. G., D. M. Reynolds, G. Mowat, and D. Paetkau. 2011. Estimating mountain goat abundance using DNA from fecal pellets. Journal of Wildlife Management 75:1527–1534.

Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. Biometrics 52:703–709.

Prigioni, C., L. Remonti, A. Balestrieri, S. Sgrosso, G. Priore, N. Mucci, and E. Randi. 2006. Estimation of European otter (*Lutra lutra*) population size by fecal DNA typing in southern Italy. Journal of Mammalogy 87:855–858.

Puechmaille, S. J., and E. J. Petit. 2007. Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. Journal of Applied Ecology 44:843–852.

Reed, J. Z., D. J. Tollit, P. M. Thompson, and W. Amos. 1997. Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. Molecular Ecology 6:225–234.

Rew, M. B., J. Robbins, D. Mattila, P. J. Palsboll, and M. Berube. 2011. How many genetic markers to tag an individual? An empirical assessment of false matching rates among close relatives. Ecological Applications 21:877–887.

Robinson, S. J., L. P. Waits, and I. D. Martin. 2009. Estimating abundance of American black bears using DNA-based capture-mark-recapture models. Ursus 20:1–11.

Roon, D. A., L. P. Waits, and K. C. Kendall. 2005. A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. Animal Conservation 8:203–215.

Ruell, E. W., S. P. D. Riley, M. R. Douglas, J. P. Pollinger, and K. R. Crooks. 2009. Estimating bobcat population sizes and densities in a fragmented urban landscape using noninvasive capture-recapture sampling. Journal of Mammalogy 90:129–135.

Ruibal, M., R. Peakall, A. Claridge, and K. Firestone. 2009. Field-based evaluation of scat DNA methods to estimate population abundance of the spotted-tailed quoll (*Dasyurus maculatus*), a rare Australian marsupial. Wildlife Research 36:721–736.

Santini, A., V. Lucchini, E. Fabbri, and E. Randi. 2007. Ageing and environmental factors affect PCR success in wolf (*Canis lupus*) excremental DNA samples. Molecular Ecology Notes 7:955–961.

Sastre, N., O. Francino, G. Lampreave, V. V. Bologov, J. M. Lopez-Martin, A. Sanchez, and O. Ramirez. 2009. Sex identification of wolf (*Canis lupus*) using non-invasive samples. Conservation Genetics 10:555–558.

Scheppers, T. L. J., A. C. Frantz, M. Schaul, E. Engel, P. Breyne, L. Schley, and T. J. Roper. 2007. Estimating social group size of Eurasian badgers (*Meles meles*) by genotyping remotely plucked single hairs. Wildlife Biology 13:195–207.

Schwartz, M. K., S. A. Cushman, K. S. McKelvey, J. Hayden, and C. Engkjer. 2006. Detecting genotyping errors and describing American black bear movement in northern Idaho. Ursus 17:138–148.

Schwartz, M. K., G. Luikart, and R. S. Waples. 2007. Genetic monitoring as a promising tool for conservation and management. Trends in Ecology & Evolution 22:25–33.

Smith, D. A., K. Ralls, A. Hurt, B. Adams, M. Parker, and J. E. Maldonado. 2006. Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analyses. Molecular Ecology 15:387–406.

Smith, P. J. 1991. Bayesian analyses for a multiple capture-recapture model. Biometrika 78:399–407.

Solberg, K. H., E. Bellemain, O. M. Drageset, P. Taberlet, and J. E. Swenson. 2006. An evaluation of field and non-invasive genetic methods to estimate brown bear (*Ursus arctos*) population size. Biological Conservation 128:158–168.

Stanley, T. R., and J. D. Richard. 2005. Software review: a program for testing capture-recapture data for closure. Wildlife Society Bulletin 33:782–785.

Stenglein, J. L., L. P. Waits, D. E. Ausband, P. Zager, and C. M. Mack. 2010. Efficient, noninvasive genetic sampling for monitoring reintroduced wolves. Journal of Wildlife Management 74:1050–1058.

Stevick, P. T., P. J. Palsboll, T. D. Smith, M. V. Bravington, and P. S. Hammond. 2001. Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. Canadian Journal of Fisheries and Aquatic Sciences 58:1861–1870.

Sundqvist, A.-K., H. Ellegren, and C. Vilà. 2008. Wolf or dog? Genetic identification of predators from saliva collected around bite wounds on prey. Conservation Genetics 9:1275–1279.

Taberlet, P., S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L. P. Waits, and J. Bouvet. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research 24:3189–3194.

Tredick, C. A., M. R. Vaughan, D. F. Stauffer, S. L. Simek, and T. Eason. 2007. Sub-sampling genetic data to estimate black bear population size: a case study. Ursus 18:179–188.

Uno, R., M. Kondo, T. Yuasa, K. Yamauchi, H. Tsuruga, H. B. Tamate, and M. Yoneda. 2012. Assessment of genotyping accuracy in a non-invasive DNA-based population survey of Asiatic black bears (*Ursus thibetanus*): lessons from a large-scale pilot study in Iwate Prefecture, northern Japan. Population Ecology 54:509–519.

Valière, N. 2002. Gimlet: a computer program for analysing genetic individual identification data. Molecular Ecology Notes 2:377–379.

Valière, N., P. Berthier, D. Mouchiroud, and D. Pontier. 2002. Gemini: software for testing the effects of genotyping errors and multitubes approach for individual identification. Molecular Ecology Notes 2:83–86.

Van Manen, F. T., M. F. McCollister, J. M. Nicholson, L. M. Thompson, J. L. Kindall, and M. D. Jones. 2012. Short-term impacts of a 4-lane highway on American black bears in eastern North Carolina. Wildlife Monographs 181:1–35.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535–538.

Vigilant, L., and K. Guschanski. 2009. Using genetics to understand the dynamics of wild primate populations. Primates 50:105–120.

Waits, J. L., and P. L. Leberg. 2000. Biases associated with population estimation using molecular tagging. Animal Conservation 3:191–199.

Waits, L. P., G. Luikart, and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology 10:249–256.

Waits, L. P., and D. Paetkau. 2005. Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. Journal of Wildlife Management 69:1419–1433.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory Rep. LA-8787-NERP, Los Alamos, New Mexico, USA.

White, G. C., and K. P. Burnham. 1999. Program Mark: survival estimation from populations of marked animals. Bird Study 46:120–139.

Williams B. K., J. D. Nichols , and M. J. Conroy, editors. 2002. Analysis and management of animal populations. Academic Press, San Diego, California, USA.

Williams, B. W., D. R. Etter, D. W. Linden, K. F. Millenbah, S. R. Winterstein, and K. T. Scribner. 2009. Noninvasive hair sampling and genetic tagging of co-distributed fishers and American martens. Journal of Wildlife Management 73:26–34.

Wilson, G. J., A. C. Frantz, L. C. Pope, T. J. Roper, T. A. Burke, C. L. Cheeseman, and R. J. Delahay. 2003. Estimation of badger abundance using faecal DNA typing. Journal of Applied Ecology 40:658–666.

Wright, J. A., R. J. Barker, M. R. Schofield, A. C. Frantz, A. E. Byrom, and D. M. Gleeson. 2009. Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. Biometrics 65:833–840.

Wright, J. A., M. R. Schofield, and R. J. Barker. 2012. Closed population genetic mark-recapture with genotyping error and heterogeneity. Proceedings of the International Statistical Ecology Conference (ISEC), 3–6 July 2012, Krokkleiva, Norway.

Yoshizaki, J., C. Brownie, K. H. Pollock, and W. A. Link. 2011. Modeling misidentification errors that result from use of genetic tags in capture-recapture studies. Environmental and Ecological Statistics 18:27–55.

Zhan, X. J., M. Li, Z. J. Zhang, B. Goossens, Y. P. Chen, H. J. Wang, M. W. Bruford, and F. W. Wei. 2006. Molecular censusing doubles giant panda population estimate in a key nature reserve. Current Biology 16: R451–R452.

Zhan, X. J., Y. Tao, M. Li, Z. J. Zhang, B. Goossens, Y. P. Chen, H. J. Wang, M. Bruford, and F. W. Wei. 2009. Accurate population size estimates are vital parameters for conserving the giant panda. Ursus 20: 56–62.

Associate Editor: Emily Latch.

Chapter 4

Non-Invasive Genetic Mark-Recapture as a Means to Study Population Sizes and Marking Behaviour of the Elusive Eurasian Otter (*Lutra lutra*)

Simone Lampa

Jean-Baptiste Mihoub

Bernd Gruber

Reinhard Klenke

Klaus Henle

Submitted to PLOS ONE on 1st of October 2014

4.1 Abstract

Non-invasive genetic capture-mark-recapture (CMR) methods became a very important tool to estimate population parameters, such as population size and sex ratio, of elusive and rare species. The Eurasian otter (*Lutra lutra*) is such a species of management concern and is increasingly studied using faecal-based genetic sampling. For reliable results, the marking behaviour of otters has to be taken into account to avoid biased sex ratios or population size estimates. Using 2132 otter faeces of a wild otter population in Upper Lusatia (Saxony, Germany) collected over a period of six years (2006–2012), we applied genetic CMR analyses to study the marking behaviour and to gain estimates of population sizes and sex ratios. We detected a sex difference in the marking behaviour of otters with jelly samples being more often defecated by males and placed actively exposed on frequently used marking sites. Since jelly samples are of higher DNA quality, it is important to not only concentrate the sampling exclusively on this kind of samples or marking sites and to invest in sufficiently high numbers of repetitions of non-jelly samples to ensure an unbiased sex ratio. For population size estimation, we used closed population CMR models that account for genetic misidentification and behavioural responses, as otters seemed to react to the handling or removing of their spraints. We obtained the first precise abundance estimate with confidence intervals for Upper Lusatia (e.g. in 2012: $\widehat{N} = 20 \pm 2.1$, 95% CI = 16–25) and showed that spraint densities are not a reliable index for abundances. We could demonstrate that if minks live in sympatry with otters and have comparably high densities, a non-negligible number of supposed otter samples are actually of mink origin. This could severely bias results of otter monitoring if samples are not genetically identified.

4.2 Introduction

Elusive species play an important role in conservation, especially if they contribute to conflicts that may have consequences to biodiversity conservation beyond their protection. However, elusive species are difficult to study with conventional methods and therefore we often lack demographic information that is an important prerequisite for appropriate conflict management. For such species faeces can provide relevant biological information (Kohn & Wayne 1997). Especially in conjunction with genetic techniques, such as microsatellite genotyping, it is possible to individually identify the originator and to use this information in capture-mark-recapture (CMR) models. The so called non-invasive genetic CMR became a very powerful tool since its first application in the 1990ies (Höss et al. 1992; Taberlet & Bouvet 1992) to study rare and elusive species without direct handling (Lukacs & Burnham 2005b; Marucco et al. 2011).

The basic principle of this approach is that non-invasively collected samples (e.g. faeces) are genotyped at multiple molecular loci (e.g. microsatellites). This multilocus genotype is then treated as a molecular individual mark. Matching genotypes are considered to belong to the same individual and are classified as recaptures. Non-matching genotypes indicate newly captured animals. Hence, for each sampling occasion, all individuals are determined to be either captured (coded as 1) or not captured

(coded as 0), resulting in individual capture histories that are used for CMR analyses. Non-invasive genetic CMR opens up the possibility to obtain estimates of population size, sex ratio, survival, migration, fecundity, or population growth (Lukacs & Burnham 2005b).

However, there are several difficulties that must be overcome, such as low success rates and genotyping errors (Pompanon et al. 2005; Lampa et al. 2013). Genotyping errors can either result in erroneously assigning a sample to a wrong individual, because they appear to have the same genotype, or can create new so far unknown but "false individuals" (ghost individuals) by only one single loci being mistyped. The latter is more likely and can lead to overestimated population sizes (Creel et al. 2003; Lampa et al. 2013).

For unbiased estimates it is also required that all individuals have a reasonable chance of being collected (Lampa et al. 2013). Hence, when using faeces as DNA source, the marking behaviour of the target species has to be understood well to avoid biased results through marking differences in e.g. sex, age, social, or reproductive status (Marucco et al. 2011). In wolves, for example, dominant individuals have an increased capture probability due to higher marking rates and preferences for marking sites that are easier to find for collectors (Marucco et al. 2011). Also transient tigers rarely defecate on regular travel routes – where collectors usually search for faeces – to avoid detection by the resident and thus are virtually undetectable through faeces (Mondol et al. 2009). Consequently, when non-invasive genetic CMR is applied behavioural variations between individuals have to be compensated through the study design, laboratory process and/or parameter estimation methods.

The Eurasian otter (*Lutra lutra*) is an elusive and conflict-laden species. It has suffered dramatic declines in Europe due to hunting and man-made changes in its aquatic habitats (e.g. canalisation, water pollution, prey decline) (Kruuk 2006; Ruiz-Olmo et al. 2008). This resulted in protective legislations throughout Europe. Following these protection activities, otters increased in densities and recolonised former haunts in Europe during the last decades. This evoked conflicts with fishermen because the otter's main prey is fish (Ruiz-Olmo et al. 2008). However, we are still lacking important information for conflict management, such as actual population sizes, in most areas of Europe (Kruuk 2006).

Since otters are elusive and mainly nocturnal, they are difficult to (live-)trap (Kruuk 2006). So far, they were either counted directly (Kruuk & Moorhouse 1991; Ruiz-Olmo et al. 2001), or indirectly assessed by counting their holts (Kruuk et al. 1989), tracks (Ruiz-Olmo et al. 2001; Sulkava 2007; Garcia-Diaz et al. 2011), or faeces (Mason & Macdonald 1987; Balestrieri et al. 2011). Otter faeces, so-called spraints, are particularly suitable to study the species, because otters use them for intraspecific communication and produce daily up to 30 spraints (Kruuk 1992, 2006). According to Kruuk (2006), all members of a population regardless of their sex, reproductive status, or age defecate in nearly equal rates. Spraints are placed on frequently visited conspicuous terrestrial sites at specific locations throughout the home range (e.g. rocks, trunks, under bridges, at junctions of water channels). These marking sites and thus the spraints can be easily detected by collectors and therefore became the

"standard survey method" (Mason & Macdonald 1987) mainly to map otter distributions but also to receive rough estimates of population sizes (see Ruiz-Olmo et al. 2001 for a review).

There are contrasting opinions whether spraint counts can be used as an index of abundance. Lanszki et al. (2008) found a positive correlation between relative spraint density and relative numbers of otter genotypes in an area and concluded that spraint counts are suitable as such an index. Similarly, Guter et al. (2008) found a positive correlation between number of spraints and number of otter visits in latrines but Calzada et al. (2010) criticised their study because they were not able to distinguish between individuals and could hence not tell whether all visits and spraint samples were deposited by a single individual. Other researchers also advised against the use of spraint density as an index of population sizes because of temporal, spatial, and individual sprainting variations (Kruuk et al. 1986; Ruiz-Olmo et al. 2001; Chanin 2003).

In recent studies, researchers used otter spraints for non-invasive genetic capture-mark-recapture (CMR) analyses to estimate population size (Arrendal et al. 2007; Hajkova et al. 2009; Bonesi et al. 2013). Although the sex ratio of otter populations is likely to be even or slightly female-biased (Sidorovich 1991; Ansorge et al. 1997; Kruuk 2006), most studies employing non-invasive genetic sampling found a male bias in their sampling (Dallas et al. 2003; Kalz et al. 2006; Arrendal et al. 2007; Janssens et al. 2008; Hajkova et al. 2009; Lanszki et al. 2010; Bonesi et al. 2013). Therefore, Bonesi et al. (2013) queried whether non-invasive sampling is appropriate to estimate population size and sex ratio of otters. They suggested differences in marking behaviour according to sex, social, or reproductive status as possible reasons and encouraged further research on these issues. Furthermore, Brzezinski and Romanowski (2006) found that the sprainting intensity increases if spraints are previously removed. This raises the question, whether non-invasive genetic CMR is also affected by such a reaction.

Here, we present the results of a faecal-based non-invasive genetic CMR study on a wild otter population in Eastern Saxony, Germany, over a period of six years. To get a better understanding of the marking behaviour, we first investigated the characteristics and intensity of spraint deposition and effects of sex and season. Subsequently, we estimated yearly population sizes and sex ratios. Using these population size estimates, we examined whether spraint densities are correlated with number of genotypes or estimated population sizes and could serve as an index for otter abundances. In a second manuscript (Lampa et al., chapter five), we further estimated survival and migration rates and present analyses on dispersal and spatial use.

4.3 Methods

4.3.1 Ethics Statement

The field sampling did not involve capturing or handling of the protected otters. Therefore, we did not require permits or approvals. The accessed land is private and required permission from the fish farmers, although the pond areas are commonly used by the local population for walks or as passage.

4.3.2 Study Area

The study area is located in the Upper Lusatian heath and pond landscape in the eastern part of Saxony, Germany (51°20′N, 14°19′E). Upper Lusatia covers about 5000 ha of ponds (Myšiak et al. 2013). The tradition to build ponds and to use them for fish farming started already in the 13th century (Böhnert et al. 1996). Fish are harvested each autumn, followed by a wintery drainage of the ponds. Three-year-old fish are sold, whereas spawning and young fish (1–2 years) are reinserted to smaller and deeper wintering ponds. In spring, summer ponds are filled with water again and stocked with fish. Besides the commercial function, the ponds offer an important habitat for many endangered species, such as the Eurasian otter. Due to fish production, the Upper Lusatia is believed to host one of the biggest and most viable otter populations in Central Europe (Ansorge 1994; Klenke 1996; Ansorge et al. 1997).

The study area consisted of one single pond (7.6 ha) and seven pond areas, each comprising 8–13 ponds of varying size (0.36–39.6 ha) (Fig. 4.1). In total, the study area included 64 ponds with an overall water surface of 505 ha. All ponds are connected by a complex system of ditches and streams and framed by naturally vegetated embankments that are partly used as agricultural roads. Islands, extensive reed belts, and heavily vegetated peninsulas can serve as resting sites for otters and induce heterogeneous structures. The pond areas are surrounded by pasture, cropland, forest, roads and urban areas.

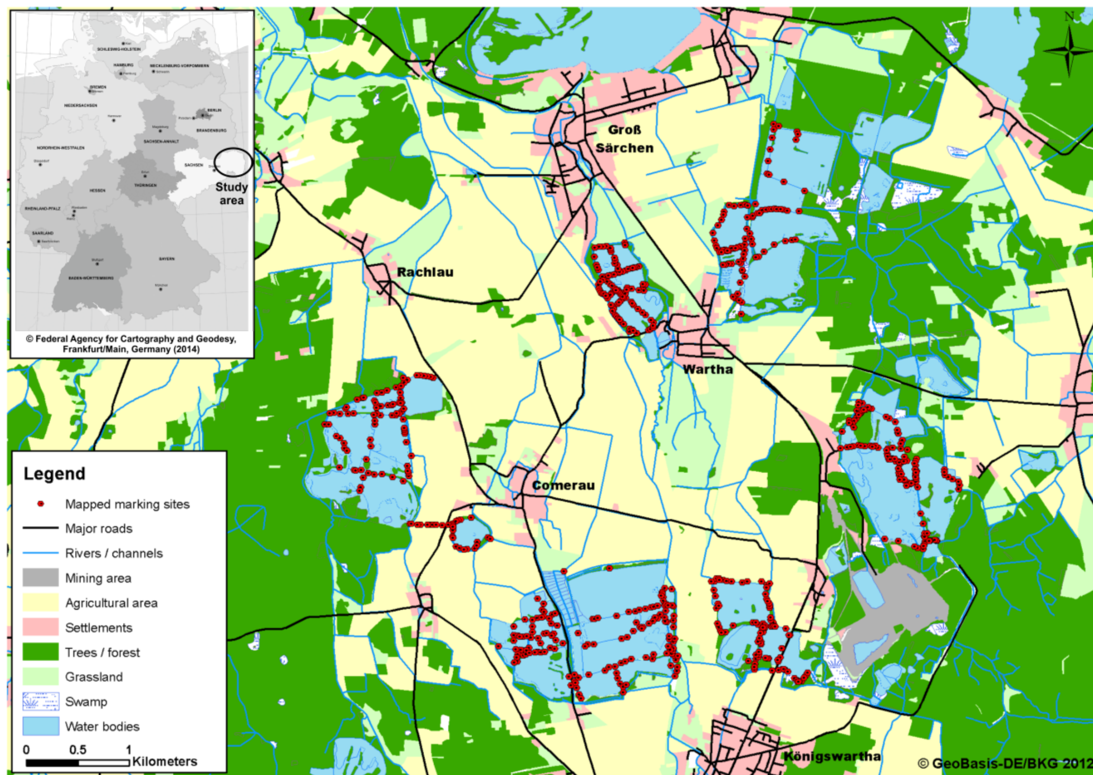

Figure 4.1 Study area map with recorded otter marking sites. Location of otter marking sites (red hexagons) in seven pond areas and one single pond in the Upper Lusatia (Saxony, Germany), where we searched for fresh faeces for genetic capture-mark-recapture (CMR) analyses (2006–2012). Main land use types of the surrounding area are outlined.

4.3.3 Sampling and Microsatellite Genotyping

From 2006 to 2012 (except 2009, missing sampling year), faecal collection was done on five consecutive days just before (March 2006, 2010, 2011, 2012) or just after (April 2007, May 2008) fish relocation into summer ponds. The chosen sampling months (March–May) are considered to be off-peak seasons for otter reproduction in Eastern Germany (Hauer et al. 2002b).

In each year, all ponds filled with water were included in the sampling. The number of ponds varied over years, due to the seasonally and yearly differing water regime management (Tab. 4.1). Each annual faecal collection started with a pre-sampling day on which we recorded active otter marking sites and marked already dropped spraints to facilitate recognition of fresh spraints the next day. In the morning of the following five days, all freshly deposited samples were collected from known or newly discovered marking sites. For each sample, we recorded location of marking site, size category of sample (small, medium, large), its degree of sliminess (spraint, spraint plus mucus, jelly), its exposure level (actively exposed (e.g. scratch piles), passively exposed (e.g. stones, roots, sticks, grass tussock), or non-exposed), and total number of old/fresh samples found on the marking site (1–2, 3–4, > 4). For each fresh spraint, the external layer containing sloughed gut cells was wiped off with a cotton stick. Cotton sticks were placed in a separate sterile 10 ml cryovial (Biozym Scientific, Hessisch Oldendorf, Germany) and either extracted on the day of collection (year 2006) or stored at –80°C until extraction in 1.8 ml buffer ASL (Qiagen, Hilden, Germany) (years 2007–2012).

DNA was extracted from all samples employing the QIAamp® DNA Stool Mini Kit (Qiagen), starting with either adding warm ASL buffer (70°C) to samples (year 2006) or warming up samples to 70°C. All samples were then vortexed and incubated for 2 min at room temperature before proceeding with step four in the manufacturer's protocol. DNA extracts were afterwards stored at –20°C. We followed all precautions recommended by Lampa et al. (2013) to rigorously prevent cross-contamination during extraction and amplification.

Extracted samples were genotyped using seven microsatellite markers (Lut435, Lut457, Lut604, Lut615, Lut701, Lut733, Lut914; Dallas & Piertney 1998; Dallas et al. 2000; Dallas et al. 2002) and sexed with markers Lut-SRY (Dallas et al. 2000) and DBY7Ggu (Hedmark et al. 2004). The latter was designed for wolverines (*Gulo gulo*) but also amplifies in male otters (Hedmark et al. 2004; Koelewijn et al. 2010). To enhance comparability of DNA fragments, the pigtail 'GTTGCTT' was added to the 5′-end of reverse primers to generate a poly(A) tail at the 3′-end. Polymerase chain reaction (PCR) products were separated and visualised in an ABI PRISM® 3100 Genetic Analyser and analysed using ABI PRISM® GeneMapper™ Software V.3.7 (Applied Biosystems, Darmstadt, Germany). The nine loci were multiplexed in three primer sets, hereafter referred to as M1, M2, M3 (M1: Lut 457, 615, 733; M2: Lut 435, 604, 701; M3: Lut 914, SRY, DBY7Ggu). Samples of the year 2006 were amplified following a variation of the multiplex pre-amplification (Bellemain & Taberlet 2004; Piggott et al. 2004), in which two consecutive PCR reactions are carried out for each primer set (see Lampa et al. 2008), to increase genotyping success rates and to lower genotyping error rates. For samples of 2007–

2012, we were able to gain the same success with only one single PCR using a more sensitive polymerase enzyme with high-fidelity and hot-start technique (AmpliTaq Gold® 360 DNA Polymerase, Applied Biosystems) under the following conditions: 25 µl reaction volumes consisted of 3 µl DNA extract, 12.5 µl AmpliTaq Gold® 360 Master Mix, 0.6 µM of each primer, and HPLC-water to the total volume. The hot-start Taq polymerase required an initial denaturation of 95°C for 10 min, followed by 45 cycles of 95°C for 30 sec, 58°C (M1, M2) or 56°C (M3) for 1 min, and 72°C for 30 sec, ending with a final extension at 72°C for 7 min.

Because otter faecal samples from our study area have fairly high genotyping error rates and low genotyping success rates (Lampa et al. 2008, 2013), the genotypes after one PCR per locus contained too many errors. Hence, it was crucial to repeat amplifications generating hereby a consensus genotype. To minimise costs and efforts, we followed a screening approach that consists of five amplification steps after that low-quality samples were removed according to certain thresholds (Lampa et al. 2013). The first amplification step was also used to screen the dataset for non-target species (e.g. mink). After the fifth amplification step, all samples that generated a genotype at all but one or two loci were repeated until a reliable genotype could be assigned to the missing markers (up to 27 repeats). For these additional steps, we partly employed the pre-amplification approach described above and/or used a G/C-Enhancer buffer (included in the AmpliTaq Gold® 360 Master Mix) to increase success rates and to lower genotyping error rates.

The generated consensus genotypes were compared to each other; equal genotypes were scored as belonging to the same individual. Similar genotypes that mismatched at one or two alleles were re-amplified three times at the locus in question to ensure that this was not due to genotyping errors. All successfully genotyped samples were then amplified with the primer set M3 to identify sex. Individuals were identified as males after three sightings of the targeted peak. If all samples of an individual showed no PCR signal after three amplifications, we sexed this individual as a female. Individuals with less than three samples were six times amplified if no targeted peak was recorded to ensure that these samples derived from a female otter.

The six datasets of each year were subsequently checked for still extant genotyping errors with Programme DROPOUT (McKelvey & Schwartz 2005) that determines probably erroneous samples (EB-test) or loci (DCH-test). Actual genotyping error rates were calculated following Broquet and Petit (2004) by comparing scored genotypes with the consensus genotype (see also Lampa et al. 2013). Amplification success rates were calculated by dividing the number of positive PCRs (PCRs showing at least one of the expected alleles) by the number of conducted PCRs, while genotyping success rates depict the number of successfully genotyped samples relative to the number of extracted otter samples. Mean expected heterozygosities (H_e) and sample size corrected probabilities of identity (PI), as well as PIs for siblings (PI_{sib}) were computed over all six loci using software CERVUS 2.0 (Marshall et al. 1998) and GIMLET 1.3.3 (Valière 2002), respectively. All calculations were done for each year separately and an overall mean is provided.

4.3.4 Marking Behaviour

Spraints can either be food remains or a jelly like substance from the intestine, both with or without anal gland secretions (Trowbridge 1983; Kruuk 2006; Kean et al. 2011). For a better understanding of the otter marking behaviour, we first assessed whether spraint sliminess, amount, and exposure, as well as marking site utilisation were affected by sex. For this purpose, we pooled all successfully genotyped samples from all years and conducted a Pearson's chi-squared test for each of the four spraint characteristics. To correct for the multiple testing problem, p-values were adjusted following the Bonferroni-Holm correction (Holm 1979).

To see if males and females defecate at similar rates, we compared the number of deposited spraints per individual first over all years taking the mean number of samples per individual applying a Mann-Whitney-U-test. Taking the actual deposited number of scats per individual, we further tested each year separately for sex differences using two-sample permutation tests implemented in the R-package *exactRankTests* (Hothorn & Hornik 2013). To account for alpha error accumulation, p-values were adjusted according to Bonferroni-Holm procedure.

To test for seasonal differences in the marking behaviour, we compared the three sampling months March, April, and May regarding faecal size and sliminess. For this comparison, we considered the first three years (2006–2008). Because we could not extract all collected samples in 2006 but in 2007 and 2008, we used all yearly collected samples excluding only samples from other species than otter. For both spraint characteristics, we compared each year with each other employing Pearson's chi-squared tests and adjusted the p-values for these six comparisons following the Bonferroni-Holm method.

Furthermore, we were interested in whether the three different spraint types are more or less often placed exposed and on frequently used marking sites and whether the latter have more or less often exposed samples. Using Kendall rank correlation coefficients, we tested for correlations between the sliminess, exposure level, and number of spraints found on the respective marking site, respectively. For this, we pooled all samples that showed at least one expected otter allele (sure otter samples). P-values were adjusted for the three correlations following the Bonferroni-Holm procedure.

The statistics performed in this chapter are done in the R environment (R Development Core Team; www.r-project.org).

4.3.5 Population Size Estimation

We estimated population sizes for each year using closed population CMR models (Pollock et al. 1990). These models require that birth, death, or migration between sampling occasions is negligible. Because our study area was large and we sampled on five consecutive days outside the main reproductive period, these assumptions are very likely met (compare otters biology (Kruuk 2006).

Since it is unlikely that all genotyping errors were completely eliminated from the datasets (Lampa et al. 2013), we employed the error-incorporating misidentification model from Lukacs and Burnham

(2005a) (hereafter L&B estimator) implemented in Program MARK (White & Burnham 1999). The L&B estimator adds to each closed population model available in MARK the misidentification parameter α – the probability of a correct classification. An alpha close to 1 indicates a low probability of still extant genotyping errors.

We estimated separately for each year the population size ($\widehat{N}$), conditional capture (p) and recapture (c) probability, probability of a correct classification (α), and number of genotypes never captured (f_0). We fitted a variety of models to the data that incorporated no capture variation (M_0), individual (M_h), behavioural M($_b$), or daily varying (M_t) catchability and combinations thereof (M_{bh}, M_{th}, M_{tb}). Since we observed a daily increase in the number of collected samples that peaked in the third or fourth sampling day and mostly decreased on the fifth day, we tested if this pattern was introduced by already sampled otters that displayed a daily changing recapture rate (c_1, c_2, c_3, c_4), while the probability to be newly captured (p) remained constant. Each model was fitted with and without a sex difference.

According to the MARK help file (White & Burnham 1999) individual heterogeneity (p_i) is difficult to be separated from misidentification (α), incorporating both can lead to inconclusive results. Whenever p_i and α were only poorly estimable, we dropped it from the candidate model set. Models were adjusted for correct parameter counts where confounding or estimates at the boundary required it.

We ranked models employing corrected Akaike's Information Criterion (AIC_c) that accounts for small sample sizes (Sugiura 1978; Hurvich & Tsai 1989). Using normalised AIC_c weights, reflecting the likelihood of a model (Burnham & Anderson 2002), we calculated a weighted average for all parameter estimates ($\widehat{N}$, p, c, α, f_0). If supported models had unidentifiable parameters, a weighted average estimate for the unidentifiable parameter was calculated by dropping the respective model, but not for estimates of identifiable parameters. The model weighted average capture and recapture probabilities were weighted once more by the respective weighted average pi-value (heterogeneity parameter) and summarised for each day to receive a daily re/capture probability. Using the obtained weighted average population sizes of each year, we calculated population densities per water area (in ha), per km shoreline, and for the total area studied.

Finally, we wanted to test the hypothesis that spraint densities are good indicators for otter densities. Similar as in Lanszki et al. (2008), we used a linear regression to check whether yearly numbers of genotyped scats per ha explain yearly numbers of genotyped individuals per ha or yearly numbers of estimated individuals per ha.

4.4 Results

4.4.1 Sampling and Microsatellite Genotyping

Out of 2132 collected faecal samples, 2001 were extracted (Tab. 4.1). After the first three amplifications with the multiplex trio M1 several samples could be identified as non-otter samples, being either mink (*Neovison vison*) or from other unknown species (Tab. 4.1). By using reference mink samples from an animal park in Leipzig, Germany, we found that Lut457 and Lut615 were

monomorphic (with 120bp and 95bp, respectively) and Lut733 polymorphic (142, 146, 150, or 154bp) in minks, all showing much shorter PCR products than the expected otter alleles. This resulted in fewer, namely 1822, potential otter samples. Since some of these samples (30.1%) did not produce any PCR product at all, they may also belong to other species. Hence the numbers of samples for which we recorded at least one expected otter allele decreased to 1273 (Tab. 4.1).

Table 4.1 Results of faecal-based genetic CMR samplings (2006–2012) from a wild otter population living in pond areas in Upper Lusatia (Saxony, Germany). Water surface: sum of all ponds filled with water. Sure otter samples: samples for which we recorded at least one expected otter allele. Sure mink samples: samples that were identified as mink. Unknown samples: samples that did not produce any PCR product. Genotyped samples: samples successfully genotyped on seven microsatellites and successfully sexed.

Sampling time	Water surface (ha)	Active marking sites	Collec-ted samples	Extrac-ted samples	Sure otter samples	Sure mink samples	Other species	Un-known samples	Geno-typed samples	Geno-types
27–31 Mar 2006	261	130	356	257	199	7	0	51	121	22
23–27 Apr 2007	399	92	282	270	211	7	0	52	134	30
26–30 May 2008	449	87	198	196	136	7	1	52	96	22
22–26 Mar 2010	294	172	381	367	204	50	3	110	130	21
28–1 Mar/Apr 2011	366	173	461	459	239	57	7	156	138	26
27–31 Mar 2012	360	159	454	452	284	33	7	128	159	24
Total	**505**	**384**	**2132**	**2001**	**1273**	**161**	**18**	**549**	**778**	**84**
Mean			**355.3**	**333.5**	**212.2**	**26.8**	**3**	**91.5**	**129.7**	**24.2**

We were able to obtain complete multilocus genotypes for 778 samples (Tab. 4.1), with a mean genotyping success rate over the years of 44.2% considering all potential otter samples (range: 34.9% (2011) – 51.1% (2008)) or 62.1% considering only verified otter samples (range: 57.7% (2011) – 70.6% (2008)). The mean amplification success rate for the autosomal markers over all samples, loci, and years amounted to 79.9% (range: 75.5% (2006) – 83.6% (2008)). The two gonosomal markers showed amplification success rates over all years of 87.3% for Lut-SRY (range: 83.1% (2010) – 91.8% (2012)) and of 54.2% for DBY7Ggu (range: 40.9% (2010) – 80% (2007)). The mean expected heterozygosity over the years amounted to 0.51 (range: 0.49 (2011) – 0.54 (2006)), whereas the observed heterozygosity reached on average 0.6 (range: 0.54 (2011/2012) – 0.65 (2006)). The probability that two different individuals share the same genotype (PI) was sufficiently low. The theoretical unbiased PI ranged between 5.3×10^{-5} (2006) and 1.6×10^{-4} (2010) and the PI_{sib} between 1.6×10^{-2} (2006) and 2.4×10^{-2} (2011).

Genotyping error rates over all years amounted to 48.9% with an AD rate of 45.1% (range: 39.3% (2012) – 48% (2006)) and a FA rate of 3.8% (range: 2.9% (2006) – 4.6% (2012)). The two tests in Programme DROPOUT (McKelvey & Schwartz 2005) indicated that no locus had significantly more errors than any other locus (DCH-test) and there was thus no need to drop any locus. However, in 2006, 2007, and 2011 artificially 4–5 new individuals were produced compared to only two new individuals in the other years. Also the numbers of mismatching loci in the EB-test showed a bimodal distribution in 2006 and too many 1MM and 2 MM pairs in 2007 and 2011. Hence, although we

amplified each sample on average 5.6 times per loci (range 3–27), datasets are likely to still contain genotyping errors. Since further replications would probably not eradicate all errors, it was necessary to employ analysis methods that can incorporate genotyping errors.

The genotyped samples could be pooled to 79 distinct genotypes out of which five dyads showed different sexes resulting in 84 different individuals (43 ♂, 41 ♀). Out of these, 46 individuals (27 ♂, 19 ♀) were only found in one of the six years, 21 (10 ♂, 11 ♀) in two years (not necessarily consecutive years), 11 (5 ♂, 6 ♀) in three years, and 6 (1 ♂, 5 ♀) individuals were found in four years.

4.4.2 Marking Behaviour

Testing size, sliminess, exposedness, and marking site utilisation for differences in sex revealed that only sliminess was significantly different between males and females (Pearson's chi-squared test: χ^2 = 9.6, df = 2, $p_{\text{-adjusted}}$ = 0.0082). Males significantly defecated more often jelly samples and less often spraints than females (Fig. 4.2).

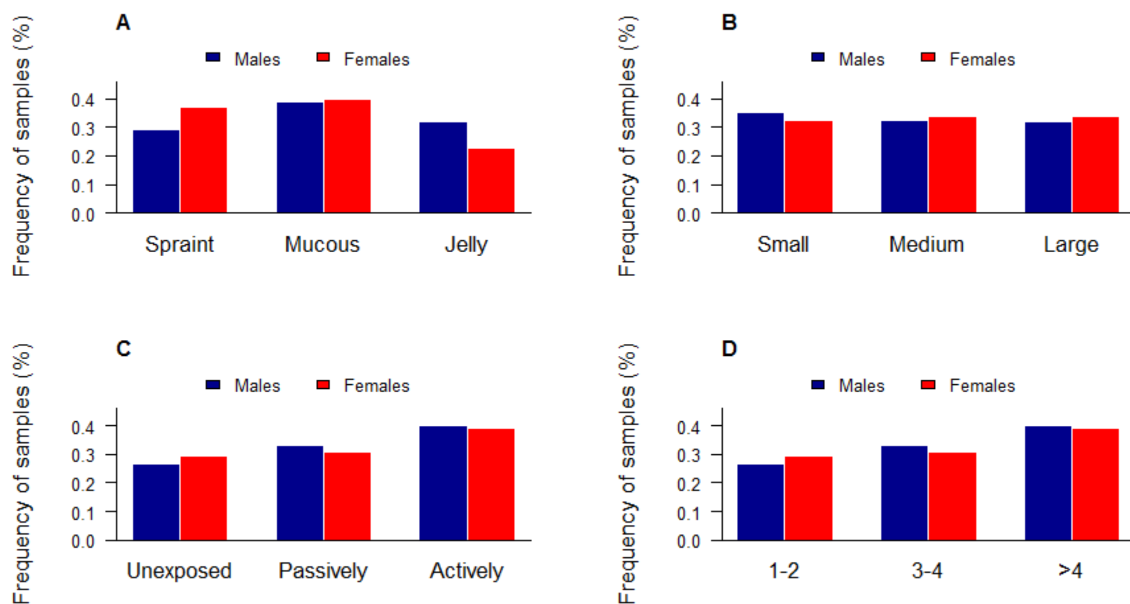

Figure 4.2 Sex differences in otter marking behaviour. Frequency of genotyped otter samples regarding (A) their sliminess (spraint, spraint plus mucus, jelly samples), (B) their size (small, medium, large), (C) their level of exposedness (non-exposed, passively exposed, actively exposed), and (D) the number of otter faeces at the specific marking site (1–2, 3–4, >4 samples); all four separated by sex. Only sliminess showed a significant sex difference in a Pearson's chi-squared test (χ^2 = 9.6, df = 2, p-adjusted = 0.0082).

The maximum number of scats deposited by one individual within a yearly sampling period amounted to 26. Within one night, individuals defecated on average 1.76 spraints with a maximum of 11. Both maxima were generated by males. However, taken over all years sex had no significant effect on the number of deposited scats (U-test: W = 971.5, p = 0.4189; $\text{mean}_{\text{-males}}$ = 4.9, $\text{median}_{\text{-males}}$ = 4, $\text{mean}_{\text{-females}}$ = 4.7, $\text{median}_{\text{-females}}$ = 4.5). Hence, there were also no significant differences within a year (permutation tests: $p_{\text{-}2006}$ = 0.51; $p_{\text{-}2007}$ = 0.46; $p_{\text{-}2008}$ = 0.22; $p_{\text{-}2010}$ = 0.083; $p_{\text{-}2011}$ = 0.28; $p_{\text{-}2012}$ = 0.97).

Comparing the three different sampling months (March, April, May) revealed that the size of samples varied significantly between March and April (Pearson's chi-squared test: $\chi^2 = 17.1$, df = 2, $p_{-adjusted} = 0.0008$) and between April and May ($\chi^2 = 9.7$, df = 2, $p_{-adjusted} = 0.02$), with April having more small but less medium and large samples than March or May. Regarding sliminess, the samples in March significantly differed to samples in April and May (Pearson's chi-squared test: March ~ April: $\chi^2 = 21.95$, df = 2, $p_{-adjusted} = 8.6 \times 10^{-05}$; March ~ May: $\chi^2 = 22.5$, df = 2, $p_{-adjusted} = 7.7 \times 10^{-05}$). They consisted more often of spraints with mucus (March: 49.3%, April: 32%, May: 0.28%) and less often without mucus (March: 32.9%, April 49.8%, May: 48.4%). No year differed in the number of jelly samples (March: 17.9%, April: 18.1%, May: 23.4%).
The correlations between sliminess, exposedness, and numbers of samples in a respective marking site showed that the more slime a sample consists of the more often it is placed exposed (more often actively than passively), whereas less slimy spraints are more often deposited in a non-exposed way (Kendall's tau = 0.087, z = 3.47, $p_{-adjusted} = 0.0011$). On marking sites that were not used the days before, we found less often jelly samples than on marking sites with at least five old/fresh spraints (Kendall's tau = 0.063, z = 2.52, p = 0.012). When correlating the exposedness with the number of samples on a marking site, the results showed that the more samples are deposited on a marking site the more likely they are actively exposed (Kendall's tau = 0.16, z = 6.42, $p_{-adjusted} = 4.23 \times 10^{-10}$).

4.4.3 Population Size Estimation

In some years we had to drop individual heterogeneity models from the candidate model set because heterogeneity was confounded with misidentification (Tab. 4.2). All models with sex-dependent parameters (p_i, p, c, α, f_0) showed no significant difference in a likelihood-ratio test compared to the respective model without the sex effect and were always ranked lower with ΔAIC_c between 3.9 and 29 (mean = 12.1). Thus, these models were dropped from the candidate model set. The model and p_i (within year capture heterogeneity) weighted average capture probabilities (p) were relatively high for each year (0.48–0.75; mean = 0.57), whereas the model and p_i weighted average recapture probabilities (c) were even higher (0.54–0.79; mean = 0.65). Except of year 2010, where we found equal but very high re/capture rates, the recapture probability was always higher than the capture probability, with differences between 0.011–0.23 (mean = 0.08) (Tab. 4.2).
The average misidentification parameter α ranged between 0.73 and 0.95 (mean = 0.85), indicating that each year's dataset still harboured ghost individuals and hence genotyping errors. The population size estimates ($\widehat{N}$) of all models for a particular year were very similar, even for those having AIC_c weights < 0.01. The model weighted average population size using AIC_c weights for each year ranged between 15 (2010) and 26 (2011) individuals (mean = 21) (Tab. 4.2). In four years (2007/10/11/12), we had more females than males with sex ratios, as a male to female mean, ranging between 0.67 and 0.88. In 2008 the sex ratio equalled 1 and in 2006 we found more males than females with a sex ratio of 1.2.

Table 4.2 Closed population models with misidentification run in Programme MARK to estimate yearly population sizes for males and females of a wild otter population in Upper Lusatia (Saxony, Germany). Numbers of modelled parameters (K), differences in AIC_c (ΔAIC_c), AIC_c model weights (w_i), and estimated population sizes with standard errors ($\hat{N}$ ± SE) are shown for all models, except of models were heterogeneity and misidentification were confounded (–). Highest ranked models according to AIC_c are highlighted in bold. For each year, we specified the weighted average population size for males ($\hat{N}_{males}$), females ($\hat{N}_{females}$), and the entire population ($\hat{N}_{total}$) together with their standard errors (SE) and 95% confidence intervals (CI). Furthermore, pi and model weighted average capture (p) and recapture (c) probabilities and model weighted misidentification parameters (α) are given with their standard error (SE).

		2006			2007			2008			2010			2011			2012		
Model	K	ΔAIC_c	w_i	$\hat{N}$ ± SE	ΔAIC_c	w_i	$\hat{N}$ ± SE	ΔAIC_c	w_i	$\hat{N}$ ± SE	ΔAIC_c	w_i	$\hat{N}$ ± SE	ΔAIC_c	w_i	$\hat{N}$ ± SE	ΔAIC_c	w_i	$\hat{N}$ ± SE
M_0	3	**0.00**	**0.44**	**19 ± 2.6**	5.81	0.02	24.0 ± 3.5	6.06	0.04	19.4 ± 2.2	**0.00**	**0.49**	**15.4 ± 2.1**	**0.00**	**0.33**	**25.1 ± 1.8**	10.83	0.00	21.5 ± 2.2
M_h	5	2.73	0.11	20 ± 2.8	**0.00**	**0.39**	**27.9 ± 5.5**	**0.00**	**0.82**	**21.9 ± 4.4**	3.62	0.08	15.5 ± 2.2	–	–	–	–	–	–
M_b	4	0.33	0.37	18 ± 2.6	1.97	0.15	23.5 ± 4.4	6.16	0.04	18.9 ± 2.2	2.16	0.17	15.4 ± 2.1	1.26	0.17	24.6 ± 1.8	3.12	0.12	20.6 ± 2.1
M_t	7	7.44	0.01	18.6 ± 2.7	7.38	0.01	22.7 ± 3.3	13.68	0.00	19.6 ± 2.2	2.13	0.17	15.4 ± 2.1	0.38	0.27	24.3 ± 1.7	1.41	0.27	20.4 ± 2.1
M_{bh}	7	5.42	0.03	18.7 ± 2.9	3.53	0.07	27.8 ± 7.1	–	–	–	–	–	–	–	–	–	–	–	–
M_{th}	13	15.85	0.00	18.9 ± 2.6	8.48	0.01	26.3 ± 6.0	–	–	–	10.36	0.00	16.4 ± 2.0	2.39	0.10	23.8 ± 1.5	–	–	–
M_{tb}	11	9.82	0.00	18.5 ± 2.6	1.93	0.15	22.8 ± 3.3	4.48	0.09	19.2 ± 2	11.55	0.00	15.4 ± 2.1	7.23	0.01	24.4 ± 1.5	4.77	0.05	20.3 ± 2.1
$M_{tb_constrained}$	7	5.16	0.03	18.5 ± 2.6	1.30	0.21	24.9 ± 4.6	9.02	0.01	19.2 ± 2	3.57	0.08	15.4 ± 2.1	1.97	0.12	24.4 ± 1.7	**0.00**	**0.56**	**20.3 ± 2.1**
$\hat{N}_{males}$ ± SE (CI)		10.2 ± 1.5 (7.3–13.1)			12 ± 2.6 (7–17.1)			10.7 ± 2.1 (6.6–14.9)			6.6 ± 0.9 (4.8–8.4)			10.4 ± 0.8 (8.9–11.9)			8.2 ± 0.8 (6.5–9.8)		
$\hat{N}_{females}$ ± SE (CI)		8.5 ± 1.2 (6.1–11)			13.7 ± 2.8 (8.1–19.2)			10.7 ± 2.1 (6.6–14.9)			8.8 ± 1.2 (6.4–11.2)			14.2 ± 1 (12.1–16.2)			12.2 ± 1.2 (9.8–14.7)		
$\hat{N}_{total}$ ± SE (CI)		18.7 ± 2.7 (13.4–24)			25.7 ± 5.4 (15.1–36.3)			21.5 ± 4.2 (13.2–29.8)			15.4 ± 2.1 (11.2–19.6)			24.6 ± 1.8 (21–28.1)			20.4 ± 2.1 (16.3–24.5)		
p ± SE		0.5 ± 0.12			0.48 ± 0.28			0.53 ± 0.16			0.75 ± 0.09			0.59 ± 0.12			0.56 ± 0.14		
c ± SE		0.57 ± 0.13			0.62 ± 0.29			=0.54 ± 0.16			0.75 ± 0.1			0.63 ± 0.12			0.79 ± 0.1		
α ± SE		0.85 ± 0.12			0.81 ± 0.14			0.93 ± 0.11			0.73 ± 0.1			0.95 ± 0.07			0.82 ± 0.08		

Using average population sizes, otter densities in our study area ranged from 0.048 (2008) to 0.072 (2006) otters per ha pond (mean = 0.06), from 0.34 (2008) to 0.48 (2006/2007) otters per km pond shore (mean = 0.42), and from 0.004 (2010) to 0.007 (2007) otter per ha area regarding the entire study area (36 km²) (mean = 0.0058).

A linear regression between yearly numbers of genotyped samples per ha and yearly numbers of different genotypes per ha showed an almost significant relationship ($R^2 = 0.62$, df = 4, p = 0.063, Fig. 3). Whereas, yearly numbers of genotyped samples per ha and yearly numbers of estimated individuals per ha showed no relationship ($R^2 = 0.24$, df = 4, p = 0.33, Fig 4.3).

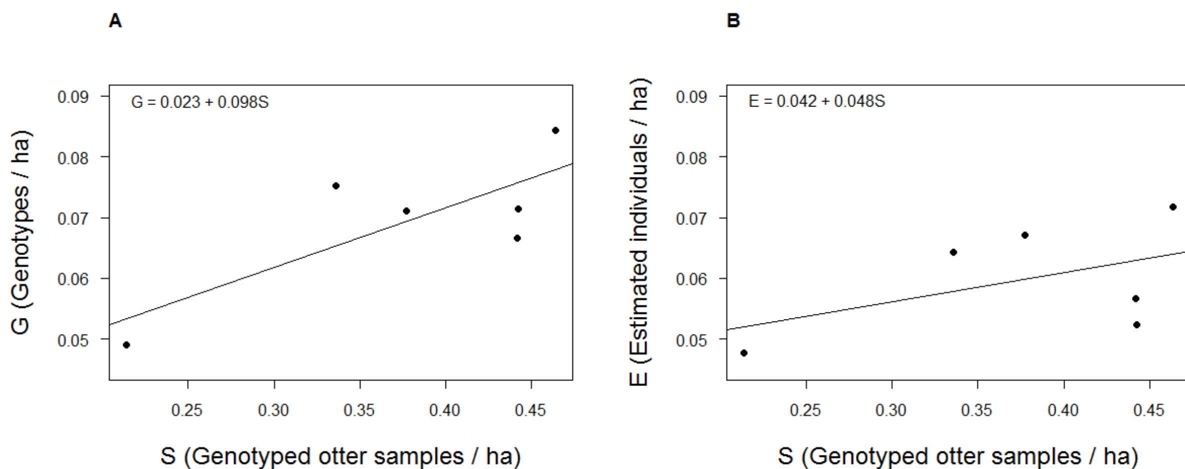

Figure 4.3 Relationship between spraint densities and otter numbers. Linear regressions between number of genotyped otter samples per ha (S) and (A) number of genotypes per ha (G) ($R^2 = 0.62$, df = 4, p = 0.063) and (B) number of estimated individuals (E) ($R^2 = 0.24$, df = 4, p = 0.33), respectively. Equations for both regressions are offered.

4.5 Discussion

4.5.1 Sampling and Microsatellite Genotyping

Since most of the otter-specific microsatellites from Dallas and Piertney (1998) can also be used to detect other mustelids, such as the mink, we were able to distinguish between minks and otters (except of those samples that did not produce a single peak).

Here it is remarkable that although we did not change our sampling design or the way of sampling, the number of collected mink scats was about two to six-fold higher in the years 2010–2012 compared to 2006–2008. For the same period of time, we were able to receive numbers of harvested minks (minks per trapnights – MPT) for one of our pond areas (100 ha) that clearly demonstrated an increase in minks: $MPT_{2008} = 0.028$; $MPT_{2009} = 0.021$; $MPT_{2010} = 0.091$ (kindly provided by A. Lehmann). For comparison, a saturated mink population in ca. 120 ha of the river Thames amounted to MPT = 0.04 using live-traps and including recaptures (Yamaguchi & MacDonald 2003). This implies that contrary to most studies stating that high otter densities are likely to entail a decline in mink densities (Bonesi & Macdonald 2004b; Bonesi et al. 2006; McDonald et al. 2007), the mink proliferated quite well in

our study area despite high otter densities. Similarly, Harrington et al. (2009) found that mink abundances remained relatively high while otter densities raised.

Bonesi and Macdonald (2004a) stated that mink may persist in the presence of otters when terrestrial prey is abundant. The Upper Lusatian pond landscape is known for a high diversity in amphibians, reptiles, water birds, and small mammals (Böhnert et al. 1996). However, most of the mink scats were collected because they contained fish remains, making them more similar to otter spraints. If minks coexist with otters, Bueno (1996) found that minks prey on smaller fishes than otters, which might well be so for our study area. Beside mink scats containing fish remains, we also unintentionally collected mink scats that looked like otter jelly samples. Dunstone (1993) already pointed out that mink can produce a jelly-like secretion. The mink samples were not only collected by students but also by expert collectors. The same difficulty was already noted by Harrington et al. (2010). In their study not a single supposed mink sample collected by experts was of mink origin; rather they belonged to pine martens (47%), foxes (41%), otters (6%), polecats (3%), or stoats (3%). In our study, fresh mink samples were found on typical otter marking sites, sometimes next to fresh otter samples from the same night. This implies that otter monitoring solely relying on otter spraints without genetically determining the species run the risk of overestimating abundance or occupancy if minks are present.

Our microsatellites were only moderately variable regarding observed heterozygosity and had low numbers of alleles. This is consistent with other studies on otters in Europe (Hajkova et al. 2007; Janssens et al. 2008; Mucci et al. 2010). Although the loci achieved acceptable low theoretical unbiased PIs to be able to distinguish between unrelated individuals (see Lampa et al. 2013), we had five dyads that had identical genotypes at the autosomal markers but different sexes. In two cases both individuals of the dyad were either found dead subsequently or in several years or by a high number of samples within a year (≥ 9) and are hence likely to exist and to be closely related (e.g. siblings). For the remaining three dyads, one sex (2 ♂, 1 ♀) was only represented by a single sample in a given year and could thus be an erroneously sexed sample. Since further repetitions could not prove this and since it applied to both sexes, we treated the found genotypes to be real ones.

The genotyping error rate (GER) was quite stable over the six sampling years (range: 0.44 (2012) – 0.51 (2006)), but fairly high compared to other otter studies that used the same way of calculation (Hung et al. 2004: GER = 31.9%; Hajkova et al. 2009: GER = 20.9%; Koelewijn et al. 2010: GER = 17.3%; Bonesi et al. 2013: GER = 18.1%). One reason might be the comparable high number of repetitions (up to 26 times) to gain increased genotyping success rates. Because of this high error rate and the low genotyping success, we followed a rigorous protocol including various contamination preventions during extraction and amplification, a screening approach to exclude low quality samples, and the generation of consensus genotypes via high numbers of repetitions (Lampa et al. 2013). Although those steps minimised errors they could not save us from having still undetected errors in the consensus genotypes. Also the two tests in DROPOUT and the misidentification parameter α indicated that errors might still be present in the yearly datasets. Therefore, it is crucial to use population size

estimators that account for genotyping errors if they cannot be entirely removed (Creel et al. 2003; Marucco et al. 2011; Lampa et al. 2013). All further here implemented statistical tests are less sensitive to ghost individuals as either individual identification was not relevant (e.g. correlations between sample characteristics) or if relevant we only differentiated between sexes. Since there was no significant difference in the number of single samples – that are potential ghost individuals – between both sexes and since re/capture probabilities were equal between males and females, the number of ghost individuals should be evenly distributed among sexes. Thus, we regard the results of the tests for marking behaviour as trustworthy.

4.5.2 Sex Differences in Marking Behaviour

Spraint is used for intraspecific communication, but there are different opinions about what information is transmitted to other otters. Kruuk (1992, 2006) postulated that spraints play a major role in resource partitioning, meaning the use of a resource is advertised by markings, and that it has probably no function in territory defence or sexual communication. Albeit, he admits that "spraints have the potential for carrying many other messages" and that "their exact information content will not be known to us for a long time to come". In contrast, Kean et al. (2011) demonstrated that volatile compounds from anal gland secretions differed in age and for adults also in sex and with reproductive status, suggesting a function in sexual communication. In line with these findings, we showed that although the number of markings did not significantly vary between sexes, jelly samples were more frequently defecated by males and placed exposed on previously used marking sites with several old/fresh scats. This indicates that especially jelly samples have a special role either in sexual communication or for another sex-dependent function, such as social status as found for river otters (Rostain et al. 2004). A function in sexual communication was also postulated by Remonti et al. (2011). Kruuk (1992) stated that such a function would require differences in sprainting behaviour and rates between the sexes and a seasonality synchronised with the breeding season. Although births occur throughout the year in our study area, there is a peak in summer months (Hauer et al. 2002b). With a gestation period of 61–74 days (Kruuk 2006), a mating peak should then be in spring. In our study, the amount of anal gland secretions on faeces decreased in later spring (April/May), whereas the number of jelly samples slightly increased. This, together with the found sex difference, could be another indication for a function of spraints in sexual communication.

4.5.3 Sex ratio

The true sex ratio of otter populations is so far unknown. Sidorovich (1991) found an almost equal (only slightly male-biased) sex ratio of new born cups (♂/♀ = 1.125) and at the age of three months (♂/♀ = 1.09). Since females have lower mortality rates than males (Lampa et al., chapter five), a female biased sex ratio is to be expected and was observed by Kruuk (2006) (♂/♀ = 0.83). However, most studies employing non-invasive genetic sampling (Dallas et al. 2003; Kalz et al. 2006; Arrendal

et al. 2007; Janssens et al. 2008; Hajkova et al. 2009; Lanszki et al. 2010; Bonesi et al. 2013) found more males both in number of samples and individuals. Therefore, Bonesi et al. (2013) questioned the usefulness of non-invasive sampling to estimate population size and sex ratios of otters. We only found more males in year 2006, but an even sex ratio in 2008 and more females in the remaining four sampling years. When comparing the 95% confidence intervals (CI) of the estimated number of individuals per sex (Tab. 4.2), the last two years (2011–2012) showed non-overlapping CIs and thus a female bias. Hence, our results seemed to better reflect the likely natural sex ratio. For 2006, we possibly found more males because we had to cull 99 samples and although those were randomly chosen regarding their origin, we preferred analysing samples with higher success rates, hence jelly samples (Hajkova et al. 2006; Lampa et al. 2008), which more likely are deposited by males (this study). In all other years, we extracted all collected samples and tried to genotype also lower quality samples by persistent repetitions, only dropping samples with no chances to gain a complete multilocus genotype. This might explain the balanced or female-biased sex ratios.

Furthermore, we found that jelly samples were more often placed exposed on more frequently used marking sites with several other faeces. These „hot spots“ are usually larger and more prominent, thus easier to find (e.g. markings sites under bridges). Hence, in some studies a preference of such marking sites might have also resulted into male-biased sex ratios. Therefore, we agree with Bonesi et al. (2013) that non-invasive genetic sampling on otters has to account for their marking behaviour to gain information about sex ratios. Our results indicate that it could be crucial to not drop too many low quality samples, but to invest in replications increasing the overall genotyping success and the numbers of females successfully genotyped, and to include all kinds of marking sites in a study design, also less frequently used sites, to minimise the risk of collecting only a fraction of a population.

4.5.4 Behavioural Response of Sampled Otters

Compared to the capture rates, we observed higher recapture rates in almost all sampling years – except of in 2010 were both rates were comparably high. This could be due to a changed sampling protocol in 2010: larger faeces were first sampled with a cotton swab for genetic analyses and then entirely taken for hormone analyses. In all other years faecal samples were not removed. As otters reuse their marking sites for many years and also daily (Kruuk 2006), higher recapture rates could be collector-induced if they searched more intensely on known marking sites or if they found more samples after a settling-in period (e.g. first 1–2 days). However, 71.1% of the individuals either never reused marking sites (45.9%) or reused one marking site at maximum twice within the five sampling days (25.2%). We also found no difference in the sampling patterns (e.g. settling-in period) between expert collectors and students. Another possibility is that already collected otters reacted on the frequent treatment of their spraints with an increased marking intensity. Such a behavioural response is called “trap-happiness”. It is known that otters use spraints for intraspecific communication (Kruuk 1992, 2006) and so it could well be that they will notice if somebody handled and thus altered their

markings. This could put them on the alert resulting in a higher marking intensity. Such behaviour was also found by Brzezinski and Romanowski (2006), who conducted an experimental approach and found higher sprainting intensity on sites where spraints were previously removed. Removing spraints in 2010 may have disturbed the intraspecific communication such that also unsampled individuals increased their marking intensity or at least used marking sites that were seemingly free of any usage because of previous faecal removing. This is reasonable as the same marking site was used by up to six different individuals within five sampling days (Lampa et al., chapter five). Regardless of whether the behavioural effect is collector- or otter-induced, it is important to account for this when estimating population size of otters (i.e. by including M_b), otherwise the results can be severely biased.

4.5.5 Population Size Estimates

Comparing the number of genotyped and estimated individuals, each year had one to six more genotyped than estimated individuals. If the actual number was not underestimated, we captured most resident individuals, which can be explained by the high sampling intensity.

Most studies estimating otter densities were conducted at rivers, streams, or ditches (Sidorovich 1991; Hung et al. 2004; Prigioni et al. 2006; Lanszki et al. 2008; Ruiz-Olmo et al. 2011), some at lakes or coasts (Erlinge 1968; Kalz et al. 2006; Kruuk 2006), but only a few in fish pond landscapes (Hajkova et al. 2009; Lanszki et al. 2010) (Tab. 4.3). While densities seem to be lower at rivers and lakes than in fish pond landscapes (Tab. 4.3), one needs to bear in mind that comparability is limited because of different methods and water body shapes. Two studies, that also investigated fish pond landscapes employing non-invasive genetic methods, obtained higher estimates per total area (Hajkova et al. 2009) or per km pondside (Lanszki et al. 2010; Tab. 4.3). Besides differences in pond sizes and overall landscape structures, methodological reasons could also account for this difference, because neither Hajkova et al. (2009) nor Lanszki et al. (2010) accounted for genotyping errors. The former used an estimation method, CAPWIRE (Miller et al. 2005), that does not account for genotyping errors. The latter counted the number of genotypes without employing population size estimators. If we would have used the same approaches, our densities would have been larger and comparable to both studies (0.006–0.009 otter per ha area using CAPWIRE; 0.35–0.56 otter per km pond shore using number of genotypes).

For Upper Lusatia, Ansorge (1994) reported densities of 0.001–0.0013 adults per ha area. These estimates are derived on the basis of expert knowledge and only referred to adult otters, whereas our estimates included all age classes. Adding juveniles and subadults (≤ 2 years) to Ansorge's (1994) guesstimate, that comprise about 38–69% of the population (Erlinge 1968; Ansorge et al. 1997), population density would increase to up to 0.002 otters per ha area, about half of our estimates (Tab.4.3). The guesstimates of Ansorge (1994) apply to the early 1990ies, a time period during which otters in Upper Lusatia were believed to still increase in density (Klenke et al. 2013). For the period covered by us (2006–2012), there is no indication that the density is still growing.

Table 4.3 Otter densities of different studies including information on studied habitats and on employed methods. For a comparison, we included results of this study (bold).

Study	Otter per ha area	Otter per ha water area	Otter per km shoreline	Habitat studied	Method used
This study	**0.004–0.007**	**0.048–0.072**	**0.34–0.48**	**Fish ponds**	**Non-invasive genetic CMR**
Ansorge (1994)	0.001–0.0013 (adults)			Fish ponds	Expert knowledge
Erlinge (1968)		0.007–0.01	0.33–0.5 [1]; 0.2 [2]	Lakes [1], rivers [2]	Tracking footprints and spraints
Hajkova et al. (2009)	0.0076–0.0081[1]		0.22–0.26[2]	Fish ponds[1], rivers[2]	Non-invasive genetic CMR
Hung et al. (2004)			1.5–1.8	Rivers	Non-invasive microsatellite genotyping (MNA)
Kalz et al. (2006)	0.0016	0.013	0.21	Lakes, rivers	Non-invasive microsatellite genotyping (MNA)
Koelewijn et al. (2010) (reintroduced pop.)	0.0025–0.0034			Lakes, ponds, rivers	Non-invasive genetic CMR
Kruuk et al. (1989)			0.5–0.7	Coastal habitat	Census of otter holts
Lanszki et al. (2008)			0.17	Rivers, backwater	Non-invasive microsatellite genotyping (MNA)
Lanszki et al. (2010)		0.018–0.046	0.35–1.2	Fish ponds	Non-invasive microsatellite genotyping (MNA)
Prigioni et al. (2006)			0.18–0.2	Rivers	Non-invasive microsatellite genotyping (AC)
Ruiz-Olmo et al. (2011)		0.015–0.063	0.07–0.26	Rivers	Direct census
Sidorovich (1991)			0.02–0.4	Rivers, backwater	Direct census, tracking of footprints

CMR – capture-mark-recapture analyses
MNA – minimum number alive (no estimation only number of genotypes)
AC – accumulation curve (e.g. Kohn et al. 1999)

4.5.6 Spraint Densities as Index of Otter Numbers

It has been argued that spraint density can be used as an index of abundance for comparison of populations in time or in space (Mason & Macdonald 1987). Hence, it was applied in several studies (see Reuther et al. 2000 for a review). A non-invasive genetic study even found a significant positive relation between spraint density and number of genotypes per area (Lanszki et al. 2008). In our study this relationship also was close to significance. However, when relating the spraint density with the number of estimated individuals this positive correlation vanished. Even when comparing only the four sampling years (2006, 2010–2012) where we always sampled end of March, there was no relationship between number of individuals and samples ($R^2 = 0.02$, df = 2, $p = 0.87$). This can be explained by the removal of ghost individuals, which was not the case in the study by Lanszki et al. (2008), who used the number of genotypes. It is natural that the more samples one collects in an area or a period, the more ghost individuals will be in the dataset and thus the more genotypes one will have. Hence, in line with other authors (Kruuk et al. 1986; Chanin 2003), we caution against the extrapolation of otter spraint densities to relative abundances. An extrapolation is even more

precarious to use given that a) mink faeces (if present) can be easily confused with otter spraints (see discussion above), b) number of samples vary seasonal (see discussion above), c) sampling rate (collector-induced) or marking intensity (otter-induced) can increase during several-day sampling periods (see discussion above), and d) one marking site is used by up to six individuals (Lampa et al., chapter five).

4.5.7 Conclusion

Faeces are a valuable source to gain information about population sizes and sex ratios via the use of genetic mark-recapture when potential error sources are carefully addressed and the marking behaviour of the target species is taken into account. We illustrated how sex differences in the marking behaviour can influence non-invasive genetic CMR, because high DNA quality jelly samples were more often defecated by males than by females and placed exposed on frequently used marking sites that are easier to find for collectors. Hence, it is crucial to not only concentrate on sampling jelly samples or on prominent marking sites. Furthermore, we recommend investing in high genotyping success rates by sufficient numbers of repetitions to ensure unbiased sex ratios and decreased genotyping error rates. Because of either collector-induced varying sampling intensity or a behavioural response of otters on spraint handling and removing, researchers should employ models that can account for a behavioural effect to receive unbiased estimates. Even when using high quality samples, researchers should use CMR models that incorporate genotyping errors to avoid overestimates, since it is difficult to completely exclude genotyping errors (Lampa et al. 2013). Our study further shows that faecal densities are not a reliable index for otter abundances because of variability in marking behaviour and because of the risk of confusion with mink faeces even by experts. Similar problems may exist for other elusive species. Therefore, we strongly recommend testing the reliability of faecal densities as index of abundance with genetic CMR methods before using them for monitoring elusive species.

4.6 Acknowledgements

We would like to thank J. Schmalzl for her great assistance in sample extraction and sex identification of the enormous number of samples. Also H. Bayerl is thanked for his help in collecting and extracting the samples of the first year. We thank A. Tschierschke, U. Hempel, B. Bauch, D. Dick, M. Gerisch, W. Dimond, E. Kumar and those students of the University of Leipzig, who participated in the practical training course, for collecting faecal samples. J. Jähnig and J. Harjung are thanked for their help in updating the Access and GIS databases, as well as M. Hoehn and K. Schwerdtner-Máñez for their support in mapping marking sites. We are grateful to A. Lehmann who offered data about his annual hunting bags on minks in our study area. Thanks to E. Küster and the Department of Bioanalytical Ecotoxicology and to W. Durka and the Department of Community Ecology for giving access to the laboratories and for helping with several challenges. We kindly acknowledge private landowners for their permission to access the ponds. This work was supported by the Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE). S. Lampa was funded by the German Science Foundation (DFG), grant GR 1960/2-1.

Chapter 5

Non-Invasive Genetic Mark-Recapture as a Means to Study Population Dynamic and Spatial Use of Eurasian Otters (*Lutra lutra*) in a fish pond landscape

Simone Lampa

Jean-Baptiste Mihoub

Bernd Gruber

Reinhard Klenke

Klaus Henle

Submitted to PLOS ONE on 1st of October 2014

5.1 Abstract

We used non-invasive genetic capture-mark-recapture (CMR) to receive data about survival, migration, sex-biased dispersal, and spatial use of a wild otter population. This population, located in the Upper Lusatia (Saxony, Germany) – a region dominated by commercially used fish ponds – was studied over a period of six years (2006–2012). Overall, we collected 2132 otter faeces, generated 778 multilocus genotypes employing eight microsatellite loci, and found 84 distinct individual genotypes. Using misidentification robust design models in Programme MARK, we found higher survival rates for females ($\Phi = 0.82 \pm 0.07$) than for males ($\Phi = 0.71 \pm 0.08$) and a higher probability to leave and enter the study area for males ($\gamma''_{mean} = 0.31$; $(1 - \gamma')_{mean} = 0.31$) than for females ($\gamma''_{mean} = 0.27$; $(1 - \gamma')_{mean} = 0.26$). Males also showed a significantly higher mobility within our study area both within a sampling year and between sampling years. Comparing male-male and female-female relatedness (R) revealed a probable male-biased dispersal with closer relatedness among females ($R_{mean} = 0.2$ or 0.3) than among males ($R_{mean} = 0.15$ or 0.17). The estimated activity range indices of sub-/adults were larger for males (mean = 26.1 ha; median = 9.4 ha) than for females (mean = 10.9 ha; median = 7.3 ha). Employing a linear mixed-effect model (LME), we demonstrated that male activity range indices were significantly larger and increased stronger with number of marking sites (or alternatively age) compared to females. Overlaps in activity ranges were frequently found both between same-sex and opposite-sex dyads. The extent of overlap positively correlated with relatedness for same-sex dyads, but negatively for opposite-sex dyads. The unusual high proportion of activity range overlaps between same-sex dyads could hint to spatial and/or social structures that are specially adapted to highly productive fish pond systems.

5.2 Introduction

Effective management and conservation of elusive, rare, and threatened animal species require accurate estimates of abundance or population dynamic parameters, such as survival, migration, or dispersal. However, this information is difficult to obtain. Here, non-invasive genetic sampling in combination with capture-mark-recapture (CMR) opened up new possibilities. Non-invasively collected samples, such as hair or faeces, serve as DNA source to generate multilocus genotypes for individual identification. Genotypes are either gained by amplifying several microsatellite loci or nucleotide polymorphisms (SNP). Once individuals are genetically tagged, repeated sampling enables to track individuals in time, producing capture-recapture histories that can be analysed with CMR models. With the aid of non-invasive genetic CMR a bunch of information on elusive species can be gained, such as population sizes and trends over time, survival, migration, growth rate, or fecundity (Lukacs & Burnham 2005b).

The Eurasian otter (*Lutra lutra*) is one example for an elusive, rare, and threatened species that is of conservation concern. The species suffered a massive decline in Europe since the end of the 19th century caused by hunting and man-made changes of their aquatic habitat such as water pollution and

habitat destruction involving a decrease of their prey species, mainly fish (Kruuk 2006; Ruiz-Olmo et al. 2008). As a result, the otter is nowadays strictly protected under international legislation and conventions (Ruiz-Olmo et al. 2008) and started to rise and to expand again throughout Europe (Reuther 2004; Kruuk 2006). However, still little is known about actual population sizes and their changes in areas where previous estimates were made (Kruuk 2006). Also precise estimates on survival and migration rates or on dispersal are still scarce.

Moreover, most studies on otters were carried out on rivers, lakes, or coastal habitats, but only little is known about otters living mainly on fish ponds. Landscapes dominated by fish ponds used for commercial fish farming are important habitats for otters, which functioned as haven during their massive decline (Kranz 2000). Naturally, otter densities can become very high in these areas evoking conflicts with fish farmers (Kranz 2000; Klenke et al. 2013). Because of this conflict and because otter populations in fish pond systems can function as a source for recolonisation of surrounding, uninhabitat areas, otters in fish pond systems are of considerable interest in applied conservation.

Regarding survival, so far only one study estimated survival rates based on non-invasive genetic data for otters living on lakes and rivers in southern Sweden (Arrendal 2007). Other studies inferred mortality rates from carcase sampling by constructing life tables (Kruuk & Conroy 1991; Ansorge et al. 1997). Life table construction from carcases require rather strict assumptions, such as reflecting the true structure of living populations, which often is not the case (Hauer et al. 2002a).

Males seemed to be much more on the move than females, at least in coastal habitats (Kruuk 2006) and on lakes (Erlinge 1967). However, no quantitative data underlay these statements and for other aquatic habitats (e.g. fish ponds) this issue remains unexplored. A sex bias was also postulated for dispersal. Most studies based this supposition either on faecal distributions in a study area (Janssens et al. 2008; Koelewijn et al. 2010) or on telemetry data (Quaglietta et al. 2013) or visual observations (Kruuk 2006) of a few individuals. Only one study employed genetic methods and found a negative correlation between relatedness and geographical distance for females but not for males and deduced a male bias in dispersal (Quaglietta et al. 2013).

Several studies were carried out to estimate sizes of home ranges or core areas (Erlinge 1967; Kruuk & Moorhouse 1991; Hung et al. 2004; Quaglietta et al. 2014), but all on rivers, lakes, or coastal areas and little is known about fish pond systems. Because of the concentration of water bodies in a small space, fish pond systems may force otters to a changed spatial use also in terms of home range sharing.

Since otters use their faeces, so-called spraints, for intraspecific communication, they tend to deposit it on conspicuous points throughout their home range (Kruuk 1992). These often well-established marking sites are used by all members of a population regardless of their sex, age, or status (Kruuk 2006) and can be easily found by collectors. Hence, otter faeces can act as a suitable non-invasive DNA source for the application of microsatellite genotyping and subsequent CMR analyses.

Here, we present the results of a faecal-based non-invasive genetic CMR study on a wild otter population in Eastern Saxony, Germany, over a period of six years. While genetic analyses, marking

behaviour, and estimates of population sizes are discussed elsewhere (Lampa et al., chapter four), we used exactly the same data focusing on estimates of apparent survival and temporary emigration using Programme MARK (White & Burnham 1999). We further investigated sex differences in dispersal and spatial use and tested whether overlaps in spatial use are correlated with the degree of relatedness.

5.3 Methods

5.3.1 Ethics Statement

Since the field sampling did not involve capturing or handling of the protected otter, we did not require permits or approvals. Although local people commonly used the pond areas for walks or as passages, it is private land and access permits were obtained from the fish farmers.

5.3.2 Study Area

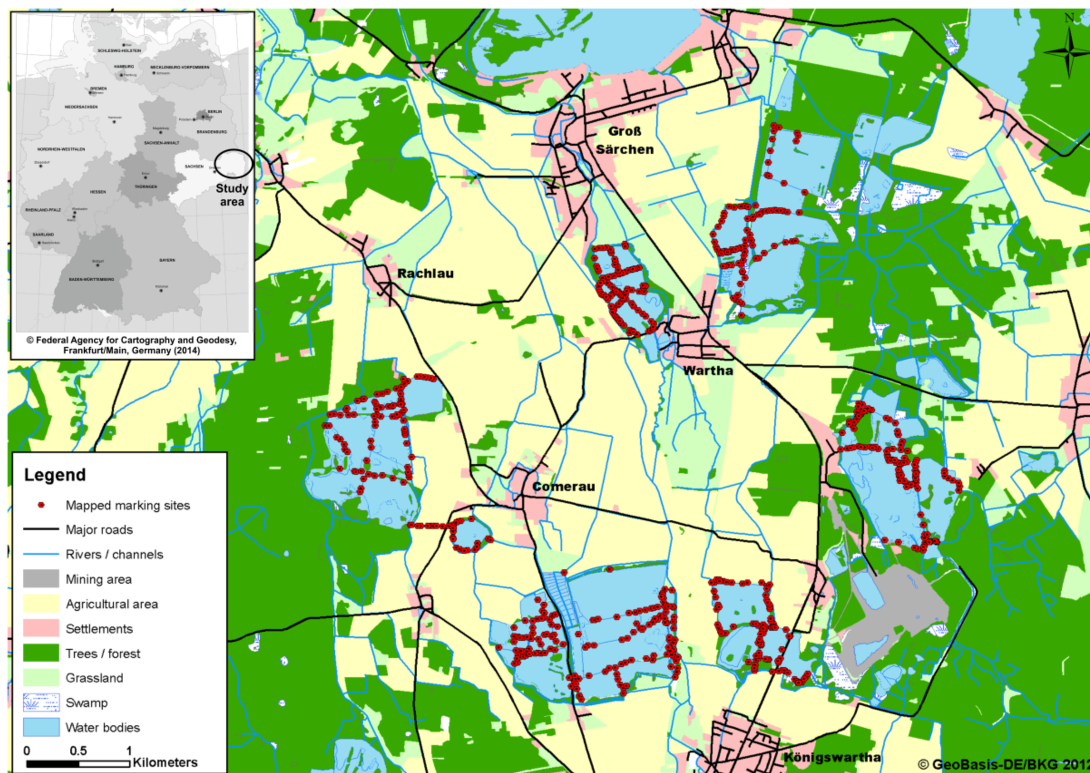

Figure 5.1 Study area map with recorded otter marking sites. Location of otter marking sites (red hexagons) in seven pond areas and one single pond in Upper Lusatia (Saxony, Germany), where we searched for fresh faeces for genetic capture-mark-recapture (CMR) analyses (2006–2012). Main land use types of the surrounding area are outlined.

Our study area is located in the Upper Lusatian heath and pond landscape in Eastern Saxony, Germany (see Lampa et al., chapter four for a detailed description). Upper Lusatia is characterised by hundreds of ponds covering about 5000 ha, mostly used for fish farming (Myšiak et al. 2013). Therefore, the region hosts one of the biggest and most viable populations of otters in Central Europe (Ansorge et al. 1997), with densities of up to 4–7 otters per 10 km² (Lampa et al., chapter four). The study area (51°20′N, 14°19′E) consisted of seven pond areas, each comprising 8–13 ponds of varying size (0.36–39.6 ha), and one single pond (7.6 ha) (Fig. 5.1). The overall water surface amounted to 505 ha. All

ponds (n = 63) are connected by a complex system of ditches and streams and framed by naturally vegetated embankments that are partly used as agricultural roads. Islands, extensive reed belts, and heavily vegetated peninsulas can serve as resting sites for otters and induce heterogeneous structures. Because of fish harvesting in autumn and subsequent wintry drainage of larger summer ponds that are only refilled and re-stocked with fish in spring again, otters are concentrated in wintertime on the fewer but deeper winter ponds that contain fish. Due to this seasonally and yearly differing water regime, the water area covered in the survey varied in each sampling year: 2006 (261 ha), 2007 (399 ha), 2008 (449 ha), 2010 (294 ha), 2011 (366 ha), and 2012 (360 ha).

5.3.3 Sampling and Microsatellite Genotyping

Because a detailed description of the sampling and microsatellite genotyping is given in Lampa et al. (chapter four), we only summarise the main points here:

Faecal collection was conducted on five consecutive days in each of our six sampling years from 2006 to 2012 with one year missing (2009). The sampling occurred either in late winter (March; 2006, 2010, 2011, 2012) or in spring (April 2007, May 2008), both seasons considered to be off-peak seasons for the reproduction in Eastern Germany (Hauer et al. 2002b). Annual faecal collections started with an initial day to get to know the marking sites that were previously mapped and to mark old faeces to facilitate recognition of fresh spraints the next day. On the following five consecutive days all faecal and anal jelly samples were collected from tagged marking sites and from new sites, which were established during that week. Overall, we collected 2132 fresh (from previous night) faecal samples. The external layer of each sample was wiped off with a cotton swab and stored in buffer ASL (Qiagen, Hilden, Germany) at –80°C (see Lampa et al., chapter four for details). Out of these, 2001 samples were extracted using the QIAamp® DNA Stool Mini Kit (Qiagen). Remaining samples had to be culled because they either contained too little spraint, were contaminated during sampling or extraction, or could not be extracted due to capacity reason, as was the case in 2006.

Extracted samples were genotyped with seven microsatellites (Lut435, Lut457, Lut604, Lut615, Lut701, Lut733, Lut914; Dallas & Piertney 1998; Dallas et al. 2000, 2002) and sexed with markers Lut-SRY (Dallas et al. 2000) and DBY7Ggu (Hedmark et al. 2004). After excluding samples that derived from minks or other species (n = 179) and after removing samples that did not produce any PCR product at all (n = 549), we counted 1273 assured otter samples (with at least one expected otter allele) with the following distribution over the years: 199 (2006), 211 (2007), 136 (2008), 204 (2010), 239 (2011), and 284 (2012). Because otter faecal samples from our study area have high genotyping error rates and low genotyping success rates (Lampa et al. 2008, 2013, chapter four), we generated a consensus genotype applying a screening approach that consists of five amplification steps after that low-quality samples were removed according to certain thresholds (Lampa et al. 2013). Complete multilocus genotypes were obtained for 778 samples. We re-amplified genotypes that mismatched at only one or two alleles and checked for still extant genotyping errors with Programme Dropout

(McKelvey and Schwartz, 2005). Actual genotyping error rates were calculated for each year following (Broquet & Petit 2004) and amounted to, on average, 48.9% with an allelic dropout rate of 45.1% (range: 39.3% (2012) – 48% (2006)) and a false allele rate of 3.8% (range: 2.9% (2006) – 4.6% (2012)). Genotyped samples could be pooled to 79 distinct genotypes out of which five dyads showed different sexes resulting in 84 different individuals (43 ♂, 41 ♀). Out of these, 46 individuals (27 ♂, 19 ♀) were only found in one year and 38 (16 ♂, 22 ♀) in up to four of the six years (see Lampa et al., chapter four).

5.3.4 Apparent Survival and Temporary Emigration

For estimation of apparent survival (Φ) and temporary emigration (γ´´ – probability of temporarily emigrating from the study area between two sampling occasions; γ´ – probability of remaining outside the study area), we used robust design models (Pollock 1982; Kendall et al. 1995, 1997) implemented in Programme MARK (White & Burnham 1999). The robust design consists of closed population models for the “secondary sampling sessions” – here, the five sampling occasions of each year – and of open population models for the “primary sampling sessions” – here, the six sampling years. Closed population models assume no birth, death, emigration, or immigration between sampling occasions. For these secondary sampling sessions, we employed the misidentification model of Lukacs and Burnham (2005a) that incorporate genotyping errors, because it is very likely that genotyping errors are still present in the dataset (Lampa et al. 2013, chapter four). These models estimate the population size ($\hat{N}$), the conditional capture (p) and recapture (c) probability, the probability of a correct genotype classification (α), and the number of genotypes never captured (f_0) for each sampling year separately. The secondary sampling sessions are nested in the primary sampling sessions for which apparent survival (Φ) and temporary emigration (γ´, γ´´) can be estimated.

For the robust design, we used a full-likelihood approach as implemented in MARK. To reduce the number of parameters to be estimated, each secondary model was constrained with the respective most parsimonious model ascertained in Lampa et al. (chapter four), namely model M_0 (2006, 2010, 2011), M_h (2007, 2008), and model $M_{tb_constraint}$ (2012). For those years were two models fitted equally well the data (e.g. in 2006 $M_0 \approx M_b$), we tested both models and chose the one with lower AIC_c. Since parameter f_0 was in each year < 1, we constrained it to be equal for all years to further reduce the number of parameters. We fitted a variety of robust design models to the data incorporating time-dependent (t), water surface area-dependent (ha), sex-dependent (sex), or constant (.) parameterisation for apparent survival and temporary emigration, including interactions (time × sex) (ha × sex). Using the most general model (Φ (time × sex), γ (time × sex)) and the most reduced one (Φ (.), γ (.)), we initially evaluated the movement patterns of temporary emigration. Temporary emigration were constrained to account for no movement (γ´´ = γ´ = 0), completely random movement (γ´ = γ´´), first-order Markovian movement ($\gamma'_k = \gamma'_{k-1}$ and $\gamma''_k = \gamma''_{k-1}$) – where the availability depends on the state in which an individual was the year before – or even flow movement (γ´´ = 1 – γ´) – where the

probability of moving from observable to unobservable is the same as moving from unobservable to observable. For parameter identifiability, robust design models require equal survival probabilities for observable and unobservable animals (Kendall et al. 1997). Since our study area is located in a landscape with same characteristics, it is a reasonable assumption that otter survival rates inside and outside the study area are equal.

Models were adjusted for correct parameter counts where confounding or estimates at the boundary required it. Following Lukacs and Burnham (2005a), all parameters were modelled using sine link function. We ranked models using corrected Akaike's Information Criterion (AIC_c) that accounts for small sample sizes (Sugiura 1978; Hurvich & Tsai 1989). For those models that together comprised 99% of the support in the data, we calculated weighted averages for all parameters using normalised AIC_c weights – the likelihood of a model (Burnham & Anderson 2002). To evaluate which of the variables are most important for apparent survival and temporary emigration, we summed up AIC_c weights for the respective variable considering all models in the candidate model set (Burnham & Anderson 2002).

While the estimation of temporary emigration accounts for movements between the study area and the surrounding area from one year to another, we also wanted to study movement differences between sexes within our study area. For this purpose, we first counted for each individual how many consecutive years it was found in the same pond area and compared males and females using a one-sided asymptotic two-sample permutation test employing the R package *exactRankTests* (Hothorn & Hornik 2013). The question behind is whether one sex is more philopatric than the other. We secondly assessed whether movements to other pond areas within a sampling year or between sampling years are more often conducted by one sex employing Fisher's exact tests. In these analyses, we could not correct for genotyping errors, as was done in the robust design models. Genotyping errors can lead to so far unknown but not existing genotypes that are classified as new individuals. These so called "ghost individuals" are mainly presented by genotypes that are only found in one single sample and never again. Since there was no significant difference in the number of single samples between both sexes (Lampa et al., chapter four), it is unlikely that ghost individuals influenced the between sex tests. Statistical analyses here and in the following (unless otherwise specified) were done in the R environment (R Development Core Team; http://www.r-project.org/).

5.3.5 Dispersal

When using nuclear marker, such as microsatellites, sex-biased dispersal can be tested either by comparing genetic differentiation between sexes through e.g. F_{ST} (genetic distance values), by applying assignment tests, or by comparing male-male with female-female relatedness (Freeland 2005). While the latter can be applied within populations, the former two analyses are performed among populations. Since we studied one population that cannot be separated reliably in subpopulations, we are only left with the comparison of relatedness (R) between sexes.

Here, Wang (2011) cautioned against the use of conventional R-values when genotyping errors and/or inbreeding are present. Genotyping errors are certainly present in our data and inbreeding is possible because low genetic diversity and small N_e are usually associated with high inbreeding (Freeland 2005). We found low genetic diversity (Lampa et al., chapter four), as found in other European otter populations (Dallas et al. 2002; Mucci et al. 2010), and Koelewijn et al. (2010) showed that effective population size can be very small ($N_e = 0.30\ N_{total}$). Hence, we calculated relatedness for each sex accounting for a) genotyping errors and b) genotyping errors and inbreeding, by employing the triadic likelihood estimator (TrioML) in Programme COANCESTRY (Wang 2007, 2011). Significant differences between the sexes were ascertained using Mann-Whitney-U-tests in R. To decide whether inbreeding is present or not, COANCESTRY estimates individual inbreeding coefficients. The TrioML-estimator requires a frequency of an incorrect genotype for each locus. Here, we used locus-specific average genotyping error rates that we calculated following Broquet and Petit (2004) (see Lampa et al., chapter four).

5.3.6 Spatial Use

Since we sampled each individual for five subsequent days each year, the area spanned by individually used marking sites represents a minimum area of their activity within the sampling period and can be used as an index for their activity range. Individual activity ranges may change over years, because individuals might shift and rearrange their territories due to environmental factors, such as water regime changes, which happen usually twice a year, or individual characteristics (e.g. sex, age). To identify the main biological factors shaping activity ranges, we intended to test the effect of sex, water surface area, the number of years an individual was known (as a proxy for age), and season on activity range indices. As an index for activity ranges we calculated minimum convex polygons (MCP) that allowed us to include each individual with at least three different sampling points per year. The use of MCPs is criticised, among others, for their sensitivity to the number of locations (Laver & Kelly 2008). To account for this, we included number of locations per MCP as a covariate in the analysis. However, we emphasise that calculated MCPs are not reflecting actual home range sizes but rather an index of activity ranges or area visited for which MCPs can be accepted (Laver & Kelly 2008).

We used ESRI ArcGIS Desktop version 10.1 to calculate MCPs. Since individuals with at least three samples are less likely “ghost individuals”, we deemed this analysis to be not severely biased by genotyping errors. We considered all activity ranges over years (41 ♂, 50 ♀) and accounted for individuals that were recorded in up to three years by including the individual as a random effect in a linear mixed-effect model (LME) implemented in the R package *nlme* (Pinheiro et al. 2014). During initial data exploration, the number of locations per MCP turned out to be correlated with the number of years an individual was known (the proxy for age) (Pearson's product-moment correlation = 0.22, $t = 2.09$, $df = 89$, $p = 0.039$). Since the former is a reliable measurement instead of a proxy, we only included number of locations in the model. Season could also not be included as an explanatory

variable as it was highly correlated with water surface area (Kendall's tau = –0.70, Z = –7.5369, p = 4.8×10^{-14}). However, activity range indices were similar in both seasons ($\text{median}_{\text{-spring}}$ = 7.91 ha, $\text{median}_{\text{-winter}}$ = 8.17 ha). A visual examination of scatterplots with activity range indices against either number of locations per MCP or water surface area for each sex indicated the potential presence of an interaction between the covariates sex and number of locations per MCP. Hence, the most general model included activity range indices (ha) as a response variable with water surface area (ha), sex, number of locations per MCP, and the interaction between the latter two as covariates (fixed effect), and the individual as a random effect. Since males had a larger variance in activity range indices than females, we fitted a heteroscedastic model using the varIdent-function for the covariate sex as part of the R package *nlme* (Pinheiro et al. 2014). We fitted all possible reduced parameter models using a maximum-likelihood estimator (ML), ranked them for their fit by using Akaike's Information Criterion with sample size correction (AIC_c) employing the R package *AICmodavg* (Mazerolle 2013), and selected the best fitting model. The final model was re-fitted with a restricted maximum-likelihood (REML) estimator and validated based on graphical inspections of the residuals (Zuur et al. 2009). Significances of fixed effects were assessed by computing an analysis of variance table.

Since otters in fish pond systems might have another spatial organisation than e.g. otters inhabiting linear home ranges along rivers, we were interested in whether activity ranges overlapped and whether overlapping individuals have a higher degree of relatedness. For this purpose, we chose all individuals that were at least subadults to avoid biased results by cups accompanying their mother. This was done by including individuals that either were known for ≥ 2 years, had known age at time of sampling because they were found dead in subsequent years and could be aged, or that were sampled for the first time but had a relatedness < 0.4 to females in their pond area and close-by ponds. Since the relatedness for parent-offspring is 0.5 (Freeland 2005), those otters were probably not juveniles accompanying their mother. We then tested for all same-sex and opposite-sex dyads collected in the same pond area whether the degree of relatedness correlates with the extent of overlap using Kendall's rank correlation. The degree of relatedness was calculated using the TrioML-estimator (Wang 2007). As an extent of overlap, we computed the overlap area relative to each individual activity range and averaged between the two percentage values. We further compared these averaged relative extent of overlaps between male-male, female-female, and male-female dyads whether any group has larger or more frequent overlaps using Mann-Whitney-U-tests and Bonferroni correction for multiple tests.

To receive information on the marking site use, we investigated by how many individuals a marking site was used, at what time, and to which degree those individuals were related with each other.

5.4 Results

5.4.1 Apparent Survival and Temporary Emigration

The best-supported assumption for the movement pattern of temporal migration was even flow movement with $\Delta AIC_c > 2$ compared to all other movement patterns (Tab. 5.1). As a result, we only considered even flow movement models for inference on apparent survival and temporary emigration.

Table 5.1 Robust design models run in Programme MARK to find the most parsimonious movement pattern for temporary emigration/re-immigration (γ´´, γ´) of otters in Upper Lusatia (Saxony, Germany) sampled between 2006–2012. We used the most general (Φ (time × sex) γ (time × sex)) and the most reduced model (Φ (.) γ (.)) to compare even flow movement (γ´´ = 1 – γ´), Markovian movement ($γ´_k = γ´_{k-1}$ and $γ´´_k = γ´´_{k-1}$), random movement (γ´ = γ´´), and no movement (γ´´ = γ´ = 0). Models are ranked according to their AIC_c values. Further parameters provided are the differences between AIC_c of the candidate model to the best fitting model ($ΔAIC_c$), AIC_c model weights (w_i), the model likelihood (derived by dividing the AIC_c weight of the respective model by the AIC_c weight of the best model), and the number of modelled parameters (K).

Model	AIC_c	$ΔAIC_c$	w_i	Likelihood	K
Φ (.) γ (.) even flow movement	696.865	0	0.72	1	23
Φ (.) γ (.) Markovian movement	699.017	2.153	0.245	0.341	24
Φ (.) γ (.) random movement	702.914	6.049	0.035	0.049	23
Φ (time × sex) γ (time × sex) even flow movement	716.852	19.987	0.00003	0.000	41
Φ (time × sex) γ (time × sex) Markovian movement	731.336	34.471	0.00000	0.000	45
Φ (time × sex) γ (time × sex) random movement	733.306	36.441	0.00000	0.000	41
Φ (.) γ (.) no movement	747.761	50.896	0.00000	0.000	22
Φ (time × sex) γ (time × sex) no movement	762.128	65.263	0.00000	0.000	31

Table 5.2 Candidate model set of robust design models run in Programme MARK to estimate apparent survival (Φ) and temporary emigration (γ) of otters in Upper Lusatia (Saxony, Germany) sampled between 2006–2012. Movement pattern used to model temporary emigration is even flow movement (γ´´ = 1 – γ´). Parameters were modelled either to be constant (.) or to vary with water surface area (ha), with sex (sex), or both (ha × sex). Models are ranked according to their AIC_c values. Further parameters provided are the differences between AIC_c of the candidate model relative to the best fitting model ($ΔAIC_c$), AIC_c model weights (w_i), the model likelihood (derived by dividing the AIC_c weight of the respective model by the AIC_c weight of the best model), and the number of modelled parameters (K).

Model	AIC_c	$ΔAIC_c$	w_i	Likelihood	K
Φ (ha × sex) γ (.)	695.236	0.000	0.129	1.000	24
Φ (sex) γ (.)	695.526	0.290	0.112	0.865	24
Φ (ha × sex) γ (ha)	696.022	0.785	0.087	0.675	24
Φ (sex) γ (ha)	696.047	0.811	0.086	0.667	24
Φ (ha × sex) γ (sex)	696.581	1.345	0.066	0.510	25
Φ (ha) γ (.)	696.607	1.371	0.065	0.504	23
Φ (.) γ (.)	696.865	1.629	0.057	0.443	23
Φ (sex) γ (sex)	696.875	1.639	0.057	0.441	25
Φ (ha) γ (sex)	697.211	1.975	0.048	0.373	24
Φ (sex) γ (ha × sex)	697.217	1.980	0.048	0.372	25
Φ (ha × sex) γ (ha × sex)	697.221	1.984	0.048	0.371	25
Φ (.) γ (ha)	697.398	2.162	0.044	0.339	23
Φ (ha) γ (ha)	697.404	2.168	0.044	0.338	23
Φ (.) γ (sex)	697.512	2.276	0.041	0.321	24
Φ (ha) γ (ha × sex)	697.823	2.586	0.035	0.275	24
Φ (.) γ (ha × sex)	697.879	2.642	0.034	0.267	24

When considering all 36 fitted models, the variable time was the least important predictor for both parameters with summed AIC_c weights of 0.0097 for apparent survival and 0.14 for temporary

emigration. Thus, we dropped this variable from the candidate model set to reduce the number of models (Tab. 5.2). In this candidate model set, the most important variable for apparent survival according to summed AIC_c weights was sex followed by water surface area (Tab. 5.3). The model weighted average apparent survival rates slightly varied over years with higher rates for larger water surface areas (range: Φ_{males} = 0.68–0.73; $\Phi_{females}$ = 0.79–0.84) and with on average 11% higher rates for females than for males (mean: Φ_{males} = 0.71 ± 0.08; $\Phi_{females}$ = 0.82 ± 0.07; Tab. 5.4).

Table 5.3 Sum of AIC_c weights ($w_{i\text{-sum}}$) for those models out of the 16 fitted robust design models of the candidate model set (Tab. 5.2) that contained the respective variable for either apparent survival (Φ) or temporary emigration (γ).

Variable	$w_{i\text{-sum}}$ (Φ)	$w_{i\text{-sum}}$ (γ)
Sex	0.632	0.426
Water surface area	0.521	0.377
Constant	0.177	0.362

Table 5.4 Model weighted average apparent survival rates (Φ) and temporary emigration rates (γ´´ – probability to leave the study area; γ´ – probability to stay outside the study area) each for males and females of a wild otter population in Upper Lusatia (Saxony, Germany). Candidate models used for averaging can be found in Tab. 5.2. For each parameter the average estimate, standard error (SE), and 95% confidence interval (CI) is provided, as well as a mean over all years.

Period	Average Φ_{males} ± SE (CI)	Average $\Phi_{females}$ ± SE (CI)	Average γ''_{males} ± SE (CI)	Average $\gamma''_{females}$ ± SE (CI)	Average γ'_{males} ± SE (CI)	Average $\gamma'_{females}$ ± SE (CI)
2006–2007	0.71 ± 0.09	0.83 ± 0.07	0.31 ± 0.08	0.26 ± 0.07		
	(0.52–0.85)	(0.63–0.93)	(0.17–0.49)	(0.14–0.42)		
2007–2008	0.73 ± 0.09	0.84 ± 0.08	0.30 ± 0.09	0.25 ± 0.08	0.70 ± 0.09	0.75 ± 0.08
	(0.51–0.87)	(0.62–0.95)	(0.16–0.49)	(0.13–0.42)	(0.51–0.84)	(0.58–0.87)
2008–2010	0.68 ± 0.08	0.79 ± 0.07	0.33 ± 0.08	0.28 ± 0.07	0.67 ± 0.08	0.72 ± 0.07
	(0.52–0.81)	(0.62–0.89)	(0.20–0.49)	(0.17–0.44)	(0.51–0.80)	(0.56–0.83)
2010–2011	0.70 ± 0.08	0.81 ± 0.07	0.32 ± 0.08	0.27 ± 0.07	0.68 ± 0.08	0.73 ± 0.07
	(0.52–0.84)	(0.63–0.92)	(0.18–0.49)	(0.15–0.42)	(0.51–0.82)	(0.58–0.85)
2011–2012	0.70 ± 0.08	0.81 ± 0.07	0.32 ± 0.08	0.27 ± 0.07	0.68 ± 0.08	0.73 ± 0.07
	(0.52–0.84)	(0.63–0.92)	(0.19–0.49)	(0.16–0.42)	(0.51–0.81)	(0.58–0.84)
Mean	**0.706**	**0.817**	**0.314**	**0.265**	**0.685**	**0.735**

For temporary emigration, the cumulative support for water surface area, sex, and constant parametrisation (independent of any covariate) received similar support (Tab. 5.3). However, the model weighted average of temporary emigration and re-immigration only slightly varied over the years with a maximum difference of 0.03 (Tab. 5.4). The sex difference amounted to 0.05 with a higher probability to leave the study area for males, but a lower probability to stay outside the study area (Tab. 5.4).

The tests for movement patterns within the study area revealed that the number of consecutive years an individual was found in the same pond area significantly differed between males and females (one-sided asymptotic permutation test: $T = 72$, $p = 0.016$). Males were more often found in only one

sampling year (31 ♂, 22 ♀), females in several consecutive years (two years: 9 ♂, 10 ♀; three years: 3 ♂, 6 ♀; four years: 0 ♂, 3 ♀). Consequently, males are the ones that significantly change more often pond areas within a sampling year (7 ♂, 0 ♀; one-sided Fisher's exact test: $p = 0.0071$) and between sampling years (10 ♂, 2 ♀; one-sided Fisher's exact test: $p = 0.016$).

5.4.2 Dispersal

Regarding sex-biased dispersal, we calculated the relatedness (R) among females and among males employing the TrioML-estimator in Programme COANCESTRY that accounts for genotyping errors and inbreeding. Here, we received highly significant differences between the sexes regardless of whether only genotyping errors (GE) were incorporated (U-test: $W = 412867$, $p = 7.63 \times 10^{-6}$) or genotyping errors and inbreeding (GEI) (U-test: $W = 446333$, $p = 2.25 \times 10^{-14}$). Females were always closer related to each other ($mean_{GE} = 0.20$, $mean_{GEI} = 0.30$) than males to each other ($mean_{GE} = 0.15$, $mean_{GEI} = 0.17$) or the total population ($mean_{GE} = 0.17$, $mean_{GEI} = 0.24$). The mean inbreeding coefficient over all individuals amounted to 0.11 and was much larger than the median with 0.004.

5.4.3 Spatial Use

Activity range indices ranged from 0.0061 ha – a female found on three closed by marking sites on a ditch – to 169.1 ha – a male found in three different pond areas, about 4 km apart from each other – with a median of 7.26 ha (mean = 10.9 ha) for females and 9.39 ha (mean = 26.1 ha) for males.

A full model that allowed different variances for each sex was better supported than a full model without this adjustment ($\Delta AIC_c = 37.9$). Hence, all reduced models containing the covariate sex were fitted as heteroscedastic model. The best fitting LME model included sex, number of locations per MCP, and the interaction between both covariates and had 22% more support in the data than the next best fitting model – the full model (Tab. 5.5). Although these two models only had a difference of $\Delta AIC_c = 1.01$, the covariate water surface area in the full model had no significant effect ($p = 0.76$) on activity range indices. The best fitting model produced residuals meeting homogeneity and normality assumptions and no influential observations were found following Zuur et al. (2009). When residuals were plotted against covariates, no clear patterns emerged. According to the best model, sex ($p = 0.009$), number of locations per MCP ($p = 0.0002$), and the interaction sex × number of location ($p = 0.025$) had a significant effect on activity range indices. Males had larger areas than females and with increasing number of locations male area indices increased stronger than those of females (Fig. 5.2).

Table 5.5 Linear mixed-effect models (LME) with individual as random effect to test for the influence of sex, number of location per MCP (points), or water surface area (area) on the size of otter activity range indices sampled in Upper Lusatia (Saxony, Germany) between 2006–2012. Models including sex as covariate were fitted as heteroscedastic models using the varIdent function of the R package nlme (Pinheiro et al., 2014). Models are ranked according to their AIC_c values. Further parameters provided are the differences between the AIC_c of the candidate model relative to the best fitting model (ΔAIC_c), AIC_c model weights (w_i), the cumulative AIC_c model weights (Cum w_i), and the number of modelled parameters (K).

Model	AIC_c	ΔAIC_c	w_i	Cum w_i	K
activity range ~ sex + points + sex:points	813.90	0	0.55	0.55	7
activity range ~ sex + area + points + sex:points	814.91	1.01	0.33	0.88	8
activity range ~ sex + points	817.00	3.10	0.12	1	6
activity range ~ sex	830.92	17.01	0	1	5
activity range ~ sex + area	833.12	19.22	0	1	6
activity range ~ points	861.26	47.36	0	1	4
activity range ~ area + points	862.65	48.75	0	1	5
activity range ~ 1	875.73	61.83	0	1	3
activity range ~ area	877.84	63.94	0	1	4

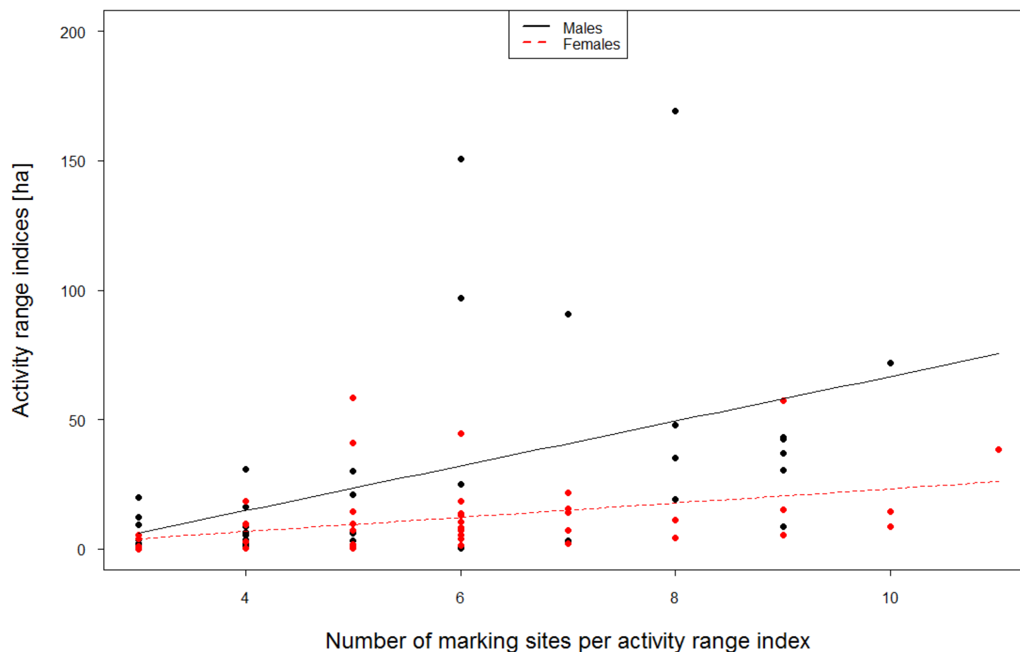

Figure 5.2 Sex and number of marking sites affects the size of activity ranges. Changes of activity range indices for males (black line) and females (red dashed line) with increasing number of marking sites that were used to calculate activity range indices predicted by a heteroscedastic linear mixed-effect model (LME) with the individual as random effect. Black dots signify the input data for males, red dots for females.

Regarding the overlap of sub-/adult activity ranges, we found 15 male-male dyads, 17 female-female dyads, and 34 male-female dyads that stayed in the same pond area in one sampling year. For the sampling year 2006, no dyad could be analysed, since we could not distinguish between juveniles and adults. The extent of overlaps between the three groups (male-male, female-female, male-female) showed no significant differences (U-tests: $p > 0.6$). Out of the 15 male-male dyads seven had no

overlap in their activity range and eight overlapped by up to 81.8% (mean = 34%; median = 26%). Non-overlapping males had a lower degree of relatedness (mean = 0.11; median = 0) than overlapping males (mean = 0.21; median = 0.09), but the relationship between percentage overlap and relatedness was only marginally significant (Kendall's rank correlation: tau = 0.38, Z = 1.67, p = 0.095). For the 17 female-female dyads only four were not overlapping, the remaining overlapped by up to 66.8% (mean = 24%; median = 19%). The more the activity ranges overlapped, the higher was the degree of relatedness (Kendall's rank correlation: tau = 0.65, Z = 3.34, p = 0.00083). When examining opposite-sex dyads that were found in the same pond area, 10 out of 34 dyads showed no overlap, whereas 24 dyads overlapped by up to 72.3% (mean = 25%; median = 15%). An overlapping male-female dyad had lower degree of relatedness (mean = 0.09; median = 0) than non-overlapping dyads (mean = 0.18; median = 0.16). The negative correlation between overlap and relatedness was significant (Kendall's rank correlation: tau = –0.33, Z = –2.43, p = 0.015).

Over the six years we found 384 active otter marking sites, with a density of 0.29 per ha (range: 0.15 (2008) – 0.38 (2010)). The same marking site was used by up to six different individuals within five consecutive sampling days (mean = 1.35), but only 2–3 individuals visited the same marking site in the same night. One small peninsula (80 m long) was even used by eight different individuals within the five sampling days (year 2007: 4 ♂, 4 ♀). We knew that five (3 ♂, 2 ♀) out of these eight individuals were at least subadults and presumable adults. Two of the three males were full siblings or parent-offspring (R = 0.5), both being unrelated to the third male, which was only represented by one single sample and could hence be a ghost individual. The two females that shared the peninsula were unrelated (R = 0). One of the two females was closely related to two of the three possible juveniles.

5.5 Discussion

5.5.1 Apparent Survival

The most important variables that explained apparent survival were water surface area and sex. The water surface area during our sampling time reflects the number and size of ponds that were available for otters five months (when sampled in March) or 1–2 months (when sampled in April/May) before sampling. It is hence reasonable to assume that survival increases the more ponds are available for fishing. However, the differences between years were only small, probably because even a smaller number of ponds provided sufficient resources to otters.

There was a clear difference in sex, with males having lower apparent survival than females. Since apparent survival is a product of true survival (S) and fidelity (F) – the probability of remaining in the population – a lower apparent male survival could either be produced by a lower true survival or by a higher permanent emigration of males or both. In almost all carcase sampling studies there was a male bias and hence the conclusion of a higher male mortality (Ansorge et al. 1997; Hauer et al. 2002a; Dallas et al. 2003; Kruuk 2006; Koelewijn et al. 2010). A higher male mortality is also reasonable if males have higher temporary emigration and re-immigration rates, larger home ranges, and are the

ones that disperse more often. All three seemed to be the case for our studied population (discussed below). Probably it is reasonable to argue that a higher dispersal rate and a higher true mortality resulted in a lower apparent survival rate of males.

To our knowledge, there is only the study of Arrendal (2007) providing estimates of apparent survival for Eurasian otters. Investigating lakes and rivers in southern Sweden, she also found a sex difference (Φ_{males} = 0.51; $\Phi_{females}$ = 0.79) with considerable lower apparent survival rates for males. However, the confidence interval (CI) of the male survival rate ranged between 0.15–0.98. Such a wide CI shows that the estimates are not very reliable. Using life tables constructed from carcases sampled in our study region, Ansorge et al. (1997) developed a population model to calculate mortality rates for each age. The mean mortality rate over all ages amounted to 0.36 and corresponds to a true survival rate of S = 0.64. Since our apparent survival rate ($\Phi_{Male-Female-Mean}$ = 0.75) is a product of true survival and permanent emigration (F) and because we have to assume that dispersal (1 – F) will not be 0 (F < 1), our true survival rate will even be higher than 0.75. However, the data of Ansorge et al. (1997) derived from carcases sampled between the years 1980–1995. In this period, the otter population increased in size (Klenke et al. 2013) and started to expand (Reuther 2004), but deriving mortality estimates from carcases-based life tables require a stationary age distribution (Caughley 1966). Hence, their estimates are less reliable and probably not comparable to our data collected between 2006 and 2012.

Kruuk (2006) stated that otters, unlike most other mammals, have a gradual increase of mortality rate with age (after the first year) and that they have a remarkably short life expectancy of 3–4 years. He also reported a comparably high mean annual mortality rate of 31% for females. However, our estimates for apparent survival show only moderate annual mortality rates. A study on 11 relocated wild-caught otters (Sjöåsen 1996) also found a comparably high survival rate of 0.79 using a Kaplan-Meier estimator with telemetry data. Two studies on river otters (*Lontra canadensis*), that are supposed to have similar mortality rates than the Eurasian otter (Kruuk 2006), also provided survival rates: Guertin et al. (2012) estimated an apparent survival of Φ = 0.889 and Bowyer et al. (2003) found true survival rates of S > 0.8, both studying populations living in coastal water. These survival rates are similar to our results and indicate that our survival rate could be a reasonable estimation or is slightly underestimated.

5.5.2 Temporary Emigration

The knowledge about temporary emigration of otters and differences between sexes is rather restricted. Ansorge et al. (1997) stated that "nothing is known about the migration of otters in the Upper Lusatia region". For coastal habitat Kruuk (2006) asserted that "males were much more erratic than females". Also on lakes Erlinge (1967) claimed that males "are more on the move and travel larger areas". But there are no concrete numbers available yet. Our study is the first estimating temporary migration parameters for an otter population. A movement pattern with balanced temporary emigration rates (males = 31.4%, females = 26.5%) and re-immigration rates (males = 31.5%; females

= 26.7%) was most supported. Our study area is located within a larger landscape with similar characteristics and should not differ in suitability or attractiveness, which is a parsimonious explanation for similar emigration and re-immigration probabilities. Males had a 5% higher probability to leave or enter the study area. This sex difference had a cumulative evidence of 42.6%, which underpins the assumption of a male-biased temporary migration.

When looking at movements within our study area, we also found that males changed pond areas significantly more often than females within a year and between years. Consequently, females were found to have a significantly stronger affinity to their pond areas with 46.3% staying in a pond area for 2–4 years, whereas 72.1% of the males were not re-sampled in their pond area in the subsequent year. The higher mobility of males within a year may be explained by significantly larger male home ranges. In line with this interpretation, our activity range index was larger for males than for females.

5.5.3 Dispersal

To infer on different levels of dispersal between sexes, we calculated relatedness (R) employing the TrioML-estimator (Wang, 2007, 2011) that incorporates genotyping errors and inbreeding. In all analyses, R-values among females were significantly higher than among males. Although a mean inbreeding coefficient of 0.11 is not very low, the median of 0.004 indicates that only a few individuals contributed to the comparably high mean (five ind. > 0.50 out of which three ind. > 0.90). Those five individuals likely were ghost individuals showing a heterozygous deficit due to high allelic dropout rates. Though our studied population is fairly small, which could have also resulted in high inbreeding rates. In any case, our results suggest that dispersal may be male-biased. Indeed, new and re-emerging alleles were only seen in male genotypes when they were recorded for the first time.

Our results are consistent with the few other studies published on dispersal. Kruuk (2006) observed family groups where the female juveniles took over the home range of their mother or settled close by in subsequent years for breeding. The same was inferred from spraint distributions in a reintroduced otter population (Koelewijn et al. 2010). Also, Janssens et al. (2008) detected five males and no females in a recolonisation area and concluded that this suggests male-biased dispersal. Quaglietta et al. (2013) found via radiotracking of subadults (5 ♂, 2 ♀), that no female dispersed but three males settled in other areas than the natal one. They also reported a significant negative correlation between relatedness and geographical distance for females but not for males and deduced that this could be explained by male-biased dispersal.

5.5.4 Spatial Use

Otters deposit their spraints throughout their home ranges (Kruuk, 1992). Hence, their markings can be used to get information about their home ranges. Since most studies on otters are conducted on rivers or coastal habitats, home range sizes are usually specified as linear stretches (e.g. Erlinge 1967; Kruuk & Moorhouse 1991; Hung et al. 2004; Quaglietta et al. 2014). Only a few provided data on the

area of activity (Kruuk 2006; Ó Néill et al. 2009) and there is no study for fish pond systems. On lakes and streams, males used 63 ha area of water, females 34 ha (Kruuk 2006), whereas on a river system males and females used on average 30.2 ha and 16.8 ha of water, respectively (Ó Néill et al. 2009).

The activity range indices we calculated cannot be compared with home range sizes because of the short study period per year and because shifts in spatial use prevented combining data across years. Therefore, our mean activity range indices based on 3–11 samples per year are smaller than the numbers offered by Kruuk (2006) or Ó Néill et al. (2009). However, just like them, our activity ranges showed a significantly larger spatial range for males than for females. This difference was due to seven males that changed pond areas during our sampling period. Removing those males from the data, resulted in a final homoscedastic model with number of locations per MCP as the only significant covariate ($p < 0.0001$) and, thus, equal activity range indices for males (mean = 12.2 ha; median = 6.2 ha) and females (mean = 10.9 ha; median = 7.26 ha). Since the remaining males could not all be juveniles accompanying their mother, they either included other pond areas outside our study area in their activity ranges, or only some males had these larger activity ranges. Other studies reported that male home ranges expanded with age (Arrendal 2007) or at sexual maturity (Sjöåsen 1997) and that only resident adult males had larger ranges than females (Kruuk & Moorhouse 1991). We could show that activity range indices increased with number of locations especially for males. However, this covariate positively correlated with our proxy age. Replacing number of locations per MCP with age resulted in the same final model with similar significances. Hence, activity range indices also increased with increasing age. Since residency, sexual maturity, and paternity are only attained at a certain age, these factors might explain the found pattern.

Regarding the overlap of activity ranges, it is known that several female ranges usually overlap with one larger male range (Erlinge 1968; Kruuk 2006; Quaglietta et al. 2014). Quaglietta et al. (2014) even reported male-female dyads sharing resting sites and spending a considerable amount of time together. For same-sex dyads there are opposing statements. Erlinge (1968) detected territorial behaviour and aggression primarily between individuals of the same sex and Quaglietta et al. (2014) found home ranges of unrelated same-sex dyads were separated by buffer areas. On the other hand, Kruuk (2006) observed home range overlaps among females and assumed these females were closely related with each other. We compared the extent of overlapping activity ranges between male-male, female-female, and male-female dyads that were at least subadults and correlated the extent of overlap with their degree of relatedness. Activity ranges of sub-/adults in the same pond area overlapped with a chance of 76.5% between females, 53.3% between males and 70.6% between males-females. None of the three groups differed in the frequency that overlaps occurred or in the relative size of overlapping activity ranges. For same-sex dyads, we found a positive correlation between the degree of relatedness and the extent of overlapping activity ranges that was significant for females, as assumed by Kruuk (2006). For males in contrast, the positive correlation was just not significant at the 5 % level. This difference between the genders once more indicates that females tend to stay close by their natal area

also for breeding, while there is a male-biased dispersal. Opposite-sex dyads overlapped in the same extent than same-sex dyads, but showed a significant negative correlation between relatedness and extent of overlap. Such differences in relatedness among males and females are also known from other species, result from gender differences in movement and dispersal, and ultimately reduce the risk of inbreeding (Lange et al. 2013).

Overall, it could be shown that same marking sites were used by up to six individuals within a sampling year and that activity ranges of sub-/adults in the same pond area overlapped frequently with high percentage of overlaps. This may indicate that otters in fish pond systems have a more condensed spatial organisation, with smaller home ranges and higher resource partitioning. Probably because the energy requirements can be covered by smaller home ranges due to higher fish densities in fish pond systems compared to systems dominated by rivers, lakes, or coastal water. If core areas overlap more frequently, encounters will also happen more often. This would foster the recent conclusion of Quaglietta et al. (2014) that the social behaviour of otters is more flexible than previously thought.

5.5.5 Conclusion

Our results demonstrated that non-invasive genetic mark-recapture can be used to study population trends, sex ratios, and marking behaviour (Lampa et al., chapter four), as well as population dynamic and spatial use of a small population, despite the presence of ghost individuals. With a relatively short sampling period each year, we received sufficiently precise estimates of apparent survival and temporary emigration and gained information on dispersal and spatial use in fish pond systems.

We found that apparent survival is higher for females than for males, probably due to higher true survival and less dispersal in females. The higher mobility of males is also reflected in higher temporary emigration/re-immigration rates, higher dispersal rates, and larger activity range indices. This should lead to a higher positive correlation between the degree of relatedness and the extent of overlapping activity ranges in females than in males, as observed in our study. The negative correlation of relatedness and overlapping activity range indices in opposite-sex dyads suggests further behavioural mechanisms to reduce the risk of inbreeding. The high proportion of activity range overlaps may indicate that spatial and maybe also social structures are specially adapted to highly productive fish pond systems.

5.6 Acknowledgements

We would like to thank J. Schmalzl for her great assistance in sample extraction and sex identification of the enormous number of samples. Also H. Bayerl is thanked for his help in collecting and extracting the samples of the first year. We thank A. Tschierschke, U. Hempel, B. Bauch, D. Dick, M. Gerisch, W. Dimond, E. Kumar and the students of the University of Leipzig, who participated in the practical training course, for collecting faecal samples. J. Jähnig and J. Harjung are thanked for their help in updating the Access and GIS databases, as well as M. Hoehn and K. Schwerdtner-Máñez for their support in mapping marking sites. Thanks to E. Küster and the Department of Bioanalytical Ecotoxicology and to W. Durka and the Department of Community Ecology for giving access to the laboratories and for helping with several challenges. We kindly acknowledge private landowners for their permission to access the ponds. This work was supported by the Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE). S. Lampa was funded by the German Science Foundation (DFG), grant GR 1960/2-1.

Chapter 6

Discussion

6.1 Overview of Research Outline and Main Results

In the previous chapters, I have first investigated how the combination of microsatellite genotyping and capture-mark-recapture (CMR) methods can be optimised and customised for the use of Eurasian otter faecal samples. To this end, I initially conducted experiments to find a suitable extraction method, to optimise the PCR amplification for otter faeces, and to determine differences in amplification success rates for increasing storage times for three types of faecal samples (Chapter two). The main finding of this exploration was that a pre-amplification approach increased amplification success rates by 11% and reduced genotyping errors by 53%. I further presented a multiplex PCR protocol that is more time- and cost-efficient.

These findings served as a basis for the next step of the work, where I extracted, amplified, and genotyped the samples from the first sampling year (2006). In doing so, I conducted a preservation experiment and compared four methods to minimise genotyping errors, five methods to detect and quantify genotyping errors, and five methods to subsequently estimate population sizes using the generated multilocus genotypes (Chapter three). These comprehensive analyses gave rise to a review discussing pros and cons along each step of non-invasive genetic CMR analyses. Main outcomes were that high genotyping error rates lead to a severely flawed dataset if no consensus genotypes are formed and yield an overestimated population size if remaining genotyping errors are not incorporated into the estimation method.

The guidelines for non-invasive genetic CMR outlined in chapter three were then used for collecting, extracting, genotyping, and analysing samples of five additional sampling years (2007–2012), with one exception: the pre-amplification approach that consisted of two consecutive PCR steps was replaced by a single PCR using a more sensitive polymerase enzyme with high-fidelity and hot-start technique. This modified PCR protocol gained comparable success rates and speeded up the PCR amplification and was applied to all samples of 2007–2012. However, the pre-amplification approach was still used to deal with difficult markers or samples having particularly low DNA quality and quantity. The generated multilocus genotypes of all six years (2006–2012) could then be used for population analyses of the Eurasian otter.

In chapter four, I examined seasonal and sex differences in the marking behaviour and assessed whether different faecal types are deposited in a specific manner. I further estimated population sizes of each sampling year employing misidentification closed population models and calculated yearly sex ratios. Main findings of this chapter were that jelly samples with higher genotyping success rates are more often defecated by males and placed exposed on frequently used marking sites, making them easier to find for collectors. Thus, when non-invasive genetic CMR is applied on otters it is crucial to avoid concentrating only on this kind of samples or only on prominent marking sites, in order to receive unbiased estimation of sex ratios. Furthermore, population size estimators should account for a behavioural effect that is either collector- or otter-induced. Finally, I demonstrated that otter faecal densities cannot be used as an index for abundances.

In chapter five, the same multilocus genotypes were then used to estimate apparent survival and temporary emigration/re-immigration, to test for sex-biased dispersal, and to identify sex differences in spatial use. Apparent survival was about 11% higher for females than for males. Males showed a higher mobility with 5% higher temporary emigration/re-immigration rates, a male-biased dispersal, and larger activity range indices. This resulted in a lower positive correlation of relatedness and proportion of activity range overlaps in males than in females, while opposite-sex dyads showed a negative correlation here. Activity ranges of sub/adults frequently overlapped and with high proportions, which could hint to a condensed spatial organisation in fish pond landscapes.

6.2 Key Findings

Key finding 1: How non-invasive genetic CMR can be successfully applied on otters

Non-invasive genetic CMR methods opened up new possibilities to receive information about animal species that were previously difficult to attain. For the Eurasian otter, this method is only applicable since the development of 13 microsatellites (Dallas & Piertney 1998) and was first tested for the efficacy on faecal samples in 2003 in comparison with tissue samples (Dallas et al. 2003). Since then, nine studies applied non-invasive genetic sampling (Hung et al. 2004; Kalz et al. 2006; Prigioni et al. 2006; Hajkova et al. 2007; Ferrando et al. 2008; Janssens et al. 2008; Lanszki et al. 2008, 2010; Quaglietta et al. 2013) and five applied non-invasive genetic CMR (Arrendal et al. 2007; Bjorklund & Arrendal 2008; Hajkova et al. 2009; Koelewijn et al. 2010; Bonesi et al. 2013) to receive information about Eurasian otters. However, researchers reported considerable genotyping error rates (Hung et al. 2004; Janssens et al. 2008), unusually high population size estimates (Hung et al. 2004; Hajkova et al. 2009), sex ratios that are probably biased (Bonesi et al. 2013), and uninformative estimates due to extremely wide confidence intervals (Arrendal 2007; Bjorklund & Arrendal 2008).

One key problem when using otter faeces is the low template DNA quantity and quality leading to low success rates, genotyping errors, and contamination susceptibility. Low success rates diminish the power of the method; if not enough samples (per day, individual, or area) are successfully genotyped, either analyses are not feasible or results will be uninformative. Genotyping errors and amplified contaminant DNA either produce not existing false (ghost) individuals, or samples appear to mistakenly belong to another already known individual, while the real individual is missed. The former is more likely and lead to overestimated population sizes, underestimated recapture and survival rates (Creel et al. 2003), and potentially biased sex ratios if genders differ in the probability to produce false individuals. A bias in sex ratio can also be introduced if researchers focus the sampling on higher DNA quality jelly samples and on prominent marking sites (e.g. under bridges), because males defecate more often jelly samples that are more often placed on prominent marking sites. Sampled otters might also react with an increased sprainting rate on the disturbance through the sampling, introducing hereby a behavioural response effect. Hence, for the application of non-invasive genetic CMR on otters it is crucial to increase genotyping success rates, to decrease genotyping errors,

to get sufficient knowledge about their marking behaviour, and to compensate individual sprainting variations through an appropriate study design, laboratory procedure, and parameter estimation methods. Here, I offer an otter-specific step-by-step protocol for each required step that includes recommendations compiled from the results of chapters two, three, and four, as well as from the literature:

Table 6.1 Simple step-by-step guide for non-invasive genetic capture-mark-recapture (CMR) analyses on Eurasian otters using faecal samples. Recommendations given derived from own results and from the literature.

Analysis step	Recommendations
Sampling design	Choose a short sampling period to avoid violence of the closure assumption (e.g. five consecutive days)
	Avoid periods with high migration, mortality, or birth rates
	Do not concentrate primarily on jelly samples but collect all kinds of samples (spraints, spraint plus mucus, jelly samples)
	Include all kinds of marking sites, not only prominent ones (e.g. under bridges)
	Check each potential marking site on each sampling day
	Collect as many samples as possible
	Train collectors 1–2 days before sampling to decrease effects (e.g. varying sampling rate) of the settling-in period
Sampling technique	Only take fresh samples
	Decrease potential behavioural response of otters by taking only parts of faeces (not entire sample) with e.g. cotton sticks and avoid modifying the marking site
Preservation	Do not use merely freezing over longer periods of time
	Use either storage buffers or the first lyse buffer of the employed extraction kit and store at –80°C
Extraction	Use silica-based extraction methods to increase success rates (e.g. commercial kits)
	Prevent cross-contamination rigorously
Microsatellite genotyping	Use engineered polymerase enzymes (e.g. hot-start technique)
	Choose only few (< 10), but highly variable and short (< 200–300 bp) markers
	Use low retention plastic tubes in all laboratory steps
	Prevent cross-contamination rigorously
	Generate consensus genotypes via several repetitions (e.g. three for homozygous, two for heterozygous genotypes)
	Discard (very) low-quality samples according to comparably relaxed thresholds (e.g. following the screening approach)
	Calculate or estimate allelic dropout and false allele rates
	Check consensus genotypes for one or two mismatches with other genotypes and verify those genotypes via repetitions
	Check dataset for still existing errors (e.g. using Programme DROPOUT)
	Use pre-amplification approach for difficult markers or samples
Population size estimation	Check on basis of biological information and/or statistically whether the assumption of closure is likely to be met
	Check for equal capture probability and consider biological information to select an appropriate model accounting for variations in catchability
	Take into account that some errors are still undetected
	Prefer the use of an error-incorporating estimation model
	Consider accounting for a behavioural response
	Assess if/how assumptions of the model are violated
	Do not accept the population size estimation uncritically

Key finding 2: Faecal monitoring cannot serve as a simple alternative to non-invasive genetic CMR in order to receive reliable population size estimates.

Most otter monitoring schemes are based on the search for indirect otter signs such as tracks in snow or mud, feeding traces, or faeces. Apart from monitoring that are deliberately conducted in winter during good snow conditions to find tracks, the most commonly sign of an otter are their faeces. The "standard survey method" makes use of these indirect signs that are comparably easy to find (Mason & Macdonald 1987). Here, 600–1000 m of riverbanks and waterways are surveyed for signs to differentiate between a positive or negative site that is interpreted as otter presence or absence in that given site. If repeated at several sites, this approach can give information about the distribution of otters, but it was also suggested to be used as a method to estimate relative abundances (Mason & Macdonald 1987). While applied in several studies, the use of the method for abundance estimates was criticised for its incapacity to account for temporal, spatial, and individual variations in sprainting behaviour (Kruuk et al. 1986; Ruiz-Olmo et al. 2001; Chanin 2003). Non-invasive genetic sampling breathed new life into the debate. A study conducting non-invasive genetic sampling on otters (Lanszki et al. 2008) presented a positive correlation between spraint densities and numbers of otter genotypes per area.

Applying a linear regression between yearly numbers of genotyped samples per ha and yearly numbers of genotypes per ha, I also obtained an almost significant relationship, which could not be found when the number of genotypes per ha were replaced by the numbers of estimated individuals per ha. This discrepancy is caused by false individuals that are still contained in the dataset when using number of genotypes. The more samples are collected and genotyped, the more ghost individuals will be in the dataset increasing the number of genotypes. Thus, a correlation between number of genotypes and spraint densities will be biased by ghost individuals and should not be applied, even if the monitoring is always conducted at the same time of the year. Additionally, I could demonstrate that the presence of American mink (*Neovison vison*) can render otter faecal monitoring useless for obtaining abundance indices. When minks live in sympatry with otters, a significant proportion of apparent otter samples can be of mink origin, even when sampled by experts. When the faecal monitoring is conducted over several days, an otter-induced increased marking intensity or a collector-induced increased sampling rate could further bias results. Finally, I could show that in fish pond systems one marking site can be used by up to six different individuals. Thus, the number of samples at a marking site is not related to the number of individuals.

As a result, I would not recommend using the "standard survey method" or a slightly modified faecal monitoring without genetic sample identification to receive population size estimates.

Key finding 3: Current status of the protected otter population in Upper Lusatia

In my dissertation, I was able to offer the first estimates with confidence intervals for population sizes and apparent survival rates of one of the presumably biggest otter populations in Europe over a

comparably long period of six years. Actual otter numbers are extremely important for fish pond systems, since these areas are strongholds for otters in Europe but are also hot spots for conflicts between humans, namely fishermen and fish farmers, and otters. The otter population in Upper Lusatia is likely to be a source for the spread into adjacent areas and for the connectivity of still fragmented populations in Saxony or neighbouring federal states.

I found relatively high otter densities compared to other habitat types, such as lakes or rivers, although one has to bear in mind that different measurements (per ha area, per ha water area, per km shoreline), methods, and water body shapes hamper comparisons. Two studies on Hungarian (Lanszki et al. 2010) and Czech (Hajkova et al. 2009) fish pond systems that also used non-invasive genetic sampling found slightly higher otter densities. However, I received comparable results when using their estimating approaches.

For Upper Lusatia, Ansorge (1994) reported expert guesses of the early 1990ies that are about half of my density estimates. The survival rate, based on life tables constructed from Upper Lusatian carcases (Ansorge et al. 1997), was also about 11% lower than the mean apparent survival rate estimated in this study. This either indicates an increase in the population size in Upper Lusatia with higher survival rates or suggests that otherwise previous studies have underestimated both. The carcases used by Ansorge et al. (1997) were sampled between the years 1980–1995. Otters in Upper Lusatia are believed to have continuously increased in densities since the 1950ies with a stepper increase from the mid 1980ies on (Klenke et al. 2013). In the 1990ies the otter also started to expand its distribution range in entire Germany (Reuther 2004). Furthermore, the Federal State of Saxony started to grant compensation payments to pond owners in 1995, which might have also raised the acceptance of otters by stakeholders (Myšiak et al. 2013). Thus, it might well be that the population indeed increased in size and in their survival rates. However, life table construction from carcases require rather strict assumptions, such as stationary age distributions (Caughley 1966) or that samples reflect the true structure of the living population, which often is not the case (Hauer et al. 2002a). For this reason, estimates derived from life tables are less reliable and difficult to compare with my data collected between 2006 and 2012.

Despite constantly increased road mortality rates in Upper Lusatia within the past 10–15 years (Zinke 1991, 2000; Hauer et al. 2000, 2002a), the survival rate was moderately high and there is no indication for a decrease in densities for the period 2006–2012. This suggests that the otter population in the Upper Lusatia is in a good condition. Here, international protection statuses and policy instruments of the state Saxony to support environmentally sound pond fisheries and habitat renaturation have certainly contributed to this, as was also shown by Klenke et al. (2013).

Key finding 4: Spatial use of otters in fish pond systems

Activity range sizes of females were roughly half of the size of a male range. This is in line with previous studies conducted on lakes and rivers (Erlinge 1967; Kruuk 2006; Ó Néill et al. 2009) or in

marine systems (Kruuk & Moorhouse 1991; Kruuk 2006). However, overlaps in activity ranges seemed to be different in fish pond systems compared to other water bodies. I detected an equally high proportion of overlaps in activity ranges between opposite-sex and same-sex dyads, regardless of whether male-male or female-female. While overlaps between sexes were also reported by others (Erlinge 1968; Kruuk 2006; Quaglietta et al. 2014), overlaps in same-sex dyads were only mentioned by Kruuk (2006). He differentiated between home ranges and "core areas", though, and found only for the former overlaps between same-sex dyads. The here calculated activity range indices are rather comparable to "core areas" that did not overlap along a coastal habitat. Kruuk (2006) supposed a close relationship between overlapping female-female dyads, but unknown relationships between male-male dyads. For overlapping females, I could show that they were indeed closer related to each other with larger overlaps. Although a similar trend seemed to occur in males, it was not significant. Hence, also unrelated males overlapped by up to 55.5%, whereas unrelated females only overlapped in up to 3% of their activity range. It seems that adult females have smaller but definite activity ranges that are only shared with close relatives, whereas males have much larger but potentially less delimited (or less defended) activity ranges that overlap with female ranges but also with other, even unrelated, male ranges. Since fish pond systems offer high food abundances almost all year long but in a small and limited area, the spatial organisation of otters might be more condensed with smaller activity/home ranges and with higher non-food resource partitioning. If core areas overlap more frequently, encounters will also happen more often. This would foster the recent conclusion of Quaglietta et al. (2014) that the social behaviour of otters is more flexible than previously thought. It could well be that the social behaviour and hence marking behaviour changes depending on otter densities. Higher densities are known to cause increased home range overlaps e.g. in felids (Nielsen & Woolf 2001) or ursids (Dahle & Swenson 2003).

6.3 Limitations and Methodological Constraints

As stated in key finding 1, one main difficulty in non-invasive microsatellite genotyping are genotyping errors. Although there are several approaches that help minimising, detecting, and quantifying genotyping errors (Miller et al. 2002; Frantz et al. 2003; McKelvey & Schwartz 2005), the final consensus genotypes may still harbour errors (Marucco et al. 2011). Therefore, I decided to use the misidentification model from Lukacs and Burnham (2005a) (in the following named L&B estimator) that estimates the probability of a correct sample classification (α) and corrects the population size estimation with this probability. Several authors criticised the estimator and offered alternative estimators that incorporate genotyping uncertainty (Wright et al. 2009; Link et al. 2010; Yoshizaki et al. 2011). However, none of the alternative estimators is yet implemented in a script/software and none of them account for behavioural effects or is applicable to estimate survival and temporary emigration.

Whatever error-incorporating estimator is employed, on can criticise that it may either still overestimate when error rates exceed the capacity of the estimator, or underestimate if there are actual no errors. The latter is rather unlikely, because all four suggested error-incorporating estimators result in reliable estimations when genotyping error rate is close to 0 (Lukacs & Burnham 2005a; Wright et al. 2009; Link et al. 2010; Yoshizaki et al. 2011). One approach to handle the first criticism could be to count the number of individuals that must surely exist, because they were encountered several times (e.g. ≥ 3 samples) and to treat this as a conservative *lower bound* for population size. For the six sampling years this *lower bound* (range = 11–19; mean = 16) was usually within the lower 95% confidence interval of the L&B estimates (Tab. 6.2). For an *upper bound* one could use an estimation received by a conventional population size estimator not correcting for errors, such as closed population models in MARK (White & Burnham 1999) or the R-package CAPWIRE (Pennell et al. 2013) (Tab. 6.2). Except for 2010, also these *upper bounds* were included in the respective 95% confidence interval of the L&B population size estimation (Tab. 6.2).
Although the L&B estimation seemed to be the best compromise and the most reliable estimation method, the true population size and hence the true sex ratio remains unknown.

Table 6.2 Results of population size estimations for each sampling year using different approaches. Genotypes: number of individual genotypes; Counted minimum: number of individuals with at least three samples; L&B estimation: weighted average population size using closed population models with misidentification (Lukacs & Burnham 2005a); MARK estimation: weighted average population size using closed population models without misidentification (Otis et al. 1978); CAPWIRE estimation: population size estimation using the most supported model of the R-package CAPWIRE (Pennell et al. 2013) – the heterogeneity model (TIRM).

Sampling year	Genotypes	Counted minimum	L&B estimation ± SE (CI)	MARK estimation ± SE (CI)	CAPWIRE estimation (CI)
2006	22	16	19 ± 2.7 (13.4–24)	22 ± 1.5 (19.6–25.4)	23 (22–25)
2007	30	17	26 ± 5.4 (15.1–36.3)	33 ± 4.5 (24.5–42.1)	31 (30–32)
2008	22	11	21 ± 4.2 (13.2–29.8)	22 ± 1.1 (20.1–24.3)	23 (22–24)
2010	21	15	15 ± 2.1 (11.2–19.6)	21 ± 0.0 (20.9–21)	21 (21–22)
2011	26	19	25 ± 1.8 (21–28.1)	26 ± 0.0 (25.9–26)	26 (26–27)
2012	24	17	20 ± 2.1 (16.3–24.5)	24 ± 0.9 (22.5–25.8)	24 (24–25)
Mean	**24**	**16**	**21**	**25**	**25**

In those conducted analyses where individual identification was important (sex differences in marking behaviour, dispersal, and spatial use), genotyping errors, mainly ghost individuals, could have biased results. Therefore, I tried to reduce the risk by including individuals with at least three samples (analyses on spatial use) and by assuring that the number of single samples – potential ghost individuals – is evenly distributed among sexes. Nevertheless, analyses can still be influenced by genotyping errors.
Regarding activity range indices, I deliberately avoided to use the name *home range (HR)* because for reliable HR estimates it would have been required to a) sample for a longer period of time than only five days, b) include only one sample per night reducing the risk for serial autocorrelation, c) use a

method other than the very controversial minimum convex polygon approach (e.g. kernel density estimation (Worton 1989)) , or to d) correct HRs for pond sizes since the size and shape of a HR is dependent on the pond geometry. However, for the specific purpose of testing for sex differences in size and extent of overlap, I deemed my approach reliable.

For further improvement of the analyses, it would have been interesting to include data on fish densities of each pond/pond area of each year. With these data I could have tested whether population size, temporary migration, marking behaviour, or spatial use are affected by changes in fish density and if so how. Theoretically, every fish pond owner knows the quantity of fishes that were inserted (usually in October to wintering ponds and in April to summering ponds) and the quantity that were harvested (usually in October from summering ponds and in April from wintering ponds). However, it is rather difficult to convince them to offer this information and would have taken too long for the purpose of this dissertation, if I would have succeeded at all. Secondly, in each year several ponds contained only wild fish that immigrated through the connected ditches and streams. For these, it would have been rather difficult to assign a respective fish density. Furthermore, not all ponds were regularly drained and harvested, meaning at the time of sampling I would have merely known what quantity of fish was inserted several months ago, but not the actual fish density. Hence, such analyses would have been possible only for those ponds that were just stocked with fish or close to harvest.

6.4 Suggestions for Further Research

This work provides the first series of abundance estimations over a comparably long period of six years for a population of high conservation interest. Since the Upper Lusatian population is the main source population for the expansion in Saxony and has high conflict potentials with humans, further monitoring is recommended. Future monitoring may not only focus on determining census population sizes, but may also estimate additional parameters, such as effective population sizes (N_e; Wright 1931) or population growth rates (λ; Pradel 1996) to offer valuable long-term information on status and trend. N_e is an adequate indicator for the viability of a population (Kirk & Freeland 2011) because it decreases with increasing loss of genetic diversity through genetic drift or with increasing inbreeding (Wang 2005). Populations with low N_e, and hence decreased genetic variability and heterozygosity, are less able to adapt to environmental changes (Kirk & Freeland 2011). A reliable estimation of N_e would require more and higher polymorphic microsatellite loci than used in this study (e.g. ≥ 10), plus several other prerequisites, such as age-specific survival or number of adults in the population or samplings of several generations, dependent on the applied N_e-estimator (see Luikart et al. 2010 for a review). While N_e could help in evaluating any negative genetic effect on the population, population growth rate could help in determining the actual development of a population (increasing, decreasing, equilibrium). The parameter can be estimated using Programme MARK, also in combination with the robust design, but it is not yet possible to include the misidentification model here. Genotyping errors may also bias the estimation of N_e. Hence, if non-invasive genetic sampling is

used to obtain these parameters, researchers have to cope with genotyping errors and either find strategies how to deal with potential ghost individuals in the dataset (e.g. by using only those individuals with a certain number of samples) or develop estimators that can incorporate genotyping errors. Generally, such error-accounting estimators for population parameters are desirable for studies investigating elusive and threatened species through non-invasive genetic sampling and should receive special attention in the future.

Future monitoring of the here studied population are also worthwhile because of the recently increased mink density. It would be expedient to estimate the mink's population size, to monitor their population trend, and to investigate whether minks are affecting e.g. the marking behaviour or spatial use of otters. Such information is important to better understand the sympatric occurrence of native otters and invasive minks and to understand which niche separations are developed to avoid or reduce interspecific competition.

Although I contributed to a better understanding of movement patterns and spatial use of otters in fish pond systems, it would be important and desirable to directly compare various otter habitats within the same landscape and within one continuous otter population. Such a direct comparison would be possible in the Upper Lusatia where some areas are characterised by lakes (mostly former opencast mines) and rivers instead of fish ponds. In this context, future research should also investigate whether otters adjust home range sizes and overlaps in dependence on population densities.

Finally, since the otter is recolonising its former haunts in Germany, it is crucial for management purposes to know when it will arrive in which part of the country. Here, non-invasive genetic CMR analyses could be conducted either at the border of the expansion range or by sequentially shifting the monitoring area from the source population to the expansion border to analyse and understand the spatially and temporally complex recolonisation process. Using thereby derived information, the future expansion range and routes could be predicted by the use of individual-based spatially explicit models, such as the ones illustrated in Wiegand et al. (2004) or Bocedi et al. (2014). Individual-based models simulate life-history events of an individual; the sum of all individual life-histories represents population dynamics (Grimm & Railsback 2005). The combination with spatially explicit models means that the individual is associated with a location in geometrical space (Dunning et al. 1995). Hence, individual-based spatially explicit models enable to track individuals over space and time mimicking individual behaviour. The otter-specific population dynamic parameters I have estimated can be used to parameterise such an individual-based spatially explicit model to relate otter demographics explicitly to the landscape in which it lives and to predict the extent of future expansion and possible expansion routes.

6.5 Suggestion for Otter Conservation

Until now, the Federal State of Saxony is the only state in Germany using damage compensation schemes to reimburse fish farmers for their economic loss (Klenke et al. 2013). Since 1995, Saxony

gives financial support to pond owners for "pond maintenance", for extra stocked fish that is considered as food for otters, for compensation in cases of hardship (if the actual otter damage exceeds €1000 per ha per year), and for measures that avoid otter damages on fish, such as otter fences (Schwerdtner & Gruber 2007; Klenke et al. 2013). These compensation payments appeared to have fulfilled their purposes in raising the acceptance of the conflict-laden species, because the studied population in Upper Lusatia seems to be in a good condition and might have even increased in size from the 1990ies on (see key finding 3). Such reimbursement payments could also be an important tool for those German federal states where the otter arrived within the last years, increasingly causing conflicts with the locals, such as in Bavaria (Sachteleben et al. 2010; Sage 2012). Here, fishermen and anglers are not used to otters and there is no contact person in case of damage and no management plan for the species yet, leading to illegal hunting and killing in some areas (Bayerl, pers. comm.; Sage 2012). Here, a reimbursement scheme could promote the species' acceptance and help establishing a sustainable population which may spread further in the federal state.

Although the Saxon compensation schemes appear to serve the purpose, there is the interest that this public money is invested in a cost-effective way and there are some indications for overcompensations (cf. Klenke et al. 2013). Furthermore, the compensations in cases of hardship are only paid if fisheries prove the occurrence of otters by observations, tracks, or fish remains (Schwerdtner & Gruber 2007), but the latter two could nowadays also derive from minks. This would co-finance a further spread of the invasive species. For an optimisation of the payment scheme, spatially differentiated population sizes are an indispensable prerequisite, as was stated by Klenke et al. (2013). With the results of this study population size estimates are now available, at least for a part of Upper Lusatia.

Regarding the current spread and increase of the mink populations, it is advisable to further monitor the process of co-existence between the two species, to investigate in population sizes of the mink, and to observe whether the mink has any impact on one of the most dense otter population in Germany. Especially the development of reliable methods to distinguish between the occurrence and damage caused by minks and otters are needed to effectively conserve otters in their core area.

As mentioned already above, it seems that the population in Upper Lusatia serves as a source population for the recolonisation of adjacent areas. Even if Upper Lusatian otters probably do not require further conservation concern, it is different for individuals at the border of the expansion range. I demonstrated that male survival rates were lower than those of females, probably because of higher male mobility through male-biased temporary emigration and dispersal and larger territories. Since expanding otter populations have an increased road-kill risk (Chanin 2003), which is often biased towards males (Hauer et al. 2002a; Koelewijn et al. 2010), migrating individuals would benefit from direct protection measurements, such as otter-friendly bridges with ledges, otter tunnels under roads, otter-proof fences for guidance along roads, or signposts along roads that dissect water bodies.

*   *   *

Summary

After decades of population decline and range contraction, the Eurasian otter (*Lutra lutra*) is now expanding in Germany and other countries in Europe. For sound conservation strategies it is crucial to determine population demographic parameters, such as population size, sex ratios, survival, and migration rates. However, the estimation of such parameters for an elusive species is challenging and knowledge here is still scarce with regard to otters, especially in fish pond landscapes. Landscapes dominated by fish ponds are strongholds for otters in Europe but are also hot spots for conflicts between fish farmers and otters.

On that account, the goal of this dissertation was to estimate population sizes for a population of high conservation interest, to examine whether and how non-invasive genetic capture-mark-recapture (CMR) using otter faeces can be successfully applied, and to gain knowledge about population dynamics and behaviour on this threatened species in a fish pond system. The present study was conducted in the Upper Lusatian heath and pond landscape in Eastern Saxony, Germany. This region is characterised by about 5000 ha pond that are used for fish farming. Here, otters never got extinct and occur in relatively high densities. In a study area including 64 ponds (505 ha), overall 2132 samples were collected over a period of six sampling years (2006–2008; 2010–2012), each consisting of five consecutive sampling days. Samples were extracted and amplified at seven microsatellite loci and two sex markers to generate multilocus genotypes for individual identification.

Using these samples, I first optimised the sampling, preservation, extraction, and amplification of otter faecal DNA that is known for low success rates and high genotyping error rates. These methodological optimisations are not only useful for otters but can generally be applied to studies employing non-invasive genetic sampling and can help deriving more reliable microsatellite genotypes.

There are recent studies arguing that impacts of genotyping errors on population demographic analyses can be reduced to an acceptable level through accurate laboratory procedure and data quality controls or that genotyping errors can even be neglected. In contrast, I demonstrated that even a rigorous lab procedure with more PCR repetitions than usual and subsequent error checks may not completely eliminate errors with certainty. Remaining errors will lead to severe biases in subsequent analyses, especially in population size estimations. Here, I illustrated that it is advisable to use error-incorporating estimation models. Employing the misidentification model implemented in Programme MARK, I estimated the population size for the first sampling year and compared the estimator with other conventional and error-incorporating methods. In doing so, I could provide a step-by-step protocol for non-invasive genetic CMR studies to achieve reliable estimates of population sizes in the presence of high genotyping error rates.

By using this step-by-step protocol for the remaining five sampling years, I demonstrated that non-invasive genetic CMR can successfully be applied on otters when their marking behaviour is taken into account, namely the male bias in defecating high DNA-quality jelly samples, preferentially on

prominent marking sites, and a potentially increased marking intensity of already collected individuals. I illustrated that contrary to the non-invasive genetic CMR approach, faecal densities cannot be used as an index of otter abundances, as suggested in the literature. This is mainly due to seasonal and individual differences in marking behaviour, but also due to invasive American minks (*Neovison vison*), whose faeces can be easily confused with otter spraints.

The abovementioned misidentification model was also used to determine apparent survival and temporary migration rates in a robust design approach. Apparent survival was 11% higher for females than for males. One reason might be the detected higher male mobility with 5% higher temporary emigration/re-immigration, higher dispersal rate, and larger activity range indices compared to females. I found a high proportion of activity range overlaps that were negatively correlated with relatedness for female-male dyads, which reduces inbreeding risk, and positively correlated for same-sex dyads, with a higher correlation among females. Such high proportions of overlap in activity ranges were not reported from other habitats and could hint to a different spatial use of otters in fish pond systems or to density-dependent changes of activity ranges and overlaps, since I found comparably high otter densities. Constantly high densities from 2006 till 2012 and moderately high apparent survival rates indicate a thriving population in Upper Lusatia that probably benefited from the damage compensation schemes of the Federal State of Saxony. These compensation schemes reimburse fish farmers for otter damages.

In summary, this research provided a step-by-step protocol for non-invasive genetic CMR studies with high genotyping error rates, guidelines of how to successfully apply this protocol on otters, and contributed to an improved understanding and increased knowledge of movement patterns, spatial use, marking behaviour, and demographic parameters for use in otter conservation practice.

Zusammenfassung

Nachdem die Bestände des Eurasischen Fischotters (*Lutra lutra)* in Europa über Jahrzehnte zurückgegangen sind und sich seine Verbreitungsareale stark verkleinert haben, erholen sich die Bestände in jüngster Zeit und breiten sich in Deutschland und anderen europäischen Ländern wieder aus. Für ein nachhaltiges Management und den Schutz der semi-aquatisch lebenden Art werden vertrauenswürdige Schätzungen populationsdynamischer Parameter wie z.B. Populationsgröße, Geschlechterverhältnisse, Überlebens- oder Migrationsraten benötigt. Eine verlässliche Schätzung dieser Parameter ist jedoch schwierig, vor allem für schwer erfassbare Arten wie den Fischotter. Deshalb ist unser Wissen über den Otter dahingehend sehr limitiert, besonders in Gebieten mit vielen Fischteichen. Landschaften, die von Fischteichen geprägt sind, stellen besonders wichtige Refugien für den Otter dar, häufig verbunden mit hohen Populationsdichten. Zeitgleich ist das Konfliktpotential zwischen Mensch und Fischotter in solchen Gebieten besonders hoch.

Aufgrund dessen war es das Ziel dieser Dissertation, die sogenannte nicht-invasive genetische Fang-Wiederfang-Methode erfolgreich auf den Otter anzuwenden und für diesen zu optimieren, um damit die Populationsgröße und andere Parameter, sowie das Markierungsverhalten einer, an Fischteichen lebenden, Otterpopulation zu bestimmen. Die Untersuchungen wurden in der Oberlausitzer Heide- und Teichlandschaft im Osten von Sachsen (Deutschland) durchgeführt. Diese Region ist geprägt von insgesamt ca. 5000 ha Fischteichen, wurde fortwährend von Ottern besiedelt und ist bekannt für seine hohen Fischotterdichten. Innerhalb von sechs Untersuchungsjahren (2006–2008; 2010–2012), die jeweils aus fünf aufeinanderfolgenden Sammeltagen bestanden, wurden 2132 Kotproben in einem Gebiet mit 64 Teichen (505 ha) gesammelt. Die DNA der Kotproben wurde isoliert und mit sieben Mikrosatellitenmarkern und zwei Geschlechtsmarkern amplifiziert, um für jede Probe einen individuellen Genotyp zu erhalten.

Mit Hilfe dieser Proben wurden zuerst die methodischen Schritte, d.h. Probennahme und -konservierung, sowie Extraktion und Amplifizierung der DNA für Otterkotproben optimiert. Otterkotproben sind bekannt für geringe Erfolgsraten und hohe Genotypisierungsfehlerraten. Die Optimierrungen können jedoch nicht nur auf den Otter angewandt werden, sondern helfen auch bei anderen Arten höhere Erfolgsraten und verlässlichere Genotypen zu erhalten.

In einigen Studien wird behauptet, dass die negativen Auswirkungen von Genotypisierungsfehlern durch ein akkurates Arbeiten im Labor zu einem akzeptablen Maß reduziert werden könnten oder man die Fehler sogar ganz ignorieren könnte. In meiner Arbeit konnte ich allerdings aufzeigen, dass auch ein rigoroses Laborprotokoll mit einer hohen Anzahl von PCR-Wiederholungen und anschließenden Fehlerkontrollen nicht alle Genotypisierungsfehler beseitigen kann und selbst diese wenigen Fehler zu schwerwiegenden Verzerrungen der Ergebnisse führen, besonders bei Populationsgrößenschätzungen. Deshalb ist es ratsam Methoden zu verwenden, die Genotypisierungsfehler in die Schätzungen mit einbeziehen. Eine dieser Methoden ist das Fehlbestimmungs-Modell im Computerprogramm MARK, mit

dem ich die Populationsgröße des ersten Sammeljahres bestimmt habe und mit Ergebnissen anderer konventioneller und fehlereinbeziehender Schätzer verglichen habe. Dadurch war es möglich eine Schritt-für-Schritt-Anleitung für die nicht-invasive genetische Fang-Wiederfang-Methode zu erstellen, die bei hohen Genotypisierungsfehlerraten eine zuverlässige Populationsgrößenschätzung ermöglicht.
Diese Anleitung wurde im Anschluss für die verbleibenden fünf Sammeljahre verwandt, wodurch gezeigt werden konnte, dass sie erfolgreich auf den Otter angewendet werden kann, sofern man sein Markierungsverhalten berücksichtigt – nämlich, dass qualitativ höherwertige Jelly-Proben mehr von Männchen und in hochfrequentierten Markierungsstellen abgelegt werden und, dass möglicherweise bereits beprobte Tiere mit einer höheren Markierungsintensität auf das Sammeln ihres Kots antworten.
Im Gegensatz zur non-invasiven genetischen Fang-Wiederfang-Methode kann die Dichte der Kotproben in einem bestimmten Gebiet nicht dazu verwendet werden, um Aussagen über die Populationsgröße zu treffen. Dies steht im Gegensatz zu einigen Studien, die diese Methode für Abundanzschätzungen nutzten. Saisonale und individuelle Markierungsunterschiede, aber auch die Anwesenheit des invasiven Amerikanischen Minks (*Neovison vison*), dessen Kot leicht mit Otterkot zu verwechseln ist, machen diese Methode für Abundanzschätzungen unbrauchbar.
Das oben genannte Fehlbestimmungs-Modell wurde auch verwendet, um Überlebensraten und temporäre Migrationsraten mit Hilfe des *robust design* zu bestimmen. Die Überlebensrate der Weibchen war 11% höher als die der Männchen. Ein Grund dafür könnte die höhere männliche Mobilität sein, die sich durch eine um 5% größere temporäre Migrationsrate, erhöhte Dispersionsraten und größere Aktivitätsräume der Männchen äußerte. Überlappungen zwischen den Aktivitätsräumen waren häufig und prozentual großflächig. Dabei zeigte sich eine negative Korrelation zwischen der Überlappungsgröße (prozentual zur Gesamtgröße) und dem Verwandtschaftsgrad bei gegengeschlechtlichen Paaren, was wahrscheinlich der Reduzierung des Inzuchtrisikos dient, während bei gleichgeschlechtlichen Paaren eine positive Korrelation gefunden wurde. Diese war für Weibchen wesentlich höher als für Männchen. Solch häufige Überlappungen der Aktivitätsbereiche wurden von keinem anderen Habitat berichtet und könnten auf eine andersartige räumliche Nutzung der Fischteiche hinweisen oder aber auf eine dichteabhängige Raumnutzung, was bei den geschätzten hohen Dichten durchaus plausibel wäre. Die von 2006 bis 2012 relativ gleichbleibend hohen Otterdichten und die mittelhohe Überlebensrate deuten auf eine vitale Population in der Oberlausitz hin, die wahrscheinlich von den Schadensersatzregelungen des Freistaates Sachsen profitiert. Dieses Kompensationsprogramm erstattet den Fischzüchtern durch den Otter entstandene ökonomische Schäden.
Zusammengefasst wurde im Rahmen dieser Dissertation eine schrittweise Anleitung für nicht-invasive genetische Fang-Wiederfang-Studien mit hohen Fehlerraten erstellt und Empfehlungen gegeben, wie diese Anleitung erfolgreich auf den Otter angewendet werden kann. Zudem wurde ein Beitrag zum besseren Verständnis des Markierungsverhaltens, der Bewegungsmuster und Raumnutzung des Fischotters geleistet und wichtige demografische Parameter bestimmt, die für den Schutz des Fischotters unerlässlich sind.

Acknowledgements

This dissertation would not have been possible without the support, patience, and guidance of the following people to whom I owe enormous thanks (please forgive me if I forget anyone):

I would like to express my deepest gratitude to my advisors Prof. Dr. Klaus Henle, Prof. Dr. Bernd Gruber, and Prof. Dr. Stefan Halle. K. Henle always supported me through all those years, regardless in which direction I went. He gave me not only financial support, which are not small potatoes when working with genetics, but always had time for me, helped in sample collection, was the best manuscript reviewer, both being critically, diligent but also tried to keep my writing style. He gave me the opportunity to develop my own individuality and self-sufficiency by allowing me to work with a great independence. B. Gruber has brought me into this topic and inducted me into the world of population ecology and statistics. Thanks to him I had the very comfortable situation of being financed by the DFG. He supported me a lot, especially in the early stage, helped me with the field work and became a friend to me. Discussions with him were always very fruitful and extremely motivating. S. Halle warmly welcomed me in his research group, was always accessible, even on short notice, and still supported me when I decided to change my PhD topic. I greatly appreciate his informal and friendly way to supervise me and his help in pushing forward the project application.

I would like to extend my appreciation to Prof. Dr. Günter Köhler and Prof. Dr. Niko Balkenhol who kindly agreed to review my thesis. I also sincerely thank Dr. Reinhard Klenke who connected me with otter experts, the biosphere reserve "Oberlausitzer Heide- und Teichlandschaft", and the "Naturschutzstation Neschwitz". He showed me how to find otter spraints, helped in sample collection and took over the project lead when B. Gruber left for Australia. Many thanks also to Dr. Marion Hoehn who showed me how to work in a genetic laboratory and helped in mapping marking sites. Likewise, I am very grateful to Dr. Jean-Baptiste Mihoub who was of great help in the last stage of the project and spent afternoons discussing statistical analyses and how to structure the last two papers, helping me out of the fix that way.

I wish to express my appreciation to Dr. Walter Durka, Dr. Eberhard Küster, and Dr. Sabine Kleinsteuber for their confidence to give me free access to the laboratories and helping with several challenges. Similarly, I sincerely thank Ute Lohse for her continual help and willingness to always find a timeframe in the stuffed sequencer schedule to run my plates. Helmut Bayerl is particularly thanked for his great assistance in collecting samples and helping with DNA extraction in the first sampling year. I am greatly thankful to Julia Schmalzl who helped me with DNA extraction and sex identification of numerous samples and made the time in the lab more pleasant and cheerful. I am also very grateful to all field helpers who collected otter scats with me or mapped marking sites, namely Alica Tschierschke, Bianca Bauch, Dr. Michael Gerisch, Daniela Dick, Dr. Wendy Dimond, Emmelianna Kumar, Ulf Hempel, Anett Valentin, Dr. Kathleen Schwerdtner-Máñez, and numerous students from the University of Leipzig. Many thanks to the "Naturschutzstation Neschwitz" for

hosting us in such a nice way, to the biosphere reserve “Oberlausitzer Heide- und Teichlandschaft” for joining the project in the first sampling year, to the pond owners for granting access to the ponds, to Julia Jähnig and Julia Harjung for updating databases, to Prof. Dr. Hermann Ansorge and Diana Jeschke from the Natural History Museum Görlitz and to Olaf Zinke from the Natural History Museum Kamenz for sending numerous otter tissue samples of dead otters found in my study area, to Alexander Lehmann for providing his annual hunting bags on minks, to Dr. Janine Wright for running my data with her model, to members of the Institute of Applied Ecology, University of Canberra, for proofreading the first manuscript, to several anonymous reviewers and the editors A.R. Hoelzel, B. Goossens, and E. Latch for greatly improving the first and second manuscripts.

The dissertation was written at the Helmholtz Centre for Environmental Research (UFZ) and I truly appreciated the working environment and atmosphere at UFZ. I was proud to be a member of this “large family”, had by now the best time of my life, learned so much, and got to know many friendly and cooperative people, also because I had the great opportunity to work within several departments. Most of my best friends I found during my time at UFZ and I am deeply indebted to all of them for their encouragements and friendship. To one of them, Prof. Dr. Andrea Carminati, I owe special thanks for acting as my official confidant. I also would like to thank the members of the Department of Community Ecology, Department of Environmental Microbiology, all past and current members of the Department of Bioanalytical Ecotoxicology and last, but certainly not least, I owe special thanks to my home department – Department of Conservation Biology:

Hier gebührt besonderer Dank meinen „besseren Bürohälften“ Annegret, Elke und Henning für die Hilfe bei vielen Fragen, für das Aufbauen in schwierigen, stressigen Zeiten, für das Gegenlesen von Textteilen und für viele lustige Momente. Auch bei NaNaWu und meinen Kollegen Christiane, Micha, Franzi, Anja, Guy, Lisa und JB möchte ich mich herzlich bedanken, dass ich sie immer wieder mit Fragen bombardieren konnte und sie Teile der Dissertation überprüften. Darüber hinaus bedanke ich mich bei allen derzeitigen und ehemaligen Mitgliedern des Departments Naturschutzforschung für die schöne Zeit, das kollegiale Verhältnis, die Unterstützung, die sehr unterhaltsamen Mittags- und Kaffeerunden, die Weihnachtsfeiern und Ausflüge, das gemeinsame Basteln von Doktorhüten und Abschiedsgeschenken, gemeinsam durchlebte Workshops und vieles mehr und für das Ertragen meines lauten Lachens über den Gang. Bei Marlies Uhlig möchte ich mich besonders dafür bedanken, dass sie sich immer wieder um mich kümmerte, ein Auge darauf hatte, dass es vertraglich weitergeht und sie mir bei der Prüfung der Literatur half.

Meinen Eltern bin ich ewig dankbar, dass sie mir beigebracht haben mich durchzubeißen und niemals aufzugeben, mir die Liebe zu Natur mitgaben, mich immer unterstützten und an mich glaubten und mir die Freiheit gaben meiner Passion zu folgen.

Meinem Mann, Sebastian, danke ich zutiefst für die emotionale Unterstützung, die Aufmunterungen, seine Geduld vor allem in den stressigen Zeiten, in denen er mir den Rücken freihielt und sich um alles kümmerte und dass er immer für mich da war.

References

Adams JR, Kelly BT, Waits LP (2003) Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). *Molecular Ecology* **12**, 2175–2186.

Adams JR, Waits LP (2007) An efficient method for screening faecal DNA genotypes and detecting new individuals and hybrids in the red wolf (*Canis rufus*) experimental population area. *Conservation Genetics* **8**, 123–131.

Almeida D, Copp GH, Masson L*, et al.* (2012) Changes in the diet of a recovering Eurasian otter population between the 1970s and 2010. *Aquatic Conservation-Marine and Freshwater Ecosystems* **22**, 26–35.

Amos B, Pemberton J (1992) DNA fingerprinting in non-human populations. *Curr Opin Genet Dev* **2**, 857–860.

Ansorge H (1994) Zur Situation des eurasischen Fischotters *Lutra lutra* Linné, 1758 im Raum Oberlausitz-Sachsen. *Säugetierkundliche Informationen* **3**, 617–622.

Ansorge H, Schipke R, Zinke O (1996) Altersstruktur und Reproduktion in der Oberlausitz. In: *Artenschutzprogramm – Fischotter in Sachsen* (ed. Geologie SLfUu), pp. 27–30. Materialien zu Naturschutz und Landschaftspflege, Radebeul.

Ansorge H, Schipke R, Zinke O (1997) Population structure of the otter, *Lutra lutra*. Parameters and model for a Central European region. *Zeitschrift für Säugetierkunde – Mammalian Biology* **62**, 142–151.

Arandjelovic M, Guschanski K, Schubert G*, et al.* (2009) Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. *Molecular Ecology Resources* **9**, 28–36.

Arandjelovic M, Head J, Rabanal LI*, et al.* (2011) Non-Invasive Genetic Monitoring of Wild Central Chimpanzees. *Plos One* **6**.

Arrendal J (2007) Conservation genetics of the Eurasian otter in Sweden. *Acta Universitatis Upsaliensis* **278**, 60 pp.

Arrendal J, Vila C, Bjorklund M (2007) Reliability of noninvasive genetic census of otters compared to field censuses. *Conservation Genetics* **8**, 1097–1107.

Balestrieri A, Remonti L, Prigioni C (2011) Detectability of the Eurasian otter by standard surveys: an approach using marking intensity to estimate false negative rates. *Naturwissenschaften* **98**, 23–31.

Banks SC, Piggott MP, Hansen BD, Robinson NA, Taylor AC (2002) Wombat coprogenetics: enumerating a common wombat population by microsatellite analysis of faecal DNA. *Australian Journal of Zoology* **50**, 193–204.

Bayes MK, Smith KL, Alberts SC, Altmann J, Bruford MW (2000) Testing the reliability of microsatellite typing from faecel DNA in the savannah baboon. *Conservation Genetics* **1**, 173–176.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources* **9**, 1279–1301.

Beja PR (1996) Seasonal breeding and food resources of otters, *Lutra lutra* (Carnivora, Mustelidae), in south-west Portugal: A comparison between coastal and inland habitats. *Mammalia* **60**, 27–34.

Bellemain E, Swenson JE, Tallmon O, Brunberg S, Taberlet P (2005) Estimating population size of elusive animals with DNA from hunter-collected feces: Four methods for brown bears. *Conservation Biology* **19**, 150–161.

Bellemain E, Taberlet P (2004) Improved noninvasive genotyping method: application to brown bear (*Ursus arctos*) faeces. *Molecular Ecology* **4**, 519–522.

Berry O, Sarre SD, Farrington L, Aitken N (2007) Faecal DNA detection of invasive species: the case of feral foxes in Tasmania. *Wildlife Research* **34**, 1–7.

BfN (2012) *Fischotter (Lutra lutra) – Verbreitung des Fischotters inklusive Hinweise auf Schwerpunktvorkommen.* http://www.ffh-anhang4.bfn.de/

Bhagavatula J, Singh L (2006) Genotyping faecal samples of Bengal tiger *Panthera tigris tigris* for population estimation: A pilot study. *Bmc Genetics* **7**.

Bjorklund M (2005) A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes* **5**, 676–679.

Bjorklund M, Arrendal J (2008) Demo-genetic analysis of a recovering population of otters in Central Sweden. *Animal Conservation* **11**, 529–534.

Blejwas KM, Williams CL, Shin GT, McCullough DR, Jaeger MM (2006) Salivary DNA Evidence Convicts Breeding Male Coyotes of Killing Sheep. *Journal Of Wildlife Management* **70**, 1087–1093.

Bocedi G, Palmer SCF, Pe'er G*, et al.* (2014) RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods in Ecology and Evolution* **5**, 388–396.

Böhnert W, Buchwald RG, Reichoff L (1996) *Analyse und Bewertung des Landschaftsraumes – Fauna* Biosphärenreservat Oberlausitzer Heide- und Teichlandschaft, Mücka.

Bonesi L, Chanin P, Macdonald DW (2004) Competition between Eurasian otter *Lutra lutra* and American mink *Mustela vison* probed by niche shift. *Oikos* **106**, 19–26.

Bonesi L, Hale M, Macdonald DW (2013) Lessons from the use of non-invasive genetic sampling as a way to estimate Eurasian otter population size and sex ratio. *Acta Theriologica* **58**, 157–168.

Bonesi L, Macdonald DW (2004a) Differential habitat use promotes sustainable coexistence between the specialist otter and the generalist mink. *Oikos* **106**, 509–519.

Bonesi L, Macdonald DW (2004b) Impact of released Eurasian otters on a population of American mink: a test using an experimental approach. *Oikos* **106**, 9–18.

Bonesi L, Palazon S (2007) The American mink in Europe: Status, impacts, and control. *Biological Conservation* **134**, 470–483.

Bonesi L, Strachan R, Macdonald DW (2006) Why are there fewer signs of mink in England? Considering multiple hypotheses. *Biological Conservation* **130**, 268–277.

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

Boulanger J, Proctor M, Himmer S*, et al.* (2006) An empirical test of DNA mark-recapture sampling strategies for grizzly bears. *Ursus* **17**, 149–158.

Boulanger J, White GC, McLellan BN*, et al.* (2002) A meta-analysis of grizzly bear DNA mark-recapture projects in British Columbia, Canada. *Ursus* **13**, 137–152.

Brinkman TJ, Person DK, Chapin FS, Smith W, Hundertmark KJ (2011) Estimating Abundance of Sitka Black-Tailed Deer Using DNA From Fecal Pellets. *Journal Of Wildlife Management* **75**, 232–242.

Brinkman TJ, Person DK, Schwartz MK*, et al.* (2010a) Individual identification of Sitka black-tailed deer (*Odocoileus hemionus sitkensis*) using DNA from fecal pellets. *Conservation Genetics Resources* **2**, 115–118.

Brinkman TJ, Schwartz MK, Person DK, Pilgrim KL, Hundertmark KJ (2010b) Effects of time and rainfall on PCR success using DNA extracted from deer fecal pellets. *Conservation Genetics* **11**, 1547–1552.

Broquet T, Ménard N, Petit E (2007) Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. *Conservation Genetics* **8**, 249–260.

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

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* **20**, 1004.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Current Opinion in Genetics & Development* **3**, 939–943.

Brzezinski M, Romanowski J (2006) Experiments on sprainting activity of otters *(Lutra lutra*) in the Bieszczady Mountains, southeastern Poland. *Mammalia* **70**, 58–63.

Buchan JC, Archie EA, Van Horn RC, Moss CJ, Alberts SC (2005) Locus effects and sources of error in noninvasive genotyping. *Molecular Ecology Notes* **5**, 680–683.

Bueno F (1996) Competition between American mink *Mustela vison* and otter L*utra lutra* during winter. *Acta Theriologica* **41**, 149–154.

Burnham KP, Anderson DR (1998) *Model selection and inference: A practical information theoretic approach* Springer-Verlag, New York.

Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: A practical information-theoretic approach*, 2nd ed. edn. Springer-Verlag, New York.

Calzada J, Delibes-Mateos M, Clavero M, Delibes M (2010) If drink coffee at the coffee-shop is the answer, what is the question? Some comments on the use of the sprainting index to monitor otters. *Ecological Indicators* **10**, 560–561.

Carss DN, Parkinson SG (1996) Errors associated with otter *Lutra lutra* faecal analysis .1. Assessing general diet from spraints. *Journal Of Zoology* **238**, 301–317.

Caughley G (1966) Mortality Patterns in Mammals. *Ecology* **47**, 906–918.

Chang ZF, Liu ZJ, Yang JY, Li M, Vigilant L (2012) Noninvasive genetic assessment of the population trend and sex ratio of the Shennongjia population of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*). *Chinese Science Bulletin* **57**, 1135–1141.

Chanin P (2003) Monitoring the otter *Lutra lutra*. . In: *Conserving Natura 2000 Rivers Monitoring Series* (ed. Skinner A). English Nature, Petersborough, UK.

Chao A, Huggins RM (2005) Modern closed-population capture-recapture models. In: *Handbook of capture-recapture analysis* (eds. Amstrup SC, McDonald TL, Manly BFJ), pp. 58–87. Princeton University Press, Princeton, New Jersey, U.S.A.

Chao A, Yang H-C (2003) Program CARE-2 (for capture-recapture part.2). In: *Program and user's guide*. http://chao.stat.nthu.edu.tw/.

Chao A, Yip PSF, Lee SM, Chu WT (2001) Population size estimation based on estimating functions for closed capture-recapture models. *Journal of Statistical Planning and Inference* **92**, 213–232.

Clavero M, Prenda J, Delibes M (2003) Trophic diversity of the otter (*Lutra lutra* L.) in temperate and Mediterranean freshwater habitats. *Journal of Biogeography* **30**, 761–769.

Conroy JWH, French DD (1991) *Seasonal patterns in the sprainting behaviour of otters (Lutra lutra L.) in Shetland*, Hankensbüttel.

Coxon K, Chanin P, Dallas J, Sykes T (1999) The use of DNA fingerprinting to study the population dynamics of otters (*Lutra lutra*) in Southern Britain: A feasibility study, Technical Report W202. In: *Research and Development*. Environmental Agency.

Creel S, Spong G, Sands JL*, et al.* (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology* **12**, 2003–2009.

Cullingham CI, Curteanu M, Ball MC, Manseau M (2010) Feasibility and Recommendations for Swift Fox Fecal DNA Profiling. *Journal Of Wildlife Management* **74**, 849–859.

Dahle B, Swenson JE (2003) Home ranges in adult Scandinavian brown bears (*Ursus arctos*): effect of mass, sex, reproductive category, population density and habitat type. *Journal Of Zoology* **260**, 329–335.

Dallas JF, Bacon PJ, Carss DN*, et al.* (1999) Genetic diversity in the Eurasian otter, *Lutra lutra*, in Scotland. Evidence from microsatellite polymorphism. *Biological Journal of the Linnean Society* **68**, 73–86.

Dallas JF, Carss DN, Marshall F*, et al.* (2000) Sex identification of the Eurasian otter *Lutra lutra* by PCR typing of spraints. *Conservation Genetics* **1**, 181–183.

Dallas JF, Coxon KE, Sykes T*, et al.* (2003) Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. *Molecular Ecology* **12**, 275–282.

Dallas JF, Marshall F, Piertney SB, Bacon PJ, Racey PA (2002) Spatially restricted gene flow and reduced microsatellite polymorphism in the Eurasian otter *Lutra lutra* in Britain. *Conservation Genetics* **3**, 15–29.

Dallas JF, Piertney SB (1998) Microsatellite primers for the Eurasian otter. *Molecular Ecology-Primer Notes* **7**, 1247–1263.

Deuter R, Pietsch S, Hertel S, Müller O (1995) A method for preparation of fecal DNA suitable for PCR. *Nucleic Acids Research* **23**, 3800–3801.

DeWoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6**, 951–957.

Dreher BP, Winterstein SR, Scribner KT*, et al.* (2007) Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bear. *Journal Of Wildlife Management* **71**, 2684–2693.

Dunning JB, Jr., Stewart DJ, Danielson BJ*, et al.* (1995) Spatially Explicit Population Models: Current Forms and Future Uses. *Ecological Applications* **5**, 3–11.

Dunstone N (1993) *The mink* T. & A.D. Poyser Ltd, London, UK.

Durbin LS (1998) Habitat selection by five otters *Lutra lutra* in rivers of northern Scotland. *Journal Of Zoology* **245**, 85–92.

Ebert C, Knauer F, Spielberger B, Thiele B, Hohmann U (2012) Estimating wild boar *Sus scrofa* population size using faecal DNA and capture-recapture modelling. *Wildlife Biology* **18**, 142–152.

Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic – mapping with trimetric and tetrameric tandem repeats. *American Journal Of Human Genetics* **49**, 746–756.

Eggert LS, Eggert JA, Woodruff DS (2003) Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. *Molecular Ecology* **12**, 1389–1402.

Elmeros M, Madsen AB (1999) On the reproduction biology of otters (*Lutra lutra*) from Denmark. *Zeitschrift Fur Saugetierkunde-International Journal of Mammalian Biology* **64**, 193–200.

Erlinge S (1967) Home Range of the Otter *Lutra lutra* L. in Southern Sweden. *Oikos* **18**, 186–209.

Erlinge S (1968) Territoriality of the Otter *Lutra lutra* L. *Oikos* **19**, 81–98.

Ernest HB, Penedo MCT, B.P.; M, Syvanen M, Boyce WM (2000) Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. *Molecular Ecology* **9**, 433–441.

Farrell LE, Romant J, Sunquist ME (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology* **9**, 1583–1590.

Ferrando A, Lecis R, Domingo-Roura X, Ponsa M (2008) Genetic diversity and individual identification of reintroduced otters (*Lutra lutra*) in north-eastern Spain by DNA genotyping of spraints. *Conservation Genetics* **9**, 129–139.

Flagstad O, Hedmark E, Landa A*, et al.* (2004) Colonization history and noninvasive monitoring of a reestablished wolverine population. *Conservation Biology* **18**, 676–688.

Flagstad O, Roed K, Stacy JE, Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology* **8**, 879–883.

Frantz AC, L.C.; P, Carpenter PJ*, et al.* (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. *Molecular Ecology* **12**, 1649–1661.

Frantz AC, Schaul M, L.C.; P*, et al.* (2004) Estimating population size by genotyping remotely plucked hair: the Eurasian badger. *Journal of Applied Ecology* **41**, 985–995.

Frantzen MAJ, Silk JB, Fergusen JWH, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423–1428.

Freeland JR (2005) *Molecular Ecology* John Wiley & Sons Ltd, Chichester, UK.

Garcia-Diaz P, Arevalo V, Lizana M (2011) Comparison of track and direct observation estimations for assessing abundance of the Eurasian otter, *Lutra lutra*. *Folia Zoologica* **60**, 37–42.

Garton EO, Horne JS, Aycrigg JL, Ratti JT (2012) Research and Experimental Design. In: *The Wildlife Techniques Manual* (ed. Silvy NJ), pp. 1–40. The Johns Hopkins University Press, Baltimore.

Gazey WJ, Staley MJ (1986) Population estimation from mark-recapture experiments using a sequential Bayes algorithm. *Ecology* **67**, 941–951.

Gerloff U, Schlötterer C, Rassmann K*, et al.* (1995) Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Molecular Ecology* **4**, 515–518.

Goossens B, Chikhi L, Utami SS, de Ruiter J, Bruford MW (2000) A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Conservation Genetics* **1**, 157–162.

Gorman ML, Jenkins D, Harper RJ (1978) Anal scent sacs of otter (*Lutra lutra*). *Journal Of Zoology* **186**, 463–474.

Gorman ML, Kruuk H, Jones C, McLaren G, Conroy JWH (1998) The demography of European otters *Lutra lutra*. *Symposia of the Zoological Society London* **71**, 107–118.

Gosling LM (1982) A Reassessment of the Function of Scent Marking in Territories. *Zeitschrift für Tierpsychologie* **60**, 89–118.

Grimm V, Railsback SF (2005) *Individual-based modeling and ecology* Princeton University Press.

Guertin DA, Ben-David M, Harestad AS, Elliott JE (2012) Fecal genotyping reveals demographic variation in river otters inhabiting a contaminated environment. *The Journal of Wildlife Management* **76**, 1540–1550.

Guichoux E, Lagache L, Wagner S*, et al.* (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources* **11**, 591–611.

Guschanski K, Vigilant L, McNeilage A*, et al.* (2009) Counting elusive animals: Comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. *Biological Conservation* **142**, 290–300.

Guter A, Dolev A, Saitz D, Kronfeld-Schor N (2008) Using videotaping to validate the use of spraints as an index of Eurasian otter (*Lutra lutra*) activity. *Ecological Indicators* **8**, 462–465.

Hajkova P, Pertoldi C, Zemanova B*, et al.* (2007) Genetic structure and evidence for recent population decline in Eurasian otter populations in the Czech and Slovak Republics: implications for conservation. *Journal Of Zoology* **272**, 1–9.

Hajkova P, Zemanova B, Bryja J*, et al.* (2006) Factors affecting success of PCR amplification of microsatellite loci from otter faeces. *Molecular Ecology Notes* **6**, 559–562.

Hajkova P, Zemanova B, Roche K, Hajek B (2009) An evaluation of field and noninvasive genetic methods for estimating Eurasian otter population size. *Conservation Genetics* **10**, 1667–1681.

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, New York.

Hansen H, Ben-David M, McDonald DB (2008) Effects of genotyping protocols on success and errors in identifying individual river otters (*Lontra canadensis*) from their faeces. *Molecular Ecology Resources* **8**, 282–289.

Harrington LA, Harrington AL, Hughes J, Stirling D, Macdonald DW (2010) The accuracy of scat identification in distribution surveys: American mink, *Neovison vison*, in the northern highlands of Scotland. *European Journal of Wildlife Research* **56**, 377–384.

Harrington LA, Harrington AL, Yamaguchi N*, et al.* (2009) The impact of native competitors on an alien invasive: temporal niche shifts to avoid interspecific aggression? *Ecology* **90**, 1207–1216.

Harris RB, Winnie J, Amish SJ*, et al.* (2010) Argali Abundance in the Afghan Pamir Using Capture-Recapture Modeling From Fecal DNA. *Journal Of Wildlife Management* **74**, 668–677.

Hauer S, Ansorge H, Zinke O (2000) A long-term analysis of the age structure of otters (*Lutra lutra*) from eastern Germany. *Zeitschrift Fur Saugetierkunde – International Journal of Mammalian Biology* **65**, 360–368.

Hauer S, Ansorge H, Zinke O (2002a) Mortality patterns of otters (*Lutra lutra*) from eastern Germany. *Journal Of Zoology* **256**, 361–368.

Hauer S, Ansorge H, Zinke O (2002b) Reproductive performance of otters *Lutra lutra* (Linnaeus, 1758) in Eastern Germany: low reproduction in a long-term strategy. *Biological Journal of the Linnean Society* **77**, 329–340.

Hausknecht R, Gula R, Pirga B, Kuehn R (2007) Urine — a source for noninvasive genetic monitoring in wildlife. *Molecular Ecology Notes* **7**, 208–212.

Hayakawa S, Takenaka O (1999) Urine as another potential source for template DNA in polymerase chain reaction (PCR). *American Journal of Primatology* **48**, 299–304.

Hedmark E, Ellegren H (2006) A test of the multiplex pre-amplification approach in microsatellite genotyping of wolverine faecal DNA. *Conservation Genetics* **7**, 289–293.

Hedmark E, Flagstad O, Segerstrom P*, et al.* (2004) DNA-based individual and sex identification from wolverine (*Gulo gulo*) faeces and urine. *Conservation Genetics* **5**, 405–410.

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

Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65–70.

Höss M, Kohn M, Paabo S, Knauer F, Schroder W (1992) Excrement analysis by PCR. *Nature* **359**, 199–199.

Hothorn T, Hornik K (2013) exactRankTests: Exact distribution for rank and permutation tests, R package version 0.8-27, URL http://cran.r-project.org/web/packages/exactRankTests/index.html, [accessed 26th Nov 2013].

Huggins RM (1989) On the statistical analysis of capture experiments. *Biometrika* **76**, 133–140.

Hung CM, Li SH, Lee LL (2004) Faecal DNA typing to determine the abundance and spatial organisation of otters (*Lutra lutra*) along two stream systems in Kinmen. *Animal Conservation* **7**, 301–311.

Hurvich CM, Tsai CL (1989) Regression and time-series model selection in small samples. *Biometrika* **76**, 297–307.

Idaghdour Y, Broderick D, Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustards. *Conservation Genetics* **4**, 789–792.

Inoue E, Inoue-Murayama M, Takenaka O, Nishida T (2007) Wild chimpanzee infant urine and saliva sampled noninvasively usable for DNA analyses. *Primates* **48**, 156–159.

Jacob G, Debrunner R, Gugerli F, Schmid B, Bollmann K (2010) Field surveys of capercaillie (*Tetrao urogallus*) in the Swiss Alps underestimated local abundance of the species as revealed by genetic analyses of non-invasive samples. *Conservation Genetics* **11**, 33–44.

Jansman HAH, Chanin PRF, Dallas JF (2001) Monitoring otter populations by DNA typing of spraints. *IUCN Otter Specialist Group Bulletine* **18**, 12–19.

Janssens X, Fontaine MC, Michaux JR*, et al.* (2008) Genetic pattern of the recent recovery of European otters in southern France. *Ecography* **31**, 176–186.

Jenkins D (1980) Ecology of Otters in Northern Scotland: I. Otter (*Lutra lutra*) Breeding and Dispersion in Mid-Deeside, Aberdeenshire in 1974–79. *Journal of Animal Ecology* **49**, 713–735.

Johansson MP, McMahon BJ, Hoglund J, Segelbacher G (2012) Amplification success of multilocus genotypes from feathers found in the field compared with feathers obtained from shot birds. *Ibis* **154**, 15–20.

Johnson PCD, Haydon DT (2007) Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. *Genetics* **175**, 827–842.

Jurisch C, Geidezis L (1997) Minimum passage rate of fishes through the digestive tract of otters *Lutra lutra* Linne, 1758 (Mustelidae). *Mammalia* **61**, 123–126.

Kalz B, Jewgenow K, Fickel J (2006) Structure of an otter (*Lutra lutra*) population in Germany – results of DNA and hormone analyses from faecal samples. *Mammalian Biology* **71**, 321–335.

Kean EF, Muller CT, Chadwick EA (2011) Otter Scent Signals Age, Sex, and Reproductive Status. *Chemical Senses* **36**, 555–+.

Kendall WL, Nichols JD, Hines JE (1997) Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* **78**, 563–578.

Kendall WL, Pollock KH, Brownie C (1995) A likehood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics* **51**, 293–308.

Kirk H, Freeland JR (2011) Applications and Implications of Neutral versus Non-neutral Markers in Molecular Ecology. *International Journal of Molecular Sciences* **12**, 3966–3988.

Klenke R (1996) Ergebnisse der Erfassung von Fischotternachweisen von 1993 bis 1995. In: *Artenschutzprogramm – Fischotter in Sachsen* (ed. Geologie SLfUu), pp. 12–17. Materialien zu Naturschutz und Landschaftspflege, Radebeul.

Klenke RA, Ring I, Schwerdtner-Máñez K*, et al.* (2013) Otters in Saxony: A Story of Successful Conflict Resolution. In: *Human – Wildlife Conflicts in Europe. Fisheries and Fish-eating Vertebrates as a Model Case.* (eds. Klenke RA, Ring I, Kranz A*, et al.*). Springer, Berlin, Heidelberg.

Knapp SM, Craig BA, Waits LP (2009) Incorporating Genotyping Error Into Non-Invasive DNA-Based Mark-Recapture Population Estimates. *Journal Of Wildlife Management* **73**, 598–604.

Koelewijn HP, Perez-Haro M, Jansman HAH*, et al.* (2010) The reintroduction of the Eurasian otter (*Lutra lutra*) into the Netherlands: hidden life revealed by noninvasive genetic monitoring. *Conservation Genetics* **11**, 601–614.

Kohn M, Knauer F, Stoffella A, Schröder W, Pääbo S (1995) Conservation genetics of the European brown bear – a study using excremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology* **4**, 95–103.

Kohn MH, Wayne RK (1997) Facts from feces revisited. *Tree* **12**, 223–227.

Kohn MH, York EC, Kamradt DA*, et al.* (1999) Estimating population size by genotyping faeces. *Royal Society* **266**, 657–663.

Kranz A (2000) Otters (*Lutra lutra*) increasing in Central Europe: from the threat of extinction to locally perceived overpopulation? *Mammalia* **64**, 357–368.

Kruuk H (1992) Scent marking by otters (*Lutra lutra*) – signaling the use of resources. *Behavioral Ecology* **3**, 133–140.

Kruuk H (2006) *Otters – ecology, behaviour and conservation* Oxford University Press, Oxford, New York.

Kruuk H, Conroy JWH (1991) Mortality of Otters (*Lutra lutra*) in Shetland. *Journal of Applied Ecology* **28**, 83–94.

Kruuk H, Conroy JWH, Glimmerveen U, Ouwerkerk EJ (1986) The use of spraints to survey populations of otters *Lutra lutra*. *Biological Conservation* **35**, 187–194.

Kruuk H, Conroy JWH, Moorhouse A (1987) Seasonal reproduction, mortality and food of otters (*Lutra lutra* L.) in Shetland. *Symposia of the Zoological Society London* **58**, 263–278.

Kruuk H, Conroy JWH, Moorhouse A (1991) Recruitment to a Population of Otters (*Lutra lutra*) in Shetland, in Relation to Fish Abundance. *Journal of Applied Ecology* **28**, 95–101.

Kruuk H, Moorhouse A (1991) The spatial organization of otters (*Lutra lutra*) in Shetland. *Journal Of Zoology* **224**, 41–57.

Kruuk H, Moorhouse A, Conroy JWH, Durbin L, Frears S (1989) An estimate of numbers and habitat preferences of otters *Lutra lutra* in Shetland, UK. *Biological Conservation* **49**, 241–254.

Lampa S, Gruber B, Henle K, Hoehn M (2008) An optimisation approach to increase DNA amplification success of otter faeces. *Conservation Genetics* **9**, 201–210.

Lampa S, Henle K, Klenke R, Hoehn M, Gruber B (2013) How to overcome genotyping errors in non-invasive genetic mark-recapture population size estimation—A review of available methods illustrated by a case study. *The Journal of Wildlife Management* **77**, 1490–1511.

Lange R, Gruber B, Henle K, Sarre SD, Hoehn M (2013) Mating system and intrapatch mobility delay inbreeding in fragmented populations of a gecko. *Behavioral Ecology* **24**, 1260–1270.

Lanszki J, Hidas A, Szentes K*, et al.* (2010) Genetic structure of otter (*Lutra lutra*) populations from two fishpond systems in Hungary. *Mammalian Biology* **75**, 447–450.

Lanszki J, Hidas A, Szentes K*, et al.* (2008) Relative spraint density and genetic structure of otter (*Lutra lutra*) along the Drava River in Hungary. *Mammalian Biology* **73**, 40–47.

Laver PN, Kelly MJ (2008) A critical review of home range studies. *Journal Of Wildlife Management* **72**, 290–298.

Lee SM, Chao A (1994) Estimating population-size via sample coverage for closed capture-recapture models. *Biometrics* **50**, 88–97.

Li R, Fan W, Tian G*, et al.* (2010) The sequence and de novo assembly of the giant panda genome. *Nature* **463**, 311–317.

Liebich HG (1999) *Funktionelle Histologie der Haussäugetiere: Lehrbuch und Farbatlas für Studium und Praxis*, 3 edn. Schattauer, Stuttgart, Germany.

Link WA, Yoshizaki J, Bailey LL, Pollock KH (2010) Uncovering a Latent Multinomial: Analysis of Mark-Recapture Data with Misidentification. *Biometrics* **66**, 178–185.

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics* **11**, 355–373.

Lukacs PM (2010) Closed population capture-recapture models. In: *Program MARK – A gentle introduction* eds. Cooch E, White GC), p. 509.

Lukacs PM, Burnham KP (2005a) Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. *Journal Of Wildlife Management* **69**, 396–403.

Lukacs PM, Burnham KP (2005b) Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology* **14**, 3909–3919.

Macdonald SM, Mason CF (1987) Seasonal marking in an otter population. *Acta Theriologica* **32**, 449–462.

Manly BFJ, Amstrup SC, McDonald TL (2005) *Capture-recapture methods in practice* Princeton University Press, Princeton and Oxford.

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

Martín-Gálvez D, Peralta-Sanchez JM, Dawson DA*, et al.* (2011) DNA sampling from eggshell swabbing is widely applicable in wild bird populations as demonstrated in 23 species. *Molecular Ecology Resources* **11**, 481–493.

Marucco F, Boitani L, Pletscher DH, Schwartz MK (2011) Bridging the gaps between non-invasive genetic sampling and population parameter estimation. *European Journal of Wildlife Research* **57**, 1–13.

Marucco F, Vucetich LM, Peterson RO, Adams JR, Vucetich JA (2012) Evaluating the efficacy of non-invasive genetic methods and estimating wolf survival during a ten-year period. *Conservation Genetics* **13**, 1611–1622.

Mason CF, Macdonald SM (1986) *Otters: ecology and conservation* Cambridge University Press, New York.

Mason CF, Macdonald SM (1987) The use of spraints for surveying otter *Lutra lutra* populations: An evaluation. *Biological Conservation* **41**, 167–177.

Mazerolle MJ (2013) AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c), R package version 1.35, URL http://cran.r-project.org/web/packages /AICcmodavg/index.html, [accessed 15th January 2014].

McDonald RA, O'Hara K, Morrish DJ (2007) Decline of invasive alien mink (*Mustela vison*) is concurrent with recovery of native otters (*Lutra lutra*). *Diversity and Distributions* **13**, 92–98.

McKelvey KS, Schwartz MK (2004) Genetic errors associated with population estimation using non-invasive molecular tagging: Problems and new solutions. *Journal Of Wildlife Management* **68**, 439–448.

McKelvey KS, Schwartz MK (2005) DROPOUT: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. *Molecular Ecology Notes* **5**, 716–718.

Melero Y, Plaza M, Santulli G*, et al.* (2012) Evaluating the effect of American mink, an alien invasive species, on the abundance of a native community: is coexistence possible? *Biodiversity and Conservation* **21**, 1795–1809.

Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics* **160**, 357–366.

Miller CR, Joyce P, Waits LP (2005) A new method for estimating the size of small populations from genetic mark-recapture data. *Molecular Ecology* **14**, 1991–2005.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. *Ecological Applications* **10**, 283–294.

Miquel C, Bellemain E, Poillot C*, et al.* (2006) Quality indexes to assess the reliability of genotypes in studies using noninvasive sampling and multiple-tube approach. *Molecular Ecology Notes* **6**, 985–988.

Mondol S, Karanth KU, Kumar NS*, et al.* (2009) Evaluation of non-invasive genetic sampling methods for estimating tiger population size. *Biological Conservation* **142**, 2350–2360.

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 versus*). *Molecular Ecology* **10**, 1835–1844.

Mowat G, Paetkau D (2002) Estimating marten Martes americana population size using hair capture and genetic tagging. *Wildlife Biology* **8**, 201–209.

Mowat G, Strobeck C (2000) Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. *Journal Of Wildlife Management* **64**, 183–193.

Mucci N, Arrendal J, Ansorge H*, et al.* (2010) Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. *Conservation Genetics* **11**, 583–599.

Murphy MA, Kendall KC, Robinson A, Waits LP (2007) The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. *Conservation Genetics* **8**, 1219–1224.

Murphy MA, Waits LP, Kendall KC (2000) Quantitative evaluation of fecal drying methods for brown bear DNA analysis. *Wildlife Society Bulletin* **28**, 951–957.

Murphy MA, Waits LP, Kendall KC*, et al.* (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. *Conservation Genetics* **3**, 435–440.

Myšiak J, Schwerdtner-Máñez K, Ring I (2013) Comparative analysis of the conflicts between carp pond farming and the protection of otters (*Lutra lutra*) in Upper Lusatia and South Bohemia. In: *Human – Wildlife Conflicts in Europe. Fisheries and Fish-eating Vertebrates as a Model Case.* (eds. Klenke RA, Ring I, Kranz A*, et al.*). Springer, Berlin, Heidelberg.

Nielsen CK, Woolf A (2001) Spatial organization of bobcats (*Lynx rufus*) in southern Illinois. *American Midland Naturalist* **146**, 43–52.

Nsubuga AM, Robbins MM, Roeder AD*, et al.* (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Molecular Ecology* **13**, 2089–2094.

Ó Néill L, Veldhuizen T, Jongh A, Rochford J (2009) Ranging behaviour and socio-biology of Eurasian otters (*Lutra lutra*) on lowland mesotrophic river systems. *European Journal of Wildlife Research* **55**, 363–370.

Ó Néill L, Wilson P, de Jongh A, de Jong T, Rochford J (2008) Field techniques for handling, anaesthetising and fitting radio-transmitters to Eurasian otters (*Lutra lutra*). *European Journal of Wildlife Research* **54**, 681–687.

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical-inference from capture data on closed animal populations. *Wildlife Monographs*, 7–135.

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

Palomares F, Godoy A, Piriz A, O'Brien J, Johnson WE (2002) Faecal genetic analysis to determine the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. *Molecular Ecology* **11**, 2171–2182.

Palsboll PJ, Allen J, Berube M*, et al.* (1997) Genetic tagging of humpback whales. *Nature* **388**, 767–769.

Paxinos E, Mcintosh C, Ralls K, Fleischer R (1997) A noninvasive method for distinguishing among canid species: amplification and enzyme restriction of DNA from dung. *Molecular Ecology* **6**, 483–486.

Pennell MW, Stansbury CR, Waits LP, Miller CR (2013) Capwire: a R package for estimating population census size from non-invasive genetic sampling. *Molecular Ecology Resources* **13**, 154–157.

Petit E, Valière N (2006) Estimating population size with noninvasive capture-mark-recapture data. *Conservation Biology* **20**, 1062–1073.

Piggott MP, Banks SC, Stone N, Banffy C, Taylor AC (2006) Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. *Molecular Ecology* **15**, 81–91.

Piggott MP, Bellemain E, Taberlet P, Taylor AC (2004) A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. *Conservation Genetics* **5**, 417–420.

Piggott MP, Taylor AC (2003) Extensive evaluation of faecel preservation and DNA extraction methods in Australian native and introduced species. *Australian Journal of Zoology* **51**, 341–355.

Pinheiro J, Bates D, DebRoy S, Sarkar D, Team RC (2014) nlme: linear and nonlinear mixed effects models, R package version 3.1-117, URL http://CRAN.R-project.org/package =nlme, [accessed 4th Sept 2014].

Pledger S (2000) Unified Maximum Likelihood Estimates for Closed Capture–Recapture Models Using Mixtures. *Biometrics* **56**, 434–442.

Pollock KH (1982) A capture-recapture design robust to unequal probability of capture. *Journal Of Wildlife Management* **46**, 752–757.

Pollock KH, Nichols JD, Brownie C, Hines JE (1990) Statistical inference for capture-recapture experiments. *Wildlife Monographs*, 1–97.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847–859.

Poole KG, Reynolds DM, Mowat G, Paetkau D (2011) Estimating Mountain Goat Abundance Using DNA From Fecal Pellets. *Journal Of Wildlife Management* **75**, 1527–1534.

Pradel R (1996) Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* **52**, 703–709.

Prigioni C, Remonti L, Balestrieri A*, et al.* (2006) Estimation of European otter (*Lutra lutra*) population size by feacal DNA typing in Southern Italy. *Journal Of Mammalogy* **87**, 855–858.

Puechmaille SJ, Petit EJ (2007) Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. *Journal of Applied Ecology* **44**, 843–852.

Quaglietta L, Fonseca VC, Hajkova P, Mira A, Boitani L (2013) Fine-scale population genetic structure and short-range sex-biased dispersal in a solitary carnivore, *Lutra lutra*. *Journal Of Mammalogy* **94**, 561–571.

Quaglietta L, Fonseca VC, Mira A, Boitani L (2014) Sociospatial organization of a solitary carnivore, the Eurasian otter (*Lutra lutra*). *Journal Of Mammalogy* **95**, 140–150.

Quaglietta L, Martins BH, de Jongh A, Mira A, Boitani L (2012) A Low-Cost GPS GSM/GPRS Telemetry System: Performance in Stationary Field Tests and Preliminary Data on Wild Otters (*Lutra lutra*). *Plos One* **7**, 10.

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

Reed JZ, Tollit DJ, Thompson PM, Amos W (1997) Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Molecular Ecology* **6**, 225–234.

Remonti L, Balestrieri A, Smiroldo G, Prigioni C (2011) Scent marking of key food sources in the Eurasian otter. *Annales Zoologici Fennici* **48**, 287–294.

Reuther C (2004) *On the way towards an otter habitat network Europe (OHNE) – method and result s of an area assessment on the European and the German level* GN-Gruppe Naturschutz GmbH, Hankensbüttel.

Reuther C, Dolch D, Green R, *et al.* (2000) Surveying and monitoring distribution and population trends of the Eurasian otter (*Lutra lutra*). *Habitat* **12**, 1–148.

Rew MB, Robbins J, Mattila D, Palsboll PJ, Berube M (2011) How many genetic markers to tag an individual? An empirical assessment of false matching rates among close relatives. *Ecological Applications* **21**, 877–887.

Ripple WJ, Estes JA, Beschta RL, *et al.* (2014) Status and Ecological Effects of the World's Largest Carnivores. *Science* **343**, 151–+.

Robinson SJ, Waits LP, Martin ID (2009) Estimating abundance of American black bears using DNA-based capture-mark-recapture models. *Ursus* **20**, 1–11.

Roeder AD, Archer FI, Poinar HN, Morin PA (2004) A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes* **4**, 761–764.

Roon DA, Waits LP, Kendall KC (2005) A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. *Animal Conservation* **8**, 203–215.

Rostain RR, Ben-David M, Groves P, Randall JA (2004) Why do river otters scent-mark? An experimental test of several hypotheses. *Animal Behaviour* **68**, 703–711.

Ruell EW, Riley SPD, Douglas MR, Pollinger JP, Crooks KR (2009) Estimating Bobcat Population Sizes And Densities In A Fragmented Urban Landscape Using Noninvasive Capture-Recapture Sampling. *Journal Of Mammalogy* **90**, 129–135.

Ruibal M, Peakall R, Claridge A, Firestone K (2009) Field-based evaluation of scat DNA methods to estimate population abundance of the spotted-tailed quoll (*Dasyurus maculatus*), a rare Australian marsupial. *Wildlife Research* **36**, 721–736.

Ruiz-Olmo J, Batet A, Manas F, Martinez-Vidal R (2011) Factors affecting otter (*Lutra lutra*) abundance and breeding success in freshwater habitats of the northeastern Iberian Peninsula. *European Journal of Wildlife Research* **57**, 827–842.

Ruiz-Olmo J, Delibes M, Zapata SC (1998) External morphometry, demography and mortality of the otter *Lutra lutra* (Linneo, 1758) in the Iberian Peninsula. *Galemys* **10**, 239–251.

Ruiz-Olmo J, Loy A, Cianfrani C, *et al.* (2008) *Lutra lutra.* www.iucnredlist.org

Ruiz-Olmo J, Saavedra D, Jiménez J (2001) Testing the surveys and visual and track censuses of Eurasian otters (*Lutra lutra*). *Journal Of Zoology* **253**, 359–369.

Sachteleben J, Simlacher C, Keller T, *et al.* (2010) Distribution of the European otter in Bavaria – status quo in 2008. *ANLIEGEN NATUR – Zeitschrift für Naturschutz* **34**, 3–8.

Sage W (2012) Der Fischotter *Lutra lutra* am "Unteren Inn" – Situation und Ausblick. *Mitt. Zool. Ges. Braunau* **10**, 271–279.

Saiki RK, Gelfand DH, Stoffel S, *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. *Science* **239**, 487–491.

Santini A, Lucchini V, Fabbri E, Randi E (2007) Ageing and environmental factors affect PCR success in wolf (*Canis lupus*) excremental DNA samples. *Molecular Ecology Notes* **7**, 955–961.

Santos-Reis M, Santos R, Antunes P, *et al.* (2013) Reconciliation of the conflict between otters and fish farmers. In: *Human-Wildlife conflicts in Europe – Fisheries and fish-eating vertebrates as a model case* (eds. Klenke RA, Ring I, Kranz A, *et al.*). Springer, Berlin, Heidelberg.

Sastre N, Francino O, Lampreave G, *et al.* (2009) Sex identification of wolf (*Canis lupus*) using non-invasive samples. *Conservation Genetics* **10**, 555–558.

Scheppers TLJ, Frantz AC, Schaul M, *et al.* (2007) Estimating social group size of Eurasian badgers *Meles meles* by genotyping remotely plucked single hairs. *Wildlife Biology* **13**, 195–207.

Schwartz MK, Cushman SA, McKelvey KS, Hayden J, Engkjer C (2006) Detecting genotyping errors and describing American black bear movement in northern Idaho. *Ursus* **17**, 138–148.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution* **22**, 25–33.

Schwerdtner K, Gruber B (2007) A conceptual framework for damage compensation schemes. *Biological Conservation* **134**, 354–360.

Sidorovich VE (1991) Structure, reproductive status and dynamics of the otter population in Byelorussia. *Acta Theriologica* **36**, 153–161.

Sidransky D, Tokino T, Hamilton SR, *et al.* (1992) Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. *Science* **256**, 102–105.

Sjöåsen T (1996) Survivorship of captive-bred and wild-caught reintroduced European otters *Lutra lutra* in Sweden. *Biological Conservation* **76**, 161–165.

Sjöåsen T (1997) Movements and Establishment of Reintroduced European Otters *Lutra lutra*. *Journal of Applied Ecology* **34**, 1070–1080.

Smith DA, Ralls K, Hurt A, *et al.* (2006) Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analyses. *Molecular Ecology* **15**, 387–406.

Smith PJ (1991) Bayesian analyses for a multiple capture-recapture model. *Biometrika* **78**, 399–407.

Solberg KH, Bellemain E, Drageset OM, Taberlet P, Swenson JE (2006) An evaluation of field and non-invasive genetic methods to estimate brown bear (*Ursus arctos*) population size. *Biological Conservation* **128**, 158–168.

Stanley TR, Richard JD (2005) Software Review: A program for testing capture-recapture data for closure. *Wildlife Society Bulletin* **33**, 782–785.

Stenglein JL, Waits LP, Ausband DE, Zager P, Mack CM (2010) Efficient, Noninvasive Genetic Sampling for Monitoring Reintroduced Wolves. *Journal Of Wildlife Management* **74**, 1050–1058.

Stevick PT, Palsb, oslash, *et al.* (2001) Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1861–1870.

Sugiura N (1978) Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics – Theory and Methods*, 13–26.

Sulkava R (2007) Snow tracking: a relevant method for estimating otter *Lutra lutra* populations. *Wildlife Biology* **13**, 208–218.

Sundqvist A-K, Ellegren H, Vilà C (2008) Wolf or dog? Genetic identification of predators from saliva collected around bite wounds on prey. *Conservation Genetics* **9**, 1275–1279.

Taberlet P, Bouvet J (1992) Bear conservation genetics. *Nature* **358**, 197–197.

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

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

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Tree* **14**, 323–327.

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463–6471.

Tikel D, Blair D, Marsh HD (1996) Marine mammal faeces as a source of DNA. *Molecular Ecology* **5**, 456–457.

Tredick CA, Vaughan MR, Stauffer DF, Simek SL, Eason T (2007) Sub-sampling genetic data to estimate black bear population size: a case study. *Ursus* **18**, 179–188.

Trowbridge BJ (1983) *Olfactory communication in the European otter (Lutra lutra)*, University of Aberdeen.

Tschirch W (1995) Neue biomedizinische Methoden für die Feldforschung – dargestellt am Beispiel des Fischotters (*Lutra lutra*). In: *Methode der feldökologischen Säugetierforschung (Methods in mammalian field ecology)* (eds. Stubbe M, Stubbe A, Heidecke D). Martin-Luther-Universität, Halle/Saale, Germany.

Uno R, Kondo M, Yuasa T, *et al.* (2012) Assessment of genotyping accuracy in a non-invasive DNA-based population survey of Asiatic black bears (*Ursus thibetanus*): lessons from a large-scale pilot study in Iwate prefecture, northern Japan. *Population Ecology* **54**, 509–519.

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377–379.

Valière N, Berthier P, Mouchiroud D, Pontier D (2002) GEMINI: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83–86.

Van Manen FT, McCollister MF, Nicholson JM, *et al.* (2012) Short-Term Impacts of a 4-Lane Highway on American Black Bears in Eastern North Carolina. *Wildlife Monographs* **181**, 1–35.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538.

Vigilant L (1999) An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. *Biol. Chem.* **380**, 1329–1331.

Vigilant L, Guschanski K (2009) Using genetics to understand the dynamics of wild primate populations. *Primates* **50**, 105–120.

Waits JL, Leberg PL (2000) Biases associated with population estimation using molecular tagging. *Animal Conservation* **3**, 191–199.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* **10**, 249–256.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection. *Journal Of Wildlife Management* **69**, 1419–1433.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* **10**, 506–513.

Wang J (2005) Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 1395–1409.

Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetics Research* **89**, 135–153.

Wang J (2011) coancestry: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11**, 141–145.

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of Ursids. *Molecular Ecology* **6**, 1091–1097.

Welsch U (2006) *Lehrbuch Histologie*, 2 edn. Urban & Fischer, München, Germany.

White GC, Anderson DR, Burnham KP, Otis DL (1982) Capture-recapture and removal methods for sampling closed populations. *Los Alamos National Laboratory Rep. LA-8787-NERP, New Mexico, USA*, pp.235.

White GC, Burnham KP (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study* **46**, 120–139.

Wiegand T, Knauer F, Kaczensky P, Naves J (2004) Expansion of brown bears (*Ursus arctos*) into the eastern Alps: a spatially explicit population model. *Biodiversity and Conservation* **13**, 79–114.

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press San Diego, California, USA.

Williams BW, Etter DR, Linden DW, *et al.* (2009) Noninvasive Hair Sampling and Genetic Tagging of Co-Distributed Fishers and American Martens. *Journal Of Wildlife Management* **73**, 26–34.

Wilson GJ, Frantz AC, Pope LC, *et al.* (2003) Estimation of badger abundance using faecal DNA typing. *Journal of Applied Ecology* **40**, 658–666.

Worton BJ (1989) Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology* **70**, 164–168.

Wright JA, Barker RJ, Schofield MR, *et al.* (2009) Incorporating Genotype Uncertainty into Mark-Recapture-Type Models For Estimating Abundance Using DNA Samples. *Biometrics* **65**, 833–840.

Wright JA, Schofield MR, Barker RJ (2012) *Closed population genetic mark-recapture with genotyping error and heterogeneity*. Krokkleiva, Norway

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97–159.

Yamaguchi N, MacDonald DW (2003) The Burden of Co-Occupancy: Intraspecific Resource Competition and Spacing Patterns in American Mink, *Mustela vison*. *Journal Of Mammalogy* **84**, 1341–1355.

Yoshizaki J, Brownie C, Pollock KH, Link WA (2011) Modeling misidentification errors that result from use of genetic tags in capture-recapture studies. *Environmental and Ecological Statistics* **18**, 27–55.

Zhan XJ, Li M, Zhang ZJ, *et al.* (2006) Molecular censusing doubles giant panda population estimate in a key nature reserve. *Current Biology* **16**, R451–R452.

Zhan XJ, Tao Y, Li M, *et al.* (2009) Accurate population size estimates are vital parameters for conserving the giant panda. *Ursus* **20**, 56–62.

Zinke O (1991) Die Todesursachen der im Museum der Westlausitz Kamenz von 1985–1991 eingelieferten Fischotter *Lutra lutra* (L., 1758). *Veröffentlichungen des Museums der Westlausitz*, 57–63.

Zinke O (2000) Die Todesursachen der im Museum der Westlausitz Kamenz von 1992–1998 eingelieferten Fischotter *Lutra lutra* (L., 1758). *Veröffentlichungen des Museums der Westlausitz*, 87–92.

Zuur AF, Ieno EN, Walker NJ, Savaliev AA, Smith GM (2009) *Mixed effects models and extensions in ecology with R* Springer, New York

Ehrenwörtliche Erklärung

Hiermit erkläre ich, Simone M. Lampa, geboren am 24.03.1980 in Hünfeld, an Eides statt,

- dass mir die geltende Promotionsordnung der Biologisch-Pharmazeutischen Fakultät bekannt ist;
- dass ich die vorliegende Dissertation selbst angefertigt habe, keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe;
- dass alle Personen, die mich bei der Auswahl und Auswertung der Daten bzw. bei der Erstellung der Manuskripte unterstützt haben als Co-Autoren genannt sind oder in der jeweiligen Danksagung der einzelnen Manuskripte aufgeführt sind;
- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und das Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen;
- dass ich die Dissertation nicht bereits zuvor als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe;
- dass ich nicht die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, 24. Oktober 2014

Simone M. Lampa

Appendix to Chapter Three

Supplemental material to: Lampa, S., K. Henle, R. Klenke, M. Hoehn, and B. Gruber. 2013. How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation – A Review of Available Methods Illustrated by a Case Study. Journal of Wildlife Management, 77(8), 1490–1511. Available online at: http://onlinelibrary.wiley.com/doi/10.1002/jwmg.604/suppinfo

Table S1. Simple step-by-step guide for non-invasive genetic capture-mark-recapture (CMR) studies, particularly for samples with high genotyping error rate and low amplification success rate. Here, we summarize the recommendations found in the literature or derived from our own results and we contrast commonly used or available methods addressing the problem or task by offering their pros (👍) and cons (👎).

	Task & Requirements	Methods Addressing the Task & Requirement[a]
Before and in the Field	Closed population (Alternative: open population models)	Choose a short sampling period relative to the biology (e.g., turnover rate) of target species
		Perform a simulation (e.g., in MARK (White and Burnham 1999)) to estimate the number of required sampling occasions
		Choose a large study area relative to the species territory size
		Avoid periods with high migration, mortality, or birth
		If possible, take only fresh samples
	Equal capture probability (is difficult to meet, but should be maximized)	Choose non-invasive material (e.g., hairs, feces, urine, saliva, feathers) that is deposited by all members of the population or collect several kinds of samples
		Choose a sampling regime within the study area (e.g., transects, follow trails, sampling points) where chances are highest to collect all members of the population
		Choose a time period where all individuals have a chance of being sampled regarding their sex, age, reproductive status
		Get high re/capture rate by collecting as many samples as possible
	Sampling technique (must fit to preservation & extraction; perform a pilot study if possible; strictly avoid cross-contamination)	Dry samples (e.g., hairs, feathers) should be taken entirely with new gloves for each sample or disposable collection tools
		Moist samples (e.g., feces, urine, saliva) can be taken entirely (caution: may alter species behavior) or parts of it (e.g., surface scrap) with a swap or other disposable collection tools
Get High Amount of Template DNA	Preservation (must fit to extraction; perform a pilot study if possible)	Freezing (i.e., –20°C; –80°C) 👍 Quick; works well for dry samples 👎 Risk of decreasing amplification success rate, especially for moist samples; equipment needed; only shortly shippable
		Drying (e.g., oven, silica gel) 👍 Cheap; can be done almost everywhere in the field; shippable 👎 Risk of decreasing amplification success rate, especially for moist samples, but works for some feces (e.g., Piggott and Tayler 2003)
		Buffers (e.g., ethanol, DET buffer, RNAlater solution at room temp. or –80°C) 👍 Seems to work well for many moist samples 👎 Extra chemicals needed; extra handling before extraction required; often not shippable
		Lyse buffer of extraction (room temp. or –80°C) 👍 No extra costs; first extraction step is done; possible for all kinds of samples 👎 Often not shippable; has not been tested yet on various kinds of samples
	Extraction (perform a pilot study if possible; prevent cross-contamination rigorously)	Phenol/chloroform extraction 👍 Effective on a wide range of samples and for very long DNA fragments (i.e., Kbps) 👎 Hazardous chemicals; time-consuming; old-fashioned

		Resin-based extraction (e.g., Chelex®) ↑ Cheap and quick ↓ Low DNA purity; can inhibit PCR; DNA degradation with increasing storage time
		Silica-based extraction (e.g., commercial kits) ↑ High DNA quality; effective on a wide range of samples; less PCR inhibitors ↓ Expensive; time-consuming
Microsatellite Genotyping	Minimizing genotyping errors (amplification pilot study must be performed to get best PCR conditions; prevent cross-contamination rigorously)	Use engineered polymerase enzymes
		Choose only few (< 10), but highly variable and short (< 200–300 bp) markers
		Use low retention plastic tubes (in all lab steps)
		Multiple-tubes approach (Taberlet et al. 1996) ↑ Reliable consensus genotype following a worst-case approach ↓ Time-consuming; cost-intensive; DNA extract-consuming
		Comparative multiple-tubes approach (Frantz et al. 2003) ↑ Reliable consensus genotype with less repetitions; less time-, cost-, and extract-consuming than multiple-tubes approach ↓ Still relatively time-consuming, cost-intensive and extract-consuming
		Multiplex pre-amplification (Bellemain and Taberlet 2004; Piggott et al. 2004) ↑ Increased amplification success; decreased genotyping error rate ↓ Relatively time-consuming and cost-intensive
		Discard low-quality samples using quantitative PCR (Morin et al. 2001) ↑ Minimizes time and costs required for genotyping; consensus genotype achievable with fewer repetitions ↓ Method itself is expensive and requires additional effort to the genotyping analysis; removing low-quality samples might induce individual capture heterogeneity
		Discard low-quality samples using mtDNA analysis (Kohn et al. 1999) ↑ Minimizes time and costs required for genotyping; consensus genotype achievable with fewer repetitions; recognition of non-target species ↓ Additional effort and costs to the genotyping analysis; removing low-quality samples might induce individual capture heterogeneity
		Quality control approach (Paetkau 2003) ↑ Minimizes time and costs required for genotyping; no consensus genotype required ↓ Not rigorous enough for high genotyping error rates; removing low-quality samples might induce individual capture heterogeneity
		Screening approach (Lampa et al. 2013) ↑ Minimizes time and costs required for genotyping; reliable consensus genotypes ↓ Still relatively time-consuming, cost-intensive and extract-consuming; removing low-quality samples might induce individual capture heterogeneity
	Detection and quantification of genotyping errors	Calculation of AD & FA following Broquet and Petit (2004) ↑ Actual values of AD & FA for the dataset; standardized calculation ↓ Requires consensus genotype, hence multiple repetitions
		Automated calculation of AD & FA using program GIMLET (Valière 2002) ↑ Fast calculation of the actual AD & FA values for the dataset ↓ Requires consensus genotype; cannot handle varying numbers of repetitions per loci within a sample
		Calculation of the quality index (Miquel et al. 2006) ↑ Standardized calculation; easy and quick ↓ No information about type and number of errors within a sample or locus; not commonly used
		MLE of AD and FA using program PEDANT (Johnson and Haydon 2007) ↑ No consensus genotype required ↓ AD & FA rate only for loci not for samples
		MLE of required PCR repetitions using program RELIOTYPE (Miller et al. 2002) ↑ Received number of repetitions are given for each sample; minimizes time and costs required for genotyping

Microsatellite Genotyping		👎 Accounts only for AD; AD must be evenly distributed; requires two positive PCRs per locus
		Simulation of required PCR repetitions using program GEMINI (Valière et al. 2002) 👍 Can be done in advance; minimizes time and costs required for genotyping 👎 Requires known error rates and heterozygosity; received number of repetitions are not specified for loci or samples
		Testing for errors in the dataset using program MICRO-CHECKER (Van Oosterhout et al. 2004) 👍 Can be used after the first positive PCR; recognizes error-prone loci 👎 Cumbersome for many loci or large populations; no indication which samples harbor errors; detects only systematic errors, no random one
		Testing for errors in the dataset using program DROPOUT (McKelvey and Schwartz 2004, 2005) 👍 Recognizes error-prone loci and samples; can be used at each state of the analysis; minimizes time and costs required for genotyping 👎 Assumes equal capture probability among individuals; requires sufficiently large tag size (preferably > 8); if used without forming consensus genotypes the per-locus error rate should not exceed 0.25
		Matching approach (Creel et al. 2003) or other correction methods (Bellemain et al. 2005) 👍 Does not require a complete elimination of all errors; biological and spatial information can be used 👎 Calibration of the threshold for the matching approach is difficult; can lead to underestimated population sizes
Estimating Population Size	Test of closure	Use of tests such as CLOSETEST (Stanley and Richard 2005) or Pradel's (1996) recruitment model following Boulanger et al. (2002) 👍 Test statistics or key figures can be used to accept or reject the assumption 👎 Individual capture heterogeneity and genotyping errors leading to ghost individuals can cause incorrect rejection of closure
		Assess closure on the basis of biological information 👍 Cannot be confounded with individual heterogeneity or the presence of ghost individuals 👎 Arguments are always contradictable as decision is only based on logical reasoning
	Test for equal capture probability	The model selected should account for the biology of the target species, the employed sampling design, and which management decisions will be based on the model
		Simulation test (Puechmaille and Petit 2007) 👍 Data-specific simulation 👎 Additional effort required; requires an estimate of population size
		Model selection algorithm in program CAPTURE (White et al. 1982) 👍 No additional effort to population size estimation; selection between models accounting for several kinds of varying capture probability (M_0, M_h, M_b, M_t, M_{tb}, M_{th}, M_{bh}) 👎 Low power; not appropriate for small populations, for low capture probability, or data with genotyping errors; model selection choose most frequently M_0 for error-free data and M_h for data having errors
		LRT in program CAPWIRE (Miller et al. 2005) 👍 No additional effort to population size estimation 👎 Misses some types of individual capture heterogeneity; decision is only made between M_0 and M_h
		AIC-based model selection in program MARK (White and Burnham 1999) 👍 No additional effort to population size estimation; selection between models accounting for several kinds of varying capture probability (M_0, M_h, M_b, M_t, M_{tb}, M_{th}, M_{bh}) and several model systems 👎 Can fail if assessed models suffer structural deficits
	Conventional estimation models (there are many more models available, but so far not widely used in non-invasive genetic CMR studies)	Program CAPWIRE (Miller et al. 2005) 👍 Incorporates multiple captures of an individual within a sampling occasion; appropriate for small populations; model selection tool available 👎 Severe overestimation if genotyping errors are still present; only M_0

Estimating Population Size		and M_h available
		Program CAPTURE (White et al. 1982) 👍 Offers several models accounting for varying capture probability; model selection tool available 👎 Severe overestimation if genotyping errors are still present
		Program MARK (White and Burnham 1999) 👍 Offers several model systems, each with models accounting for varying capture probability; model selection tool available 👎 Severe overestimation if genotyping errors are still present
	Models accounting for genotyping errors	Corrected Lincoln-Peterson estimator (Stevick et al. 2001) 👍 Specially developed for individual identification using poor quality photographs; useful for data with two sampling sessions 👎 Assumes equal capture probability; with individual capture heterogeneity estimator underestimates population sizes; accounts only for two sampling sessions; requires known false negative rate; no software available implementing the method
		GUAVA approach (Knapp et al. 2009) 👍 Accounts for the shadow effect; useful for data with two sampling sessions; software available 👎 Assumes equal capture probability; with individual capture heterogeneity estimators underestimate population sizes; accounts only for two sampling sessions; requires genotyping error rate, allele frequency and HWE; should not be used if target population was never studied before; extant error rate in the data should not markedly exceed 6%
		Mis-indentification models in MARK (Lukacs and Burnham 2005) → termed L&B estimator 👍 Works well if misidentification is ≤ 5%; offers probability of a correct classification; offers several models accounting for varying capture probability; model selection tool available; software implementing the method available 👎 Still extant error rate should not exceed 5%; biased for low capture probabilities; not correctly accounting for misidentification process; some assumptions may not always be true: a) errors are not repeated, b) errors cannot lead to an existing genotype, c) finite mixture of true and false genotypes with same initial detection probabilities, d) probability of correctly identifying an individual on first capture equals the proportion of true genotypes; requires a certain amount of data to separate misidentification from heterogeneity
		M_t-based approaches of Yoshizaki et al. (2011) 👍 Works well if misidentification is between 0–10%; mimics a more realistic misidentification process; multiple captures of individuals per sampling occasion are possible 👎 No other variance in catchability than time is considered; biased for low capture probabilities if misidentification is close to 1; requires the same assumptions like the L&B estimator; has not yet been applied by others and performance still needs to be evaluated using field data; no software or script available implementing the method
		Bayesian method of Link et al. (2010) 👍 Works well if no errors are present; mimics a more realistic misidentification process; does not require assumptions c) and d) of L&B estimator; multiple captures of individuals per sampling occasion are possible; allows to include additional knowledge using informative priors 👎 No other variance in catchability than time is considered; requires assumption a) and b) like the L&B estimator; extensions of the model are needed; requires more computing time; requires knowledge on the statistic process; has not yet been applied by others or compared to other methods (except with Yoshizaki et al. (2011)); no software or script available implementing the method
		Bayesian method of Wright et al. (2009,2012) → termed Wright-model 👍 Works well if no errors are present; accounts for time variance and individual heterogeneity; does not require assumptions a) and b) of L&B estimator; multiple captures of individuals per sampling occasion are possible; allows to include additional knowledge using informative priors; R-Script available upon request

		⓿ Does not account for behavioral response; requires more computing time; requires knowledge on the statistic process; has not yet been applied by others or compared to other methods

[a] PCR = polymerase chain reaction, AD = allelic dropout, FA = false allele, HWE = Hardy-Weinberg equilibrium

Literature Cited

Bellemain, E., J. E. Swenson, O. Tallmon, S. Brunberg, and P. Taberlet. 2005. Estimating Population Size of Elusive Animals with DNA from Hunter-Collected Feces: Four Methods for Brown Bears. Conservation Biology 19:150–161.

Bellemain, E., and P. Taberlet. 2004. Improved Noninvasive Genotyping Method: Application to Brown Bear (*Ursus Arctos*) Faeces. Molecular Ecology 4:519–522.

Boulanger, J., G. C. White, B. N. McLellan, J. Woods, M. Proctor, and S. Himmer. 2002. A Meta-Analysis of Grizzly Bear DNA Mark-Recapture Projects in British Columbia, Canada. Ursus 13:137–152.

Broquet, T., and E. Petit. 2004. Quantifying Genotyping Errors in Noninvasive Population Genetics. Molecular Ecology 13:3601–3608.

Creel, S., G. Spong, J. L. Sands, J. Rotella, J. Zeigle, L. Joe, K. M. Murphy, and D. Smith. 2003. Population Size Estimation in Yellowstone Wolves with Error-Prone Noninvasive Microsatellite Genotypes. Molecular Ecology 12:2003–2009.

Frantz, A. C., P. L.C.;, P. J. Carpenter, T. J. Roper, G. J. Wilson, R. J. Delahay, and T. Burke. 2003. Reliable Microsatellite Genotyping of the Eurasian Badger (Meles Meles) Using Faecal DNA. Molecular Ecology 12:1649–1661.

Johnson, P. C. D., and D. T. Haydon. 2007. Maximum-Likelihood Estimation of Allelic Dropout and False Allele Error Rates from Microsatellite Genotypes in the Absence of Reference Data. Genetics 175:827–842.

Knapp, S. M., B. A. Craig, and L. P. Waits. 2009. Incorporating Genotyping Error into Non-Invasive DNA-Based Mark-Recapture Population Estimates. Journal Of Wildlife Management 73:598–604.

Kohn, M. H., E. C. York, D. A. Kamradt, G. Haught, R. M. Sauvajot, and R. K. Wayne. 1999. Estimating Population Size by Genotyping Faeces. Royal Society 266:657–663.

Lampa, S., K. Henle, R. Klenke, M. Hoehn, and B. Gruber. 2013. How to Overcome Genotyping Errors in Non-Invasive Genetic Mark-Recapture Population Size Estimation – A Review of Available Methods Illustrated by a Case Study. Journal of Wildlife Management 77: 1490–1511.

Link, W. A., J. Yoshizaki, L. L. Bailey, and K. H. Pollock. 2010. Uncovering a Latent Multinomial: Analysis of Mark-Recapture Data with Misidentification. Biometrics 66:178–185.

Lukacs, P. M., and K. P. Burnham. 2005. Estimating Population Size from DNA-Based Closed Capture-Recapture Data Incorporating Genotyping Error. Journal Of Wildlife Management 69:396–403.

McKelvey, K. S., and M. K. Schwartz. 2004. Genetic Errors Associated with Population Estimation Using Non-Invasive Molecular Tagging: Problems and New Solutions. Journal Of Wildlife Management 68:439–448.

McKelvey, K. S., and M. K. Schwartz. 2005. Dropout: A Program to Identify Problem Loci and Samples for Noninvasive Genetic Samples in a Capture-Mark-Recapture Framework. Molecular Ecology Notes 5:716–718.

Miller, C. R., P. Joyce, and L. P. Waits. 2002. Assessing Allelic Dropout and Genotype Reliability Using Maximum Likelihood. Genetics 160:357–366.

Miller, C. R., P. Joyce, and L. P. Waits. 2005. A New Method for Estimating the Size of Small Populations from Genetic Mark-Recapture Data. Molecular Ecology 14:1991–2005.

Miquel, C., E. Bellemain, C. Poillot, J. BessiÈRe, A. Durand, and P. Taberlet. 2006. Quality Indexes to Assess the Reliability of Genotypes in Studies Using Noninvasive Sampling and Multiple-Tube Approach. Molecular Ecology Notes 6:985–988.

Morin, P. A., K. E. Chambers, C. Boesch, and L. Vigilant. 2001. Quantitative Polymerase Chain Reaction Analysis of DNA from Noninvasive Samples for Accurate Microsatellite Genotyping of Wild Chimpanzees (*Pan Troglodytes Versus*). Molecular Ecology 10:1835–1844.

Paetkau, D. 2003. An Empirical Exploration of Data Quality in DNA-Based Population Inventories. Molecular Ecology 12:1375–1387.

Piggott, M. P., E. Bellemain, P. Taberlet, and A. C. Taylor. 2004. A Multiplex Pre-Amplification Method That Significantly Improves Microsatellite Amplification and Error Rates for Faecal DNA in Limiting Conditions. Conservation Genetics 5:417–420.

Piggott, M. P., and A. C. Taylor. 2003. Extensive Evaluation of Faecel Preservation and DNA Extraction Methods in Australian Native and Introduced Species. Australian Journal of Zoology 51:341–355.

Pradel, R. 1996. Utilization of Capture-Mark-Recapture for the Study of Recruitment and Population Growth Rate. Biometrics 52:703–709.

Puechmaille, S. J., and E. J. Petit. 2007. Empirical Evaluation of Non-Invasive Capture-Mark-Recapture Estimation of Population Size Based on a Single Sampling Session. Journal of Applied Ecology 44:843–852.

Stanley, T. R., and J. D. Richard. 2005. Software Review: A Program for Testing Capture-Recapture Data for Closure. Wildlife Society Bulletin 33:782–785.

Stevick, P. T., Palsb, oslash, P. J. ll, T. D. Smith, M. V. Bravington, and P. S. Hammond. 2001. Errors in Identification Using Natural Markings: Rates, Sources, and Effects on Capture-Recapture Estimates of Abundance. Canadian Journal of Fisheries and Aquatic Sciences 58:1861–1870.

Taberlet, P., S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L. P. Waits, and J. Bouvet. 1996. Reliable Genotyping of Samples with Very Low DNA Quantities Using Pcr. Nucleic Acids Research 24:3189–3194.

Valière, N. 2002. Gimlet: A Computer Program for Analysing Genetic Individual Identification Data. Molecular Ecology Notes 2:377–379.

Valière, N., P. Berthier, D. Mouchiroud, and D. Pontier. 2002. Gemini: Software for Testing the Effects of Genotyping Errors and Multitubes Approach for Individual Identification. Molecular Ecology Notes 2:83–86.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. Molecular Ecology Notes 4:535–538.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-Recapture and Removal Methods for Sampling Closed Populations. Los Alamos National Laboratory Rep. LA-8787-NERP, New Mexico, USA:pp.235.

White, G. C., and K. P. Burnham. 1999. Program Mark: Survival Estimation from Populations of Marked Animals. Bird Study 46:120–139.

Wright, J. A., R. J. Barker, M. R. Schofield, A. C. Frantz, A. E. Byrom, and D. M. Gleeson. 2009. Incorporating Genotype Uncertainty into Mark-Recapture-Type Models for Estimating Abundance Using DNA Samples. Biometrics 65:833–840.

Wright, J. A., M. R. Schofield, and R. J. Barker. 2012. Closed Population Genetic Mark-Recapture with Genotyping Error and Heterogeneity. *in* International Statistical Ecology Conference. Krokkleiva, Norway.

Yoshizaki, J., C. Brownie, K. H. Pollock, and W. A. Link. 2011. Modeling Misidentification Errors That Result from Use of Genetic Tags in Capture-Recapture Studies. Environmental and Ecological Statistics 18:27–55.