

Non-Invasive Genetic Typing in the Study of Badger (*Meles meles*) Ecology



by

Alain Camille Frantz

Department of Biology and Environmental Science

University of Sussex

United Kingdom

Thesis submitted for the degree of Doctor of Philosophy (D.Phil.)

2004

Declaration

I hereby declare that this thesis has not been submitted in whole or in part, either in the same or different form, for a degree or diploma or any other qualification at this or any other university. All the work described in this thesis was carried out by me in the School of Life Sciences at the University of Sussex or at the field study sites or laboratories in Luxembourg. Where other sources are referred to this is indicated.

This thesis was first submitted to the University of Sussex in March 2004. The final version was handed in in July 2004.

Alain Frantz
September 2004
Roeser, Luxembourg

© Alain Frantz 2004, except contents of articles referred to on page 24; copyright with respective journals.

All rights reserved.

No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the written permission of the author.

Contact address: alainfrantz@yahoo.co.uk

Cover picture: Badger hair on barbed wire. Photo: Alain Frantz

To my parents
& Marguerite Klein

.

.

.

“Though such matters may be
interesting, it may be as well for you
and me not to have the reputation
of too frequently and too minutely
examining faecal products.”

Dr William Wollaston in a letter to
the Reverend William Buckland

Acknowledgements

I was supported by the *Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche*, Luxembourg, through a *Bourse de Formation-Recherche*. Guy Colling and Christian Ries (*Fondation Flore-Faune*) managed the *Bourse the Formation-Recherche*. The following institutions made financial contributions towards the running cost of the study: *Musée National d'Histoire Naturelle*, Luxembourg, the *Sheffield Molecular Genetics Facility* at the *University of Sheffield*, the *University of Sussex*, the *Institute of Immunology* at the *Laboratoire National de Santé*, Luxembourg, The *Direction des Eaux et Forêts*, Luxembourg, the *Central Science Laboratory*, UK.

I wish to express my sincere thanks to (in alphabetical order) Fränk Adam, Wim Ammerlaan¹, Adil Baghli², Georges Bechet², Trevor Beebee³, Romain Bei², Petra Carpenter⁴, David Coltman⁴, Deborah Dawson⁴, Richard Delahay⁵, Kristien Erven⁴, Camille Frantz, Laurent Frantz, Christine Goedert², Jean-Michel Guinet², Pierre Hentges³, Sylvie Hermant², Myriam Heuertz, David Hitchin³, Tom Hoffmann², Sarah Jackson, Ady Krier⁶, Andy Krupa⁴, Gordon Luikart⁷, Duncan McKnight, Craig Miller⁸, Mick M Mulders¹, Michael Pocock⁹, Katrina Service¹⁰, Jean-Marie Sinner⁶, Nathaniel Valière¹¹, Tania Walisch², Gavin Wilson⁵, Fränk Wolff¹².

I would like to thank Laurent Schley⁶, Edmée Engel² and Guy Colling² for their enthusiasm and support of the project. During the research I visited the laboratories of Terry Burke⁴, Guy Colling² and Claude P Muller¹. I would like to thank them for their hospitality. I would like to thank Fred Fack¹ for his help in setting up the laboratory work in Luxembourg.

Finally, special thanks to Lisa Pope⁴ for her patient and priceless help with the laboratory work at the start of the project, to Michel Schaul for his enthusiasm and invaluable help with the field work in Luxembourg and to Tim Roper³, for his support and supervision of the project as well as for his invaluable help, ideas and advice throughout the study.

Affiliations

¹ Institute of Immunology, Laboratoire National de Santé, Luxembourg; ² Musée National d'Histoire Naturelle, Luxembourg; ³ University of Sussex, UK; ⁴ University of Sheffield, UK; ⁵ Central Science Laboratory, UK; ⁶ Administration des Eaux et Forêts, Luxembourg; ⁷ Université Joseph Fourier, Grenoble, France; ⁸ University of Idaho, USA; ⁹ University of Bristol, UK; ¹⁰ University of Oxford, UK; ¹¹ Université Claude Bernard, Lyon, France; ¹² University of Durham, UK.

Table of Contents

Chapter 1 General Introduction	11
1.1. Project Aims	12
1.2. Non-Invasive Genetic Typing	12
1.2.1 How does non-invasive genetic typing work?	13
1.2.2. Technical challenges associated with non-invasive genetic typing	15
1.2.3. Examples of applications of non-invasive genetic typing	17
1.3. The Eurasian Badger	20
1.4. Objectives of the Study and Contents of the Thesis	23
1.5. Related Publications	24
Chapter 2 Reliable Microsatellite Genotyping of Badger Faecal DNA	25
2.1. Introduction	26
2.1.1. Accurate ‘genetic profiles’	26
2.1.2. The multiple-tubes approach	26
2.2. Materials and Methods	27
2.2.1. Study area	27
2.2.2. Sample collection and preservation	28
2.2.3. DNA extraction	28
2.2.4. PCR amplification	29
2.2.5. Comparison of storage and extraction methods	30
2.2.6. Probability of identity	30
2.2.7. Comparative multiple-tubes approach	31
2.3. Results	35
2.3.1. Comparison of storage and extraction methods	35
2.3.2. Probability of identity	35
2.3.3. Comparative multiple-tubes approach and assessment	37
2.4. Discussion	40
2.4.1. Comparison of storage and extraction methods	40
2.4.2. Probability of identity	41
2.4.3. Comparative multiple-tubes approach and assessment	41
2.5. Summary	43
Chapter 3 Estimation of Badger Abundance by Genotyping Faecal DNA	44
3.1. Introduction	45
3.2. Materials and Methods	46
3.2.1. Baseline population data	46
3.2.2. Collection and preservation of faecal samples; laboratory procedures	46
3.2.3. Estimation of badger abundance using mark-recapture models	47

3.2.4. Evaluation of rarefaction methods using synthetic data	48
3.2.5. Estimation of badger abundance using accumulation curves	49
3.3. Results	50
3.3.1. Faecal DNA typing and baseline estimates of abundance	50
3.3.2. Sex- and age-specific defecation patterns at hinterland latrines	51
3.3.3. Direct enumeration of faecal profiles as a means of estimating badger abundance	52
3.3.4. Estimation of badger abundance using mark-recapture models	52
3.3.5. Evaluation of rarefaction equations using synthetic data	54
3.3.6. Estimation of badger abundance using accumulation curves	54
3.4. Discussion	57
3.4.1. Sex- and age-specific defecation patterns at hinterland latrines	57
3.4.2. Baseline estimates	57
3.4.3. Estimation of badger abundance using mark-recapture models	58
3.4.4. Evaluation of rarefaction methods using synthetic data	58
3.4.5. Estimation of badger abundance using accumulation curves	59
3.5. Summary	61
Chapter 4 Estimation of Badger Abundance by Genotyping Remotely Plucked Hair	62
4.1. Introduction	63
4.2. Materials and Methods	63
4.2.1. Study site	63
4.2.2. Genetic variability of Luxembourg badgers	64
4.2.3. Field methods	65
4.2.4. DNA extraction and amplification	65
4.2.5. Reliability of hair DNA typing and identification of genetic profiles of individual badgers	67
4.2.6. Counting badgers by direct observation	68
4.2.7. Estimation of badger group sizes using accumulation curves	68
4.3. Results	68
4.3.1. Genetic variability of badgers in Luxembourg	68
4.3.2. Success of hair capture	69
4.3.3. Reliability of hair DNA typing	70
4.3.4. Genetic profiles of individual badgers	72
4.3.5. Counting badgers by direct observation and baseline estimates of badger abundance	72
4.3.6. Direct enumeration of DNA profiles as a means of estimating badger abundance	73
4.3.7. Estimation of badger group sizes using accumulation curves	73

4.4. Discussion	75
4.4.1. Genetic variability of badgers in Luxembourg	75
4.4.2. Predicting the quality of DNA extracts	75
4.4.3. Hair-sampling as a practical method of estimating population size	76
4.5. Summary	79
Chapter 5 Spatial Organisation of Badgers in a Medium-Density Population in Luxembourg and Analysis of Territorial Marking by Genotyping Faecal DNA	80
5.1. Introduction	81
5.2. Materials and Methods	83
5.2.1. Capturing badgers	83
5.2.2. Accuracy of radio-tracking fixes	84
5.2.3. Autocorrelation of fixes and radio-tracking regime	85
5.2.4. Home range estimators	86
5.2.5. Data analysis	88
5.2.6. Identification of latrines and collection of faecal samples	89
5.2.7. DNA extraction and amplification	89
5.2.8. Genotyping of faecal samples	90
5.2.9. Sex-determination using molecular markers	91
5.3. Results	92
5.3.1. Spatial organisation of badgers in Luxembourg	92
5.3.2. Collection of faecal samples	97
5.3.3. Genotyping of faecal DNA samples	97
5.3.4. Sex-determination using molecular markers	97
5.4. Discussion	100
5.4.1. Spatial organisation of badgers in Luxembourg	100
5.4.2. Genotyping of faecal samples	103
5.4.3. Territorial marking	104
5.4.4. Synthesis	105
5.5. Summary	105
Chapter 6 Overview and Conclusions	107
6.1. Main Results Presented in this Study	108
6.1.1. Estimating badger abundance by non-invasive genotyping	108
6.1.2. Investigation of territorial marking by genotyping faecal DNA	113
6.2. Discussion of Main Results and Conclusions	115
6.3. Summary of Conclusions	120
References	122
Appendices	135

UNIVERSITY OF SUSSEX

Alain Camille Frantz

Thesis submitted for the degree of Doctor of Philosophy (D.Phil.)

NON-INVASIVE GENETIC TYPING IN THE STUDY OF BADGER (*Meles meles*) ECOLOGY

SUMMARY

Despite the badger's role as an agricultural pest and carrier of bovine tuberculosis, there has hitherto been no robust and cost-effective way of estimating its abundance. The main aim of my project was to develop and test methods of estimating badger abundance by genetic profiling of DNA extracted from faecal samples and remotely plucked hairs.

Since faecal DNA is often degraded and repeated amplifications are required to obtain reliable genotypes, I developed a protocol that minimised the required number of amplifications by directing replication at similar, but not identical, profiles. Estimates of abundance generated from faecal profiles were lower than the number of animals known to be present in a high-density UK population where reliable independent estimates of abundance were available.

Hair samples were remotely plucked from badgers using barbed wire hair traps near five target setts in a population in Luxembourg. All DNA extracts that contained amplifiable DNA produced 100% accurate genetic profiles in a single round of amplification. During a pilot study aimed at estimating population density, hair samples were collected daily from five social groups during a three-week period. The absolute number of genetic profiles obtained corresponded well with independent estimates of social group size obtained by direct observation. Given the accuracy and reliability of the technique, DNA genotyping of remotely plucked hair could form the basis of a cost-effective and widely applicable methodology of estimating badger abundance.

Finally, the spatial organization of a medium-density badger population in Luxembourg was analysed using a combination of radio-tracking and faecal DNA genotyping. The pattern of space use indicated that the badgers were territorial. By comparing faecal DNA samples collected from boundary latrines with reference profiles obtained from captured badgers, I found that all members of a social group defecated in latrines and, hence, appear to contribute to territorial defence.

Chapter 1

General Introduction

1.1. Project Aims

The aim of my project was to develop a methodology for non-invasive genetic sampling of badgers (*Meles meles*), and to apply this methodology in order to estimate badger population density and to answer questions about badger social structure and territorial behaviour.

This chapter consists of three further sections. Firstly, I review the literature concerning the use of non-invasive genetic typing in the management and ecology of wild mammals. Secondly, I describe some of the management problems that have arisen in relation to badgers, and summarise the literature concerning the social and territorial behaviour of badgers. Finally, I list the objectives of the project and provide a summary of the contents of the remainder of this thesis.

1.2 Non-Invasive Genetic Typing

Conservation and wildlife managers often require demographic, behavioural, genetic and life-history information about a species of interest in order to assess its conservation status or in the context of disease and pest control. Most organisms, however, are difficult to detect with certainty and cannot easily be observed, especially if they are rare, cryptic, small, arboreal or fossorial. Furthermore, it is often difficult to capture these animals and, if the target species is sensitive or endangered, trapping and marking should not be considered since the risks of negatively affecting the survival and disturbing the dynamics of the population of the study animals is too great (Greenwood 1996, Kohn & Wayne 1997).

In the past, some of these elusive species have been studied by means of field signs. Faeces, in particular, have proved to be useful as they can provide information about the relative abundance, feeding biology, reproductive status, territory size and reproductive status of the defecators (Putman 1984, Kohn & Wayne 1997). However, mainly because of uncertainties about the identity of the defecator and the species it belongs to, there has always been a degree of uncertainty and error associated with these studies (e.g., Kendall *et al.* 1992, Hansen & Jacobsen 1999, Davison *et al.* 2002, Wilson & Delahay 2001). An alternative might be the use of radio-tracking, or genetic analyses of tissue or blood samples, but these approaches inevitably run into the problems associated with trapping (see above).

In recent years, a tool has been added to the armoury of the wildlife manager that has made it possible to solve some of the problems associated with the study of rare, endangered or cryptic species: non-invasive genetic typing. Because the technique is based on the Polymerase Chain Reaction (PCR; Saiki *et al.* 1985, Mullis & Faloona

1987, Saiki *et al.* 1988), which allows the amplification of DNA from minute amounts of source material, it is no longer necessary to acquire large amounts of fresh tissue for genetic analyses. Consequently, DNA extracted from source materials that can be obtained without catching the target animal, such as faeces or hair follicles, can now be used to generate a genetic profile unique to the animal from which the sample originated.

1.2.1 HOW DOES NON-INVASIVE GENETIC TYPING WORK?

Most comprehensive studies involving non-invasive genotyping have been based on faeces and shed or plucked hair samples because these are plentiful and relatively easy to collect in the field (for a review see Piggott & Taylor 2003b). In the case of faecal samples, DNA is extracted from epithelial cells shed from the intestinal lining (Höss *et al.* 1992). Plucked hair containing a whole follicle gives rise to better DNA extracts than shed (telogen) hairs and a number of studies have shown that freshly plucked human and chimp hair can provide enough DNA to allow straightforward genetic analyses (Higuchi *et al.* 1988, Gagneux *et al.* 1997, Morin *et al.* 2001, but see Goossens *et al.* 1998). Less frequently used non-invasive DNA sources include sloughed skin (Amos *et al.* 1992, Bricker *et al.* 1996), urine preserved in snow (Valière & Taberlet 2000), buccal cells from chewed food remnants (Sugiyama *et al.* 1993, Takenaka *et al.* 1993) and, in the case of birds, feathers (Smith *et al.* 1992, Segelbacher 2002), and eggshells (Pearce *et al.* 1997).

The PCR, which is at the basis of non-invasive genotyping, allows the amplification of an equivalent region of DNA from different individuals or species. The sequences of the fragments flanking the region to be amplified need to be known so that specific oligonucleotide primers can be designed that can be annealed on opposite strands and extended towards each other using a thermostable DNA polymerase. A PCR cycle consists of denaturing the target DNA duplexes, annealing of the primers and polymerisation of new double-stranded segments. Repeated cycles of synthesis and denaturation result in an exponential increase in the number of segments replicated. A host of specific DNA regions or genetic markers can be PCR-amplified from non-invasive DNA samples, making it possible to determine the species, individual identity and gender of the animal from which the sample originated.

Mitochondrial DNA (mtDNA) is an extra-nuclear genome that is maternally inherited, haploid and non-recombining (Moritz 1994). Because each cell has a large number of mitochondria, amplification of specific segments of mtDNA is relatively straightforward, even from non-invasively collected, degraded DNA samples (Morin *et al.* 1994, Kohn *et al.* 1995). Because of moderate intraspecific variation, PCR-amplification and DNA sequencing of the mtDNA cytochrome *b* gene can

unequivocally determine the species from which a sample was obtained (Kocher *et al.* 1989, Moritz 1994, Verma & Singh 2003). Furthermore, the mtDNA control region has enough variability to be used in intraspecific phylogeographic studies (Moritz 1994).

It is possible to identify the gender of animals by PCR typing of non-invasive samples (Reed *et al.* 1997, Woods *et al.* 1999, Dallas *et al.* 2000, Bradley *et al.* 2001, Eggert *et al.* 2003). These studies are based on the simultaneous amplification of a Y-chromosome fragment, usually the testis-determining factor gene (*SRY*), and a region on the X-chromosome or some other part of the nuclear genome included as a positive control. Samples that amplify the positive control without amplifying the *SRY* fragment are scored as females, while those that amplify both fragments are scored as males. Even though relatively little sequence information is known for the sex chromosomes of most animals, the sequences appear to be highly conserved so that it is possible to design “universal” primers that are applicable in many species of mammals (Aasen & Medrano 1990, Griffiths & Tiwari 1993).

Recent developments in molecular genetics have created new and powerful genetic markers that can be used to solve problems in many ecological disciplines: microsatellite loci (Goldstein & Schlötterer 1999). Microsatellite loci, which occur in high numbers in any eukaryotic genome, are DNA sequences that consist of short repetitive arrays typically less than five base pairs in length (*e.g.*, [AC]*n*), with a high variability due to different repeat numbers. They are embedded in unique sequences that allow specific PCR primer to be designed. By amplifying a sufficient number of different microsatellite loci from a DNA sample, it is possible to generate a multi-locus genotype unique to the specific animals (Palsbøll 1999, Taberlet & Luikart 1999). This is particularly useful for non-invasively obtained DNA extracts. The technique therefore has potential applications in population inventories, estimation of local abundances and in wildlife forensics. Because the markers are codominantly inherited, with alleles from both parents observable in the offspring, microsatellites can also be used for studying paternity and kinship, genetic variation, population genetic structure and gene flow (Queller *et al.* 1993, Kohn & Wayne 1997).

Since microsatellites are rapidly evolving non-coding regions, they can typically only be amplified from DNA samples originating from the same or a closely related species. This means that, when working with a new model organism for which no microsatellite primers have been sequenced, it might be possible to use primers that have been developed for another species in the family (*e.g.*, Schlötterer *et al.* 1991, Coote & Bruford 1996, Engel *et al.* 1996, Primmer *et al.* 1996, Slate *et al.* 1998). Significantly, however, because of the high specificity of microsatellite loci, it is unlikely that DNA from bacteria or dietary components, which would otherwise

interfere with the analyses, would be amplified (Morin *et al.* 2001, but see Bradley & Vigilant 2002).

1.2.2. TECHNICAL CHALLENGES ASSOCIATED WITH NON-INVASIVE GENETIC TYPING

Given the rapid technological advances in molecular ecology, non-invasive genetic typing appears to have a massive potential in answering questions about the biology and ecology of elusive animals. Due to the high copy numbers present in each cell, the amplification of mtDNA fragments from non-invasive samples is relatively simple. Unfortunately, the nuclear DNA obtained from non-invasive samples, especially shed hair and faeces, is often of very low quantity and quality (*i.e.*, degraded) and the extracts contain PCR inhibitors (Taberlet *et al.* 1999). These drawbacks create a host of significant problems that need to be fully addressed in any investigation based on genotyping non-invasively collected material.

Because of the degraded nature of some non-invasive DNA samples, it is necessary to choose PCR primers that amplify microsatellite loci shorter than 300 base pairs (Frantzen *et al.* 1998). This necessity can cause problems if not enough short microsatellite loci are available for a study species. Researchers working on chimpanzees (*Pan troglodytes*) and gorillas (*Gorilla gorilla gorilla*) have circumvented the problem by using human microsatellite markers to design primers that produce short fragments (Clifford *et al.* 1999, Bradley *et al.* 2000). In most other families, for which no sequence information will be available, this is not feasible and the only way around the problem is the identification of new, shorter loci.

Since the amounts of amplifiable non-invasive DNA added to a PCR are often in the picogram range (Taberlet *et al.* 1996, Morin *et al.* 2001), the sensitivity of the reaction needs to be adjusted accordingly by increasing the number of PCR cycles (Ramekers *et al.* 1997, Taberlet & Luikart 1999). As a consequence of increased PCR sensitivity, the risk of laboratory contamination is also increased. Therefore, preparation and extraction of samples must be performed under rigorous conditions of cleanliness (Gill *et al.* 2000). Furthermore, it is imperative to separate pre- and post-PCR reactions, to avoid presence of concentrated DNA extracts in the pre-PCR room, to use aerosol-resistant filter tips and to continuously monitor for DNA contamination via reaction negatives (Taberlet & Luikart 1999).

The most important problem associated with non-invasive DNA, however, is the high prevalence of genotyping errors. While it is relatively easy to obtain a PCR product from picogram amounts of DNA, it often happens that only one allele of a heterozygous individual is amplified and detected (Gerloff *et al.* 1995, Taberlet *et al.* 1996, Gagneux *et al.* 1997, Goossens *et al.* 1998, Taberlet & Luikart 1999, Taberlet *et al.* 1999, Segelbacher 2002). This type of error, called allelic dropout, is

primarily caused by stochastic sampling. At low copy numbers, a molecule that is amplified by chance during the early rounds of the PCR is likely to be preferentially amplified (Taberlet & Luikart 1999). Taberlet *et al.* (1996) and Morin *et al.* (2001) found that when working with between 17 and 100 picograms of template DNA, a concentration frequently found in non-invasive DNA samples, there was a very high risk of allelic dropout. Another problem associated with non-invasive DNA is false alleles. These artefacts generated during the amplification process can be misinterpreted as true alleles, but they are less frequent than allelic dropouts and often show unusual patterns or produce a combination of three alleles that can be recognised as erroneous (Taberlet *et al.* 1996, Gagneux *et al.* 1997, Goossens *et al.* 1998).

As pointed out by Taberlet & Luikart (1999), efficient and reliable genotyping of non-invasive DNA samples would require determination of the DNA content of the corresponding extracts. Unfortunately, this is not possible using conventional methods, because of low quantities of template DNA and additionally, in the case of faecal DNA, because of the co-purification of prey or bacterial DNA. Recently, Morin *et al.* (2001) suggested pre-screening of extracts for DNA quantity by using a 5' nuclease assay (Holland *et al.* 1991, Livak *et al.* 1995) for real-time quantitative PCR amplification (Heid *et al.* 1996, Lie & Petropoulos 1998). This approach reduces the total number of necessary amplifications by identifying good quality extracts, by allowing the elimination of extracts of very poor quality and by redirecting stringent replication towards the remaining samples. Even though this methodology would reduce the overall number of replications necessary for reliable genotyping, the primers and fluorescent labels that the technique requires are restrictively expensive, defeating the point of an improved methodology. Furthermore, the necessary real-time PCR equipment is also very expensive and often unavailable in molecular ecology laboratories in general, and low-tech laboratories in particular.

Frequently, the suitability of an extract for microsatellite genotyping is tested by amplification of a mtDNA fragment (*e.g.*, Reed *et al.* 1997, Flagstad *et al.* 1999, Taberlet *et al.* 1999, Lucchini *et al.* 2002, Eggert *et al.* 2003, Pires & Fernandes 2003). In case of amplification failure, the whole sample is discarded because mtDNA is likely to be present in greater quantities in the extract than nuclear DNA fragments. The basic approach to the problem of obtaining reliable genotypes from non-invasive DNA is to independently amplify each sample several times until the observed genotypes are considered reliable. A number of different protocols have been proposed that aim to make the corresponding methodology as simple and efficient as possible (Navidi *et al.* 1992, Taberlet *et al.* 1996, Morin *et al.* 2001, Miller *et al.* 2002). Even with the most efficient protocol, however, genotyping errors

appear hard to avoid when working with non-invasively collected DNA samples (review in Piggott & Taylor 2003b), which consequently leads to a considerable increase in the cost and duration of any such study.

In any study that uses microsatellite loci for genetic profiling, it is possible for different individuals to have an identical profile. It is therefore important to quantify the number of loci necessary to have sufficient power to distinguish between individuals with certainty. This can be achieved by computing a probability of identity (P_{ID}) statistic. Bruford *et al.* (1998) defined P_{ID} as “the probability that two individuals in a population will share a genotype by chance”. P_{ID} is first calculated for a *single* locus and the *overall* probability of match for any two individuals in a population is obtained by multiplying the P_{ID} values across all loci that have been analysed (Paetkau *et al.* 1995).

1.2.3. EXAMPLES OF APPLICATIONS OF NON-INVASIVE GENETIC TYPING

Despite these problems, non-invasive genetic sampling, because of its great potential to obtain information about elusive animals, has been establishing itself in recent years as a standard method in wildlife research and management.

Mitochondrial DNA can be amplified and sequenced with relative ease from non-invasive samples and has been used to study the distribution, population structure and phylogeography of a number of species. Working on shed and plucked hair, Morin *et al.* (1994) analysed two mitochondrial loci of chimpanzees (*Pan troglodytes*) from 20 African sites and found evidence for historically high long-distance gene flow. Furthermore, by analysing the relationships among three allopatric subspecies, they suggested that the West African *P. t. verus* was a well-differentiated and independently evolving taxon. In order to identify a source population for the restocking of brown bears (*Ursus arctos*) in the Adamello-Brenta Nature Park in Northern Italy, Kohn *et al.* (1995) analysed the mitochondrial control regions of extracts obtained from faecal and hair samples of bears in that area as well as from other parts of Europe. They confirmed the observation by Taberlet & Bouvet (1994) that brown bears in the west and the east of Europe had different evolutionary histories and found that the Slovenian population would be best suited for restocking the Brenta population. Similar studies based on non-invasive DNA samples have been performed on the gray wolf (*Canis lupus*; Valière *et al.* 2003) and on eastern gorillas (Jensen-Seaman & Kidd 2001).

Traditionally, faecal and hair samples have been assigned to species based on their morphology (Putman *et al.* 1984, Oli 1993, Raphael 1994). Contrary to received opinion, however, faeces are frequently assigned incorrectly, particularly if the host species occurs at very low densities (Halfpenny & Biesot 1986, Davison *et al.*

2002, Palomares *et al.* 2002). The hair of closely related species is also difficult to distinguish by macro- and microscopic means (Brunner & Coman 1974, Woods *et al.* 1999, Lobert *et al.* 2001). Misidentification of faecal samples collected in the field can lead to serious errors when analysing the diet of a species. In recent years, researchers have avoided this problem by determining the host species with molecular methods. For example, amplification of the mitochondrial cytochrome *b* gene and enzymatic restriction at specific sites has allowed the separation of faecal samples from otter (*Lutra lutra*), American mink (*Mustela vison*) and polecat (*Mustela putoris*; Hansen & Jacobsen 1999). Similarly, because the jaguar (*Panthera onca*) and the puma (*Puma concolor*) are sympatric across much of their respective ranges, Farrell *et al.* (2000) needed to determine the origin of a faecal sample by amplification of a region of the cytochrome *b* gene before analysing the feeding ecology of both species. In addition to determining the identity of the defecator, it is also possible to identify prey or plant remains in scats by molecular means (Höss *et al.* 1992, Taberlet & Fumagalli 1996, Symondson 2002, Banks *et al.* 2003).

Because most traditional census methods are not readily applicable to rare or elusive species, non-invasive genotyping has proved particularly useful in population inventories and in the estimation of local abundance. Problems with the traditional methods include the facts that population counts based on trapping and marking have the potential to affect the survival of the study animals (Greenwood 1996) and disturb the dynamics of the population that is under investigation (Kohn & Wayne 1997). Furthermore, the mathematical models used to analyse mark-recapture data require large sample sizes, high capture probabilities and frequent trappings to provide reasonable population estimates (Otis *et al.* 1978, White *et al.* 1982, Rosenberg *et al.* 1995). Indirect methods based on field signs only provide estimates of relative abundance (Kendall *et al.* 1992). Using non-invasive genotyping, on the other hand, robust estimates of population sizes can be generated without catching, handling or disturbing animals.

In recent years, for example, population inventories based on DNA extracted from remotely plucked hair have become a routine procedure in North America for mustelids (Foran *et al.* 1997, Mowat & Paetkau 2002) and bear species (Woods *et al.* 1999, Mowat & Strobeck 2000, Poole *et al.* 2001, Paetkau 2003). In the case of the mustelids, DNA was obtained by means of baited glue traps, while scented barbed-wire enclosure hair traps were used to snare bear hair. In these studies, hair samples are usually pooled to increase the DNA yield and reduce the frequency of genotyping errors. The study areas were subdivided into grids in order to allow a strategic and systematic distribution of the hair traps, and estimates of the abundance of the species under investigation were obtained by applying mark-recapture analyses (Otis

et al. 1978, White *et al.* 1982) to the data.

Microsatellite genotyping of faecal samples, in addition to sex-determination by gender-specific genetic markers, has frequently been used to estimate the abundance of elusive carnivores. Similarly to analyses of feeding biology, mtDNA fragments are often amplified to determine the species of the defecator (*e.g.*, Paxinos *et al.* 1997, Palomares *et al.* 2002, Pires & Fernandes 2003). This technique has been used to obtain the minimum size of endangered populations by counting the number of different reliable microsatellite profiles obtained from dung collected in the field (Kohn *et al.* 1995, Taberlet *et al.* 1997, Ernest *et al.* 2000). Other studies apply the technique to animals that are not necessarily endangered but are very difficult to census by conventional methods, either due to logistical problems or because the species is difficult to observe in its natural habitat, as is the case for forest-dwelling and fossorial animals (Kohn *et al.* 1999, Banks *et al.* 2002, Eggert *et al.* 2003). In these studies, enough genetic profiles were obtained to estimate the true size of the population using statistical techniques such as rarefaction (Kohn *et al.* 1999, Eggert *et al.* 2003) or mark-recapture analyses. An additional advantage of estimating population size by means of plucked hair or faeces is that the genetic variability of the study population can be analysed (*e.g.*, Lucchini *et al.* 2002, Boersen *et al.* 2003, Eggert *et al.* 2003)

Non-invasive genetic sampling is also on its way to becoming a standard tool in behavioural biology, because the technique makes it possible to analyse paternity and the relationships of members of a social group without needing to catch and disturb the target animals. The technique has been used quite frequently to analyse paternity, mating patterns and relatedness in primate species (*e.g.*, Gerloff *et al.* 1999, Constable *et al.* 2001, Vigilant *et al.* 2001). Faeces are an interesting source of DNA from species that are difficult to capture. For example, faecal DNA was used in the analysis of the mating systems and social organization of wild black rhinoceros (*Diceros bicornis*; Garnier *et al.* 2001) and Asian elephants (*Elephas maximus*; Fernando & Lande 2000). Because they can easily be attributed to individual animals, Parsons *et al.* (2003) extracted DNA from bottlenose dolphin faeces to show that kin selection played a significant role in the formation of alliances between reproductive males.

Despite the scepticism of some early reviews about the practicality of this technology (*e.g.*, Taberlet & Waits 1998, Taberlet *et al.* 1999), the potential of non-invasive genetic typing simply is too great to be ignored. To demonstrate this, I searched the Web of Knowledge for the following key-words: shed hair, plucked hair, hair and DNA, non-invasive DNA, noninvasive DNA, noninvasively and DNA, faecal DNA, faeces and DNA, feces and DNA, dung and DNA, excremental PCR, and sloughed

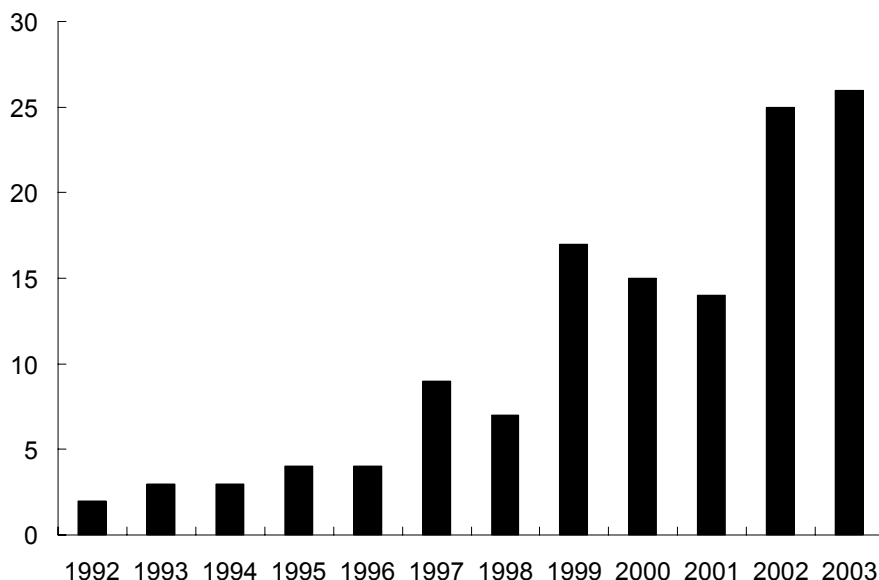


Fig. 1.1: Number of scientific publications per year that deal with non-invasive genetic sampling. The analysis was performed by means of a key-word search in the Web of Knowledge.

skin. Not considering papers that deal with the bacterial fauna of stool samples, 129 relevant papers were found, with a clear increase in the number of papers per year over the last decade (Fig. 1.1). Thus, after a slow start due to the technical problems, the methodology is quickly establishing itself as standard in behavioural studies, ecology, conservation and wildlife management.

1.3. The Eurasian Badger

The Eurasian badger is a medium-sized carnivore of the Mustelidae family (Neal & Cheeseman 1996). Despite the existence of a number of subspecies, the badger is generally considered to be the only species of the genus *Meles* (Henry *et al.* 1988, Neal & Cheeseman 1996). It is a semi-fossorial mammal that forages at night and spends daylight hours in an underground den known as a sett (Harris 1982, Neal & Cheeseman 1996, Rodríguez *et al.* 1996, Kowalczyk *et al.* 2003a). Badger setts can vary significantly in size, ranging from simple burrows with a single entrance to very long and complex tunnel systems with multiple entrances and underground chambers (Roper 1992).

Badgers are very widely distributed across the Western Palaearctic, ranging from Ireland and Spain in the West to Southern China and Japan in the East as well as from Scandinavia in the North to the Middle East in the South (Long & Killingley 1983). Despite occurring all over Europe, except in regions of very high latitude or altitude and on a number of islands, the densities of badger populations vary between regions (Griffiths & Thomas 1993, Prigioni 1999). For example, according to Griffiths & Thomas (1993, 1997), less than 0.1 badger/km² can be found in the Netherlands, Belgium and Estonia. In Ireland, Sweden and Britain, on the other hand, abundances are very high with, on average, more than 1 badger/km². Generally speaking, badger occurrence and densities are influenced by habitat type in general and the availability of food and suitable sett sites in particular (Kruuk & Parish 1982, Feore *et al.* 1991, O’Corry-Crowe *et al.* 1993, Roper 1993, Smal 1995, Feore & Montgomery 1999, but see Johnson *et al.* 2002).

Even though some authors regard badgers as earthworm specialists (*e.g.*, Kruuk 1989), it is generally accepted that the badger is an omnivorous and opportunistic feeder able to exploit any locally abundant food source (Henry *et al.* 1988, Roper 1994). This characteristic often brings the animal into conflict with humans. Indeed, the species can consume large quantities of cereals, maize (*Zea mays*) and other crops and is often considered to be an agricultural pest (Lüps 1988, Roper *et al.* 1989, Roper *et al.* 1995, Moore *et al.* 1999, Schley 2000). In certain parts of Europe, notably Britain and Ireland, the badger also conflicts with human interests because it is a wildlife reservoir of bovine tuberculosis (BTB: *Mycobacterium bovis*). The species has been the focus of much debate and research due to its alleged role in the transmission of this disease to cattle (Krebs *et al.* 1997). Bovine tuberculosis is a disease that is now judged by the government to constitute a serious problem, costing the British taxpayer £74 million a year (DEFRA, pers. comm.) and causing personal cost to farmers and welfare costs to cattle (Krebs *et al.* 1997, Agriculture Committee 1999). A recent independent review (Krebs *et al.* 1997) reported a variety of circumstantial evidence suggesting that badgers were responsible for transmission of BTB to cattle. Research activities focus presently on the development of effective strategies to reduce the prevalence of TB in cattle herds by means of targeted culling of badgers, improvements to farm bio-security and vaccination.

Because of the badger’s role as an agricultural pest and because of its implication in the transmission of bovine tuberculosis, accurate estimates of badger abundance are required. Like many mammals, however, badgers are very difficult to census accurately, owing to their semi-fossorial and nocturnal lifestyle. Estimates based on live-trapping and mark-recapture analysis can potentially provide accurate results (*e.g.*, Rogers *et al.* 1997b, Tuytens *et al.* 1999) but this approach is time-consuming,

labour-intensive and expensive. Indirect methods based on badger field signs have been shown to provide poor estimates of group size (*e.g.*, Tuytens *et al.* 2001, Wilson *et al.* 2003). Since badger groups usually share only one main sett, and main setts are relatively easy to recognise, the number of groups inhabiting a given area can be determined with a reasonable degree of accuracy by counting the number of main setts (*e.g.*, Harris *et al.* 1989, Ostler & Roper 1998). What is lacking, however, is a reliable way of estimating group size. Indeed, the UK Department for Environment, Food and Rural Affairs (DEFRA) has recently called for research into innovative ways of censusing badgers.

For a long time, badgers were believed to be a solitary species with non-overlapping home ranges (Kruuk 1989). Extensive work over the last 25 years, however, has shown that in high-density populations in the United Kingdom, Ireland and Sweden the animals live in stable social groups of up to 20 individuals that, despite being solitary foragers, inhabit and defend clearly defined territories (for reviews see Kruuk 1989, Neal & Cheeseman 1996, Krebs *et al.* 1997). In these populations, individuals use communal defecation sites throughout the territory (Kruuk 1978b). These “latrines” are thought to be sources of information transfer via scent marking and thereby to play a role in maintaining the animals’ spacing system (Roper *et al.* 1986, Roper *et al.* 1993). Latrines have been classified into two types based on their location and their “patrons”. While “hinterland” latrines are used exclusively by members of the same social group, “boundary” latrines are shared by members of the same and neighbouring groups (Roper *et al.* 1986). Furthermore, while hinterland latrines are located in the core of a territory, boundary latrines largely define the territories of neighbouring groups (Kruuk 1978b, Roper *et al.* 1986).

Because cooperation between individual badgers has rarely been observed, it has been difficult to understand the advantages of group living in this species (Woodroffe & Macdonald 2000). Researchers have therefore tried to explain the formation of non-cooperative groups in terms of group territoriality (*e.g.*, Macdonald 1983, von Schantz 1984, Kruuk & Macdonald 1985). In these models, as Woodroffe & Macdonald (1992) put it, “the rule a territory holder uses to choose its territory allows additional “satellite” animals to be accommodated at little or no cost to the primary”. If territorial behaviour was a necessary precursor of sociality, then the question arises why badgers form territories in the first place. Defence of females or food resources are usually put forward as the underlying reasons for group territoriality in badgers (Kruuk 1978a, Kruuk & Parish 1982, Macdonald 1983, Roper *et al.* 1986, Kruuk 1989, Stewart *et al.* 1997).

In order to gain a better understanding for badger territoriality, however, it is important to analyse which animals are involved in territorial defence. It has been

reported that in captivity, breeding males carried out most of the scent marking (Kruuk *et al.* 1984). Similarly, after monitoring radio-collared badgers, Roper *et al.* (1993) concluded that females visited boundary latrines on fewer occasions than males and that, therefore, territorial defence fell chiefly to the males. These results are in line with the hypothesis that territories are set up by males to defend access to females. However, Stewart *et al.* (2002) marked individual badgers with fur clips and monitored one hinterland and four boundary latrines by means of infrared video equipment for 7 to 30 days. They concluded that all age and sex classes defecated at boundary latrines; a result in line with the hypothesis that territoriality serves chiefly to defend food resources. Because of this contradiction between two apparently valid studies, further research is needed to clarify the issue of defecation at boundary latrines. Also, the study of Stewart *et al.* (2002) was performed in Wytham Woods, UK, a population where densities are atypically high (38 adults / km²).

Another problem is that, generally speaking, any hypothesis aiming to explain the social organisation of badgers in general and territory formation in particular has to take inter-population variability into account. A disproportionate number of studies have been conducted in the UK where environmental conditions are not typical of all badger populations (Johnson *et al.* 2002). Indeed, within their wide geographic distribution, badger populations appear to greatly vary in their behaviour, physiology and degree of prey specialisation, as well as in social organization, which ranges from territorial groups to a solitary lifestyle (Woodroffe & Macdonald 1993, Johnson *et al.* 2002). Also, the density of badger populations is low or moderate throughout most of Central and Eastern Europe (Griffiths & Thomas 1997). Because the conclusions and hypotheses based on studies performed in the United Kingdom might not be relevant to continental medium- and low-density populations, there is a need for studies of badger behaviour and ecology in other parts of the species' geographic range, and especially in central Europe.

1.4. Objectives of the Study and Contents of the Thesis

The main objective of my study was to apply non-invasive genetic sampling to questions relating to the management and ecology of the Eurasian badger. Specifically, I aimed to develop an accurate non-invasive method of determining social group sizes in the Eurasian badger and to analyse territorial marking in a medium-density population on the European mainland.

The ideal source material for non-invasive genotyping should not only be abundant and easy to collect, but also allow the generation of good-quality DNA extracts. In the case of badgers, the only potential sources of non-invasively collected DNA samples are faeces and hairs. Regarding DNA-quality, there was no *a priori* preferred source

of DNA since genotyping errors have been reported from shed and plucked hairs as well as from faeces (e.g., Taberlet *et al.* 1996, Gagneux *et al.* 1997, Goossens *et al.* 1998). I therefore tried to estimate badger abundance by genotyping both faecal and remotely plucked hair DNA. After developing both techniques, they were applied in the study of territorial marking in a medium-density population in continental Europe. This was done by comparing DNA profiles generated from faeces collected from boundary latrines to reliable genetic profiles obtained from captured badgers or remotely plucked hair samples.

In Chapter 2, I summarise the steps I took to obtain reliable microsatellite genotypes from faecal DNA. I compare the performance of my methodology to a standard multiple tubes approach by Taberlet *et al.* (1996) and the more recent maximum-likelihood approach by Miller *et al.* (2002). In Chapter 3, I describe a pilot study of the estimation of badger abundance using faecal DNA. Faecal samples were collected in Woodchester Park, UK, a study area for which robust independent estimates of badger abundance were available. In Chapter 4, I present a pilot study in which I estimate badger abundance by genotyping remotely plucked hair. This study involved five social groups in Luxembourg, the sizes of which were independently determined by direct observation. This chapter also contains an analysis of the variability of badger microsatellite loci in Luxembourg. Chapter 5 presents an analysis of territorial marking in my study population in Luxembourg, based on genotyping faecal and hair DNA. In addition, the spatial organisation of the badgers was determined by a radio-tracking study. Finally, in Chapter 6, I present an overview of all the results, linking different aspects and problems encountered during the study, and finish with conclusions and recommendations for further studies.

1.5. Related Publications

Some of the results contained in this thesis have been published or submitted for publication in the following papers:

- 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. *Molecular Ecology* **12**, 1649-1661.
- Wilson GJ, Frantz AC, Pope LC, Roper TJ, Burke TA, Cheeseman CL, Delahay RJ (2003) Estimation of badger abundance using faecal DNA typing. *Journal of Applied Ecology* **40**, 658-666.
- Frantz AC, Schaul M, Pope LC, Fack F, Schley L, Muller CP, Roper TJ (in press) Estimating population size by genotyping remotely plucked hair: the Eurasian badger. *Journal of Applied Ecology*.

Chapter 2

Reliable Microsatellite Genotyping of Badger Faecal DNA

2.1. Introduction

For various reasons, including the badger's role as an agricultural pest and its potential role in the transmission of bovine tuberculosis to cattle in certain parts of Europe, accurate estimates of the local abundance of badgers are required (see Chapter 1). Recent developments in molecular genetics have created new methods that allow populations to be censused through non-invasive DNA sampling, using microsatellite loci to establish a genetic profile. Because of the non-ephemeral and conspicuous nature of badger latrines (Roper *et al.* 1993, Neal & Cheeseman 1996), faeces should provide a plentiful non-invasive source material that is easy to collect and to attribute to an individual animal. The objective of this chapter was to develop a protocol for obtaining reliable microsatellite profiles from badger faeces.

2.1.1. ACCURATE 'GENETIC PROFILES'

Several studies have shown that the amplification success of microsatellite loci from faecal DNA extracts can depend on both the faecal preservation method and the extraction method (Wasser *et al.* 1997, Frantzen *et al.* 1998, Murphy *et al.* 2000). Optimal preservation techniques can vary between species and ecological conditions (Frantzen *et al.* 1998), making it necessary to perform trials for each new faecal study. Accordingly, the first aim of this chapter was to optimise a technique for extracting DNA from badger faeces. This was achieved by comparing the efficacy of three different storage methods and two extraction methods.

2.1.2. THE MULTIPLE-TUBES APPROACH

Faecal DNA extracts are generally of low quantity and quality, which causes a high prevalence of errors such as "allelic drop-out" (ADO; Gagneux *et al.* 1997, Bayes *et al.* 2000) and "false alleles" (FA; Taberlet & Luikart 1999, see also Chapter 1). The multiple-tubes approach proposed by Navidi *et al.* (1992) is the standard protocol used to obtain reliable genotypes from faecal DNA (Taberlet *et al.* 1996, Goossens *et al.* 2000, Constable *et al.* 2001, Garnier *et al.* 2001) and assumes a worst-case scenario for allelic drop-out (hereafter referred to as worst-case rule, WCR, following Miller *et al.* 2002). Assuming each allele is equally likely to drop out, in the WCR every amplification of a heterozygous locus will give rise to one allele only, or, in other words, the probability of obtaining only one of the two alleles of a heterozygote is 0.5 in each reaction. Reliable genotypes are obtained by recording an allele only if it has been observed at least twice (in at least three amplification reactions) and by only recording an individual locus as homozygous if a certain number of positive amplification reactions gave rise to the same allele (for a single locus $n \geq 8$ for 99% confidence; Miller *et al.* 2002). This approach is reliable

but very conservative, requiring large numbers of amplifications to obtain correct genotypes. In practice, few researchers working on low-concentration DNA strictly follow the WCR (Gagneux *et al.* 1997, Gerloff *et al.* 1999, Kohn *et al.* 1999).

A recent study suggested that genotype reliability should be assessed using a maximum likelihood method (hereafter referred to as maximum likelihood rule, MLR) that allowed replication to be directed at those genotypes most likely to be erroneous (Miller *et al.* 2002). The reliability of a multilocus genotype (after a specific number of initial rounds of amplification) was determined by weighting the probability of occurrence of sequential dropout errors by the probability that the locus is heterozygous multiplied across the observed homozygous loci. The dropout rate for a specific sample was estimated by determining the dropout rate that makes the observed results most likely. When allelic dropout rates were low, theoretical consideration showed that the method had the potential to reduce the number of PCR amplifications by up to 50% compared with the WCR approach. The efficiency of the method has not yet been empirically tested, however. Furthermore, the MLR is based on three assumptions: (i) both alleles at a heterozygous locus are equally likely to drop out, (ii) allelic dropout rates are even across loci, and (iii) all false alleles can be detected and eliminated from the dataset. However, in order to avoid erroneous genotypes, it is important to know whether these assumptions are met in the data set under investigation.

The second aim of this chapter was to modify the multiple-tubes approach of Taberlet *et al.* (1996) so as to achieve a reduction in the number of amplifications without a significant reduction in power. In addition, I compared in retrospect the efficiency of our new approach with that of the WCR and MLR approaches, and tested the assumptions of the MLR approach.

2.2. Materials and Methods

2.2.1. STUDY AREA

The work summarised in this chapter and the next was performed on faecal samples collected from latrines in a high-density badger population at Woodchester Park, Gloucestershire, UK. Woodchester Park is situated on the Cotteswold sandstone escarpment between 47 m and 210 m above sea level and covers 11 km². The hilly landscape consisted mostly of permanent pasture and deciduous woodland, with smaller areas of arable land, rotational grassland, scrub, coniferous and mixed woodland. A total of 36 social groups have been recorded within the study area and the locations of the corresponding territories can be found in Tuytens *et al.* (2000) and Delahay *et al.* (2000b). As a result of an ongoing long-term study involving

regular live-trapping, the population was well known in terms of abundance, life history and epidemiology (Rogers *et al.* 1997a, Rogers *et al.* 1997b, Rogers *et al.* 1999, Delahay *et al.* 2000b). During this ongoing long-term study, individuals were regularly captured using steel mesh box traps (Cheeseman & Mallinson 1979) and permanently marked for individual identification (Rogers *et al.* 1997b). At each individual's first capture, a blood or hair sample was taken for use in genotyping of the trapped badger.

2.2.2. SAMPLE COLLECTION AND PRESERVATION

Faecal material was taken from the surface of individual droppings using toothpicks and was immediately placed in 1.5 or 2.0-ml screw-cap microfuge tubes used for DNA extraction (see below), to minimise the handling of samples. In order to allow fresh faecal deposits to be identified, all droppings at the relevant latrines were dusted with builder's chalk (Stanley Tools, Connecticut, USA) on the previous day. In addition, samples were collected at first light, in order to avoid prolonged exposure to the atmosphere (Jansman *et al.* 2001).

Samples for the storage and extraction trials were collected from 16 badger droppings in July 2001. Six aliquots were taken from each sample in order to test three different storage methods in combination with two DNA extraction protocols. The storage methods were: (1) buffering in DETs (20% DMSO, 0.25 M EDTA, 100 mM Tris – pH 7.5 and NaCl to saturation; Seutin *et al.* 1991); (2) buffering in 70% ethanol; and (3) freezing at –20°C. Samples were frozen immediately after collection but had to be transported during a three hour journey from the field site to the laboratory.

The same faecal samples that were collected for estimating badger abundance (see Chapter 3) were also used for the work presented in this chapter. Each day during a 10-day period in October 2001, a sample was taken from every overnight dropping (N=53) deposited at latrines less than 30 m away from the active setts of three adjoining social groups: Parkmill (34 samples), Kennel (9 samples) and Nettle (10 samples). In accordance with the results of the storage and extraction trials (see below), aliquots of the faecal samples were stored in 70% ethanol and extracted using the Guanidine Thiocyanate (GuSCN) / silica method. To verify the results from the faecal study, hair or blood DNA was also extracted from 36 individuals that had been captured in 2000 and 2001 in the three target social groups.

2.2.3. DNA EXTRACTION

In order to avoid contamination of the faecal samples, all extractions were performed in a separate laboratory that was free of concentrated badger DNA or PCR product. Aerosol-resistant pipette tips were used in all manipulations. Negative controls were

included in each extraction to monitor contamination. Faecal samples that had been frozen or stored in a DETs solution were potentially infected with *M. bovis*. These samples were extracted in a Category 3 containment laboratory. In order to evaporate the supernatant, samples stored in 70% ethanol and DETs solution were placed overnight in a heating block at 45°C. Two faecal extraction methods were tested: the GuSCN / silica method (Boom *et al.* 1990, Höss & Pääbo 1993) and extraction with the QIAamp DNA Stool Mini kit (Qiagen). The Qiagen kit is an adaptation of the GuSCN / silica method.

For the GuSCN / silica method, between 400 and 600 mg of wet faecal material was suspended in 1 ml of extraction buffer (5 M GuSCN, 0.1 M Tris-HCl pH 6.4, 0.02 M EDTA pH 8.0 and 1.3% Triton X-100) and incubated overnight at room temperature with rotation. Extracts were then centrifuged for 10 min at 13,000 g, the supernatant was added to 20 µl of silica matrix and the mixture was vortexed and incubated for 10 min at room temperature with agitation. The silica matrix was washed twice with 500 µl of washing buffer (5 M GuSCN, 0.1 M Tris-HCl pH 6.4, 0.02 M EDTA pH 8.0) and twice with 500 µl of ethanol washing buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA and 50% ethanol). The pelleted silica was dried in a heating block at 56 °C for 15 min and the DNA was eluted by incubation with ddH₂O for 10 min in a heating block at 56 °C. The extractions with the Qiagen kit were carried out according to the manufacturer's instructions. The only modification to the recommended protocol was that, instead of 180-220 mg, between 400 and 600 mg of faecal sample was added to the ASL buffer.

DNA was extracted from hair or blood samples of the badgers captured in the three social groups under investigation. Hair samples had been stored in 70% ethanol and were extracted using a chelex protocol (Chelex 100, Bio-Rad; Walsh *et al.* 1991). At least 10 hairs were used in each extraction (Goossens *et al.* 1998). Blood samples were extracted using a slightly modified version of the phenol:chloroform method (Sambrook *et al.* 1989, Bruford *et al.* 1998).

2.2.4. PCR AMPLIFICATION

PCRs were prepared using aerosol-resistant pipette tips in a laboratory that was free of concentrated badger DNA or PCR product. Reagents were always tested for contamination by including a PCR negative control. Frantzen *et al.* (1998) have shown that amplification success of faecal DNA will be reduced for microsatellite loci with alleles longer than 300 base pairs (bp). From the 39 microsatellite loci published by Carpenter *et al.* (2003), seven loci with alleles shorter than 250bp were chosen for this study: *Mel-102*, *Mel-105*, *Mel-106*, *Mel-109*, *Mel-111*, *Mel-113*, *Mel-117*. The microsatellite loci were amplified in a 25-µl volume, each containing 5 µl of

DNA extract. The final reaction concentrations consisted of 75 mM Tris-HCl (pH 8.8), 20 mM $(\text{NH}_4)_2\text{SO}_4$, 2.5 mM MgCl_2 , 0.15 $\mu\text{g}/\mu\text{l}$ bovine serum albumine (BSA), 0.01 % of Tween, 100 μM of each dNTP, 0.2 μM of primer and 0.6 Units of *Taq* Polymerase (ABgene).

Microsatellite loci were amplified either using a touchdown profile (Don *et al.* 1991; *Mel*-102, *Mel*-106, *Mel*-109, *Mel*-111, *Mel*-113) or with a specific annealing temperature (*Mel*-105, *Mel*-117). All PCRs started with a 5-min denaturation at 95 °C. This was followed by either touchdown cycles of 95 °C for 30 s, annealing at 64-52 °C for 30 s and 72 °C for 30 s, decreasing the annealing temperature by 2 °C every other cycle for 14 cycles then 30 cycles of holding the annealing temperature at 50 °C; or 55 cycles with a specific annealing temperature (*Mel*-117: 55°C; *Mel*-105: 56 °C). PCRs ended with a final extension at 72 °C for 5 min. Reactions were performed using a Hybaid Touchdown Thermal Cycler. Primers were end-labelled with a fluorescent dye (Carpenter *et al.* 2003) and amplification products were separated on a 5% polyacrylamide gel using an ABI 377 DNA sequencer, and sized with a TAMRA- or ROX-labelled size marker with bands of known size every 50bp. All gels were analysed using GENESCAN 2.0 and GENOTYPER™ 1.1 software.

2.2.5. COMPARISON OF STORAGE AND EXTRACTION METHODS

For the six trials of combined storage and extraction methods, each of the six aliquots from each of the 16 samples was extracted. Each aliquot was amplified once with the seven different primers and the proportion of these seven reactions that produced a PCR product was calculated. To compare the six trials, a two-way analysis of variance (ANOVA) was applied on ranked data, using the Scheier-Ray-Hare extension of the Kruskal–Wallis test (Dytham 1999) with the proportion of successful amplifications (of seven) as the dependent variable and the storage methods and extraction techniques as factors. Amplifications were deemed successful if a PCR product of the expected size was present, even if the genotype may not have been reliable.

2.2.6. PROBABILITY OF IDENTITY

When using microsatellite loci to establish a genetic profile, it is possible for different individuals to have identical profiles if an insufficient number of loci have been used. Mills *et al.* (2000) showed that, in order to be useful in population size estimations, genetic profiles should consist of enough microsatellite loci to distinguish between individuals with 99% certainty. Estimating the required number of loci can be achieved by computing probability of identity (P_{ID}) statistics. For a *single* locus, the probability of identity is calculated as follows (Paetkau & Strobeck 1994, Paetkau *et al.* 1995, Paetkau *et al.* 1998):

$$P_{ID} = \underbrace{\sum_i p_i^4}_a + \underbrace{\sum_i \sum_{j>i} (2p_i p_j)^2}_b \quad (1)$$

where p_i, p_j = frequencies of i th and j th allele, respectively. Part (a) of the equation gives the sum of squares of the expected frequencies of the homozygous genotypes, while part (b) gives the same result for the heterozygous genotypes. The *overall* probability of match for any two individuals in a population is obtained by multiplying the P_{ID} values across all loci that have been analysed (Paetkau 1995).

The estimation of match probabilities assumes independence of alleles at particular loci, and thus that the genotypes are in Hardy-Weinberg equilibrium, as well as independence among alleles from different loci (Donnelly 1995). These assumptions, however, may not hold in natural populations due to such factors as population substructure, non-random mating, genetic drift and natural selection (Taberlet & Luikart 1999). It is therefore likely that P_{ID} -values obtained with the above statistics will be an underestimate of the actual match probability in natural populations (Donnelly 1995, Waits *et al.* 2001). Where there is the potential for relatives to be present in the sample, it is best to use an estimate of P_{ID} among siblings (P_{ID-Sib} ; Evett & Weir 1998, Waits *et al.* 2001):

$$P_{(ID)sib} = 0.25 + (0.5 \sum p_i^2) + [0.5(\sum p_i^2)^2] - (0.25 \sum p_i^4) \quad (2)$$

where p_i = frequency of the i th allele. The overall P_{ID-Sib} is the upper limit of the possible ranges of P_{ID} in a population and will thus provide the most conservative number of loci required to resolve all animals, including relatives.

P_{ID-Sib} was calculated using a data set of genotypes obtained from the blood or hair DNA of 36 badgers captured in 2000 and 2001 in the three target social groups. P_{ID-Sib} values were estimated using the program GIMLET v.1.3.2 (Valière 2002), after arranging loci in order of increasing value of P_{ID-Sib} . PROB-ID5 (G. Luikart unpublished) was used to estimate the observed P_{ID} (P_{ID-Obs}) by computing the proportion of all possible pairs of individuals that had identical genotypes.

2.2.7. COMPARATIVE MULTIPLE-TUBES APPROACH

Faecal samples were scored using a comparative method, based on the WCR approach (Taberlet *et al.* 1996; see Fig. 2.1). I retained the rule that an allele was accepted only if it had been recorded at least twice. However, rather than initially performing three positive PCRs, samples were amplified twice. Loci that gave rise to the same heterozygous genotype twice were then accepted. After this, a stepwise amplification was introduced until each allele was observed at least twice. This stepwise process was continued for a maximum of seven positive PCRs and

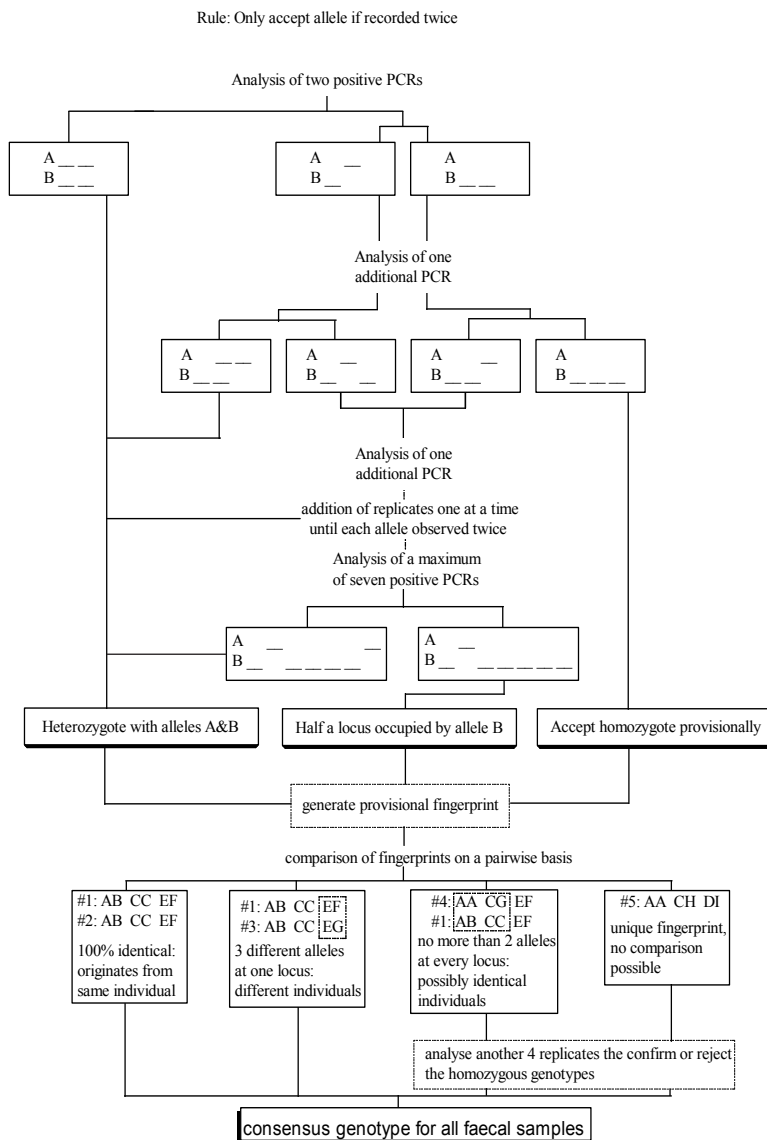


Fig. 2.1: Flow chart of the comparative multiple-tubes approach used in this study, modified from Taberlet *et al.* (1996)

contrasted with the standard multiple-tubes approach where, after the first three positive amplifications, a further four PCRs were performed as a block. In the ambiguous case where, after seven positive amplifications, a locus yielded one heterozygous result and the same homozygote for the six other reactions, I followed the suggestion of Miller *et al.* (2002) and counted it as a half-locus occupied by the allele observed in the homozygote.

The main difference between the two approaches lay in the second rule that dealt with homozygote genotypes. In my study, a number of faecal DNA extracts were expected to originate from the same defecator, allowing genetic profiles to be compared to identify typing errors. Given this, a homozygote was *provisionally* accepted after three positive PCRs gave rise to the same allele. I then compared these provisional profiles using the program GIMLET 1.3.2 (Valière 2002). Provisional profiles that were shown to be 100% identical were grouped together and classified as originating from the same individual. Incomplete profiles were only considered if a consensus genotype was obtained at, according to the P_{ID-Sib} statistics, the most informative locus (*Mel-105*). This made it possible to match them by hand to the only possible candidate group. Although incomplete profiles may have belonged to a new individual, I grouped them to matching complete ones in order to avoid the identification of nonexistent individuals. The M_h -Jackknife estimator utilized for estimating population size (see next chapter) is robust when dealing with this type of error, or “shadow effect”.

After grouping the genetic profiles, pair-wise comparisons of the different groups were then performed. If three different alleles were observed at a specific locus in a pair-wise comparison, the groups were declared different. If, however, in no case were there more than two different alleles, the profiles in the group could potentially originate from one defecator (Fig. 2.1). In this case, the potentially homozygous loci were replicated a further four times. Assuming the worst-case scenario, the probability of falsely accepting a single locus to be homozygous after seven independent replicates gave rise to the same homozygous allele is 1.6% (Taberlet *et al.* 1996, Miller *et al.* 2002). Again assuming the WCR, if there was more than one homozygous genotype in a unique profile that consisted of seven loci that had been replicated seven times, the probability of generating a false multi-locus profile due to allelic drop-out would vary from 3.1% (two homozygous genotypes) to 10.4% (seven homozygous genotypes; see Miller *et al.* 2002). However, because the WCR is unlikely to be appropriate in a real study and in order to reduce replication, seven replicate PCRs were judged to be sufficient to confirm homozygous status at the relevant loci with an acceptable amount of error.

GIMLET 1.3.2 (Valière 2002) was used to verify the accuracy of the complete faecal

profiles, and thus the power of our comparative approach, by comparing them with molecular profiles obtained from hair and blood samples of 36 badgers captured in the three target social groups. Incomplete faecal profiles were compared with the reference profiles by hand. The economy of the comparative approach relative to the WCR and MLR methods was tested by comparing the total number of reactions the three methods would require to obtain consensus multi-locus profiles.

The MLR model estimates the probability that a genotype is correct, *i.e.*, its reliability, and suggests a replication protocol if the estimate is below a certain threshold. Because the study was not designed to apply the MLR model, the technique was applied in retrospect to the data set using the program RELIOTYPE (C. Miller, unpublished). Only loci that gave rise to a consensus genotype using the comparative approach in the actual study were considered. For every locus two initial replicates were added to the input file. These replicates corresponded to the first two obtained in the actual study. Obviously false alleles were left out of the input files but were included in counts of total number of replicates performed. The reliability of the initial replicates was then estimated and a suggested replication strategy obtained. A multiple-sample correction was applied in order to limit the sample-wide number of genotype errors to <5% with 95% probability. The total number of replicates required if the suggested replication strategy had been followed to the full extent was calculated. In the laboratory, it will often be more practical to add replicates *en masse*, so that these results should provide a general idea of the performance of the MLR model with a real data set. I simulated the number of reactions needed to achieve the specified reliability given an upper confidence bound on the dropout rate of 75%, both using reliability criteria alone and using reliability criteria in addition to the condition that alleles need to be observed at least twice before being recorded.

I tested the assumptions of the MLR method for our data set on a *post hoc* basis by checking for the presence of genotyping errors in all the amplification reactions that gave rise to a consensus heterozygous genotype. Cases in which seven amplifications at a locus yielded one heterozygous result and the same homozygous during the six other reactions were excluded from this error analysis. For each locus, it was noted whether the long or the short allele did not amplify if an allelic dropout had occurred. A Mann–Whitney *U*-test was then used in SPSS 9.0 (SPSS Inc.) to test whether both alleles are equally likely to drop out. To test whether there was a difference in the allelic dropout rate among loci, a general linear mixed model (GLMM) was fitted using SPLUS 6.1 (Insightful). As the response variable was binary (drop-out/no drop-out), a binomial error structure was assumed. Locus was fitted as a fixed effect (seven-level factor) and sample was fitted as a random effect (46 levels).

2.3. Results

2.3.1. COMPARISON OF STORAGE AND EXTRACTION METHODS

A two-way analysis of variance (ANOVA) was applied on ranked data, using the Scheier-Ray-Hare extension of the Kruskal-Wallis test, with the proportion of successful amplifications as the dependent variable and the storage methods and extraction techniques as factors. Only the interaction term was significant (Table 2.1), reflecting the fact that the optimal extraction method varied between samples stored in DETs and the other storage methods. The highest amplification success rate was obtained with samples stored in 70% ethanol and extracted with the GuSCN / silica method (Fig. 2.2). The variation in amplification success between the different primers was also smallest for this treatment (Fig. 2.2). This method was therefore used for the rest of the study.

2.3.2. PROBABILITY OF IDENTITY

P_{ID-Sib} , calculated from the reliable genotypes of 36 badgers from the three social groups under investigation, predicted that the seven loci used in this study were necessary, yet sufficient, to distinguish with 99% certainty between sibling badgers (Fig. 2.3). The observed P_{ID} showed that the proportion of individuals with identical profiles dropped to zero if the five most informative loci were used.

Table 2.1: Results from an ANOVA of the ranked PCR success rate from faecal DNA obtained using three storage methods and two extraction techniques (d.f. = degrees of freedom; SS = sum of squares; MS = mean square). The ANOVA was performed using a Scheier-Ray-Hare extension of the Kruskal-Wallis test.

Source	d.f.	SS	SS/MS _{total}	P-value
Storage	2	34.204	0.045	0.978
Extraction	1	879.874	1.159	0.282
Storage x Extraction	2	5941.642	7.824	0.020
Error	90	65292.096		
Total	95	72147.816		

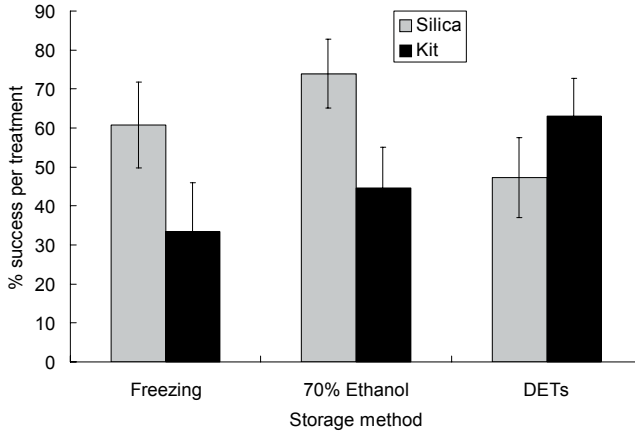


Fig. 2.2: Comparison of amplification success rates of different storage and extraction methods. Results are from 16 aliquots of faecal samples that were extracted with the GuSCN / silica method or the faecal DNA kit from frozen faeces or from faeces stored in 70% ethanol or a DETs solution. Each sample from each treatment was amplified once with seven microsatellite loci and the percentage of successful amplifications for each locus in each treatment was calculated by pooling the data across the 16 samples. For each treatment, the mean of these percentages, with its standard error, is shown.

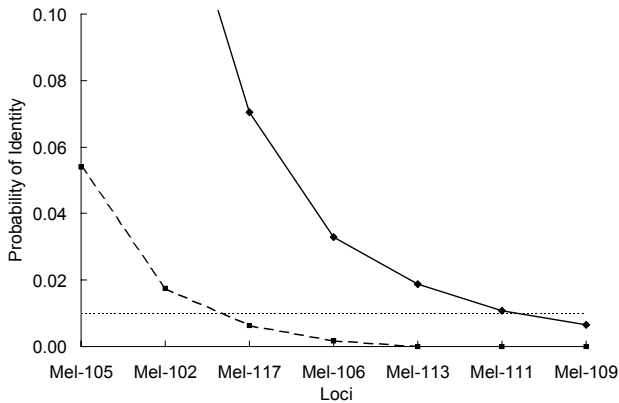


Fig. 2.3: Sibling probabilities of identity (P_{ID-Sib}) from three badger social groups at Woodchester Park. Probabilities were calculated for seven nuclear DNA microsatellite loci from a data set of 36 badgers. P_{ID-Sib} for individual loci was first calculated and the loci in the data set were arranged in order of increasing value (solid line). P_{ID-Obs} gives the proportion of all possible pairs of individuals that had identical genotypes (dashed line). The dotted 1% cut-off line represents the point where enough loci are typed to distinguish between individuals with 99% certainty.

2.3.3. COMPARATIVE MULTIPLE-TUBES APPROACH AND ASSESSMENT

Of the 53 faecal samples collected from the latrines of three social groups, DNA was obtained from 47 samples (89%). Using the comparative multiple-tubes approach, 33 of these samples gave rise to complete genetic profiles with consensus genotypes at all seven loci (Table 2.2). A further six profiles were complete but for the presence of an ambiguous case (six homozygote, one heterozygote score). All seven loci could therefore be amplified in 39 samples (74%). It was possible to reliably score the remaining eight samples for at least the most informative locus (*Mel*-105).

In order to analyse the reliability of the consensus genotypes, the complete faecal profiles were compared to reference profiles obtained from animals caught in the three social groups. There was a 100% match between 28 of the 33 complete profiles and the reference profiles. Faecal sample 42 matched with reference profile U61 except for one allele at locus *Mel*-102 (42: 199 199; U61:195 199). Using the P_{ID-sib} -statistics in GIMLET v1.3.2, the probability that these two profiles represent the same individual was calculated to be 0.985 if locus *Mel*-102 was excluded from the analysis. These two profiles were therefore classified as originating from the same animal. A further four profiles could not be matched to any reference. In order to increase the confidence that these unique profiles were not the result of allelic dropout, homozygous loci were amplified a total of seven times.

Ignoring failed reactions, a total of 1009 PCRs had to be performed to obtain 293 consensus genotypes, with an average of 3.4 reactions per locus per genotype (see Appendix I.a. for detailed results of the PCR replicates). Had the WCR approach been followed, an additional 517 positive PCRs would have had to be analysed, giving a total of 1526 reactions (5.2 reactions per locus per consensus genotype). When testing the MLR approach, a multiple-test correction was applied. For all the tests, it was found that, in order to limit the sample-wide number of genotype errors to <5% with 95% probability, each individual sample required a reliability of 98.29%. The 21 PCR replicates that were left out of the initial input file because they contained obviously false alleles were included in all counts. All RELIOTYPE output files are summarized in Appendix 1.b.

When applying the MLR model to the initial data and using reliability only as a criterion, a total of 1007 replicates would have been necessary if the replication strategy suggested by RELIOTYPE had been followed to its full extent. When, in addition to the reliability criteria, all alleles need to be observed at least twice, the number of PCR replicates required if the suggested replication strategy had been

Table 2.2: Consensus genotypes obtained using the comparative multiple-tubes approach. Genotypes containing an “F” are those cases in which an additional allele was observed once in seven amplifications, and was therefore scored as a half-locus. A dash indicates that there were insufficient positive PCRs available to derive a consensus genotype. The last column indicates with which reference individual the faecal profile could be matched.

	Sample	Consensus alleles at the microsatellite loci under investigation							Match
		Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113	
Parkmill Social Group	2	148 148	195 197	187 189	220 222	132 132	106 106	120 120	Q36
	3	148 148	195 197	187 189	220 222	132 F	106 106	120 120	
	7	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	8	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	16	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	40	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	48	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	49	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	57	148 148	195 197	187 189	220 222	132 132	106 106	120 120	X59
	1	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	20	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	25	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	28	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	29	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	43	138 142	199 199	174 193	220 222	132 138	106 116	120 120	D77
	17	138 142	199 199	174 187	220 224	130 132	106 125	120 120	
46	138 142	199 199	174 187	220 224	130 132	106 125	120 120		
47	138 142	199 199	174 187	220 224	130 132	106 125	120 120		
53	138 142	199 199	174 187	220 224	130 132	106 125	120 120	U41	
9	138 142	199 199	187 187	220 220	130 132	106 106	118 120		
41	138 142	199 199	187 F	220 220	130 132	106 106	118 120		
44	138 142	199 199	187 187	220 220	130 132	106 106	118 120		
31	138 140	195 197	174 187	222 222	132 F	106 106	120 126	-	
11	138 140	-	-	-	-	-	-		
19	138 138	197 199	174 187	220 222	132 132	106 106	120 126		
30	138 142	199 199	174 187	220 222	132 132	106 116	120 126		
52	136 144	-	-	-	-	-	-	Q66	
54	140 144	195 197	174 187	220 224	132 132	106 106	126 126	Q65	
58	138 138	199 199	174 187	222 226	132 138	106 125	120 120	H51	
6	142 142	195 199	174 174	222 224	130 132	106 127	120 126	J68	
32	142 142	195 199	174 174	222 224	130 132	106 127	120 126		
51	142 F	195 199	174 174	222 224	130 132	106 127	120 126		
4	138 142	195 199	174 174	222 224	130 132	106 106	120 126	J56	
39	138 142	195 199	174 174	222 224	130 132	106 106	120 126		
5	138 142	195 199	174 187	222 224	130 132	106 106	120 126		
33	138 142	195 199	174 187	222 224	130 132	106 106	120 126	-	
23	138 138	197 199	174 189	222 226	132 F	106 106	120 120	Q72	
55	138 138	197 199	174 189	222 226	132 132	106 106	120 120		
56	136 140	193 193	174 187	222 F	132 132	106 127	120 120		
21	138 148	195 199	-	-	-	-	-	U62	
34	138 148	195 199	187 187	222 222	130 132	106 106	120 126		
37	138 148	-	-	-	130 132	106 106	120 126		
42	144 148	199 199	174 187	222 222	130 130	106 106	126 126	U61	
36	144 148	-	-	-	130 130	106 106	126 126		
13	136 142	-	-	-	-	-	-		
26	140 148	-	-	-	-	106 106	120 126	T50	
35	148 148	-	-	-	132 132	106 106	120 120	M58	
								X30	

Table 2.3: Summary of the faecal DNA PCR errors observed, by locus and type. Data are from heterozygous genotypes at seven microsatellite loci in 47 individuals. Only PCRs in which a consensus genotype was obtained were considered. Type I errors included PCRs where three alleles were obtained, as well as cases where one or two alleles were observed but one of these was false.

Result type	Nbr. of individuals	Loci							Total
		<i>Mel-102</i>	<i>Mel-105</i>	<i>Mel-106</i>	<i>Mel-109</i>	<i>Mel-111</i>	<i>Mel-113</i>	<i>Mel-117</i>	
Correct		23	30	31	16	22	16	30	364
Error type II	Short allele missing	30	10	17	2	8	4	13	84
	Long allele missing	12	13	13	3	7	3	16	67
Error type I	Wrong genotype	3	8	8	0	10	0	14	43
Total		88	90	97	47	75	50	111	558

followed to its full extent would be 1165.*

In order to determine whether the assumptions of the MLR model were met, the errors of the replicate PCRs were analysed (Table 2.3). Allelic drop-out occurred in 27% of the amplification reactions for heterozygous genotypes and an otherwise wrong result was obtained in 8% of these reactions so that, pooling both error types, a mean error rate for heterozygous loci of 35% was obtained. Considering both homozygous and heterozygous genotypes, *i.e.*, all the PCRs, 19% of all the amplification reactions

* Please note: The version of program RELIOTYPE used in this study was not error-free. Even when the minimum number of observation per allele was set to two, the program did not recommend any further replicates in seven cases, even though one allele at the loci in question had been observed only once (The cases in question are highlighted in the output files in Appendix I). This would have led to one erroneous consensus genotype (sample 37 at locus *Mel-111*). Despite this problem, the results presented in this study provide a general idea about the performance of the maximum-likelihood model when applied to a real data set. The problem with the program needs to be addressed, however, if the technique is to be used in the future.

were erroneous. There was no difference in the dropout rate between short and long alleles (Mann–Whitney U ; $N_1 = 7$, $N_2 = 7$; $Z = -0.321$; $P = 0.805$). The GLMM indicated significant variation in error rates among loci ($P = 0.0103$) after between-sample variation was taken into account. The difference among loci was mainly due to a much greater dropout rate at locus *Mel*-102. At this locus, 47.7% of the amplifications for heterozygous genotypes experienced allelic dropout compared with values ranging from 10.6% (*Mel*-109) to 29.7% (*Mel*-106) for the other loci.

The MLR method assumes that all false alleles can be detected and eliminated from the dataset so that the only possible source of error is undetected allelic dropout events. In order to simulate an actual study when performing the MLR in retrospect, I only eliminated false alleles that would have been recognized at each specific round of replication. There were six ambiguous cases (with six homozygote, one heterozygote score; Table 2.2). Of these, one (locus *Mel*-106 in sample 56) was found to be a true heterozygote through comparison with the reference genotypes, whereas the rest were homozygotes. If all alleles were accepted on the basis of reliability this would have led to five erroneous profiles. Furthermore, the initial replicates contained four false alleles that would have remained undetected based on reliability criteria, leading to a total of nine erroneous profiles (The consensus genotypes and error that would have been obtained are shown in Appendix I.c.). If, in addition to a decision based on reliability criteria, alleles needed to be observed at least twice before being recorded, both the inconclusive cases and the false alleles should be detected using the MLR method, as they were with the comparative method (This assumes, however, that the program runs error-free; see footnote above).

Given the multiple test correction, we could be 95% sure that <5% of the multi-locus profiles were wrong because of undetected allelic drop-out. From 47 profiles we would expect errors for 2.35 genetic profiles at most. Consistent with this expectation, the consensus genotypes generated by both MLR models (*i.e.*, reliability criteria alone or with the additional requirement of observing alleles twice) did not contain any undetected allelic dropouts. The MLR model correctly indicated the need for further replication at locus *Mel*-102 of individual 42, the allelic dropout that remained undetected using the comparative approach.

2.4. Discussion

2.4.1. COMPARISON OF STORAGE AND EXTRACTION METHODS

In order to maximize the success of the faecal DNA extractions, various storage and extraction techniques were tested for their suitability for use with badger faecal DNA. Although all methods were successful, storage of faecal samples in

70% ethanol plus extraction of DNA with the GuSCN / silica method was slightly superior. As well as being the cheapest method, this combination is also safest as storage in 70% ethanol is an effective disinfectant against *M. bovis* (Seymour 1991). Murphy *et al.* (2002) found that storage of brown bear (*Ursus arctos*) faeces in 90% ethanol gave rise to the highest proportion of amplifiable DNA and had the longest post-collection longevity. Frantzen *et al.* (1998) found that storage in DETs solution was the most appropriate method of preserving faecal samples when fragments longer than 300bp were amplified, but that for shorter fragments all storage methods performed similarly. It is therefore possible that the treatments would have had more of an effect on loci >300bp in length. It is also possible that genotyping error rates varied significantly between treatments (Flagstad *et al.* 1999), but this was not tested in my study.

Of the 53 faecal samples collected from three social groups of badgers (Parkmill, Nettle and Kennel) it was possible to amplify all seven loci in 39 samples (74%), and for a further eight samples it was possible to amplify at least the most informative locus. This resulted in 47 samples (89%) with at least partially amplifiable DNA. This success rate is high in comparison with many other studies, which report successful amplification in 48–66% of faecal samples (Gerloff *et al.* 1995, Kohn *et al.* 1999, Farrell *et al.* 2000, Jansman *et al.* 2001, Lucchini *et al.* 2002). However, a success rate of 83% (*i.e.*, similar to ours) has been reported by Banks *et al.* (2002), and a rate of 93–95% has been reported by Flagstad *et al.* (1999). One reason for our success may be that only faeces that were less than a day old were analysed (Reed *et al.* 1997, Jansman *et al.* 2001, Lucchini *et al.* 2002).

2.4.2. PROBABILITY OF IDENTITY

P_{ID-Sib} statistics suggested that DNA profiles consisting of the seven loci used in our study would be sufficient to distinguish between individual badgers, including siblings, with 99% certainty. This statistic was supported by P_{ID-Obs} , which dropped to zero after the five most informative loci were used. It is possible that the number of loci required would differ between populations, and perhaps even social groups (see Banks *et al.* 2002). The former possibility will be considered in Chapter 4.

2.4.3. COMPARATIVE MULTIPLE-TUBES APPROACH AND ASSESSMENT

When testing the reliability of my comparative approach I found that in 293 consensus genotypes, one case of allelic dropout was not detected. The error arose because it was not possible to compare the genotype in question with the corresponding locus from an identical profile. Although this error would not have led to an incorrect estimate of the number of individuals present (see Table 2.2), I determined that a further modification to my approach would reduce this error. Loci in unique

profiles that after three positive amplifications were provisionally recorded as being homozygous, needed to be amplified a further four times to ensure that the genotype was scored correctly. This modification should be applied when correct profiles need to be obtained for single-sampled individuals, as would be the case when paternity analysis is performed. If I had performed this step, an additional 64 reactions would have been required, giving a total of 1073 PCR. The WCR approach would require 1526 reactions; therefore my approach appears to greatly reduce PCR effort.

A mean error rate for heterozygous genotypes of 35% was obtained with, overall, an error observed in 19% of all PCRs. Allelic dropout occurred in 27% of replication reactions for heterozygous genotypes and was therefore a more significant problem than the occurrence of false alleles. This error rate is higher than in most faecal DNA studies (e.g., Bayes *et al.* 2000, Ernest *et al.* 2000, Goossens *et al.* 2000, Constable *et al.* 2001) but similar to the error rate reported from single hairs (Gagneux *et al.* 1997, Goossens *et al.* 1998). This high error rate increases the number of PCRs required to obtain consensus genotypes but does not invalidate the use of faeces as a source of badger DNA.

I found that the MLR model, when applied in practice, would also significantly reduce the number of amplifications needed to obtain consensus genotypes compared with the WCR approach. The effectiveness of the method is dependent, however, on whether the assumptions of equal dropout rates of alleles of different sizes and across loci, as well as the detection of all false alleles, can be met in a given data set. The assumption that there is no difference in the allelic dropout rate between the longer and the shorter alleles was confirmed for our data set. This result has also been found in some other studies (Gerloff *et al.* 1995, Gagneux *et al.* 1997), but not all (Constable *et al.* 2001), and so ideally should be tested for each new data set. Although I found that the allelic dropout rate varied significantly among loci (see also Lucchini *et al.* 2002), the overall rate was estimated to be 27%, a value low enough to expect the replication strategy suggested by the MLR model to be robust, even if an upper bound of 75% on the drop-out rate was used (Miller *et al.* 2002). In future studies, it may be worth replacing the most unreliable locus (*Mel-102*) or even removing it from the analysis if the power to distinguish between different individuals would not be too greatly reduced. Contrary to the first two assumptions, I found that by relying on reliability criteria alone (such that alleles do not need to be observed twice to be recorded) a total of nine erroneous multi-locus profiles would have been obtained. My results therefore confirmed the suggestion by Miller *et al.* (2002) that it will be necessary to not rely solely on reliability criteria, but also to observe each allele at least twice before recording it.

When utilizing the MLR model with my dataset, I applied a multiple test correction

in order to be 95% certain that fewer than 5% of the multi-locus profiles contained undetected allelic dropouts. Because no allelic dropout remained undetected after following the replication strategy of the MLR model, it was concluded that more stringent reliability criteria (for example, limiting the incidence of errors to 0% with a probability of 95 or 99%) requiring more PCR replicates would not have been necessary for my study. This level of reliability should therefore be sufficient for other studies, but I recommend that the rule of observing an allele at least twice before recording always be applied when using the MLR model.

The advantage of the comparative method is that no advance knowledge of differences in the dropout rate between large and small alleles or between different loci is required for it to be used, making it better suited for studies with limited financial resources. However, the method requires a number of identical genetic profiles that can be compared with each other, which might not be achievable with other species. The MLR model may be more appropriate where faecal samples from a single individual are less likely to be replicated. The MLR method may also be better suited to large studies in which batch PCR replication, rather than single additions, is more practical, though the comparative method could be adjusted to allow batch replication. My method, however, seems better suited to small studies in which many faeces will be obtained from single individuals.

2.5. Summary

The DNA extracted from faecal samples collected in the field is often of poor quantity and quality. In order to use microsatellite profiles obtained from faecal DNA as a basis for estimating the size of badgers populations, it was necessary to optimize DNA storage and extraction techniques, and to develop a methodology that reduced the number of replicate PCRs necessary to obtain reliable genotypes. After trialling several methods, a high amplification success rate (89%) was obtained by storing faeces in 70% ethanol and using the guanidine thiocyanate/silica method for extraction. Using 70% ethanol as a storage agent had the additional advantage of it being an antiseptic. In order to obtain reliable genotypes with a limited number of amplification reactions, I devised a comparative approach in which genetic profiles were compared and replication directed at similar, but not identical, genotypes. This modified method achieved a reduction in the necessary number of PCR replicates by comparison with the standard multiple-tubes approach. Additionally, I performed the first empirical test of the new maximum-likelihood approach for replication. Using this method, a similar reduction in PCR effort would have been achieved. My comparative approach would be best suited for studies that included multiple faecal samples from each individual.

Chapter 3

Estimation of Badger Abundance by Genotyping Faecal DNA

3.1. Introduction

Microsatellite profiles generated from faecal DNA samples have formed the basis of abundance estimates in a number of studies (see Chapter 1). After collecting faecal samples in a systematic fashion and generating individual specific microsatellite profiles, the size of the population under investigation can be estimated in three ways. The minimum size of endangered populations has been estimated by counting the number of different reliable microsatellite profiles obtained from dung (Kohn *et al.* 1995, Taberlet *et al.* 1997, Ernest *et al.* 2000). After collecting faecal samples on repeated, temporally distinct occasions, mark-recapture models have been applied to a dataset of genetic profiles to assess population size (Banks *et al.* 2002). Mark-recapture studies are based on the capture, marking and release of an initial population. After resampling the population for ≥ 1 additional sessions, the ratio of unmarked animals to recaptures is used to estimate population size (White *et al.* 1982). Other studies have estimated local abundance of a species by plotting accumulation curves of the number of faeces sampled against the cumulative number of new profiles. The asymptote of the curve determined by the accumulation of unique genotypes is the estimated population size (Kohn *et al.* 1999, Eggert *et al.* 2003).

The main objective of this part of the study was to test whether microsatellite genotyping of faecal DNA could constitute a robust methodology for estimating badger abundance. The study utilised the high-density badger population at Woodchester Park, Gloucestershire, UK, where, as a result of an ongoing long-term and intensive study, independent estimates of abundance and information on the life histories of the resident badgers were available (Rogers *et al.* 1997b, Rogers *et al.* 1999, see also Chapter 2).

I collected faeces from latrines near setts (“hinterland latrines”) because Roper *et al.* (1993) have suggested that all members of a social group are equally likely to visit these latrines. My first aim was to confirm by means of genetic profiling of faecal DNA that there was indeed no age- or sex-specific defecation pattern at the hinterland latrines in my study population. This information allowed me to make inferences about the probability of “capturing” the profiles of various members of the target social groups. I then tested whether robust estimates of social group size and local abundance could be obtained either by direct enumeration of different genetic profiles or by means of one of the statistical techniques described above.

3.2. Materials and Methods

3.2.1. BASELINE POPULATION DATA

Characteristics of the study site are summarised in Chapter 2. All the independent information on badger abundance and life histories cited in this chapter was collected by the staff of the Woodchester Park research station. Seasonal live trapping was performed on a regular basis at Woodchester Park throughout the entire study area and trapping success was believed to be high. According to unpublished data by G. Wilson, for example, the number of individual badgers trapped each year between 1987 and 1997 was within 10% of the count of all the animals known to be alive, based on previous and subsequent trapping sessions. Furthermore, given the relatively low proportion of unmarked badgers amongst those found dead in the study area, the Woodchester Park staff believes that the number of captured badger closely represents the true population size.

The present pilot study involved three adjoining social groups, called 'Parkmill', 'Kennel' and 'Nettle'. All adult badgers caught at these three target groups prior to the collection of faecal samples were given individual-specific fur-clips by Woodchester Park staff. The contrast between the light under fur and the darker guard hair allowed identification of individuals in the field (Stewart & Macdonald 1997). During the pilot study, an infrared light (Tracksys Ltd), a monochrome video camera (Sanyo VCB-35721RP, Sanyo Electric Co. Ltd) and a time-lapse video recorder (Sanyo TLS-9168P; Sanyo Electric Co. Ltd) were used to record the presence of badgers at the most active sett in each of the three target territories from 19:30 to 06:00 each night. The fur-clips made it possible to identify some of the badgers that had been recorded. In addition to these observations, the identity and life history of a defecator could be revealed by comparing the faecal profiles to the reliable profiles generated from 36 blood or hair samples taken from badgers that had been captured in 2000 and 2001 in the three target territories.

3.2.2. COLLECTION AND PRESERVATION OF FAECAL SAMPLES; LABORATORY PROCEDURES

The genetic profiles generated using the comparative multiple-tubes approach summarised in Chapter 2 were used to estimate badger abundance. The corresponding faecal samples had been collected during a 10-day period in October 2001 from latrines close to the setts of the three target groups. Samples were stored in 70% ethanol and extracted using the GuSCN / silica method as described in Chapter 2. After performing a P_{ID-Sib} analysis, it was shown that genetic profiles consisting of the following seven loci would be enough to distinguish between individual badgers with 99% certainty: *Mel-102*, *Mel-105*, *Mel-106*, *Mel-109*, *Mel-111*, *Mel-113*, *Mel-*

117. The PCR conditions and reagent concentrations are summarised in Chapter 2.

3.2.3. ESTIMATION OF BADGER ABUNDANCE USING MARK-RECAPTURE MODELS

Mark-recapture models were applied to the reliable genetic profiles detailed in the previous chapter in order to assess the total number of badgers present in the three social groups. As with data from direct trapping or observation of animals, genetic profiles obtained from faecal DNA must meet the assumptions of the relevant mark-recapture models for accurate estimates to be obtained. It was important that the size of the population remained constant during the study period. Demographic closure could be reasonably assured by collecting faeces over a short period, while violation of geographical closure was minimized by collecting faecal samples from latrines close to setts well within the territorial boundary of social groups.

In a natural population, it is unlikely that individuals have equal probabilities of being captured. Three causes of variation in capture probability have been identified: behavioural responses to capture, variation over time (with constant trapability for all individuals) and individual heterogeneity (Otis *et al.* 1978). Different models of estimating population size that allow relaxation of the assumption of equal capture probability have been developed (Otis *et al.* 1978, Chao *et al.* 1992, Lee & Chao 1994). My results suggested that latrines were used equally by the two sexes and by individuals of all age groups (see below), but variation in 'capture' probability may still have occurred through other means. It is believed that changing environmental conditions, especially differences in humidity or exposure to sun or shade, have an effect on the quality of faecal DNA (Farrell *et al.* 2000, Goossens *et al.* 2000), so that there will be variation in extraction success, and hence capture probability, due to time effects. Furthermore, it is believed that the length of the interval between deposition and collection of a dropping has an effect on DNA extraction success (Dallas *et al.* 2000, Goossens *et al.* 2000, Jansman *et al.* 2001, but see Palomares *et al.* 2002). In this case, in addition to time effects, some individual heterogeneity will be introduced by differences in the time of defecation. Finally, given the importance of time of deposition and of exposure to the elements, the location of a faecal sample would be expected to affect its extraction success, introducing additional individual heterogeneity.

These considerations suggest that models allowing for individual heterogeneity, time effects and a combination of both should be used in the analysis. Furthermore, Mills *et al.* (2000) have suggested that the Jackknife estimator for model M_h (which allows for individual heterogeneity) should be used with non-invasive population size estimation because this estimator would produce the least biased results when faced with a 'shadow effect', *i.e.*, failure to identify different individuals with identical profiles.

Mark-recapture models were applied using the program CAPTURE (Otis *et al.* 1978, White *et al.* 1982). It was assumed that the population was closed over the 10-day collection period. Each collection day was considered a capture session, giving rise to ten sampling sessions with each profile assigned a '1' for a sampling session in which it was detected and a '0' when absent. Given the above considerations, estimates of badger abundance were generated using the null model (M_0 -Null) as well as models allowing for variation in capture probability due to individual heterogeneity (M_h -Jackknife, M_h -Chao), time effects (M_t -Chao, M_t -Darroch) and a combination of both (M_{th} -Chao).

3.2.4. EVALUATION OF RAREFACTION METHODS USING SYNTHETIC DATA

Due to the relatively small number of different profiles collected at each of the three setts, mark-recapture analyses were considered unlikely to produce meaningful estimates of individual group sizes. Therefore, in addition to total abundance, the size of the social groups was estimated using accumulation curves. In this method, population size corresponds to the projected asymptote of a function describing the number of samples analysed versus the cumulative number of unique genetic profiles (rarefaction curve). Three possible equations for the rarefaction curve been suggested in the literature:

(1) Kohn *et al.* (1999) used the following hyperbolic function to estimate coyote (*Canis latrans*) numbers with faecal genotyping:

$$y = ax / (b + x)$$

where y = cumulative number of genetic profiles, x = number of genotypes sampled, a = asymptote (or population size estimate), b = non-linear slope of the function. This method is referred to henceforth as "Kohn's method".

(2) Eggert *et al.* (2003) used an exponential function ("Eggert's method") to estimate African forest elephant (*Loxodonta cyclotis*) abundance from faecal DNA typing :

$$y = a (1 - e^{(bx)})$$

(3) In the manual of the program GIMLET (Valière 2002), D. Chessel suggests using the equation

$$y = a - a (1 - (1/a))^x$$

corresponding to the expectation of the number of full boxes when x balls are distributed in a boxes. This approach will be referred to as "Chessel's method".

Two studies have reported simulations aimed at determining the accuracy of the three rarefaction methods. While the GIMLET manual (Valière 2002) reported results

from limited simulations on the accuracy of Kohn's and Chessel's method, Eggert *et al.* (2003) compared the accuracy of Kohn's and Eggert's methods. While the GIMLET manual does not contain any information on the sizes of the simulated datasets, Eggert *et al.* (2003) used either 125 or 150 samples with replacement from simulated populations with 75, 150 and 300 members. In order to allow correct interpretation of the asymptote values generated by the three rarefaction curves when using the methodology to estimate the sizes of badger social groups, I performed simulations on relevant, and thus smaller, sample and population sizes.

The simulated populations contained 10 or 15 members and, using VBA programs in Excel 2000, 100 subsets of 10, 15 and 30 samples were selected randomly, with replacement. Additionally, a simulation was performed using a population of 30 members and 100 subsets of 50 randomly selected samples. It was assumed that each individual had an equal chance of being sampled and that the population was closed. Program R (Ihaka & Gentleman 1996) was used to perform the analyses of the accumulation curves using a script file generated by program GIMLET 1.3.2 (Valière 2002). Each of the 700 subsets was transformed into an input file for program R by regrouping and counting identical samples. Because the order in which the samples are added affects the shape of an accumulation curve (Colwell & Coddington 1994), Program R randomised the order of the samples in each subset 100 times and projected the asymptote for each of these randomisations using the three equations introduced above. The mean and the median of all 100 iterations of a were taken to be estimates of a for a specific sub-sample. The accuracy and variance of the results of the three rarefaction methods were investigated by calculating the average and the standard deviations of the 100 estimates of population size generated from the 100 sub-samples available for each combination of population and sample size.

3.2.5. ESTIMATION OF BADGER ABUNDANCE USING ACCUMULATION CURVES

Rarefaction methods were used to estimate the total size of the three social groups as well as the size of each group individually. Data input files for program R were generated by using GIMLET 1.3.2 to regroup and count the genetic profiles that were identical. A script file for R was again generated using GIMLET 1.3.2. Program R randomised the order of the profiles in the dataset 1000 times and for each of these randomisations the asymptote was projected using the three equations. The mean and median values of all iterations for the asymptote, a , were calculated and the distributions of the projected values for the various asymptotes were tested for normality using a one-sample Kolmogorov-Smirnov (K-S) test. The variance of the a -estimate was analysed by calculating the standard deviation (SD) and the standard error (SE) associated with the estimate, as well as the interquartile range of the results of all the iterations.

3.3. Results

3.3.1. Faecal DNA typing and baseline estimates of abundance

As already explained in the previous chapter, 47 of the 53 droppings that were collected gave rise to at least partially amplifiable DNA and 20 different genetic profiles were generated from these extracts. During 2001, 29 badgers had been live-trapped in the three social groups under investigation (though for one of these individuals a DNA profile was not available; information provided by Woodchester Park research staff). According to capture results, Parkmill was the largest of the three target groups as 12 animals had been captured at its setts in 2001, followed by Nettle with nine animals and by Kennel with eight. Of the 20 genetic profiles 15 could be matched to the 29 captured badgers (see Table 2.2.), and one genetic profile belonged to an individual trapped during 2000 but not 2001. Thus, of 20 different genetic profiles obtained from faecal DNA, 16 could be matched to known group members.

By counting the captured badgers that visited the latrines, the known badgers that were observed but did not visit a latrine and the unknown genetic profiles, it was established that at the time of the study at least 28 badgers were present in the three target groups combined (Table 3.1). Only 24 of these 28 badgers had been captured prior to my study. By including those badgers that had been captured in 2001 but that were neither observed nor visited any latrines, the maximum number of badgers present in the study groups was counted to be 34. The actual number of badgers resident in the three groups was therefore between 28 and 34 individuals. Similarly, the minimum and maximum number of badgers present in each of the three social groups was also determined (Table 3.1) Parkmill was the largest social group, followed by Nettle and Kennel.

Table 3.1: The minimum and maximum number of badgers known to be present in the three social groups as determined by a combination of capture, genetic profiling and direct observation through infra-red camera recordings.

Social Group	Captured & defecated	Captured & observed	Unknown faecal profiles	Minimum no. present	Capture only	Maximum no. present
Parkmill	7	4	3	14	2	16
Nettle	4	3	1	8	2	10
Kennel	5	1	0	6	2	8
Totals	16	8	4	28	6	34

3.3.2. SEX- AND AGE-SPECIFIC DEFECATION PATTERNS AT HINTERLAND LATRINES

The sex ratio of the animals that had been captured and that deposited faeces at latrines was 1:1 (eight males, eight females; Table 3.2). Additionally, the sex ratio of the animals that had been caught in the three target groups in 2001, but whose profile could not be identified from the faecal samples was also 1:1 (seven males, seven females). One male and one female deposited a dropping in nine and six different latrines respectively during the ten-day study period. All the other individuals deposited faeces between one and three times. G. Wilson analysed my dataset for the presence of age-specific defecation patterns. Only badgers for which age data were available, *i.e.*, which had been caught originally as cubs, were considered for

Table 3.2: Gender of the different defecators and the total number of faecal samples deposited by these 20 individuals. Only the gender of captured animals was known.

Social group	Individual	Gender	No. of faeces deposited
Parkmill	Q36	Male	9
Parkmill	X59	Female	6
Parkmill	D77	Female	3
Parkmill	U41	Female	3
Parkmill	Sample 31 & 11	-	2
Parkmill	Sample 19	-	1
Parkmill	Sample 30	-	1
Parkmill	Q66	Male	1
Parkmill	Q65	Male	1
Parkmill	H51	Male	1
Nettle	J68	Male	3
Nettle	J56	Male	2
Nettle	Sample 5 & 33	-	2
Nettle	Q72	Female	2
Nettle	U8	Female	1
Kennel	U62	Female	3
Kennel	U61	Male	2
Kennel	T50	Female	1
Kennel	M58	Female	1
Kennel	X30	Male	1

Table 3.3: Estimates from six closed mark-recapture models of the population size of three badger social groups at Woodchester Park. Estimates were generated using the program CAPTURE, with the abbreviations for each model described in full in the text.

Model	Source of variation in capture probability	\hat{N}	SE	95% CI	Average estimated recapture probabilities
M ₀ -Null	None	23	2.5	21-32	0.17
M _t -Darroch	Time	23	2.2	21-30	0.18, 0.22, 0.22, 0.13, 0.22, 0.22, 0.13, 0.18, 0.18, 0.09
M _t -Chao	Time	24	3.9	21-39	0.14, 0.18, 0.18, 0.11, 0.18, 0.18, 0.11, 0.14, 0.14, 0.07
M _h -Jackknife	Heterogeneity	26	4.1	22-40	0.15
M _h -Chao	Heterogeneity	26	5.0	22-45	0.15
M _{th} -Chao	Time*Heterogeneity	28	6.4	23-51	same as M _t -Chao

this analysis. There was no significant difference in the mean ages of those badgers that did and did not deposit faeces at latrines (Mann-Whitney $U_{13,15} = 135$, $P > 0.05$). Furthermore, it was found that badgers ranging from cubs of that year to nine-year olds used the latrines that were investigated. Given these results, no sex- or age-specific defecation patterns could be identified, so there was no reason to expect intrinsic heterogeneity in capture probability between individual badgers.

3.3.3. DIRECT ENUMERATION OF FAECAL PROFILES AS A MEANS OF ESTIMATING BADGER ABUNDANCE

Given the baseline estimates, the 20 different profiles generated from faecal DNA corresponded to ≈ 59 -71% of the total population of the three social groups. The number of different microsatellite profiles identified in Parkmill, Nettle and Kennel social groups was 10, 5 and 5 respectively (Table 2.2).

3.3.4. ESTIMATION OF BADGER ABUNDANCE USING MARK-RECAPTURE MODELS

Mark-recapture analyses of profiles were performed using program CAPTURE, assuming a closed population during the 10-day collection period and treating each day as a capture session. Of the 53 samples collected, the daily percentage of samples from which DNA could be successfully extracted varied between 75 and 100% depending on the day of collection. Point estimates of the total number of badgers in the three social groups varied between 23 individuals (95% CI: 21-30) for model M_t-Darroch, to 28 individuals (95% CI: 23-51) using model M_{th}-Chao (Table 3.3). The

Table 3.4: Estimated population sizes using three different equations for the rarefaction curve of an accumulation plot of simulated data. Either 10, 15 or 30 samples were taken from simulated populations with 10 or 15 members. One population of 30 members was simulated from which 50 samples were taken. One hundred replicates were analysed for each combination of population and sample size.

Method	Population Size	Sample Size	Median			Mean		
			Average	Range	SD	Average	Range	SD
Kohn	10	10	17.6	5.7-48.8	8.9	29.3	6.3-146.3	22.3
	10	15	15.6	7.0-27.3	4.0	22.1	7.7-94.3	11.9
	10	30	13.2	10.5-14.8	1.2	13.7	10.6-16.4	1.3
	15	10	28.9	8.2-75.1	14.6	57.1	13.2-168.9	37.7
	15	15	25.3	9.7-47.6	8.4	42.4	11.3-136.9	23.0
	15	30	22.2	16.2-29.8	2.9	24.1	16.9-41.0	4.0
	30	50	45.2	27.2-63.5	6.6	47.3	27.3-67.2	7.2
Eggert	10	10	10.6	4.3-26.7	4.6	15.2	4.6-37.2	7.8
	10	15	10.2	5.4-16.3	2.2	12.5	5.7-23.9	3.9
	10	30	10.1	8.2-10.9	0.8	10.2	8.2-11.7	0.8
	15	10	16.2	5.7-39.2	6.8	24.4	6.8-40.8	8.9
	15	15	15.4	6.8-26.5	4.4	20.9	7.6-45.4	7.9
	15	30	15.3	11.7-19.4	1.6	16.1	12.0-20.8	2.0
	30	50	29.9	19.8-39.6	3.6	30.9	19.7-41.5	4.0
Chessel	10	10	12.4	4.3-43.1	7.3	13.1	4.2-54.0	8.9
	10	15	10.4	5.4-17.9	2.4	10.5	5.4-17.9	2.5
	10	30	10.0	7.6-11.1	0.8	9.9	7.7-11.1	0.8
	15	10	20.5	5.9-43.1	10.9	23.6	6.1-57.6	15.4
	15	15	16.5	6.1-32.1	5.5	16.8	6.1-34.1	5.8
	15	30	15.2	11.6-19.0	1.6	15.2	11.5-19.0	1.6
	30	50	29.9	19.7-39.5	3.8	29.9	19.7-39.7	3.8

two models that allowed for individual heterogeneity both generated a point estimate of 26 individuals, with a slightly smaller 95% confidence interval for M_h -Jackknife (M_h -Jackknife: 22-40, M_h -Chao: 22-45; Table 3.3).

3.3.5. EVALUATION OF RAREFACTION EQUATIONS USING SYNTHETIC DATA

The main results of the simulation are summarised in Table 3.4. All three methods produced estimates closer to the actual population size with reduced variance between estimates when more samples were available for analyses. The median values of the estimates obtained for each combination of population and sample size were closer to the actual population size than the mean. This difference was not important in Chessel's method and, generally, became smaller as sample sizes got larger.

The three different accumulation methods did not perform equally well in estimating the correct population size. Kohn's method always overestimated population size, especially if the mean, rather than the median, was used to summarise the iterations. Chessel's method overestimated population size when applied to the small sample size, but gave rise to accurate results with the larger samples. Finally, the point estimates obtained by the median values of the iterations generated by using Eggert's equation for the rarefaction curve were consistently accurate and the variance of the estimate was also the smallest of all three methods. Except at the largest sample size, the mean values of the iterations obtained with this method, however, were an overestimate of the actual population size.

The distribution of the point estimates obtained using Eggert's method for 100 subsets of the six combinations of small population and sample sizes are shown in Fig. 3.1. The graphs illustrate that the point estimates became more precise as the sample sizes increased. The graphs show that as sample size increased, the proportion of individuals that needed to be sampled from the population to allow accurate estimates of population size increased as well.

3.3.6. ESTIMATION OF BADGER ABUNDANCE USING ACCUMULATION CURVES

The main results from the accumulation curve analyses are summarised in Table 3.5. Considering all the estimates, an asymptote was obtained in at least 947 iterations. The estimated population size depended on whether the mean or the median statistics were used to summarise the distribution of the values obtained for the asymptote, a . The mean produced higher values than the median, especially when Kohn's equation was used for the rarefaction curve. The K-S test showed that, in all but two cases, the frequency distribution for possible values of the asymptote did not follow normality and, consequently, the median was considered to be the appropriate summary statistic. In the two cases that formed an exception, the mean and median values

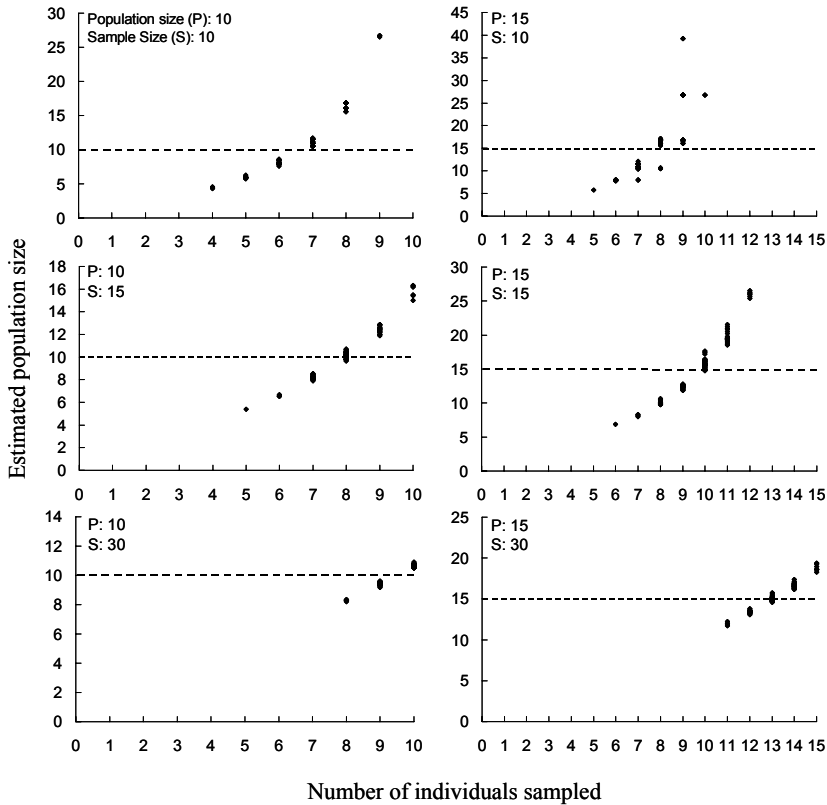


Fig. 3.1: Distribution of the median estimates obtained using Eggert's method for the 100 subsets of 10, 15 and 30 samples selected randomly, with replacement, from two simulated populations containing 10 and 15 members respectively. The x-axes correspond to the number of individuals sampled in each of the 100 subsets. The dotted line corresponds to the simulated population size.

were very close to each other.

Both Eggert's and Chessel's methods gave rise to total abundance estimates that were smaller than the minimum value of the baseline estimate. Contrarily, the point estimate obtained with Kohn's method was larger than the maximum value of the baseline estimate. Only the interquartile range obtained with Kohn's method overlapped partially with the range of 28 to 34 individuals that served as a baseline estimate.

Kohn's method gave rise to the highest estimate of the sizes of the individual social

Table 3.5: Estimates of badger abundance obtained by projecting the asymptote of a curve determined by the accumulation of unique genotypes. Three different functions, referred to as “Kohn”, “Eggert” and “Chessel”, were fitted to the accumulation plot (more detail in the main text). The second column gives the number of times the regression was iterated with the sample order randomised each time. A Kolmogorov-Smirnov (K-S) test was used to test the distributions of the projected values for the various asymptotes for normality. The mean or median values of all iterations for the asymptote represent the population estimate. The variance of the estimate was analysed by calculating the standard deviation (SD), the standard error (SE) and the interquartile ranges of the results of all the iterations.

Equation	No. of iterations	Test for normality		Min	Median	Mean	Max	Interquartile range	SD	SE
		K-S	Z	P						
All three social groups										
Kohn	1000	4.563		<0.001	23.0	35.2	38.0	141.9	31.3-41.4	11.0 0.35
Eggert	1000	4.622		<0.001	17.3	23.7	25.1	77.0	21.9-26.6	5.3 0.17
Chessel	1000	1.222		0.101	16.1	21.5	21.4	27.3	20.2-22.6	1.7 0.05
Parkmill social group										
Kohn	994	14.788		<0.001	10.2	15.2	23.3	5290.2	13.4-18.4	168.0 5.33
Eggert	993	9.964		<0.001	8.1	10.9	12.2	185.0	10.0-12.3	8.0 0.25
Chessel	1000	1.133		0.153	6.4	9.6	9.6	12.2	8.8-10.3	1.0 0.03
Nettle social group										
Kohn	995	8.047		<0.001	6.2	8.6	10.1	53.0	7.8-10.2	5.3 0.17
Eggert	995	8.255		<0.001	4.6	6.0	6.6	26.6	5.6-6.6	2.6 0.08
Chessel	1000	3.045		<0.001	4.4	6.3	6.3	7.4	5.7-6.9	0.7 0.02
Kennel social group										
Kohn	947	9.742		<0.001	7.7	10.0	16.5	107.8	9.1-15.6	18.4 0.60
Eggert	947	10.031		<0.001	5.3	6.5	9.8	57.0	6.1-9.2	9.3 0.30
Chessel	1000	3.777		<0.001	4.2	6.8	7.2	9.8	6.2-8.4	1.5 0.05

groups in all three cases. The point estimates obtained with Chessel's and Eggert's methods were similar. While in the case of Parkmill and Nettle groups, the estimate of Kohn's method fell within the range created by the independent estimates of the minimum and maximum number of badgers present in the specific groups, the estimate obtained for Kernel group was slightly higher than the maximum number of badgers known to be present. Conversely, in the case of Parkmill and Nettle, the estimates obtained using the other two methods were lower than the corresponding minimum number of badger known to be present, but in the case of Kennel, the estimate fell within the expected range. The same observation applied to the interquartile ranges associated with the estimates.

3.4. Discussion

3.4.1. SEX- AND AGE-SPECIFIC DEFECATION PATTERNS AT HINTERLAND LATRINES

The 53 faecal samples collected during the 10-day study period gave rise to 20 different profiles, of which 16 could be matched to known group members. At the time of the study, no apparent sex- or age-specific defecation pattern could be observed at the hinterland latrines of the three social groups. This result is in line with the observation by Roper *et al.* (1993) that radio-collared males and females visited hinterland latrines equally often. Thus, faecal samples collected in autumn in hinterland latrines of high-density populations were a valuable source of DNA for genetic fingerprinting. A useful number of different genetic profiles was obtained with no obvious age- or sex-specific bias, and hence no intrinsic difference in "capture" probabilities caused by differences in marking behaviour.

3.4.2. BASELINE ESTIMATES

By combining faecal profiling with direct observations and capture data, the minimum number of badgers known to be present was counted to be 28, while the maximum number of badgers that could be present was found to be 34. If the unmatched profiles belonged to visiting badgers (Christian 1994), the minimum number of animals present would only have been 24. However, even though "visitors" have been observed to copulate with members of the resident social group, it is not known whether they leave scent marks at their host's sett (Christian 1994). Furthermore, while radiotracking 14 badgers during a six-year period, Christian (1994) observed only four diurnal visits and fifteen nocturnal visits to other main setts. In my study, four visiting badgers would have been identified in a ten-day sampling period from 47 usable faecal DNA samples. Given these considerations, I believe that it is more appropriate to assume that the four unmatched profiles originated from resident badgers that had not been captured previously and that, consequently, a minimum of

28 badgers were resident in the three social groups at the time of the study.

3.4.3. ESTIMATION OF BADGER ABUNDANCE USING MARK-RECAPTURE MODELS

Too small a percentage of the badger population ($\approx 59\text{--}71\%$) was profiled from faecal DNA for direct enumeration to produce a robust estimate of the total size of the high-density population. However, the use of mark-recapture models could potentially solve the need to obtain a profile from every individual in the population. The model with the point estimate closest to the minimum count of 28 individuals was M_{th} -Chao (28 individuals), but it also had by far the largest 95% confidence interval (23-51). Model M_t -Darroch had the smallest 95% confidence interval (21-30), but this model is known to perform poorly when faced with individual heterogeneity (Otis *et al.* 1978, White *et al.* 1982). The M_h -Jackknife estimator works best with individual heterogeneity and is robust to some variation due to time as well as to the 'shadow effect' (Otis *et al.* 1978, White *et al.* 1982, Manning *et al.* 1995, Mills *et al.* 2000). This model therefore seemed the most appropriate, estimating the population size in the three social groups to be 26 individuals. The corresponding 95% CI of 22-40 animals contained the independent estimate of the range of the population size.

3.4.4. EVALUATION OF RAREFACTION METHODS USING SYNTHETIC DATA

The total size of the three social groups and of each group individually was also estimated by means of rarefaction analyses. Two studies have reported simulations on the accuracy of the projected results of all three equations for the rarefaction curve. The GIMLET manual (Valière 2002) reported results from limited simulations on the accuracy of Kohn's and Chessel's methods. The results varied depending on whether or not individuals exhibited heterogeneity of capture probability. It was predicted that the estimates generated using Chessel's method would be lower than the ones produced by Kohn's. Furthermore, in the presence of heterogeneity of capture probability amongst individuals, Chessel's method would underestimate population size while Kohn's method would generate an overestimate if a large proportion of the population had been sampled. Eggert *et al.* (2003) compared the accuracy of the estimates generated using Kohn's and Eggert's equations for the rarefaction curves. These authors took either 125 or 150 samples with replacement from simulated populations with 75, 150 and 300. It was assumed that members of the population had an equal probability of being sampled. The results suggested that, while Kohn's method significantly overestimated population size, Eggert's approach would produce consistently unbiased results. Furthermore, Eggert *et al.* (2003) showed that mark-recapture models gave similar results to a rarefaction curve that used their equation.

The simulations that I performed on smaller populations and samples complement

and supplement these results. Both my simulations and the ones by Eggert *et al.* (2003) assumed that all members of the population were equally likely to be sampled. However, even though no sex- or age-specific defecation pattern could be identified at the investigated latrines, individual heterogeneity can be introduced in faecal DNA studies because of the technical problems that do not allow every sample to be genotyped (see section 3.2). Nevertheless, the simulations provided a useful indication as to how to interpret the results obtained here. Both previous studies used the mean as their summary statistic. My simulations, however, suggested that the median values would give more accurate results. Indeed, after testing the distribution of the projected asymptote values obtained from the actual datasets for normality, it was shown that the median would be a more appropriate summary statistic than the mean. As suggested by the simulations, the mean values obtained from the actual datasets were higher than the corresponding median values, especially when using Kohn's method. The following conclusions will therefore focus on the estimates obtained by summarising the iterations with the median.

3.4.5. ESTIMATION OF BADGER ABUNDANCE USING ACCUMULATION CURVES

The value of the point estimate of the total population generated using Kohn's equation for the rarefaction curve was higher than the maximum number of badgers expected to be present, but the corresponding interquartile range overlapped partially with the range that served as a baseline estimate. Nevertheless, the simulations presented in all three studies agree that, given a specific dataset, Kohn's method overestimates population size. The point estimates generated using Chessel's and Eggert's methods were smaller than the minimum number of resident badgers. My simulation that used a combination of population and sample size that was similar to that in the actual study (a population of 30 members and 50 samples), suggested that Chessel's and Eggert's method would give rise to similar and accurate estimates. Given the simulation result that Chessel's method underestimates population size in the presence of heterogeneity, I conclude that the median estimate of 24 badgers generated by Eggert's equation was probably the least biased estimate given the available dataset. Furthermore, this value was close to the M_h -Jackknife estimate of 26. Hence, as predicted by the simulations presented in Eggert *et al.* (2003), the models that were most likely to generate the least biased results with both techniques gave rise to similar results.

The least biased point estimates generated by the two different statistical approaches were only slightly lower than the minimum number of badgers thought to be present. Considering, however, that as many as 34 badgers could feasibly be present in the three groups, neither estimate appears to be particularly robust. Neither the 95% CI of model M_h -Jackknife of 22 to 40, nor the interquartile range of the distribution of a

of 21.9-26.6, generated using Eggert's method, compared well with the range of 28 to 34 individuals taken as a baseline estimate. Overall, given the number of faecal samples collected in the three-week collection period, it was not possible to generate a robust estimate of the size of the three social groups or of a fitting range of values that could feasibly contain the correct estimate.

Rarefaction methods were also used to estimate the size of individual social groups from the corresponding faecal DNA microsatellite profiles. Kohn's method produced median estimates that corresponded well with the baseline estimates. Except in the case of Kennel social group, the other two rarefaction methods produced estimates of group size that were lower than the minimum number of badgers known to be present in the individual groups. The simulations performed using small population and sample sizes again suggested that Kohn's method would overestimate population size considerably and confirmed that the median values of the iterations produced by Eggert's method would most likely be an accurate estimate. Hence, similarly to the total group size, it was not possible to accurately estimate the size of the specific social groups given the small number of genetic profiles that had been collected.

To summarise, it was not possible to generate robust estimates of badger abundance using microsatellite genotyping of faecal DNA. More specifically, it was not possible to correctly estimate the abundance of badgers in three social groups in a high-density population by collecting faecal samples from hinterland latrines during a 10-day period. The estimates generated by Kohn's method corresponded well with the baseline data. Because, however, simulations showed that this method would overestimate population size from the available dataset of genetic profiles, the method cannot be applied to estimate the size of unknown populations.

One possibility to solve this problem would be to collect more faecal samples by either increasing the length of the collection period or by collecting samples from latrines further from the sett. Plots of the point estimates obtained from simulated datasets suggested, however, that increasing the size of the dataset only leads to improved estimates if the proportion of individuals sampled from the population could thereby also be increased. Or, in other words, increasing the proportion of individuals sampled would improve the accuracy of an estimate more than increasing the sample size. In the high-density population of Woodchester Park, however, it is difficult to envisage a procedure that would allow such an increase. Furthermore, in low-density populations, systematic collection of faecal samples near badger setts might not be feasible due to differences in the defecation patterns of the animals (Schley 2000, Hutchings *et al.* 2001, Hutchings *et al.* 2002, see Chapter 5).

Finally, one of the main reasons why genotyping of faecal DNA was envisaged as

a means of estimating badger abundance was the fact that the most reliable direct method for obtaining accurate estimates of badger abundance, live trapping and mark-recapture analysis, are for logistical reason impractical at anything other than a local scale. Despite the procedural modifications outlined in Chapter 2, the same restriction applies to faecal DNA genotyping because of the laboratory precautions and repeated amplifications that are necessary to obtain reliable genotypes. In conclusion, systematic collection and DNA profiling of faecal DNA samples is unlikely to form the basis of a robust methodology for estimating badger abundance.

3.5. Summary

The main objective of this chapter was to test whether microsatellite genotyping of faecal DNA could form the basis of a robust methodology for estimating badger abundance. Faecal samples were collected during a ten-day period from latrines close to the setts of three social groups. Forty-seven of the 53 droppings that were collected gave rise to at least partially amplifiable DNA and 20 different genetic profiles were generated from these extracts. The genotypes derived from the faecal samples were compared with those obtained from blood or hair samples from badgers live-trapped at the same setts. By combining faecal profiling with direct observations and capture data, it was possible to derive an estimate for the minimum and maximum number of badgers present. The faecal genotypes that originated from badgers with known trap histories did not reveal any apparent sex- or age-specific defecation pattern at the investigated latrines. Badger abundance was estimated by means of mark-recapture models and rarefaction curves. Simulations were performed to investigate which of three rarefaction methods would give rise to accurate results. It was not possible to generate a robust estimate of the size of the three social groups or of a fitting range of values that could feasibly contain the correct estimate. Simulations showed that the accuracy of the estimates would be improved by increasing the proportion of individuals sampled. As this would be difficult in practice, I concluded that systematic collection and DNA profiling of faecal DNA samples is unlikely to constitute a robust methodology for estimating badger abundance.

Chapter 4

Estimation of Badger Abundance by Genotyping Remotely Plucked Hair

4.1. Introduction

In the previous chapters, I described a method for estimating badger abundance using genetic profiles obtained from badger faeces and showed that systematic collection and genetic profiling of faecal DNA samples could not form the basis of a robust methodology. Too small a fraction of the population was sampled to obtain robust estimates, not only of total abundance, but also of the size of individual social groups. In addition, this method is unlikely to be cost-effective on a large scale because faecal DNA extracts were often of poor quality, and therefore required repeated amplifications in order to obtain reliable profiles.

In the present chapter I investigate whether microsatellite genotyping of DNA obtained from remotely plucked hair samples could form the basis of a reliable, accurate and cost-effective, methodology for estimating badger abundance. In a study by Sloane *et al.* (2000) on northern hairy-nosed wombats (*Lasiorninus krefftii*) plucked hair samples provided DNA of sufficient quality for repeated amplifications not to be necessary (but see Gagneux *et al.* 1997, Goossens *et al.* 1998). Recently, after performing feeding experiments near setts, Macdonald *et al.* (2002) observed that almost all the badgers present within a territory visited artificial feeding sites. These results, together with the observation that badgers leave hair on barbed-wire fences (Neal & Cheeseman 1996), suggested that a baited barbed-wire-enclosure hair trap (Woods *et al.* 1999) might be a feasible approach for remotely collecting fresh badger hair, applicable in any area regardless of population density or habitat.

Because the study was performed in a medium-density population in continental Europe, the first aim of this chapter was to test whether the chosen microsatellite loci exhibited enough variability in the local badger population to produce individual-specific genetic profiles. Secondly, I tested whether it was possible to remotely pluck hairs from badgers. Then I tested whether it was possible to obtain reliable genetic profiles from remotely plucked hair DNA without the need to repeatedly amplify microsatellite loci. Finally, I estimated the sizes of a local badger population and of individual social groups using the genetic profiles generated this way.

4.2. Materials and Methods

4.2.1. STUDY SITE

The main study site was located in the Northeast of the Grand-Duchy of Luxembourg, east of the river Ernzt Blanche and between the villages of Ermsdorf and Eppeldorf. The site covered approximately 5.4 km², was situated between 225 and 420 m above sea level and consisted of a mosaic of pasture, arable land and woodland (see Schley

2000 for further details). The study focussed on five adjoining main setts previously identified by Schley (2000), named Ermsdorf 1, Ermsdorf 2, Knäipenhecken, Bëlz and Grott respectively.

4.2.2. GENETIC VARIABILITY OF LUXEMBOURG BADGERS

Before applying a non-invasive genetic approach to the identification of individuals, it was necessary to verify whether the chosen genetic markers exhibited enough variability to produce individual-specific genetic profiles. This was done using samples of ear tissue collected opportunistically from 52 badgers killed on Luxembourg roads during 2000 and 2001. The mean distance between the locations of the road-kills was 16.8 km and the overall area from which the samples were collected, calculated as a minimum-area convex polygon, was 1325 km² (51% of the total area of Luxembourg). In addition, the profiles of 13 animals captured in the main study area in 2002 and 2003 were included in the dataset in order to increase the likelihood of including related animals in the analyses (Waits, Luikart & Taberlet 2001), giving a total of 65 samples.

DNA was extracted from tissue samples using an ammonium acetate DNA extraction protocol. Between 20 and 50 mg of tissue were added to 600 µl of lysis buffer (40 mM Tris, 20 mM EDTA, 100 mM NaCl). To this, 20 µl of Proteinase K (10 mg/ml) and 100 µl 10% SDS were added and the whole was incubated at 55°C under rotation overnight. After adding 300 µl of ammonium acetate (10 M), the extract was incubated at -80°C for 5 minutes and spun down at 13,000 g for 5 min. The supernatant was added to 0.5X cold 100% isopropanol and mixed gently. After centrifugation at 13,000 g, the isopropanol was removed and the pellet was rinsed with 70% ethanol. The DNA was re-suspended in 500 µl ddH₂O.

Microsatellite loci were tested for linkage disequilibrium using an exact test based on a Markov chain method as implemented in GENEPOP version 3.3 (Raymond & Rousset 1995). GENEPOP version 3.3 was also used to implement the exact tests of Guo & Thompson (1992) to test for deviations from the Hardy-Weinberg equilibrium at each locus. In both analyses, the Bonferroni technique was used to eliminate false assignment of significance by chance (Rice 1989). Allele numbers and frequencies, estimates for expected heterozygosity (H_e) and observed heterozygosity (H_o) were calculated for each locus using GENETIX version 4.02 (Belkhir 1996).

When using microsatellite loci to genotype remotely collected hair samples, it was important to ensure that the resulting genetic profiles were specific to each individual. I tested whether the seven loci planned for use in my study would fulfil this requirement by computing a probability of identity of siblings (P_{ID-Sib}) statistic (Evetts & Weir 1998, Waits *et al.* 2001). P_{ID-Sib} values for badgers in Luxembourg

were calculated using GIMLET 1.3.2 (Valière 2002) on the data set of 65 genotypes obtained from tissue samples taken from road-killed and captured badgers (see above). PROB-ID5 (G. Luikart unpublished) was used to estimate the observed P_{ID} (P_{ID-Obs}) by computing the proportion of all possible pairs of individuals that had identical genotypes.

4.2.3. FIELD METHODS

I set out to snare badger hair using barbed-wire enclosures, similar to those described by Woods *et al.* (1999). They were constructed less than 10 m away from each main sett by suspending a single strand of barbed wire around two or three trees, 20 cm above ground level. The stations were baited with peanuts placed under a pierced box covered with a stone in order to stop non-target species from reaching the bait. Because previous studies suggested that it would be difficult to attract Luxembourg badgers to bait (Schley 2000), peanuts were put near the setts up to four months prior to the construction of the hair-trap. Old bait was replaced weekly until the peanuts first disappeared, after which bait was renewed every two to three days. By Feb 2003, badgers were eating the bait at only four of the five setts, the exception being Grott. Rather than delaying the start of the pilot study, two additional hair traps were constructed at Grott by suspending barbed wire between stakes set either side of a clearly visible badger run.

A preliminary test of a barbed-wire enclosure was carried out at Ermsdorf 2 during Dec 8-15, 2002 and Jan 12-19, 2003. The station was baited daily and all snared hairs were collected and stored in paper envelopes at -20°C until DNA extraction. The main study (henceforth referred to as the “pilot study”) was carried out from Feb 14 to March 7, 2003. The four enclosures and the suspended barbed wire were visited daily to replace bait and collect hair samples (defined as all the hair collected from a single barb). Hair samples that did not contain any follicles were discarded immediately. Both guard hairs and under-fur were collected, stored in paper envelopes at room temperature, and extracted the same day.

Because of the small number of hairs collected at some setts (see Results) and given uncertainties about the quality of DNA to be obtained from single-hair extracts, further samples were systematically collected during March 17-24, 2003. After this period and until May 2003, hairs were collected opportunistically and extracted within three days of collection.

4.2.4. DNA EXTRACTION AND AMPLIFICATION

In order to avoid contamination, all extractions and PCRs were performed in a separate laboratory that was free of concentrated badger DNA or PCR product,

aerosol-resistant pipette tips were used and negative controls were included in each manipulation to monitor contamination. Hair samples were extracted using a chelex protocol (Chelex®100, Bio-Rad, Hercules CA, USA; Walsh *et al.* 1991). After incubating the root portion of the hairs at room temperature for 30 min in 1 ml ddH₂O, 200 µl of 5% Chelex was added to the root and mixed well. This was followed by incubation at 56°C for 30 to 45 min, mixing the samples occasionally. After checking that the hairs were immersed, the Chelex solution was boiled for 8 min. After centrifugation for 3 min at 13,000 g, the supernatant was removed and placed in a sterile tube.

As an initial test of the quality of the hair DNA extracts, I tried to amplify a 533 base-pair (bp)-product of exon-3 of the *c-myc* proto-oncogene (CMYC_seq_F1: 5'-GAAATCGATGTTGTTTCTGTG-3', CMYC_seq_R1: 5'-CAAGAGTTCCTAG-CTGTTC-3'; Smith *et al.* 2002). The sequences of these two primers are conserved among 18 eutherian mammalian species (Miyamoto *et al.* 2000, Smith *et al.* 2002) and the product proved to be amplifiable using badger DNA. The PCR conditions were the same as those used for the microsatellite loci. The *c-myc*-PCR products were visualised on 1% agarose gels.

Of the 39 microsatellite loci published by Carpenter *et al.* (2003), I chose seven loci with alleles shorter than 250 bp for the work on faecal DNA. This was done because the amplification success of faecal DNA is reduced for loci with alleles longer than 300 bp (Frantzen *et al.* 1998). Because of uncertainties about the quality of hair DNA, I adopted the same approach here. The following seven loci were used: *Mel*-105, *Mel*-106, *Mel*-109, *Mel*-111, *Mel*-113, *Mel*-116, *Mel*-117. In order to be able to run all the samples on one gel and after checking the allele range of the various loci in Luxembourg (see Table 4.1), the primers were end-labelled with the following dyes: *Mel*-105: TET, *Mel*-106: TET, *Mel*-109: HEX, *Mel*-111: HEX, *Mel*-113: TET, *Mel*-116: 6-FAM, *Mel*-117: 6-FAM (see Carpenter *et al.* 2003). The microsatellite loci were amplified in a 25 µl volume, each containing 5 µl of DNA extract. The final PCR concentrations were the same as those described in Chapter 2.

The PCR conditions used to amplify the microsatellite loci were slightly modified from those described in Chapter 2. All the microsatellites were amplified using a touchdown profile (Don *et al.* 1991). PCRs started with a 5-min denaturation at 96°C. This was followed by touchdown cycles of 96°C for 45 s, annealing at 61-56°C for 60 s and extension at 72°C for 45 s, decreasing the annealing temperature by 1°C every cycle. The touchdown was followed by 32 cycles of holding the annealing temperature at 55°C. PCRs ended with a final extension at 72°C for 5 min. Reactions were performed using a Bio-Rad iCycler (Bio-Rad, Hercules, USA). Amplification products were separated on a 5% polyacrylamide gel using an ABI

377 DNA sequencer (Applied Biosystems, Foster City, USA), and sized with a TAMRA-labelled size marker with bands of known size every 50 bp. All gels were analysed using GENESCAN Analysis 2.0 software (Applied Biosystems).

4.2.5. RELIABILITY OF HAIR DNA TYPING AND IDENTIFICATION OF GENETIC PROFILES OF INDIVIDUAL BADGERS

By following the trapping procedure outlined in Chapter 5, 13 badgers were captured at the five target setts in 2002 and 2003. A hair sample was taken from every captured animal in order to obtain reference profiles to which the profiles generated from remotely plucked hairs could be compared. The reliability of the reference profiles was ensured by including at least ten hairs in each extraction (Goossens *et al.* 1998) and by genotyping every sample twice at all the loci. After a single round of amplification, the profiles obtained from remotely collected hair DNA were tested for their reliability. Samples that gave rise to three or more alleles at any locus were thought to originate from more than one animal and were excluded from the analysis. GIMLET 1.3.2 (Valière 2002) was used to compare the reference profiles with the remote-DNA profiles and to group profiles together that were 100% identical.

Remote-DNA profiles that were observed only once and did not match any reference profiles could have been obtained for a number of reasons. Firstly, they could originate from an unknown animal that was sampled only once. Secondly, they could correspond to a multiple-individual sample that did not have more than one or two alleles at all the loci examined. Thirdly, a genotyping error could have occurred. Finally, they could be the result of a multiple-individual sample that was genotyped with errors. In order to exclude genotyping errors, the unique profiles were amplified a total of three times.

Multiple-individual samples could be a mixture of two known profiles, one known and one unknown profile, or two unknown profiles. To test the first possibility, all the available reliable single-badger profiles were compared by hand to test whether a combination could be found that would give rise to the observed unique profiles. The possibility that a unique profile was a mixture of a known and unknown profile was tested by comparing the profile in question with all the single-badger profiles on a pair-wise basis. If three different alleles were observed at a specific locus in a pair-wise comparison this possibility was excluded. The probability of a unique fingerprint originating from one rather than two unknown animals was estimated by means of the likelihood ratio of Weir *et al.* (1997). The likelihood of generating a genotype homozygous for an allele A (that has a frequency of p) from a single rather than two DNA contributors was p^2/p^4 . The likelihood of generating a heterozygous genotype with alleles A and B (with frequencies p and q) from a single rather

than two contributors was $[(p+q)^2 - p^2 - q^2] / [(p+q)^4 - p^4 - q^4]$. Likelihood ratios were multiplied across loci.

4.2.6. COUNTING BADGERS BY DIRECT OBSERVATION

In order to validate the results obtained from genotyping remotely collected hair samples, the badgers at the five target setts were counted by direct observation. This was done mainly in the spring of 2003 since at that time of year it was still possible to distinguish between adults and cubs. Observation of a sett was started approximately 2 h before sunset and continued until 1 h after sunset, using night-vision equipment if necessary. I aimed to visit each sett until badgers were observed on a minimum of six occasions.

4.2.7. ESTIMATION OF BADGER GROUP SIZES USING ACCUMULATION CURVES

In order to simulate an actual census operation, I estimated badger abundance using only the samples collected during the main three-week study period. Due to the relatively small number of different profiles collected during this period, mark-recapture analysis was unlikely to produce meaningful results. Therefore, the size of the local population and of individual social groups was estimated using accumulation curves. As described previously (Chapter 3), the technique estimates population size by projection of the asymptote of a rarefaction curve describing the number of samples analysed versus the cumulative number of unique genetic profiles. Given the simulation results presented and discussed in the previous chapter, the asymptotes of the accumulation curves were projected using Eggert's equation.

The rarefaction curve analyses were performed with Program *R* (Ihaka & Gentleman 1996), using a script and data input file generated in GIMLET 1.3.2 (Valière 2002) with the latter file being generated by regrouping and counting identical genetic profiles. Program *R* randomised the order of the profiles in the dataset 1000 times and for each of these randomisations the asymptote was projected using Eggert's equations. The distributions of the projected values for the asymptote were tested for normality using a one-sample Kolmogorov-Smirnov (K-S) test and, depending on the results of this analysis, either the mean or the median values of all iterations for the asymptote, a , were calculated. The variance of this point estimate was analysed by calculating the interquartile range of the results of all the iterations.

4.3. Results

4.3.1. GENETIC VARIABILITY OF BADGERS IN LUXEMBOURG

The suitability of the seven microsatellite loci for genetic profiling of Luxembourg badgers was tested using 65 complete fingerprints obtained from road-kills and

Table 4.1: Genetic variability of the seven microsatellite loci used in this study of Luxembourg badgers. The samples were either obtained from road-kills or from badgers captured in the study area in 2002 and 2003. N = individuals analysed; A = number of different alleles observed; H_e = expected heterozygosity; H_o = observed heterozygosity; $P_{ID-Sib}/locus$ = sibling probability of identity for individual loci. The loci in the table are arranged in order of increasing P_{ID-Sib} values. P_{ID-Sib} product = cumulative product of individual P_{ID-Sib} values. P_{ID-Obs} = proportion of all possible pairs of individuals that had identical genotypes after loci were added consecutively.

Locus	N	A	Observed allele size range (bp)	H_e	H_o	$P_{ID-Sib}/locus$	P_{ID-Sib} product	P_{ID-Obs}
<i>Mel-105</i>	65	6	138-148	0.78	0.71	0.38	3.81×10^{-1}	7.45×10^{-2}
<i>Mel-116</i>	65	10	113-136	0.75	0.77	0.40	1.51×10^{-1}	6.25×10^{-3}
<i>Mel-117</i>	65	6	174-195	0.72	0.69	0.42	6.33×10^{-2}	4.81×10^{-4}
<i>Mel-113</i>	65	6	120-132	0.68	0.68	0.45	2.82×10^{-2}	0
<i>Mel-106</i>	65	7	216-228	0.65	0.55	0.47	1.32×10^{-2}	0
<i>Mel-111</i>	65	4	132-142	0.58	0.54	0.52	6.81×10^{-3}	0
<i>Mel-109</i>	65	4	106-129	0.31	0.32	0.72	4.92×10^{-3}	0

captures (Table 4.1). Two loci (*Mel-111*, *Mel-116*) departed from Hardy-Weinberg expectations at the 0.05 level before adjustment with the sequential Bonferroni test, but not after. There was no linkage disequilibrium between any pair of loci. The average number of alleles per locus was 6.14 (SD = 2.04; range: 4-10) and mean expected heterozygosity value was 0.64 (SD = 0.16; range 0.31-0.78). The P_{ID-Sib} calculation suggested that, of the seven microsatellite loci, the six most informative loci would be sufficient to distinguish between sibling badgers with more than 99% certainty (Table 4.1). The observed P_{ID} showed that the proportion of individuals with identical profiles dropped to zero if the four most informative loci were included in a genetic profile.

4.3.2. SUCCESS OF HAIR CAPTURE

Fourteen hair samples were collected during a trial run of a barbed-wire-enclosure hair trap at Ermsdorf 2 social group. During the main three-week study period, in February and March 2003, 71 hair samples were collected at the five target sets (Table 4.2.a). Only a single guard hair was included in about a third of the extractions (Table 4.2.b). Because of uncertainties about the quality of the results, a further 28 samples were collected from March to May 2003. Overall, therefore, a total of 113 hair samples was available for testing the ease, reliability and precision of the methodology.

Table 4.2: (a) Remote collection and extraction success of badger hair sampled during the main three-week study period from Feb 14 to March 7, 2003 and during the total collection period from Feb to May 2003. At the first four setts, baited barbed wire enclosures were used to snare hairs. At Grott sett, hair was obtained by suspending barbed wire over a badger run. ED 1 = Ermsdorf 1; ED 2 = Ermsdorf 2; KH = Knäipenhecken; B = Bëlz; G = Grott. Samples were considered to be “useful” if they originated from a single animal and gave rise to amplifiable DNA. (b) The number of guard hairs included in the extractions. Results for samples collected between Feb 14 and March 7, 2003.

(a)	Badger sett					Total
	ED 1	ED 2	KH	B	G	
<i>Three-week pilot study</i>						
Hair samples collected	35	9	5	16	6	71
Useful samples	34	9	5	14	4	66
<i>Total collection period</i>						
Hair samples collected	47	24	10	20	12	113
Useful samples	44	23	10	18	10	105

(b)	No. of guard hairs / extraction				Total
	1	2 to 3	4 to 10	>10	
Samples with guard hairs only	22	13	6	10	51
Samples with guard hair and under-fur	2	4	6	0	12
Samples with under-fur only	-	-	-	-	8

4.3.3. RELIABILITY OF HAIR DNA TYPING

Of the 113 hair samples, 108 gave rise to a complete profile after a single round of amplification, with only five samples not containing any amplifiable DNA. Three profiles contained more than two alleles at one locus and were excluded from the analysis. The accuracy of the remaining 105 profiles was tested by comparing them to reliable reference profiles obtained from trapped animals (Table 4.3). There was a 100% match between 94 hair DNA profiles and the reference profiles. Of these 94 profiles, 31 were generated from a DNA extract obtained from a single guard hair.

The 11 samples that did not match any reference profiles were compared with one another to identify possible amplification errors. Four different profiles were identified. Two profiles (called “B” and “C”) were observed more than once (see Table 4.3) and were considered accurate because identical genotypes were generated from reliable samples obtained from more than ten outer guard hairs as well as from single-hair extracts. The two remaining profiles, “A” and “D”, were amplified a total

Table 4.3: Genetic profiles generated from badgers in the five social groups under investigation. Individuals whose profiles were only known through non-invasive hair capture are marked with a “-” in the second column. The frequency with which the different genetic profiles were generated from 105 remote hair DNA extracts is given in the last three columns. Period A: 14/02/03 to 07/03/03; period B: 17/03/03 to 02/05/03; period C: 09/12/02 to 25/01/03. During period C, samples were only collected from Ermsdorf 2. “-” indicates that the profile of the captured badger had not been obtained non-invasively. The order of the loci in the Table corresponds to the increasing Probability of identity values determined for the local study population (see text).

Individual	Year & status when caught	Alleles at microsatellite loci under investigation							Observations in period		
		Mel-105	Mel-116	Mel-117	Mel-113	Mel-106	Mel-111	Mel-109	A	B	C
“Ermsdorf 2” Social Group											
EMa1	2002, adult	142 144	121 136	193 193	124 130	222 224	132 138	116 116	2	1	1
EMa2	2002, adult	144 144	123 132	187 193	124 124	222 224	132 132	116 116	1	0	0
EMa3	2002, cub	144 146	121 123	193 193	124 124	222 224	132 132	116 116	5	0	2
EMa4	2002, cub	142 144	121 123	174 193	124 124	222 222	132 132	116 116	0	0	10
Profile A	-	142 144	121 132	193 193	124 130	222 222	138 138	116 116	1	0	0
“Ermsdorf 1” Social Group											
EMb1	2003, adult	142 144	132 136	193 193	124 124	218 222	132 138	116 116	19	7	na
EMb2	2002, cub	142 142	132 136	174 193	124 124	222 222	132 138	116 116	--	--	--
EMb3	2002, cub	142 142	121 136	193 195	124 130	222 224	132 132	116 116	10	2	na
Profile B	-	142 146	121 132	174 174	124 130	222 224	132 138	116 116	5	1	na
“Knäipenhecken” Social Group											
KH1	2003, adult	148 148	123 123	174 195	126 130	222 222	132 140	106 116	1	0	na
KH2	2002, cub	146 148	123 136	174 174	130 130	222 222	138 140	106 116	4	5	na
KH3	2002, cub	146 148	123 132	174 189	130 130	222 222	132 140	106 116	--	--	--
“Bëlz” Social Group											
B1	2003, adult	144 146	132 136	187 189	124 124	222 222	132 138	116 116	3	2	na
B2	2003, adult	144 146	123 132	174 191	124 124	218 222	132 140	116 116	11	2	na
“Grott” Social Group											
G1	2003, adult	142 142	123 136	174 189	120 124	222 224	132 140	116 116	2	4	na
Profile C	-	144 146	132 132	174 193	124 130	222 222	140 140	116 116	1	2	na
Profile D	-	142 144	123 132	189 193	124 124	224 224	132 140	116 116	1	0	na

of three times at all the loci. The same results were obtained all three times, so that a single round of amplification would have sufficed to generate a reliable profile.

At the start of the study, I tried to predict the quality of the extracts by amplifying a 533 bp product of the *c-myc* proto-oncogene. Only the 99 samples collected from Feb 14 onwards were considered. A band of the expected size could be obtained from 92 extracts and these samples subsequently also allowed the generation of an accurate genetic profile in a single round of amplification. Two of the remaining seven samples produced reliable genotypes. The results therefore suggest that amplification of this long product is a good predictor of the reliability of microsatellite genotyping.

To summarise, in a total of 749 positive amplifications of microsatellite loci from non-invasively collected hair DNA, there were no obvious genotyping errors. Even DNA samples obtained from single outer guard hairs provided reliable genotypes after only one round of amplification.

4.3.4. GENETIC PROFILES OF INDIVIDUAL BADGERS

From the 105 usable DNA samples collected from December 2002 to May 2003, 15 different profiles were generated (Table 4.3). Of these, eleven could be matched to known profiles from badgers captured during 2002 and 2003. As explained above, profiles “B” and “C” were scored reliably from single-hair extracts. Comparison of the reliable single-badger profiles showed that, assuming equal contribution from different genomes, the remaining profiles “A” and “D” could neither be a mixture of any two known individuals nor of a known and an unknown individual. The Weir *et al.* (1997) likelihood ratio suggested that it was 1.9×10^5 or 1.0×10^4 more likely that profiles “A” and “D” respectively originated from a single rather than two unknown individuals. These results suggest that the 15 different profiles were genuine. The genetic profiles of badgers “EMb2” and “KH3”, which had been captured as cubs in 2002 at Ermsdorf 1 and Knäipenhecken setts respectively, were not obtained from the remotely collected hair samples.

4.3.5. COUNTING BADGERS BY DIRECT OBSERVATION AND BASELINE ESTIMATES OF BADGER ABUNDANCE

Hair-DNA typing results were validated by censuses obtained from direct observation of the five setts (Table 4.4). Badgers from Ermsdorf 1, Ermsdorf 2 and Bëlz setts were observed on a minimum of six occasions. At Grott and Knäipenhecken, however, wind conditions were often unfavourable and the topography and vegetation cover of the setts made them difficult to observe. Overall, the minimum number of badgers counted in the five social groups was 13 (Table 4.4). However, it is unlikely that all the inhabitants of Grott and Knäipenhecken setts were observed and counted.

Table 4.4: Estimates of badger group sizes. Badgers were counted by direct observation of the five setts. The counts marked with (*) were based on at least six observations. The profiles were counted that had been generated from all the useful DNA samples collected and that had been generated during the three-week period of the pilot study from Feb 17 to March 7, 2003. A rarefaction curve based on Egger's equation was used to estimate the sizes of the different social groups for both collection periods. The fifth column gives the number of times the regression was iterated with the sample order randomised each time. A Kolmogorov-Smirnov (K-S) test was used to test the distributions of the projected values for the various asymptotes for normality. The median values of all iterations for the asymptote represent the population estimate, except in the case of the point estimate marked with (**), which is the mean of the various asymptote values. The variance of the estimate was analysed by calculating the interquartile ranges of the results of all the iterations.

Social group	Animals observed	No. of genetic profiles	Point estimate	No. of iterations	Test for normality		Interquartile range
					K-S Z	P	
<i>Total collection period</i>							
All 5 groups	13	15	14.3**	1000	1.148	0.143	13.8-14.8
Ermsdorf 1	3*	3	3.0	999	7.528	<0.001	3.0-3.0
Ermsdorf 2	5*	5	5.0	999	9.109	<0.001	4.5-5.3
Knäipenhecken	1	2	2.1	918	9.672	<0.001	2.0-2.4
Bëlz	2*	2	2.0	1000	10.292	<0.001	2.0-2.0
Grott	2	3	3.1	987	10.812	<0.001	3.0-3.4
<i>Three-week pilot study</i>							
All 5 groups	13	14	14.0	1000	3.309	<0.001	13.3-14.9
Ermsdorf 1	3*	3	3.0	999	7.687	<0.001	3.0-3.0
Ermsdorf 2	5*	4	4.4	923	8.062	<0.001	4.1-5.3
Knäipenhecken	1	2	2.1	1000	8.365	<0.001	2.1-2.4
Bëlz	2*	2	2.0	998	9.360	<0.001	2.0-2.1
Grott	2	3	4.6	834	11.549	<0.001	4.5-4.6

Indeed, as explained above, the genetic profiles collected from Dec 2002 to May 2003 showed that a minimum of 15 badgers were present in the study area in early 2003.

During the six months of the study, 44 useful samples were collected at Ermsdorf 1, 23 at Ermsdorf 2 and 18 at Bëlz social group (Table 4.2.a). The same number of different genetic profiles was generated from these samples as animals had been counted at the three different setts (Table 4.4). It was only possible to collect ten useful hair samples at both Grott and Knäipenhecken during the whole study (Table 4.2.a.). The genetic profiles generated from these samples suggested that at least one more animal was present than had been observed at each of these setts (Table 4.4).

4.3.6. DIRECT ENUMERATION OF DNA PROFILES AS A MEANS OF ESTIMATING BADGER ABUNDANCE

In order to simulate an actual application of the hair-trapping technique, the profiles generated from hair samples collected during the pilot study period from Feb 14 to March 7 2003 were used to estimate the badger group sizes and local abundance. Given that one sample originated from more than one individual and that four extractions did not produce any DNA, 66 profiles were available for the analyses. During the three-week study, a genetic profile was obtained from 14 of the 15 animals known to be present (Table 4.3). Even though only small numbers of samples were collected from Ermsdorf 2, Knäipenhecken and Bëlz setts (Table 4.2), the number of genetic profiles counted at all five setts corresponded well with the baseline estimates obtained by direct observation or from the larger dataset of profiles (Table 4.4). The one profile identified during the total collection period that was missing from the smaller dataset belonged to a member of the Ermsdorf 2 sett.

4.3.7. ESTIMATION OF BADGER GROUP SIZES USING ACCUMULATION CURVES

Given the simulation results presented and discussed in the previous chapter, the asymptotes of the accumulation curves were projected using Eggert's equation (Table 4.4). Considering all the estimates, an asymptote was obtained in at least 834 iterations. The K-S test showed that, in all but one case, the frequency distribution for the 1000 projected values of the asymptote did not follow normality and, consequently, the median was considered to be the appropriate summary statistic.

The rarefaction technique was first applied to the complete dataset, containing all the DNA profiles generated from samples collected between Dec 2002 and May 2003 (Table 4.4). The projected asymptotic social group sizes were identical to the number of genetic profiles identified in all five groups. The estimate of total population size of 14.3 was smaller than the total number of different genetic profiles obtained (*i.e.*, 15).

Thus, the results of the statistical technique suggested that all the individuals had been identified during the total collection period.

When the rarefaction method was applied to the dataset generated for the period between Feb 17 and March 7, 2003, the point estimate obtained for the size of the local population, as well as for setts Ermsdorf 1, Knäipenhecken and Bëlz, was identical to the respective number of profiles generated from samples collected in those setts during that period (Table 4.4). The median asymptote value obtained for Ermsdorf 2 suggested the presence of a further animal in this sett and Grott social group was estimated to include 4.6 animals.

4.4. Discussion

4.4.1. Genetic variability of badgers in Luxembourg

Previous studies, using different types of genetic markers, have reported low to moderate values for the genetic variability of Eurasian badgers both from the UK and the European mainland (Evans *et al.* 1989, da Silva *et al.* 1994, Bijlsma *et al.* 2000, Pertoldi *et al.* 2000, Domingo-Roura *et al.* 2003). Compared to these studies, the average H_e of 0.64 in Luxembourg was relatively high. However, high variability (in addition to small product size) was one of the criteria for choosing our specific loci out of the panel of 39 markers reported by Carpenter *et al.* (2003). Even though differences in sample size and area of sample collection make comparisons difficult, the overall variability of the Luxembourg microsatellite loci appears to be similar to that of a population at Woodchester Park, UK (Carpenter *et al.* 2003).

P_{ID-Sib} values confirmed that the seven microsatellite loci used in our study were sufficient to produce individual-specific profiles with more than 99% certainty, even if the individuals in question were siblings. This was supported by the fact that when at least the four most informative loci were included in a profile, no identical genotypes were observed.

4.4.2. Predicting the quality of DNA extracts

Before genotyping DNA extracts, I tried to assess their quality by amplifying a 533-bp-long fragment of exon-3 of the *c-myc* proto-oncogene (Smith *et al.* 2002). It was possible to amplify this fragment in 93% of samples, all of which also allowed error-free genotyping of our seven microsatellite loci. Amplification of the *c-myc* fragment should therefore be a good predictor of DNA quality in future studies. The primers used are conserved across 18 mammalian species (Miyamoto *et al.* 2000, Smith *et al.* 2002) and could thus also be used with work on other species.

Due to budget constraints, I did not trial any additional loci other than the seven

reported here. However, since it was possible to amplify a 533-bp-long fragment from most samples, it seems likely that microsatellite loci with alleles longer than 250 bp would also have produced reliable genotypes. If this was the case, it might be possible to use fewer but more variable loci and thereby reduce the cost of a study.

4.4.3. HAIR-SAMPLING AS A PRACTICAL METHOD OF ESTIMATING POPULATION SIZE

If genotyping of non-invasively collected hair samples is to be an effective way of estimating the size of badger populations, it needs to be reliable and accurate. Here, I assess our results with respect to these attributes.

As regards reliability, it is necessary that a high proportion of hair samples yield reliable genotypes without the need for repeated amplifications. Of 113 hair samples, five did not produce any amplifiable DNA and three originated from at least two different animals. By comparison with reliable reference genotypes of captured badgers, by comparison of samples amongst themselves and by triple amplification of unique profiles, I showed that all extracts that contained amplifiable DNA, including those (about a third of the samples) that were obtained from single guard hairs, produced 100% accurate profiles in a single round of amplifications. Thus, the DNA extracted from remotely collected badger hairs allowed error-free genotyping.

Some previous studies have been equally successful in obtaining reliable genotypes from single hairs (*e.g.*, Higuchi *et al.* 1988, Sloane *et al.* 2000) whereas others needed to pool up to 10 hair follicles (*e.g.*, Goossens *et al.* 1998, Woods *et al.* 1999, Mowat & Strobeck 2000). A comparison of these studies suggests that a possible explanation for these discrepancies might be a delay between hair plucking and DNA extraction. According to Roon *et al.* (2003), DNA quality of hair samples started to degrade after a storage period of six months and I suggest that it is important for DNA to be extracted as soon as possible after collection of the hair, as was done in my study. An advantage of the chelex method, which I used, is that it is simple enough to be carried out in the field on the day of sample collection.

It is possible that in a high-density population, the proportion of samples that are a mix of two individuals will be greater than observed here. However, even though reliable genotypes can be obtained from single-hair extracts, on a practical level it is desirable to pool hairs in order to increase DNA quantity, the amount of PCR product and, ultimately, the ease with which samples can be genotyped. In future applications of this methodology, I recommend that two extractions be performed with multiple-hair samples, one containing a single guard hair with a clearly visible follicle and the other containing all the remaining hairs. While mainly working with the pooled extracts, the corresponding single-hair extracts could always be used at a later stage to confirm or reject the profile generated using the first extract.

Only hair samples collected during the main study period from Feb 14 to March 7 were used to estimate badger abundance. The accuracy of the resultant estimates was assessed by comparison to the observed number of badgers and the number of genetic profiles obtained for the whole 6-months collection period. Overall, 13 animals were observed at the five setts whereas genotyping of all the hair samples yielded 15 profiles. It is generally suspected that direct observation leads to underestimation of population sizes (Macdonald *et al.* 1998, Tuytens *et al.* 2001), and the fact that more profiles were generated than badgers were counted by direct observation is consistent with this view. Nevertheless, at Ermsdorf 1, Ermsdorf 2 and Bëlz setts, where badgers were relatively easy to observe and where most of the hair samples were collected, the number of badgers observed corresponded exactly with the number of genetic profiles compiled from hair samples collected during the whole collection period.

During the three-week pilot study, a genetic profile was obtained from 14 of the 15 animals known to be present and the number of genetic profiles counted at all five setts corresponded well with the baseline estimates. At three setts (Ermsdorf 2, Knäpenhecken and Grott), the 18 usable samples were shown to originate from nine of the ten badgers known to be present there. It appears as though in a short period of time, a high proportion of this medium-density population was sampled, so much so that a census of the different genetic profiles gave rise to an estimate that compared well with the results from direct observations and from the complete dataset of genetic profiles generated during a 6-months period.

In order to test whether the accuracy of the estimate could be improved, a rarefaction analysis was performed on the profiles collected during the pilot study. The projected size of four social groups corresponded with the baseline estimates. The Grott dataset consisted of only four genetic profiles that originated from three different animals and the sett was estimated to have 4.6 inhabitants. However, when performing the same analysis on all the genetic profiles gathered from this sett, the sett was estimated to have 3.1 members, a value that corresponded to the baseline estimate. Generally, when the rarefaction technique was applied to the complete datasets, the projected group sizes were identical to the number of genetic profiles identified. In other words, in situations where only a few genetic profiles have been gathered from a small number of samples, more conservative and reliable estimates appear to be obtained by simply counting the number of different profiles.

The estimates of group size and total size obtained for the three-week pilot study compared well with the baseline estimates. However, the highest alternative estimate, which was the total number of different profiles collected during the study, might be a slight underestimate of the actual number of badgers that were present in the study

area. The genetic profiles of two juvenile badgers, caught in 2002 near Ermsdorf 1 and Knäipenhecken setts respectively, were not identified from the hair samples collected non-invasively in Feb/March 2003. Badger EMb2, which was captured and radiocollared in November 2002, was found dead on the 29th of March 2003 on a road 8 km linear distance away from its natal sett. Given the results from direct observation, it seems likely that this animal had already dispersed at the start of the hair-capture exercise. When analysing faecal samples collected in January 2003, the profiles of a further animal was identified (see Chapter 5). It is therefore possible that 17 badgers were present in the study area at the time of sample collection. In the period from Feb 17 to May 2, 2003, hairs were found on only seven occasions in the hair trap at this particular sett, giving rise to ten hair samples. The fact that only a few samples were collected during a relatively long sampling period suggested that the local badgers did not systematically visit the hair trap. Or to put it differently, hairs were plucked on so few occasions that possibly not all the resident badgers had been sampled.

In order to avoid negatively biased abundance estimates the methodology needs to be improved by optimising the hair collection techniques. One possibility of collecting more hair samples would be to increase the number of hair traps at individual setts. It would also be necessary to replace peanuts with more alluring bait. In my study, setts needed to be pre-baited for up to four months in order to attract badgers to the bait and at one sett (Grott) bait was never taken. In the UK, badgers are readily attracted to peanut bait (Kruuk 1989, Delahay *et al.* 2000a), so that the use of baited barbed-wire enclosures should be feasible there. An alternative approach, however, would be to suspend barbed wire over well-used badger runs, as I did at Grott sett. The fact that this technique yielded three different profiles from four usable hair samples collected at a single run suggests that it deserves further testing. It might also be possible to suspend barbed wire or double-sided adhesive tape over sett entrances (Sloane *et al.* 2000).

It is essential for the general utility of the methodology that (a) a high proportion of a badger population can be sampled in a relatively short period of time (three weeks in my case) and (b) no expensive repeated amplifications of DNA are required to obtain reliable genetic profiles. Future optimisation of hair collection should lead to larger datasets. It is also realistic to envisage that in a high-density population, more hair samples will be collected. It was shown in a previous chapter that, at large sample sizes, a high proportion of the population needed to be sampled to obtain unbiased estimates of population size. In contrast to faecal genotyping, the results presented here suggest that this requirement can be fulfilled with hair trapping. Given the reliability of plucked hair DNA, an increase in the number of samples

collected and genotyped would not lead to a disproportionate increase in the cost of the study. I conclude that, given its reliability and accuracy, genotyping of remotely plucked hairs can potentially form the basis of a novel and cost-effective approach to estimating badger abundance.

4.5. Summary

This chapter summarises an investigation of the reliability and accuracy, and hence practicality, of estimating badger abundance by genotyping DNA extracted from remotely plucked hair. Hair samples were collected near five target setts using a baited barbed-wire enclosure or barbed wire suspended over a clearly visible badger run. Of the 113 samples collected over a 6-month period, 105 originated from a single animal and gave rise to amplifiable DNA. Reliable microsatellite profiles were obtained in a single round of amplifications, even from single-hair extracts. I demonstrate that population size estimated from remotely collected hair was similar to a conservative baseline estimate. Furthermore, given the fact that a high proportion of the local badger population was sampled in a relatively short period of time, the method has the potential to form the basis of a feasible, practicable and cost-effective technique for estimating badger abundance, applicable independent of habitat characteristics and over a range of population densities.

Chapter 5

Spatial Organisation of Badgers in a Medium- Density Population in Luxembourg and Analysis of Territorial Marking by Genotyping Faecal DNA

5.1. Introduction

In previous chapters I showed that it was possible to reliably genotype DNA obtained from badger faeces (Chapter 2), and attempted to use the DNA profiles generated in this way to estimate the size of social groups of badgers (Chapter 3). However, DNA profiles obtained from badger faeces could potentially provide other information, for example about the ranging behaviour and scent-marking activity of specific individuals or sexes. In the present Chapter, I investigate these aspects of badger behaviour using a combination of DNA profiling and radio-telemetry.

Because badgers show no obvious cooperative behaviour, it has been difficult to understand the advantages of group living in this species (Woodroffe & Macdonald 2000). Researchers have therefore tried to explain the formation of non-cooperative groups in terms of group territoriality, the idea being that groups could form if a territory that was large enough to support a single individual could also support additional individuals without additional costs in terms of competition for resources (*e.g.*, Macdonald 1983, von Schantz 1984, Kruuk & Macdonald 1985; see also Chapter 1). But if territorial behaviour was a necessary precursor of sociality, then the question arises why badgers form territories in the first place. Roper *et al.* (1986) put forward the suggestion that territories are maintained by males to defend females. They supported their hypothesis, named Anti-Kleptogamy Hypothesis (AKH), by their observation that seasonal change in boundary latrine use, taken to represent territorial defence, correlated better with mating activity than with food availability.

Many authors, however, believe that territoriality originated as a means of defending food resources. If this was the case, the social and spatial organisation of badgers could simply be a consequence of a number of animals simultaneously occupying the same food-based territory (Resource Dispersion Hypothesis or RDH: see Kruuk 1978b, Kruuk & Parish 1982, Macdonald 1983, Kruuk 1989). Another example of a food-based explanation for the formation of territories is the Passive Range Exclusion Hypothesis (PRE) by Stewart *et al.* (1997). According to these authors, badgers create a gradient of food availability when they travel away from their sett to feed. Because food closest to the sett is discovered and exploited first, the probability of encountering an unexploited resource increases with increasing distance from the sett and an area of high resource availability persists between neighbouring main setts. It would not be in the interest of a specific animal to penetrate a neighbouring group's range as the available food resources would decrease but travel times increase. On the contrary, animals would optimise their feeding behaviour by foraging in the border areas. Accordingly, the existence of group ranges observed in badgers would be a result of individuals avoiding each other in the context of exploitation competition

and feeding optimisation, rather than of active territorial defence. Faeces in border latrines would therefore serve mainly to advertise individual presence and signal encounter likelihood and/or foraging pressure of badgers along a particular border (Stewart *et al.* 2001). Accordingly, all individuals, irrespective of age or sex class, should be involved in territorial defence.

In order to gain a better understanding of the underlying reasons for group territoriality in badgers, it is important to know which animals are involved in territorial defence. As noted in Chapter 1, however, the only two field studies to have addressed this issue have produced different results. Roper *et al.* (1993) concluded that males visited boundary latrines more than females, as predicted by the anti-kleptogamy hypothesis, whereas Stewart *et al.* (2002) concluded that all age and sex classes defecate at boundary latrines, as predicted by food-based hypotheses. The main objective of the work reported in this chapter was to attempt to solve this discrepancy by genotyping faecal DNA samples collected from boundary latrines in order to identify and sex the individuals involved in boundary marking.

An additional reason for the work reported here was that a disproportionate number of studies of badger behaviour have been performed in the British Isles, where the population density of badgers is generally high and where environmental conditions are not typical for all badger populations (Johnson *et al.* 2002). I therefore wanted to carry out the present study on a population of badgers in continental Europe, in a forested landscape with a lower population density of badgers. I chose Luxembourg for several reasons, namely: (1) the country is located between 49°26'52" and 50°10'58" Northern latitude, placing it in the centre of continental Europe; (2) the climate is temperate without extremes (Faber 1971); (3) the minimum overall badger population density has been estimated to be 0.37 adults per km² (Schley 2000); (4) the habitat consists of a forest-farmland mosaic typical of much of continental Europe (Schley 2000); and (5) in the period from July to October, badger diet in Luxembourg is very variable, with maize (*Zea mays*), insects, plums (*Prunus domestica*) and earthworms (*Lumbricus terrestris*) as staple foods (Schley 2000). Badgers in Luxembourg are thus potentially very interesting in terms of sociality, being representatives of a low- to medium-density population with a broad food niche.

Although my main objective was to analyse defecation patterns at boundary latrines using faecal DNA profiling, it was necessary to investigate the spatial organisation of my study population in Luxembourg first. I therefore used radio-tracking to test whether the spatial system could be considered territorial in so far as it consisted of non-overlapping group-ranges, and I carried out a latrine search to determine whether badger territory boundaries are delineated by latrines in Luxembourg, as

they are in the UK (*e.g.*, Kruuk 1989, Neal & Cheeseman 1996). I chose a study area where Schley (2000) had analysed badger diet, had already performed limited radio-tracking work, and had identified a number of badger latrines.

Because the study summarised in Chapter 4 (DNA profiling of badgers from hair samples) was carried out in the same Luxembourg study area, the genetic profiles generated from faecal DNA could be compared to reference profiles, giving information about the identities and, to a certain extent, the life-histories of the badgers that defecated at boundary latrines. Additionally, the sex of the resident animals that had not been captured or were too young to be anaesthetised was determined by molecular methods in the present Chapter.

5.2. Materials and Methods

5.2.1. CAPTURING BADGERS

The same study site was used as in the previous chapter. It was located in the Northeast of Luxembourg and was a mosaic of pasture, agricultural land and woodland. Five main setts, named Ermsdorf 1, Ermsdorf 2, Knäipenhecken, Bëlz and Grott respectively, were located within its 5.4 km² (Schley 2000).

Animals were trapped during the period April to June in 2002 and 2003 using cage traps similar to the model presented in Cheeseman & Mallinson (1979; for trapping dates see Table 5.1, below). In order to capture badgers successfully, the cage traps had to be pre-baited for up to four months using peanuts, placed under a pierced box covered with a stone in order to stop non-target species from reaching the bait. Because in 2002, the badgers did not touch the bait when food was plenty, the pre-baiting for the 2003 season was started during the winter. At Grott sett, where the animals did not touch the bait, one badger was caught by means of a free-running snare placed over a clearly visible run, following the instructions of Cheeseman & Mallinson (1979).

Box traps were checked just after dawn, and captured badgers were transferred to a holding cage and weighed. After restraining, adult badgers were anaesthetised by intra-muscular injection in the thigh-muscle of 20 mg/kg of ketamine hydrochloride ("Imalgène", Rhône Mérieux, France) using 10 ml syringes and 3 cm long stainless steel needles. When the badgers were motionless and did not respond to tactile stimuli (Pigozzi 1990), they were sexed, inspected for past and present injuries and a hair sample was taken for future genetic work. The animals were then fitted with radio-collars (Biotrack, UK; frequency band: 147 MHz) and released at the site of capture after full recovery from anaesthesia (3-4 hours). The juvenile badgers (weighing between 6 and 7.5 kg) were not anaesthetised, and hence were not sexed

or radio-collared, but were released from the holding cage after a hair sample had been taken.

5.2.2. ACCURACY OF RADIO-TRACKING FIXES

The animals were followed from a car and on foot, using a hand-held, 3-element Yagi antenna (Biotrack, UK) connected to a M57 receiver (Mariner, UK). With such a mobile tracking system, records would potentially be most accurate if increasing signal strength could be followed until the animal was observed directly. While this “homing-in” technique has been used successfully on badgers (for a review see Kruuk 1989), in the present study it proved impractical because no infrared equipment or image intensifier was available and because the tagged animals were often located in maize and cornfields, making direct observation difficult and disruptive. It was therefore necessary to determine the animal’s position by triangulation. Using this technique, the direction of the strongest signal is determined from several known locations. The locations are then plotted on a map and their intersection (the “fix”) is considered to be the position of the animal at the time when the bearings were taken.

Following Kenward (2001), the following precautions were taken in order to determine the location of an animal as accurately as possible. Bearings were taken from sites that could be clearly identified on a topographical map and that, whenever possible, were clear of rocks, buildings, power lines and woodland edges. A directional fix was only recorded on a map if identical peak signal positions were determined holding the elements of the Yagi antenna both vertically and horizontally. In open country, the tag signal has a horizontal polarization and will be picked up easiest with the antenna polarized the same way. However, horizontally polarized signals often give a diffuse peak and a bearing should be verified with the antenna polarized vertically (Kenward 2001). Since the habitat was quite heterogeneous, it was possible to use landmarks to plot the directional fix as accurately as possible on a map. If the signal was weak or the animal too far away to obtain a precise bearing, further bearings were taken after approaching the animal as near as possible without causing any disturbance. In order to further improve the estimate of an animal’s location, the second bearing was always taken as close to 90° from the first as possible and, if possible, a third bearing was taken. Finally, an effort was made to take consecutive bearings as quickly as possible. Unless when moving between feeding patches, it appeared reasonable to assume that badger movement between successive bearings would not cause major inaccuracies in the determination of an animal’s position (Kruuk 1978a, Kruuk 1989).

5.2.3. AUTOCORRELATION OF FIXES AND RADIO-TRACKING REGIME

Autocorrelation occurs when radio-tracking fixes are taken so close in space and time that they are not independent, that is, the position of an animal at time $t + \Delta t$ is a function of its position at time t , so that one can predict an animal's position based on its last location (Ackerman *et al.* 1990, White & Garrott 1990). Mainly based on simulations by Swihart & Slade (1985), it has been a long held view that home range sizes will be negatively biased if estimated using autocorrelated data sets and that autocorrelation should therefore be avoided or minimised (Swihart & Slade 1985, Worton 1987, Solow 1989, Ackerman *et al.* 1990, Harris *et al.* 1990). Many authors (Andersen & Rongstad 1989, Harris *et al.* 1990, Cresswell & Smith 1992, Zubaid & Gorman 1993, Tallmon & Mills 1994) have recommended the test for statistical independence by Swihart & Slade (1985) as the best method for determining the degree of spatial autocorrelation in a data set.

A number of investigators have realised that, more often than not, practical considerations dictate radio-tracking regimes, so that researchers try to accumulate as many fixes as possible in the available time-span irrespective of issues of autocorrelation (Dunn & Gipson 1977, Harris *et al.* 1990, Cresswell & Smith 1992). In other words, for practical reasons associated with the collection of data, the generation of an autocorrelated dataset is hard to avoid. The traditional approach to this problem has been to either sub-sample the dataset, to use the Minimum Convex Polygon (a technique of home range estimation that is not affected by autocorrelation), or to simply ignore the problem (Worton 1987, Ackerman *et al.* 1990, Harris *et al.* 1990, Cresswell & Smith 1992, Kenward 1992). Some authors have used an arbitrary time interval between fixes that they believed to be large enough to be confident about the independence of their data (*e.g.*, Rolstad & Wegge 1988, Andersen & Rongstad 1989, Schley 2000, Revilla & Palomares 2002).

Despite a reduction in sample size, the implicit assumption when sub-sampling has always been that home range estimates will be more precise due to the elimination of autocorrelation. However, in a study by Reynolds & Laundre (1990), the home range sizes of pronghorns (*Antilocapra americana*) and coyotes (*Canis latrans*) were underestimated when limiting sampling to time intervals that were statistically independent. In addition, Rooney *et al.* (1998) showed that sub-sampling of autocorrelated data sets led to severe redundancy in the data as well to the underestimation of home range sizes and rates of movement. Rooney *et al.* (1998) concluded that "what is statistically sound may have little or no biological meaning and, in fact, the collection of independent locational data may be undesirable for the correct estimation of range sizes and home range use". Recently, Otis & White (1999) and De Solla *et al.* (2002) have pointed out a conceptual error in the

simulation by Swihart & Slade (1985) that led to the conclusion that autocorrelated data will negatively bias the estimation of home range size. In this simulation, the total number of locational fixes in both autocorrelated and independent datasets remained constant, so that the independent datasets were collected over a longer period of time. This is unrealistic in an actual study where the total length of the sampling period will be the same whether the fixes collected are independent or autocorrelated. Indeed, by simulating datasets of identical total sampling period, but different and constant intervals between successive sampling occasions, De Solla *et al.* (1999) showed that, irrespective of levels of autocorrelation, the shorter the time interval between successive fixes the more accurate the estimates of the true home range. In other words, truncated dataset can be expected to be of limited biological significance.

Given the above considerations, a consensus is emerging in the literature that autocorrelation is required to sufficiently model animal movement and space use (Rooney *et al.* 1998, De Solla *et al.* 1999, Otis & White 1999). White & Garrott (1990, p. 148) have pointed out that the important question in terms of independence of fixes is whether the time interval has been sampled in a representative manner, with each moment in time having an equal chance of being sampled. If this was the case, locational fixes could be assumed to be independent within the time frame of the sample. McNay *et al.* (1994), Swihart & Slade (1997), De Solla *et al.* (1999) and Otis & White (1999) all agreed that autocorrelated data will not invalidate a number of common home range estimators if the time interval between successive fixes remains relatively constant and if inferences about space use and habitat selection are limited to the monitoring time frame.

In order to make our data set robust to the effects of autocorrelation, the following radio-tracking regime was adopted. When following one or two animals at a time, fixes of a specific animal were taken every 30 min. When more than two animals were followed, the location of an animal was recorded once per hour, with a minimum interval of 30 min between fixes. In order to analyse whether the activity period of the animals was sampled in a representative manner, the temporal distribution of the fixes was plotted. While this was only done in retrospect with the data collected from the animals captured in 2002, in 2003 the relevant graphs were updated regularly and the tracking periods were adjusted accordingly.

5.2.4. HOME RANGE ESTIMATORS

Home range has been traditionally defined as “that area traversed by the individual in its normal activities of food gathering, mating and caring for young” (Burt 1943). It does not include occasional excursions (Burt 1943) and the time frame over which the home range is determined as well as the status of the animal (*e.g.*, sex and age)

should be specified (White & Garrot 1990, Harris *et al.* 1990). Furthermore, in order to analyse interactions between individuals, it may be particularly important to identify areas of intensive use within a home range (Samuel *et al.* 1985). These “core” areas are likely to contain nest sites, refuges and the most dependable food sources (Burt 1943, Kaufman 1962, Ewer 1968). Biologists have long been trying to find optimal ways of using radiotracking fixes to describe the three main questions required from studies of spatial organisation of a target species: the size and shape of the home range, and the intensity of use of various parts of the area (Kenward 2001).

One of the earliest, simplest and most commonly used methods in the estimation of home range size is the Minimum Convex Polygon (MCP; Mohr 1947). The outer locations are connected to generate a polygon whose angles do not exceed 180° and the home range is considered to be the area within this polygon. The MCP has been widely used and, because it is the only method that is strictly comparable between studies, it should be included as one of the techniques of home range analysis (Harris *et al.* 1990). Furthermore, the home range estimate obtained using a MCP is not affected by autocorrelation of locational fixes (Harris *et al.* 1990). When working with MCPs, it is necessary to determine how many fixes are needed to define home range of the tracked animals by analysing how the area of a range changes when additional fixes are added to the dataset (Harris *et al.* 1990). The number of fixes necessary to estimate home range is obtained when, in a plot of range size versus number of locations taken, an asymptote is reached or, in other words, when adding additional locations no longer substantially increases the estimated area of the home range.

Kernel functions calculate utilisation distributions describing the relative intensity of an animal's use of areas within a defined space and then specify the home range boundary by the contour that encompasses a selected percentage of the total space used (Van Winkle 1975, Anderson 1982, Silverman 1986, Worton 1989). Seaman & Powell (1996) give a good summary of kernel density estimators and the subsequent explanations follow it to a large extent. The basis of the technique consists of placing a probability density (or kernel) over each fix in the data set, overlaying a grid over the animal's area of use and then estimating the density at each node of the grid by using information from the whole sample. The density estimate for each grid node is obtained by calculating the average of the densities of all the kernels that overlap that specific intersection. Because fixes further away from a node will contribute less than fixes close to it, areas with a high(er) number of fixes will give rise to large(r) density estimates. The advantages of the kernel estimators are that they directly produce a density estimate and that, even though they require a grid to be overlaid on the data

points, they are not sensitive to the size of the grid or its origin (Silverman 1986).

Kernel estimates are greatly affected by the width of the kernel (also called “bandwidth” or “smoothing parameter”). While nearby locations have the greatest influence if kernels are narrow, distant fixes gain in importance if kernels are wide. Choosing an appropriate bandwidth is thus an important exercise (Silverman 1986). Worton (1989) suggested that an appropriate bandwidth would be obtained by the process of the least square cross validation (LSCV), which examines a number of bandwidths and chooses the one that has the lowest mean integrated square error for the density estimate (Silverman 1986). Another issue is whether to use adaptive or fixed kernel estimators. Worton (1989) suggested that by permitting the bandwidth to vary at each node, better estimates of home range use might be obtained. This suggestion has been refuted in later reviews. Worton (1995) and Seaman & Powell (1996) both concluded that fixed kernels would produce more reliable results and, furthermore, after comparing the performance of a number of kernel estimators, the latter authors concluded that cross-validated fixed kernels gave the best results in almost all cases.

The final issue that needs consideration is how to determine the percentage of fixes that describe the core areas and thereby separate range cores from excursive areas. A method that gives a good indication of the core areas is the examination of utilisation plots (Ford & Krumme 1979, Clutton-Brock *et al.* 1982). These graphs compare the cumulative proportion of an animal’s locations and the related increase in area within the total home range. After excluding outliers, home range area should increase less sharply (Clutton-Brock *et al.* 1982) so that it should be possible to identify a point where the gradient of the slope changes and the core area will then be considered to be equal to the lower inclusion value (Harris *et al.* 1990).

5.2.5. DATA ANALYSIS

Unless otherwise noted, all the subsequent analyses were performed using the Animal Movement Analyst Extension (AMAE; Hooge & Eichenlaub 1997) to ArcView®. The size and shape of the home ranges of the animals were determined using the 100% MCP. The area of the MCP delineating all the fixes that were recorded for all the badgers caught at the same sett was calculated and defined, in accordance with Kruuk (1978b), as the range of the corresponding badger group. The sizes of the respective polygons were obtained using the “Location Statistics” option in AMAE. The graphs of home range size versus number of observations were generated using BIOTAS 1.0.2 (Ecological Software Solutions). The sizes of the areas of overlap between the different home ranges were obtained by modifying the home range polygons by hand in ArcView®.

The internal configuration of the animals' home ranges was investigated using fixed kernel techniques. As suggested by Seaman & Powell (1996), Powell *et al.* (1997), Seaman *et al.* (1998, 1999) and Powell (2000), optimal bandwidths were determined using the LSCV technique. The appropriate inclusion level for core areas was determined by examining utilisation plots (Ford & Krumme 1979, Clutton-Brock *et al.* 1982). Ten kernel density isopleths (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% & 95%) were included in the analysis.

5.2.6. IDENTIFICATION OF LATRINES AND COLLECTION OF FAECAL SAMPLES

In December 2002, the main study area was surveyed for dung pits and latrines. A total of ten latrine sites were identified, four of which were located near the main setts and classified as hinterland latrines. Their location as well as work by Schley (2000) suggested that the six remaining latrines lay on the borders of the territories associated with Ermsdorf 1, Ermsdorf 2 and Knäipenhecken. Subsequent telemetry work confirmed this assumption. Faecal samples were collected daily from these six latrines for a total of three weeks at the beginning of 2003 (Jan 15-21, Feb 15-21, March 19-24). Two aliquots of fresh faecal samples were collected at first light and stored in 1.5-ml screw-cap microfuge tubes containing 70% ethanol. Latrines were dusted using builder's chalk (Stanley Tools) in order to identify freshly deposited samples. Towards the end of the sample collection, an 11th latrine was identified on the boundary of Ermsdorf 1 and Knäipenhecken social group. The samples collected from this latrine did not, however, yield any amplifiable DNA.

5.2.7. DNA EXTRACTION AND AMPLIFICATION

In order to avoid contamination, all extractions and PCRs were performed in a separate laboratory in Luxembourg that was free of concentrated badger DNA or PCR product, aerosol-resistant pipette tips were used and negative controls were included in each manipulation to monitor contamination. DNA extractions were performed in mid-March 2003, meaning that, while the samples collected in January and February were stored for some time in 70% ethanol at room temperature, the samples collected in March were extracted on the day following collection. After evaporating the ethanol overnight in a water bath at 56°C, samples were extracted using the GuSCN / silica method described in Chapter 2. Since no rotating oven was available in the laboratory in Luxembourg, samples and extraction buffer were not incubated overnight with rotation, but were agitated for 4 h at 3.5 shakes/s using a Retsch MM200 mixer mill (Retsch GmbH, Haan, Germany). Successful extractions have been reported where faecal samples were incubated only for up to 20 min with the extraction buffer (Reed *et al.* 1997, Verma *et al.* 2003).

Genetic profiles were generated using six of the seven microsatellite loci used with hair DNA, namely *Mel-105*, *Mel-106*, *Mel-111*, *Mel-113*, *Mel-116* and *Mel-117*. As shown in

a previous chapter, these six loci were sufficient to generate individual-specific profiles with 99% certainty. The same PCR reagent concentrations and reaction conditions were used as in Chapter 4, with the exception that, when working with faecal DNA, the touchdown was followed by 39, rather than 32, cycles of amplification with the annealing temperature set at 55°C. Amplification products were separated on a 5% polyacrylamide gel using an ABI 377 DNA sequencer, and sized with a TAMRA-labelled size marker with known bands every 50 base pairs. All gels were analysed using GENESCAN Analysis 2.0 software.

5.2.8. GENOTYPING OF FAECAL SAMPLES

Because of the low quantity and quality of DNA contained in many faecal DNA extracts, amplifications have to be repeated a number of times to obtain reliable genetic profiles (Taberlet *et al.* 1996, Miller *et al.* 2002; see also Chapter 2). From previous work on the estimation of local badger abundance using plucked hairs, reliable reference profiles were available from members of the five target social groups (the available profiles are summarised in Chapter 4). Knowledge of the genetic profiles to be obtained from the faecal DNA samples made it possible to reduce the required overall number of amplifications necessary for reliable genotyping. In a first step, the samples that contained amplifiable DNA were genotyped once at all the loci. In order to test the reliability of the profiles obtained in this way and to strategically direct replication, the following procedure was adopted:

1. Using GIMLET v1.3.2 (Valière 2002), the reliable reference profiles and the provisional faecal profiles were compared and identical profiles were grouped together.
2. The faecal profiles that were 100% identical to the reference profiles were considered correct and further amplifications were unnecessary.
3. The reference profiles were grouped together with the remaining faecal samples that differed by one allele, by two alleles at two loci and by two alleles at one locus.
4. The repeated amplifications were then directed at those loci where there was a difference between the reference and faecal profile until a match was established.
5. If a profile differed at more than two alleles from the reference profiles, the comparative approach outlined in Chapter 2 was adopted until either a match or the uniqueness of the faecal profile was established.

There was one exception to the procedure explained above. The profile of faecal

sample 33 differed from a reference profile by two alleles at two loci. However, at one of the loci in question, which was the second most informative one, a total of three alleles were observed in the comparison. Therefore, the faecal profile was compared by hand with the reference profiles from Ermsdorf 1, Ermsdorf 2 and Knäipenhecken. Despite an initial three-allele difference, repeated amplifications showed that faecal sample 33 originated from animal EMb1 (see Appendix II).

5.2.9. SEX-DETERMINATION USING MOLECULAR MARKERS

The gender of the animals that had not been caught and anaesthetised had to be determined by molecular means. I aimed to simultaneously amplify a fragment of the *SRY* gene and a microsatellite locus as positive control. PCR products were visualised on a 3% agarose gel. Samples that amplified the positive control without amplifying the *SRY* fragment were be scored as females, while those that amplified both fragments ere be scored as males. Griffiths & Tiwari (1993) have reported primers for the amplification of a 216-bp-long fragment of the *SRY* gene in a range of mammal species that included the badger and the authors gave the sequence of the fragment in the different species. A shorter version of the forward primer RG4, described in Griffiths & Tiwari (1993), was used in this study: 5'-GGTCAAGCGACCCATGAACG-3'. Because a fragment shorter than 216 bp was needed for work on non-invasive samples, the sequences published in Griffiths & Tiwari (1993) were used to design a reverse primer for a shorter fragment. Because of the danger of contamination with human male DNA, the forward primer was designed with the help of L. Pope to include a mismatch with the human *SRY* sequence at its first nucleotide. The sequence of the forward primer was as follows: 5'-AAGCATTTTCCACTGGCACCCCAA-3'.

The sex-identification system was tested with DNA obtained from road-killed badgers of known gender. The PCR concentrations and conditions were the same as those used in the microsatellite amplifications described in Chapter 4. When two sets of primers were included in the PCR, the amount of water added to individual reactions was adjusted accordingly. When performing a PCR with the *SRY* primers only, a band was obtained for all the samples originating from males but no amplification product was obtained from female DNA. In order to identify a suitable multiplex system, I amplified the *SRY* fragment simultaneously with microsatellite loci Mel-106 or Mel-117 from six male and six female DNA samples (see Appendix II). Clear and correct bands were obtained only using a PCR multiplex containing the primers for the *SRY* fragment and microsatellite locus Mel-117. When applying this multiplex to DNA obtained from the badgers in the main study site, all animals of known gender were scored correctly (see Appendix II).

5.3. Results

5.3.1. SPATIAL ORGANISATION OF BADGERS IN LUXEMBOURG

A total of eight adult or sub-adult badgers was caught and radio-collared during 2002 and 2003 in the five social groups under investigation (Table 5.1). While male EMb1 had scarred wounds on its rump, the other badgers were in good body condition. For three social groups, data were available for one animal only. Three and two animals were monitored from Ermsdorf 2 and Bëlz social groups respectively. A maximum number of two badgers from each sett were followed at a time. It was possible to collect an average of 157 fixes per animal ($SD = 112$), with the smallest dataset consisting of 40 and the largest of 402 fixes.

In order to analyse whether the activity period of the animals was sampled in a representative manner, the temporal distributions of fixes were plotted. For the two animals tracked in 2002 (EMa1, EMa2), this was done in retrospect (Fig. 5.1.a). These analyses showed a peak in data collection for the later part of the night. By continually monitoring the temporal distribution of fixes in subsequent work, it was possible to avoid a substantial bias in the timing of the data collection (Fig. 5.1.a). Some graphs show paucity in the number of fixes recorded between 21:00 and 21:59 as well as 05:00 and 06:00. This arose because the corresponding badgers either emerged late or returned to the setts early, or because radio-tracking was performed mostly around the summer solstice. Overall, our data set should be robust to effects of autocorrelation.

Table 5.1: Capture and tracking details of radio-collared badgers

Animal	Sett	Sex	Date of capture	Weight (kg)	no. of fixes	Tracking period
EMa1	Ermsdorf 2	F	27/05/2002	11.5	183	29/05/02-31/07/02
EMa2	Ermsdorf 2	M	11/06/2002	13.5	145	12/06/02-07/08/02
EMa3	Ermsdorf 2	F	08/04/2003	11.0	181	11/04/03-24/07/03
EMb1	Ermsdorf 1	M	10/04/2003	12.0	74	10/04/03-15/05/03
B1	Bëlz	M	08/04/2003	12.5	76	10/04/03-15/05/03
B2	Bëlz	F	15/04/2003	11.5	402	18/04/03-27/08/03
KH1	Knäipenhecken	F	10/04/2003	10.0	40	11/04/03-24/04/03
G1	Grott	F	02/05/2003	11.5	154	05/05/03-24/07/03

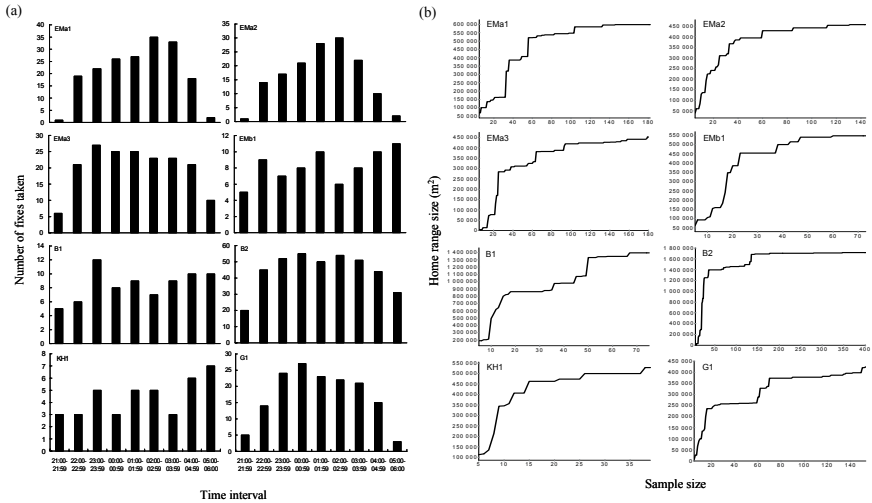


Fig. 5.1: (a) Temporal distribution of the radio fixes recorded for the eight radio-collared badgers. The minimum time interval between fixes was 30 min. (b) Cumulative home range area as a function of radio-tracking effort for eight badgers. The range size was calculated using MCP.

Incremental area plots suggested that the range sizes of all the animals, calculated as a MCP of the recorded fixes, reached an asymptote during the respective tracking period (Fig. 5.1.b). Given our tracking regime, six of the eight accumulation curves reached an asymptote after at least 60 fixes had been taken. In the case of animal G1, a minimum of 70 fixes was necessary for home range size to stabilise. It was only possible to collect 40 fixes from animal KH1 before its collar dropped off.

When estimated as a MCP that included 100% of all the fixes (Fig. 5.2), the sizes of individual home ranges varied from 42.5 ha to 171.8 ha (Table 5.2), with an average of 76.5 ha (SD = 49.9). Females had larger home ranges than their male counterparts caught at the same sett and the home ranges of the animals caught at Bêlz sett were the largest. These differences, however, were not significant (Table 5.3). Individual core areas were obtained by plotting graphs of home range area against kernel isopleth value and defined as the point at which the gradient of the slope changed (Fig. 5.3a). Core areas were equated either to the 50, 60 or 70% inclusion levels and the number of core areas within the range of the eight animals varied from one to five, with an average of 2.9 cores (Fig 5.3.b). In both these aspects, there was no

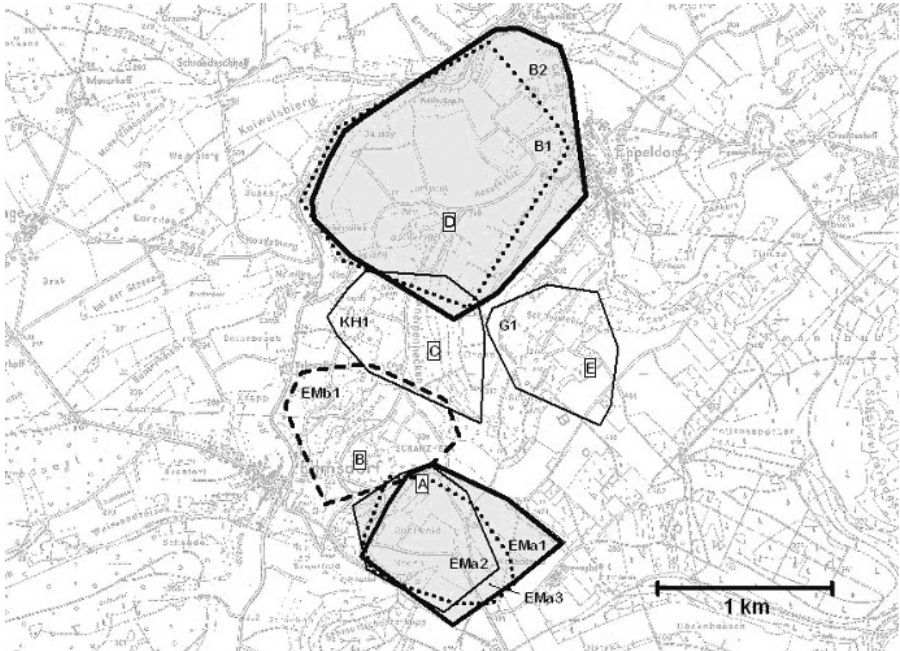


Fig. 5.2: Spatial arrangement of the home range of eight badgers in the study area in Luxembourg. The MCP was used in determining home ranges. The combinations of letters and numbers identify the individual whose home range is represented. The capital letters indicate the location of the main sett: A = Ermsdorf 2; B = Ermsdorf 1; C = Knäipenhecken; D = Bëlz; E = Grott. The area shaded in grey represents an MCP that included all the fixes that were recorded for a social group, in cases where more than one member of the group was captured.

Table 5.2: Home ranges estimates of eight Luxembourg badgers calculated using MCP and 95% Kernel analysis and determined by radio-tracking locations. Determination of the Kernel inclusion levels that was taken to represent the core area of each home range are summarised in Fig. 5.3.a.

Animal	Social Group	Home range area (ha)			% of total home range that belongs to core
		MCP 100%	Kernel 95%	Core range	
EMa1	EM2	59.8	57.0	8.1	13.5
EMa2	EM2	45.7	40.7	5.5	12.0
EMa3	EM2	45.2	42.3	4.7	10.4
EMb1	EM1	54.7	48.9	3.0	5.5
B1	Bëlz	139.5	95.2	13.7	9.8
B2	Bëlz	171.8	126.7	14.6	8.5
KH1	KH	52.9	46.5	6.7	12.7
G1	Grott	42.5	36.6	3.7	8.7

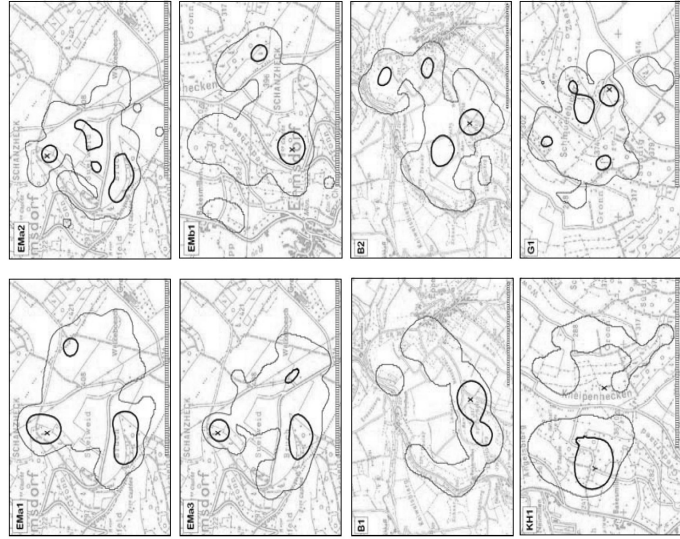
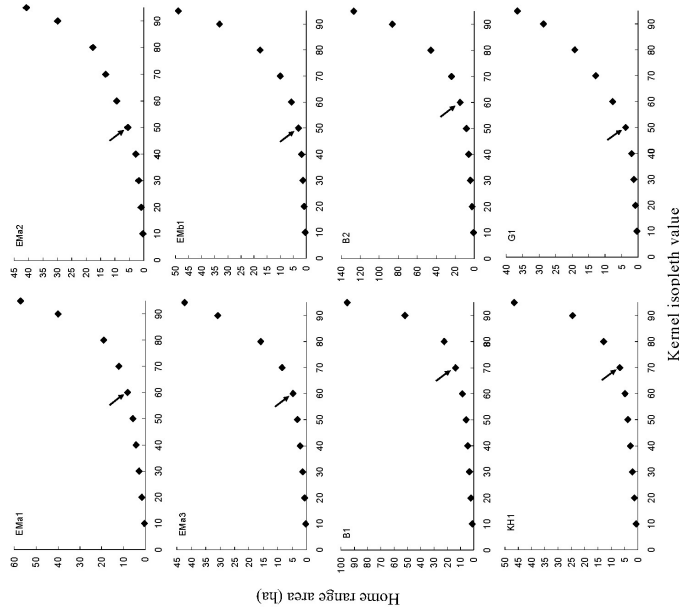


Fig. 5.3: (a) Home range areas against kernel isopleth values for eight radio-collared badger. Core ranges are defined as the inflection point of the corresponding curves (determined visually and arrowed). (b) The position of the 95% Kernel home range (thin outline) and the Kernel inclusion level representing the core of the home ranges of the eight investigated badgers (bold outline). “X” = location of the main sett. “Y” = location of outlying sett. The bar at the right-hand bottom of each map corresponds to a length of 1 km.

Table 5.3: Results from an ANOVA comparing the 100% MCP home range sizes of the different sexes and of members of different setts.

Source of variation	d.f.	Sum of squares	Mean square	<i>F</i>	<i>P</i>
Sex	1	4.4×10^{10}	4.4×10^{10}	4.098	0.292
Sett	4	1.7×10^{12}	4.2×10^{11}	39.439	0.119
Sex*Sett	1	1.9×10^{10}	1.9×10^{10}	1.743	0.413
Error	1	1.1×10^{10}	1.1×10^{10}		
Total	8	6.4×10^{12}			
Corrected total	7	1.7×10^{12}			

apparent pattern according to social group or the sexes of the radio-tracked animals. Core ranges had an average size of 7.5 ha (SD = 4.4; range: 3.0-16.4) and covered an average of 10.1% of individual 100% MCP home ranges (SD = 2.6; range: 5.5-13.5%). There was a significant positive correlation between the 100% MCP home range size and the size of the core areas ($r_s = 0.762$, $P < 0.05$), but the negative correlation between the 100% MCP home range size and the percentage of a range that belonged to the core area was not significant ($r_s = -0.143$, $P = 0.736$). Note that the 95% Kernel home range of animal KH1 does not include the main sett at which the animal was caught. After capture, the animal moved to an outlying sett where it remained until the collar fell off two weeks later.

Overlap of 100% MCP home ranges of badgers caught at the same sett varied from 66.6 to 97.6% (average 83.3%) and the largest home range recorded for both social groups contained almost completely the ranges of the other animals caught at the respective setts (>85% shared area; Fig. 5.2). In the largest overlap between neighbouring ranges, 13.8 and 8.8% of the home range of animal KH1 overlapped with the home ranges of animals B1 and B2, respectively, and 10% of the fixes recorded for animal KH1 were located in the area of overlap. Since, however, the combined number of fixes from the two other animals in this area was three (out of a total of 478) and there was little or no overlap between pairs of the remaining neighbouring ranges (<4% shared area), it can be concluded that the telemetry data collected from the eight individuals represented five exclusive group ranges.

When estimated as a MCP that included all the fixes that were recorded for each member of the two setts, total ranges associated with Ermsdorf 2 and Bëlz social groups were 68.3 and 177.2 ha respectively (Fig. 5.1). Individual home ranges covered, on average, 79% of the area of a group range. While animal EMa3 had the smallest home range compared to the group range size (66% of overlap), the home range of animal B2 corresponded almost completely to the group range (97% of overlap).

5.3.2. COLLECTION OF FAECAL SAMPLES

Faecal samples were collected on a regular basis from six latrines located near the boundaries of the group ranges of Ermsdorf 1, Ermsdorf 2 and Knäipenhecken social groups (Fig. 5.4). Telemetry results suggested that, while latrine 1 (L1) and latrine 2 (L2) were located at an intersection of the ranges associated with setts Ermsdorf 1 and Ermsdorf 2, latrines 3, 4 and 5 (L3, L4 and L5) were located near the border of the Ermsdorf 2 range but not in an area also frequented by badgers belonging to Ermsdorf 1 social group. The sixth latrine (L6) was located near an intersection of Ermsdorf 1 and Knäipenhecken ranges. L1 was by far the largest latrine, consisting of more than 30 dung pits and covering an area of about 500 m². During the three weeks of sample collection in Jan, Feb and March 2003, a total of 70 faecal samples were collected from the six latrine sites (Fig. 5.5), of which more than half were found at L1.

5.3.3. GENOTYPING OF FAECAL DNA SAMPLES

Genotyping of the 70 faecal samples proved to be rather difficult as only 40 extracts (57%) contained DNA of sufficient quality and quantity to allow microsatellite genotyping (Fig 5.5). The genetic profiles that were generated from these 40 samples are summarised in Table 5.4. Despite some effort, it was not possible to amplify locus *Mel*-105 from eight samples that nevertheless allowed amplification of the remaining five loci. Using the P_{ID-sib} -statistics in GIMLET, the probability that identical profiles consisting of these five loci represent the same individual was calculated to be 0.982. Therefore, the profiles in question were classified as being identical to the complete profiles with which they matched. Because one profile could not be matched to any reference, a second amplification was performed to confirm its accuracy. Ignoring failed reactions and given our methodology based on comparing genetic profiles obtained from faecal samples to the ones generated reliably from remotely plucked hairs, a total of 283 amplification reactions were necessary to compile the 40 profiles. This corresponds to an average of 1.2 reactions per locus per genotype. (For the complete genotyping results, please refer to Appendix II):

5.3.4. SEX-DETERMINATION USING MOLECULAR MARKERS

The gender of all the known animals from Ermsdorf 1, Ermsdorf 2 and Knäipenhecken was determined by amplification of a fragment of the *SRY*-gene in a multiplex-PCR with microsatellite locus *Mel*-117 (Table 5.5; Appendix II). The correct sex was identified in the case of the five badgers whose gender was known independently through capture. Additionally, the remaining eight known badgers consisted of one male and seven females. Only one adult male was identified in Ermsdorf 1 and Ermsdorf 2, and no adult male in Knäipenhecken.

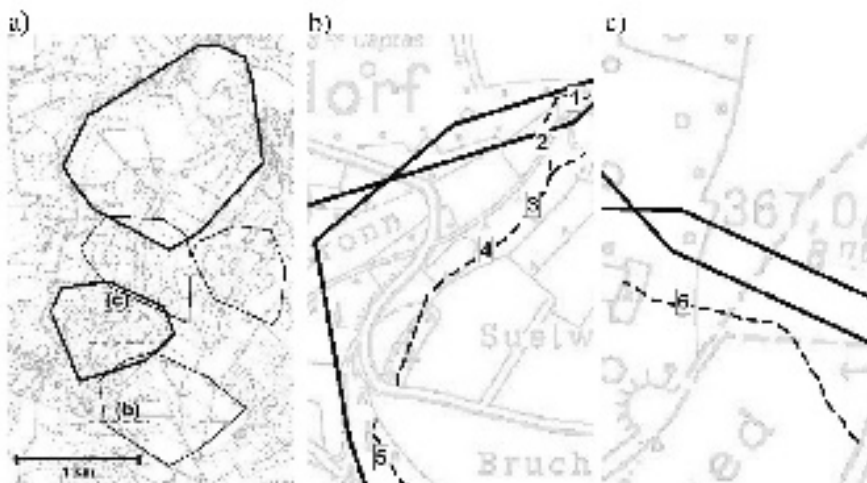


Fig. 5.4: (a) Location of the investigated latrines at the boundaries of three group ranges. (b), (c) exact location of the six latrines. Dashed lines in (b) and (c) show visible badger paths.

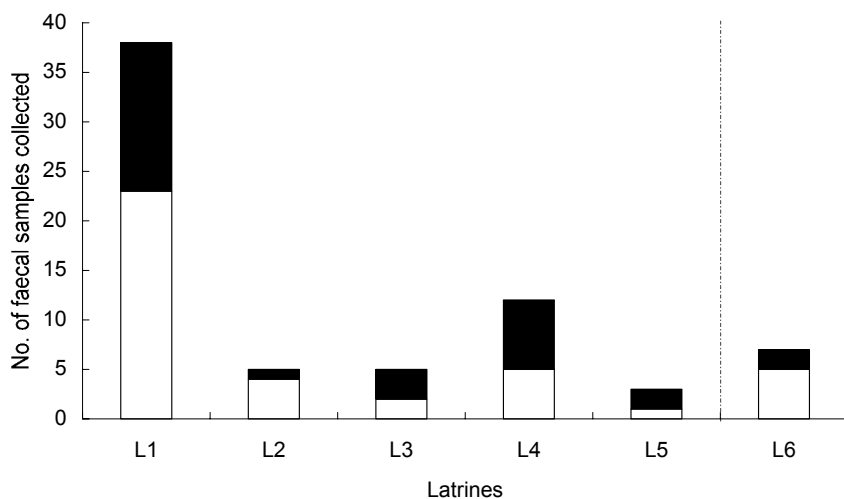


Fig. 5.5: Number and origin of collected faecal samples. The white segment of the columns indicates the number of samples collected that gave rise to amplifiable DNA

Table 5.4: Consensus genotypes generated from 40 amplifiable faecal DNA samples. The table also indicates the matching reference profile (see Chapter 4) and the number of times the faecal profile was observed. It was not possible to amplify locus *Mel-105* from eight samples that nevertheless allowed amplification of the remaining five loci. P_{ID-Sib} -statistics showed that the profiles in question could be classified as being identical to the complete profiles with which they matched.

Profile	Alleles at the microsatellite loci under investigation						Match	Observations
	<i>Mel-105</i>	<i>Mel-116</i>	<i>Mel-117</i>	<i>Mel-113</i>	<i>Mel-106</i>	<i>Mel-111</i>		
P	142 144	121 136	193 193	124 130	222 224	132 138	EMa1	4
Q1	144 144	123 132	187 193	124 124	222 224	132 132	EMa2	2
Q2	-	123 132	187 193	124 124	222 224	132 132		1
R	144 146	121 123	193 193	124 124	222 224	132 132	EMa3	7
S	-	121 123	174 193	124 124	222 222	132 132	EMa4	2
T1	142 144	132 136	193 193	124 124	218 222	132 138	EMb1	6
T2	-	132 136	193 193	124 124	218 222	132 138		2
U	142 142	132 136	174 193	124 124	222 222	132 138	EMb2	1
V	142 142	121 136	193 195	124 130	222 224	132 132	EMb3	1
W1	142 146	121 132	174 174	124 130	222 224	132 138	EMbH	8
W2	-	121 132	174 174	124 130	222 224	132 138		3
X	146 148	123 136	174 174	130 130	222 222	138 140	KH2	2
Y	142 142	121 132	174 195	124 130	222 224	132 140	-	1

5.3.5. TERRITORIAL MARKING

Due to the reduced success of the faecal DNA genotyping, the dataset available for the analysis of defecation patterns at boundary latrines was limited. For this reason, data were pooled from latrines 1 and 2, which were both located in the area of overlap between territories Ermsdorf 1 and 2, as well as from latrines 3 to 5 (Table 5.5). An unmatched profile was identified from a faecal sample collected from latrine 6, located between Ermsdorf 1 and Knäipenhecken group ranges (this has consequences for the results presented in Chapter 4 and is discussed there).

Some trends seemed to emerge from the sparse data that allowed a few careful conclusions. At L1-2, L3-5 and L6, profiles from members of the respective neighbouring social groups were found, confirming that the investigated latrines were indeed located on the territorial boundaries. More generally, and bearing in mind that only a very small number of samples were collected from latrines bordering Knäipenhecken territory, it appears that all or most members of a social group are involved at least to a certain extent in boundary defence by means of defecation. No obvious pattern could be observed that suggested a predominance of a specific sex

Table 5.5: Sex, status and frequency of observed defecation at boundary latrines of all members of the three target groups. (*) indicates sex-identification by molecular markers. The unmatched profile Y originated from a latrine at the border between Ermsdorf 1 and Knäipenhecken group ranges. Data was pooled from latrines 1 and 2, as wells as from latrines 3 to 5. Social group abbreviations: EM2 = Ermsdorf 2; EM1 = Ermsdorf 1; KH = Knäipenhecken.

Profile/Match	Social Group	Sex	Year & status when caught	Observations at latrines			
				L1-L2	L3-L5	L6	Total
EMa1	EM2	F	2002, adult	3	1	-	4
EMa2	EM2	M	2002, adult	2	1	-	3
EMa3	EM2	F	2002, cub	4	3	-	7
EMa4	EM2	F*	2002, cub	0	2	-	2
EMaD	EM2	F*	-	0	0	-	0
EMb1	EM1	M	2003, adult	6	1	1	8
EMb2	EM1	F*	2002, cub	1	0	0	1
EMb3	EM1	F*	2002, cub	1	0	0	1
EMbH	EM1	F*	-	10	0	1	11
KH1	KH	F	2003, adult	-	-	0	0
KH2	KH	M*	2002, cub	-	-	2	2
KH2	KH	F*	2002, cub	-	-	0	0
Profile Y	EM1/KH	-	-	0	0	1	1

or age group in the genetic profiles generated from boundary faeces. The profile of female EMa3, for example, caught as a cub at Ermsdorf 2 in 2002, was found more frequently in boundary latrines than the profiles the adult male and female caught in 2002 at the same sett. On the other hand, the profiles of the adult male and the unknown female were identified much more frequently at boundary latrines than those of the cubs caught at Ermsdorf 1 in 2002. Thus, it appeared as though marking behaviour at the various latrines was not individual specific, but rather that all the individuals concerned help in the defence of the whole territory.

5.4. Discussion

5.4.1. SPATIAL ORGANISATION OF BADGERS IN LUXEMBOURG

In order for issues of autocorrelation not to be relevant, the time interval between observations should be constant and the activity period of the animal should be

sampled in a representative manner (McNay *et al.* 1994, Swihart & Slade 1997, De Solla *et al.* 1999, Otis & White 1999). In order to fulfil the first criterion, fixes were taken at half-hour intervals or every hour, depending on the total number radio-collared badgers that were available for telemetry work. Because the temporal distribution of the fixes was regularly monitored, important biases in the data set could be avoided, thereby fulfilling an important condition for home range estimates to be robust and analyses of home range utilisation patterns to be meaningful.

Animals were captured and radio-tracked in the spring and summer of 2002 and 2003 and the subsequent conclusions should be considered relevant only for these two time periods. Analyses of incremental area plots suggested that all the eight home ranges are accurately represented by means of a 100% MCP *for the total period the respective animals were followed*. Given my tracking regime, a minimum of 60 fixes should be collected to obtain a robust estimate of a badger's home range. It was only possible to collect 40 fixes from animal KH1 before its collar dropped off. Even though an asymptote was reached during the two weeks when the animal was followed, comparison with the other seven graphs suggests that an additional increase in its home range size would probably have been observed had it been possible to track the animal for longer.

The average 100% MCP home range size was estimated to be 76.5 ha (SD = 49.9; range: 42.5-171.8 ha). Group ranges were estimated to be 68.3 ha for Ermsdorf 2 and 177.2 ha for Bëlz social groups. On average, individual home ranges covered 79% of the group range. Therefore, even if only a single animal could be captured at a specific sett, analysis of its telemetry data should provide a good indication of the corresponding total group range size. Similarly to other carnivores (Macdonald 1983, Kruuk & Macdonald 1985), the estimated sizes of badger home and group ranges can vary substantially between studies. Due to differences in the analyses, direct comparison of values between studies is fraught with difficulty. Nevertheless, the values reported here are in the lower end of the spectrum of home ranges (and territory sizes) reported from the medium- to low-density populations on the European mainland (see Hofmann *et al.* 2000). In studies performed in Germany (Bock 1986), Poland (Goszczyński 1994, Kowalczyk *et al.* 2003b) and Spain (Revilla & Palomares 2002), the smallest home ranges were still larger than the largest range reported in my study. Additionally, home ranges larger than 5 km² have been reported from Norway (Brøseth *et al.* 1997) and Switzerland (Graf *et al.* 1996). Contrarily to results reported from studies in Europe, the home range sizes reported in my study fall well within the range reported from populations of different densities in rural Britain (Woodroffe & Macdonald 1993, Krebs *et al.* 1997).

It is believed that the sizes of home ranges of badgers are determined by access to

resources, particularly food and den site. In areas where earthworms constitute the main food, some studies have shown that distribution of earthworm-rich patches determines the size of badger territories (Kruuk & Parish 1982, Hofer 1988, da Silva *et al.* 1993, Kowalczyk *et al.* 2003b). Studies in continental Europe have also found that the proportion of habitats rich in the animal's main food source was a determining factor in territory size (Rodríguez *et al.* 1996, Brøseth *et al.* 1997). Additionally, some authors have suggested that the location of territorial boundaries is a direct result of the location of the corresponding setts, especially if, because of geology, suitable den sites are scarce (Doncaster & Woodroffe 1993, Roper 1993).

Schley (2000) has shown that, in July and August, maize was the food eaten most frequently by badgers in my study area and cereals were shown to be an important food source. Even though earthworms were present in the scats in significant quantities, they were a far less important component of badger diet than in other regions (see Roper 1994). The study area was a mixture of pasture and arable land with cereal and maize fields in the vicinity of the setts and therefore potentially a very rich habitat. The relatively small home range size observed in my study might thus be explained by the high concentration of high quality resource patches within the study area. The fact that there was a significant positive correlation between the size of core areas and home range sizes, but not between the percentage of a home range that was the core area and the total home range size, supports this conclusion.

The appropriate percentage of fixes to be included for core range analyses by Kernel estimators, as determined by the method of Harris *et al.* (1990), varied between 50 and 70%. Core areas covered, on average, 10.1% of the corresponding 100% MCP home ranges, so that, in other words, badgers appear to concentrate their foraging efforts on the 10% of their home range that contains the best resources. Using the same technique, Revilla & Palomares (2002) defined core area as the 70% Kernel isoline without justifying their choice. Kowalczyk *et al.* (2003b) chose the 75% Kernel. Similarly to my results, core areas defined in this way represented a small fraction of the whole home range or territory.

Analysis of the overlap between the different 100% MCP home ranges suggested a pattern of mutually exclusive group ranges. While the home ranges of animals caught at the same sett overlapped to a large extent, exactly the opposite was the case for animals caught at different setts. In order to investigate whether group ranges were defended as territories by scent marking, the boundaries of the group ranges, as determined by radiotracking, were surveyed for the presence of latrines. While seven latrines were discovered on the border of the group ranges of Ermsdorf 1, Ermsdorf 2 and Knäipenhecken, no such feature was identified at the intersections of Bëlz, Knäipenhecken and Grott group ranges. When systematically

collecting faecal samples from the six latrines early in 2003, the majority of the samples were collected from a latrine (named L1), located on the intersection of the group ranges Ermsdorf 1 and Ermsdorf 2, so that this site appeared to be important in the exchange of information between members of Ermsdorf 1 and Ermsdorf 2 social groups. The total number of identified boundary latrines was rather small, even compared to other studies performed on low-density populations (Pigozzi 1990, Graf *et al.* 1996, Hutchings *et al.* 2001, Revilla & Palomares 2002; but see Kruuk & Parish 1982). However, the wounds found on male badger EMb1 can be taken as further evidence of active territorial defence by the animals.

Schley (2000) reported results of from three badgers captured and radio-tracked in 1998 and 1999 in the same study area. The size and shape of the home ranges of a male captured at Ermsdorf 2 and a female from Knäipenhecken are similar to the results for the same setts presented in my study. However, the home range of the third animal, a male captured at Knäipenhecken sett, roughly encompassed both the Knäipenhecken and Bëlz home ranges presented in my study. There are two possible explanations for the discrepancies in the home ranges between the two studies. Firstly, it is possible that, in the three years between the two studies, a large territory consisting of Knäipenhecken and Bëlz split into two separate territories. However, the boundaries of the home range of the female caught at Knäipenhecken in 1998 were very similar to the one of the female caught at that sett during my study. Therefore, the spatial organization of the badgers in the study area in Luxembourg appeared to be similar, at least in the first study, to the social organization of solitary mustelids, where the territories of males can overlap those of more than one female (Powell 1979).

The fact that the badgers in my study area appeared to inhabit group-specific home ranges, and that scent marks were found on some of the boundaries of these group ranges, suggested that, in principle, the spatial system of the Luxembourg badgers was based on territoriality. However, comparison with the results obtained by Schley (2000) suggests that the spatial system is quite flexible, with the boundaries of some territories remaining constant over the years, while others expanded or contracted. It appears that those boundaries that contained latrines remained more stable than those where no such scent-marks could be identified.

5.4.2. GENOTYPING OF FAECAL SAMPLES

The success of the faecal DNA extraction was limited as only 57% of all the available extracts allowed genotyping of all the six microsatellite loci. Compared to the high success rate reported in Chapter 2, the value of 57% falls roughly within the results obtained in most studies (Gerloff *et al.* 1995, Kohn *et al.* 1999, Farrell

et al. 2000, Jansman *et al.* 2001, Lucchini *et al.* 2002). Faeces were collected on the day of their deposition and the same storage and extraction techniques were used as in Chapter 2. The extraction success might have been reduced by the delay between sample collection and extraction (see above) or by the slight modifications in the extraction technique that became necessary due to a change in the laboratory facilities. Otherwise, the fact that the Luxembourg faeces were collected during winter, while the Woodchester Park faeces were collected during summer and autumn, might negatively affect the DNA extraction success due to differences in temperature and humidity between these seasons. In a future study, it might be worth optimising a commercially available faeces DNA kit for badger droppings that could be used as a successful standard technique, independently of laboratory facilities.

The reliable reference profiles that were generated from hair DNA allowed a substantial reduction in the number of necessary PCR reactions to produce reliable profiles. This was achieved by directing repeated amplifications at faecal genotypes that were similar, but not identical, to the ones obtained from hair DNA. In future studies on faecal badger DNA, it would be worth considering obtaining reference profiles first by remotely collected hair samples at main setts.

5.4.3. TERRITORIAL MARKING

The only known boundary latrines that yielded faecal samples were located on the intersection of Ermsdorf 1 and Ermsdorf 2 as well as Ermsdorf 1 and Knäipenhecken territories. It was difficult to draw robust conclusions from the relatively limited data set of 40 genetic profiles. One unknown, but reliably scored, profile was obtained from a faecal sample collected between Ermsdorf 1 and Knäipenhecken. The majority of the known members of Ermsdorf 1 and Ermsdorf 2 social groups were identified from the faeces without an obvious preponderance of a particular sex or age class. Only one known member of Knäipenhecken was identified from the faecal samples, but too few profiles were obtained from this social group to attach any significance to this result. Furthermore, it appears as though no badgers use specific latrines exclusively. Our observations are therefore identical to the conclusions by Stewart *et al.* (2002), who, after filming badgers at boundary latrines, concluded not only that all sex and age classes defecated at boundary latrines, but also that no badger monopolized access to specific latrines. After monitoring radio-collared badgers, Brown *et al.* (1992) and Roper *et al.* (1993) concluded that males would preferentially defecate at boundary latrines. Apart from the fact that only a few animals were radio-collared, the major shortcoming of the studies on latrine use based on telemetry alone was that they only provided information about the frequency of visits by individuals to the latrines, but not about the behaviour of the studied animals at the latrines.

5.4.4. SYNTHESIS

During my study, the badgers inhabited largely exclusive group ranges. The presence of scent marks on boundaries and wounds identified on captured badgers (see also Schley 2000) suggested that these ranges were defended as territories. Faecal DNA extracted from droppings collected from boundary latrines were genotyped to identify the individuals that were involved in territorial defence. The results suggested that most members of a social group, irrespective of their sex and age, defecated at boundary latrines and that no individual monopolises access to a specific latrine.

The AKH predicts that, during the period of data collection from Jan-March, males should mostly mark the territory to limit access to oestrous females. It therefore appeared that this crucial assumption of the AKH did not hold in our study population. In other words, my results suggest that defecation at territorial boundaries does not serve the function of advertising the presence of males in a social group. This conclusion is confirmed by the observations that badgers can only distinguish faeces from their own and neighbouring groups (Christian 1993) and that the chemical composition of anal gland secretions, sometimes deposited with faeces, contains group-, but not individual- or sex-specific information (Davies *et al.* 1988). Conversely, the faecal genotyping results appear to support the defence of food resources as the underlying reason for territoriality (*e.g.*, Kruuk 1978b-Kruuk & Parish 1982, Macdonald 1983, Kruuk 1989). The fact that very few latrines were observed in the study area makes it unlikely, however, that boundary latrines transmit information on group resource depletion as suggested by the PRE hypothesis (Stewart *et al.* 1997). I conclude that presently, no hypothesis has yet been put forward that satisfactorily explains the functional basis of territoriality, taking into account the variability of territorial and social behaviour that badgers exhibit within their wide geographic distribution.

5.5. Summary

In order to understand territoriality in badgers, I aimed to investigate boundary defence by genotyping faecal DNA collected from boundary latrines. The spatial organisation of a badger population in Luxembourg was analysed by means of radio-tracking. The telemetry data collected from eight individuals suggested the presence of five exclusive group ranges in the study area. The presence of scent marks on boundaries and bite-wounds on captured badgers suggested that these ranges were defended as territories. However, it was only possible to identify 11 latrines in the study area. Telemetry showed that seven of these latrines were located on the borders of the territories associated with three social groups. Faecal samples were collected from these seven latrines for a three-week period at the beginning of 2003. Forty samples gave rise to amplifiable DNA. The faecal profiles were

compared with reliable reference profiles generated previously from the animals in the study area (see Chapter 4). The results suggested that most members of a social group, irrespective of their sex and age, defecated at boundary latrines and that no individual monopolises access to a specific latrine. These results support the defence of food resources as the underlying function of territoriality.

Chapter 6

Overview and Conclusions

6.1. Main Results Presented in this Study

The main objective of this study was to apply non-invasive genetic typing to questions relating to badger ecology and management. Specifically, I aimed to develop an accurate non-invasive method of determining social group sizes in the Eurasian badger and to analyse territorial marking in a medium-density population on the European mainland.

6.1.1. ESTIMATING BADGER ABUNDANCE BY NON-INVASIVE GENOTYPING

The UK Department for Environment, Food and Rural Affairs (DEFRA) has recently called for research into innovative ways of censusing badgers, which, due to their nocturnal and semi-fossorial life-style, are notoriously difficult to census (Macdonald, Mace & Rushton 1998, Tuytens *et al.* 2001, Wilson *et al.* 2003a). Nevertheless, for various reasons that include the badger's role as an agricultural pest (Moore *et al.* 1999, Schley 2000, Poole *et al.* 2002) and its potential role in the transmission of bovine tuberculosis to cattle in certain parts of Europe (for a review see Krebs *et al.* 1997), accurate estimates of its local abundance are often required. I tried to estimate badger abundance by genotyping both faecal and remotely plucked hair DNA, the only two potential sources of non-invasive DNA in the badger. Regarding DNA quality, there was no *a priori* preferred source of DNA since genotyping errors have been reported from shed and plucked hairs as well as from faeces (*e.g.* Gagneux *et al.* 1997, Goossens *et al.* 1998, Taberlet *et al.* 1996).

In Chapter 2, I summarised the steps I took to obtain reliable microsatellite genotypes from faecal DNA. It has been reported frequently that, because of the degraded nature of faecal DNA, repeated amplifications of the microsatellite loci were required to obtain reliable genotypes (*e.g.* Taberlet *et al.* 1999, Taberlet & Luikart 1999, Bayes *et al.* 2000). Furthermore, the quality and quantity of DNA in a faecal extract can depend both of the faecal preservation and extraction method (Wasser *et al.* 1997, Frantzen *et al.* 1998, Murphy *et al.* 2000). Therefore, in order for faecal DNA genotyping to form the basis of an efficient and widely applicable methodology of estimating badger abundance, it was necessary to develop a genotyping protocol that limited the number of repeat amplifications necessary to obtain reliable microsatellite genotypes and also to optimise storage and extraction methods for badger faeces.

Badger faeces are an abundant and easily obtainable source material, since individuals use communal defecation sites called 'latrines'. Faecal samples were collected from latrines near the setts of three social groups from the high-density population at Woodchester Park, UK. Genetic profiles were generated by amplification of seven

microsatellite loci with alleles shorter than 250 base pairs. The likelihood of two distinct individuals in a population carrying the same genetic profile by chance, which is contingent on allelic diversity, number of loci analysed, and the percentage and degree of related individuals in a population (Waits *et al.* 2001), was shown to be less than 1% if the profiles consisted of these seven loci. After testing several methods, I obtained a high amplification success rate by storing faeces in 70% ethanol and using the guanidine thiocyanate / silica method for extraction. Using 70% ethanol as a storage agent had the additional advantage of being an antiseptic.

The multiple-tubes approach proposed by Navidi *et al.* (1992) is the standard protocol used to generate reliable genotypes from faecal DNA (Taberlet *et al.* 1996, Goossens *et al.* 2000, Constable *et al.* 2001, Garnier *et al.* 2001). Reliable genotypes are obtained by recording an allele only if it has been observed at least twice (in at least three amplification reactions) and by only recording an individual locus as homozygous if a certain number of positive amplification reactions gave rise to the same allele (for a single locus $n \geq 8$ for 99% confidence; Miller *et al.* 2000). This approach is reliable but very conservative, requiring large numbers of amplifications to obtain correct genotypes. In order to obtain reliable genotypes with fewer amplification reactions than required by this standard protocol, I devised a comparative approach in which genetic profiles were compared and replication directed at similar, but not identical, genotypes. The protocol, illustrated in Fig. 2.1, was used to genotype forty-seven faecal DNA extracts and the reliability of the consensus genotypes was tested by comparison to reference profiles obtained from animals caught near the setts from which the faecal samples had been collected. Ignoring failed reactions, an average of 3.4 reactions had to be performed per locus per genotype to obtain reliable genetic profiles. My approach was much more efficient than the standard multi-tubes approach which would have required 5.2 amplification reactions per locus per genotype to give rise to the same genetic profiles.

A recent study has suggested that genotype reliability should be assessed using a maximum likelihood method that allows replication to be directed at those genotypes most likely to be erroneous (Miller *et al.* 2002). This method estimates the probability that a genotype is correct, *i.e.*, its reliability, and suggests a replication protocol if the estimate is below a certain threshold. I assessed the efficacy of the maximum likelihood approach by applying it in retrospect to my dataset and by counting the required number of amplifications. It was shown that, in addition to the reliability criteria, all alleles needed to be observed at least twice and that an average of 4.0 amplification reactions was necessary per locus per genotype to obtain reliable profiles. I conclude that my approach would be better suited to small studies with limited financial resources than the maximum likelihood approach.

An analysis of the pattern of genotyping errors showed that 35% of all the amplifications of heterozygous loci, corresponding to 19% of all the amplification reactions (including both heterozygous and homozygous loci), gave rise to an erroneous result. Allelic dropout, or the amplification of only one allele of a heterozygous individual, was the most important source of error. Statistical analysis showed a significant variation in error rates among loci, with one locus in particular being prone to allelic dropout. Comparison with other studies showed that the error rates obtained here were quite high. It was concluded that the high error rate would increase the number of amplifications required to obtain consensus genotypes but that it did not invalidate the use of faeces as a source of badger DNA.

In Chapter 3, I tested whether microsatellite genotyping of faecal DNA could constitute a robust methodology for estimating badger abundance. I used the faecal DNA profiles generated in Chapter 2 to estimate the size of the three social groups from which the faeces were sampled. To recapitulate, the faecal samples were collected during a 10-day period from latrines near three main setts in the high-density population at Woodchester Park, UK. As a result of an ongoing long-term and intensive study, independent estimates of abundance and information on the life histories of the resident badgers were available for this study area. The genotypes derived from the faecal samples were compared with those obtained from blood or hair samples from badgers live-trapped at the same setts, thereby revealing the identity and life history of a defecator. Furthermore, all adult badgers caught at the three target groups prior to the collection of faecal samples were given individual-specific fur-clips that allowed identification of individuals in the field using infrared video equipment.

The 47 useful faecal DNA extracts gave rise to 20 different genetic profiles, of which 16 could be matched to known group members. The faecal genotypes that originated from badgers with known trap histories did not reveal any apparent sex- or age-specific defecation pattern at the investigated latrines. By combining faecal profiling with direct observations and capture data, the minimum number of badgers known to be present was counted to be 28, while the maximum number of badgers that could be present was found to be 34. Similarly, the minimum and maximum number of badgers present in each of the three social groups was also determined.

Too small a percentage of the population was sampled ($\approx 59\text{--}71\%$) to obtain robust abundance estimates by simply counting the different genetic profiles generated from faecal DNA. Because samples had been collected on repeated, temporally distinct occasions, mark-recapture models were applied to the dataset to assess population size. This statistical technique is based on the capture, marking and release of an initial population. After re-sampling the population for ≥ 1 additional sessions, the

ratio of unmarked animals to recaptures is used to estimate population size (White *et al.* 1982). Mark-recapture models allow for variation in capture probabilities due to time, trap response and individual heterogeneity (Otis *et al.* 1978). Given the absence of an apparent bias in the defecation pattern, no intrinsic difference in “capture” probabilities caused by differences in marking behaviour could be assumed. However, individuals’ capture probabilities may vary depending on where and when faeces were deposited and the quantity of DNA available from the faeces. Given these considerations, the M_h -Jackknife population size estimation model was considered to be appropriate. Additionally, this estimator produces the least biased results when faced with a ‘shadow effect’, *i.e.*, failure to identify different individuals with identical profiles (Mills *et al.* 2000). The model estimated the population size in the three social groups to be 26 individuals. The corresponding 95% CI of 22-40 animals contained the independent estimate of the range of the population size.

Due to the relatively small number of different profiles collected at each of the three setts, mark-recapture analyses were considered unlikely to produce meaningful estimates of individual group sizes. Therefore, in addition to total abundance, individual group sizes were estimated using accumulation curves. In this method, population size corresponds to the projected asymptote of a function describing the number of samples analysed versus the cumulative number of unique genetic profiles (rarefaction curve). Three possible equations for the rarefaction curve have been suggested in the literature, but simulations in my study and elsewhere have shown that the equation suggested by Eggert *et al.* (2003) would produce consistently the least biased estimates of population size. Further details of the equation are given in section 3.2. Similarly to the appropriate mark-recapture model, the total point estimate generated with the suitable rarefaction method was smaller than the minimum number of badgers known to be present. Similarly to the total group size, it was not possible to accurately estimate the size of the specific social groups using the most appropriate rarefaction technique. Simulations showed that the accuracy of the estimates would be improved by increasing the proportion of individuals sampled. As this would be difficult in practice, I concluded that systematic collection and DNA profiling of faecal DNA samples could not constitute a robust methodology for estimating badger abundance.

In Chapter 4, I investigated whether microsatellite genotyping of DNA obtained from remotely plucked hair samples could be used to estimate badger abundance. Some reports had suggested that, unlike faeces, plucked hair samples could provide DNA of sufficient quality for repeated amplifications not to be necessary (Higuchi *et al.* 1988, Gagneux *et al.* 1997, Sloane *et al.* 2000). I aimed to perform a pilot study in Luxembourg, estimating badger abundance in a study area containing five

mains setts (Schley 2000). I first tested whether the seven chosen microsatellite loci exhibited enough variability in the Luxembourg badger population to produce individual-specific genetic profiles. This was confirmed by genotyping good-quality DNA obtained from road kills and captured badgers.

Macdonald *et al.* (2002) concluded that there was no evidence for a social hierarchy amongst feeding badgers, and badgers are known to leave hair on barbed-wire fences (Neal & Cheeseman 1996). I therefore considered it feasible to remotely collect fresh badger hair by using baited barbed-wire-enclosure hair traps, similar to the ones described in Woods *et al.* (1999). It was possible to collect hair samples near four setts using a baited barbed-wire enclosure. At the one remaining sett, badgers ignored the bait and hairs had to be sampled using barbed wire suspended over a clearly visible badger run. A preliminary test of a barbed-wire enclosure was carried out at one sett during Dec 8-15, 2002 and Jan 12-19, 2003. The main study was carried out from Feb 14 to March 7, 2003. Because of the small number of hairs collected at some setts and given uncertainties about the quality of DNA, further samples were collected systematically from March 17-24, 2003 and, after this period, opportunistically until May 2003. In other words, hairs samples were collected for a period of up to six months.

All the hairs found on a barb were included in the extraction. Of the 113 samples collected over a 6-month period, 105 originated from a single animal and gave rise to amplifiable DNA. Through comparison with reliable reference genotypes of captured badgers, by comparing samples amongst themselves and by triple amplification of unique profiles, I showed that all extracts that contained amplifiable DNA, including those obtained from single guard hairs, produced 100% accurate profiles in a single round of amplification.

Direct observation of the five target setts suggested that a minimum of 13 badgers were present in the study area. The 105 usable samples collected during the total study period originated from 15 different animals though, confirming the assumption that direct observation leads to underestimation of population sizes (Macdonald *et al.* 1998, Tuytens *et al.* 2001). The utility of hair trapping in estimating population size was tested by collecting hair samples daily from Feb 14 to March 7, 2003. During this three-week period, 66 usable samples were obtained, originated from 14 of the 15 known badgers. In a short period of time, a high proportion of a medium-density population was sampled, so much so that a census of the different genetic profiles gave rise to an estimate that compared well with the results from direct observations and from the complete dataset of genetic profiles generated during a 6-month period. Rarefaction analyses did not improve the accuracy of the estimates, but even introduced an upward bias in the estimates when applied to a few genetic

profiles generated from a small number of samples.

As explained in Chapter 4, it was possible that two individuals had not been sampled so that 17 badgers were present in the study area at the time of sample collection. I concluded that, in order to avoid negatively biased abundance estimates, the methodology needed to be improved by optimising the hair collection techniques. What was considered essential for the general utility of the methodology was that (a) a high proportion of a badger population was sampled in a relatively short period of time (as at large sample sizes, a high proportion of the population needs to be sampled to obtain unbiased estimates of population size, see Chapter 3) and (b) that no expensive repeated amplifications of DNA were required to obtain reliable genetic profiles. Genotyping of remotely plucked hairs could potentially form the basis of a novel and cost-effective approach to estimating badger abundance.

6.1.2. INVESTIGATION OF TERRITORIAL MARKING BY GENOTYPING FAECAL DNA

Extensive work over the last 25 years has shown that in high-density populations in the United Kingdom, Ireland and Sweden, badgers live in stable social groups of up to 20 individuals that, despite being solitary foragers, inhabit and defend clearly defined territories (for reviews see Kruuk 1989, Neal & Cheeseman 1996, Krebs *et al.* 1997). It has been difficult to understand the advantages of group living in the badger (Woodroffe & Macdonald 2000) and researchers have tried to explain the formation of non-cooperative groups in terms of group territoriality (e.g. Macdonald 1983, von Schantz 1984, Kruuk & Macdonald 1985). If territorial behaviour was a necessary precursor of sociality, then the question arises as to why badgers form territories in the first place. Defence of food resources or of females are usually put forward in order to explain the functional basis of group territoriality in badgers (Kruuk 1978b, Kruuk & Parish 1982, Macdonald 1983, Roper *et al.* 1986, Kruuk 1989, Stewart *et al.* 1997).

In order to gain a better understanding of badger territoriality, it is important to analyse which animals are involved in territorial defence. Latrines at territorial boundaries are shared by members of the same and neighbouring groups and are therefore believed to play a role in maintaining the social spacing system (Roper *et al.* 1986). Because two valid studies on the defecation pattern of badgers at boundary latrines have come to different conclusions (Roper *et al.* 1993, Stewart *et al.* 2002), the main objective of Chapter 5 was to genotype faecal DNA samples collected from boundary latrines to identify the animals involved in boundary marking.

Because a disproportionate number of studies have been conducted in the UK, where the environmental conditions are not typical for other badger populations, and because any hypothesis aiming to explain the social spacing of badgers has to

take inter-population variability into account, this part of the study was performed in study area in Luxembourg, in continental Europe. Because the work on remotely plucked hair was performed in the same area, genetic profiles generated from faecal DNA could be compared to reference profiles, giving information about the identities and, to a certain extent, the life-histories of the badgers that defecated at boundary latrines. Additionally, the sex of the resident animals that had not been captured or were too young to be anaesthetised was determined by molecular methods.

In order to achieve the main objective of analysing defecation patterns at boundary latrines, it was necessary to first investigate the spatial organisation of my study population in Luxembourg. A total of eight adult or sub-adult badgers were caught and radio-collared in 2002 and 2003 near the five setts in the study area. All the animals were radio-tracked during spring and/or summer. For three setts, data was available for one animal only. An average of 157 fixes was collected per animal ($SD = 112$), with the smallest dataset consisting of 40 and the largest one of 402 fixes. Incremental area plots showed that the range sizes of all the animals, calculated as a MCP of the recorded fixes, reached an asymptote during the respective tracking period. When estimated as MCP that included 100% of all the fixes, the sizes of individual home ranges varied from 42.5 ha to 171.8 ha, with an average of 76.5 ha ($SD = 49.9$). Analysis of the overlap between the different 100% MCP home ranges suggested a pattern of mutually exclusive group ranges. While the home ranges of animals caught at the same sett overlapped to a large extent, exactly the opposite was the case for animals caught at different setts.

The study area was searched for latrines to test whether boundaries of group ranges would be demarcated with faeces. Six boundary latrines, situated between three group ranges, were identified. The fact that the badgers in my study area appeared to inhabit group-specific home ranges and that scent marks were found on some of the boundaries of these group ranges suggested that, in principle, the spatial system of the Luxembourg badgers was based on territoriality. By collecting faecal samples daily from boundary latrines for a total of three weeks at the beginning of 2003, a total of 70 samples were collected for analysis of defecation patterns. The success of the faecal DNA extraction was limited as only 57% of all the available samples gave rise to amplifiable DNA. This value was rather low compared to the high success rate reported in Chapter 2. The reliable reference profiles generated from hair DNA allowed a substantial reduction in the number of necessary PCR reactions to produce correct faecal profiles. This was achieved by directing repeated amplifications at faecal genotypes that were similar, but not identical, to the ones obtained from hair DNA.

Given the small number of 40 genetic profiles, it was impossible to draw robust

conclusions about territorial marking. Nevertheless, some trends seemed to emerge. The results suggested that most members of a social group, irrespective of their sex and age, defecate at boundary latrines and that no individual monopolises access to a specific latrine. If territories were set up by males to defend females, one would expect that males would be more likely to defecate at boundary latrines than females. The apparent absence of a bias in defecation pattern suggests that territories play a role in defence of food resources, insofar as food is a critical resource for all members of a social group, irrespective of age or sex.

6.2. Discussion of Main Results and Conclusions

My work clearly showed that non-invasive genetic typing could play a significant role in the study of badger ecology. It proved possible to generate accurate microsatellite profiles of individual badgers from faecal samples collected in the field as well as from hairs remotely plucked by a hair capture device. This, in turn, made it possible to address the two main aims of this study, which were to develop an accurate non-invasive method of determining social group sizes in the Eurasian badger and to analyse territorial marking by genotyping faecal DNA.

Badger abundance was estimated by genotyping both faecal and remotely plucked hair DNA, the only two potential sources of non-invasive DNA in this species. The main motivation behind this work was that prior to this study, there had been no accurate and cost-effective method of estimating badger abundance. The most reliable direct method, live trapping and mark-recapture analysis, is, for logistical reasons, impractical at anything other than a local scale. Additionally, trapping can be abusive to an animal's welfare and could result in abnormal behaviour on part of the animals. Both faecal and hair DNA have been reported to be susceptible to genotyping errors and the usual remedy against this problem has been to repeat amplification a sufficient number of times to obtain reliable profiles. In order for genotyping of non-invasively collected DNA samples to be useful as a general and widely applicable methodology to estimate badger abundance, DNA extracts needed to be of such quality that they could be genotyped in a relatively straightforward way, minimising the required number of repeat amplifications. Even though a certain expertise and investment would still be needed for the laboratory work, a non-invasive approach giving rise to unbiased abundance estimates would then be cheaper and require less manpower than live trapping. If, on the other hand, it would not be possible to obtain reliable genotypes with a limited number of amplification reactions, the cost-effectiveness of any such non-invasive DNA study would be reduced to the level where, similar to live tracking, the methodology would only be feasible on a local scale.

In Chapter 2, I reported the steps I took to obtain reliable microsatellite genotypes from faecal DNA and devised an approach that minimised the required number of repeats. The protocol, which I named the ‘comparative multiple tubes’ approach, proved to be more efficient than the standard multiple tubes approach in generating reliable genetic profiles. The maximum likelihood approach introduced by Miller *et al.* (2002), which I tested for the first time empirically in this study, proved to also significantly reduce the required number of repeats as compared to the standard technique. The comparative approach does not require any advance knowledge as to the error rate of a faecal study, which makes it better suited for studies with limited financial resources that probably include multiple faeces from each individual. I believe, therefore, that the comparative approach makes an important contribution to the field of molecular scatology.

The development of an effective methodology for scoring faecal DNA samples, however, does not in itself make faeces a suitable source of DNA for population censuses. For this to be the case, a majority of faecal DNA extracts needs to systematically yield DNA of sufficient quality and quantity to allow straightforward generation of microsatellite profiles. In the faecal DNA study reported in Chapter 2, 89% of the samples allowed the generation of a full profile consisting of seven loci or of a partial profile in which at least the most informative locus could be scored. In Chapter 5, however, only 57% of the faecal samples collected gave rise to at least partially amplifiable DNA. Therefore, even though in the first study a relatively high proportion of samples yielded usable DNA, faeces generally appear to be an unreliable source of DNA. Additionally, 19% of all the amplification reactions in Chapter 2 were shown to be erroneous and, even using the efficient comparative multiple tubes approach, an average of 3.4 amplification reactions were necessary per locus per genotype to obtain reliable genetic profiles (fewer amplifications were required in Chapter 5, but this was because reference profiles were available from remotely plucked hair samples). It certainly could not be stated that genotyping of faecal DNA extracts was a straightforward matter.

The results obtained with DNA extracted from fresh remotely plucked hair samples were in stark contrast to the faecal DNA results. A total of 113 hair samples was remotely collected during a six-month period. Of these, 93% originated from single contributors and gave rise to DNA of sufficient quality to amplify all the seven microsatellite loci. Even though about a third of the 105 extracts originated from single hairs, all samples were genotyped 100% accurately at all the loci in just a single round of amplification. It is possible that in future applications of this technique, some erroneous genotypes might be generated. However, freshly plucked hairs do not seem to vary as much in the quantity of DNA they yield as do faecal

samples. Indeed, Higuchi *et al.* (1988) have obtained around 500 ng of DNA from a single freshly plucked human hair. Badger hair follicles do not differ markedly in size from human hair. I therefore suggest that DNA extracts obtained from freshly plucked badger hair will systematically contain enough DNA to allow error-free genotyping, especially if they were obtained from more than hair.

To conclude, I have developed a relatively efficient approach for obtaining reliable microsatellite genotypes from faecal DNA. The high error rate associated with genotyping of faecal DNA, however, would substantially reduce the cost-effectiveness of a census technique based on this methodology. Fresh remotely plucked hairs, on the other hand, yielded DNA of sufficient quality and quantity to allow simple and error-free microsatellite genotyping.

Other than a providing good-quality DNA extracts, a high proportion of the population needs to be sampled by non-invasive genotyping in order for a technique to be useful in the estimation of badger densities. This requirement was tested in a pilot study for both faecal and hair DNA. Faecal samples were collected during a ten-day period from latrines near the setts of three social groups in the high-density population at Woodchester Park, UK. Genetic profiles of too small a percentage of the known population were sampled ($\approx 59\text{--}71\%$) to obtain robust abundance estimates either by simply counting the different profiles or by applying statistical techniques to the dataset. Simulations showed that, rather than by collecting more samples, the estimate could only be improved by sampling a higher proportion of the population. Hair samples were remotely plucked during a three-week period near five setts in a medium-density population in Luxembourg. During this short period, a high proportion of the population was sampled, so much so that a census of the different genetic profiles gave rise to an estimate that compared well with the results from direct observations and from the complete dataset of genetic profiles generated during a six-months period.

The hair technique allowed the sampling of a higher proportion of the population than the collection of faeces. However, the fact that the two studies were performed at populations of different densities makes a direct comparison of the effectiveness of the two approaches difficult. A valuable methodology ideally should be applicable across a range of different badger densities. It is hard to envisage ways of sampling adequately high proportions of high-density populations by collection of faeces. In addition, systematic collection of faecal samples near badger setts might not be feasible in low-density populations, due to differences in the defecation patterns of the animals (Schley 2000, Hutchings *et al.* 2001).

Even though it was possible to remotely pluck hairs from a substantial proportion of

a medium-density population, this achievement might be difficult to emulate in other populations. Because, according to the Woodchester Park staff, a high proportion of their high-density population could be trapped using baited cage traps and because in the UK badgers are readily attracted to peanut bait (Kruuk 1989, Delahay *et al.* 2000a), it appears reasonable to assume that the same success could be achieved with baited barbed wire hair traps in other high-density populations. In populations of lower density, badgers appear to be difficult to bait and trap in cages (Rodríguez *et al.* 1996, Graf *et al.* 1996, Hofmann *et al.* 2000). In those populations, suspending barbed wire over well-used badger runs might be an effective alternative to baited hair traps. The fact that in my study this technique yielded three different profiles from four usable hair samples collected at a single run suggests that it deserves further testing. Given these considerations, further research will be required into identifying a generally applicable and effective method of capturing hair from a large proportion of a local population. Given the reliability of plucked hair DNA, an increase in the number of samples collected and genotyped would not lead to a disproportionate increase in the cost of the study.

My results showed that genotyping of remotely plucked hair did not suffer from the drawbacks associated with the application of the technique to faecal DNA. Furthermore, baited barbed wire enclosures or suspension of barbed wire over clearly visible badger runs allowed easy collection of hair samples from most members of the target groups. As I demonstrated, population size estimates generated from remotely collected hair were similar to a conservative baseline estimate, suggesting that this method could form the basis of a feasible, practicable and cost-effective technique of estimating badger abundance. Further research into optimising hair capture techniques should make the technique applicable over a range of population densities and independent of habitat characteristics.

Given the fact that, contrarily to faecal DNA, plucked hair DNA extracts were easy to genotype and that a high proportion of the local population was sampled, I suggest that wildlife researchers working with badgers consider remotely plucked hair rather than faeces as a source of non-invasive DNA. However, it should be emphasised that the reliability of faecal DNA is species-dependent (Piggott & Taylor 2003a) and that it can be more practical and less disruptive to the target species to genotype faeces rather than plucked hair DNA (*e.g.* Gerloff *et al.* 1999, Vigilant *et al.* 2001, Garnier *et al.* 2001, Eggert *et al.* 2003). When considering hair-trapping for other species, however, it is worth pointing out that work on population inventories in North America has shown that it may constitute a feasible approach even if the population to be investigated is small and distributed over a large area (Foran *et al.* 1997, Woods *et al.* 1999, Poole *et al.* 2001, Mowat & Paetkau 2002, Paetkau 2003).

Since it was possible to reliably genotype DNA obtained from badger faeces, I wanted to test whether this technique could potentially provide other information than population censuses. Given the general assumption that latrines situated on territorial boundaries play an important role in the maintenance of the spatial system of badgers (Roper *et al.* 1993, Stewart *et al.* 2002), faecal DNA profiling could be used in questions relating to scent-marking activity of specific individuals or sexes. Indeed, understanding territorial marking in badgers could potentially give important clues about the underlying basis of territoriality in this solitary forager. Identifying the animals that defecate at territorial borders could, I believed, help to clarify whether defence of food resources or of females was a more likely explanation of territoriality (Kruuk 1978b, Kruuk & Parish 1982, Macdonald 1983, Roper *et al.* 1986, Kruuk 1989, Stewart *et al.* 1997).

Telemetry showed that the badgers in my study area inhabited group-specific home ranges, and latrines were found on some of the boundaries of these group ranges. In principle, therefore, the spatial system of the Luxembourg badgers was based on territoriality. Seventy faecal samples were collected from six latrines located on the boundary of three group ranges. As already mentioned, only 57% of these samples could be satisfactorily profiled. However, a substantial reduction in the number of amplifications necessary to obtain reliable genotypes (1.2 reactions per locus per genotype) was achieved by comparing faecal profiles to the reliable reference profiles generated from hair DNA and by directing repeated amplifications at faecal genotypes that were similar, but not identical, to the ones obtained from hair DNA. The magnitude of this reduction strongly suggests that when genotyping faecal DNA in future studies on the behavioural ecology of badgers, it is worth establishing a database of reliable profiles by genotyping hair remotely plucked at the target setts. More generally, the same applies to behavioural studies on other species where it is possible to remotely collect hair DNA samples and where there is a good chance of obtaining a matching faecal sample.

Given the fact that only 40 genetic profiles were obtained from faecal samples collected from latrines situated at the boundary of three group ranges, it was difficult to draw robust conclusions from the limited dataset. The results suggested, nevertheless, that most members of a social group, irrespective of their sex and age, defecated at boundary latrines and that no individual monopolised access to a specific latrine. It was concluded that these observations would support the defence of food resources as the underlying reason for territoriality. Despite the paucity of the data, the results showed that faecal DNA profiling has the potential to provide detailed information about territorial defence, of a kind that would be difficult to obtain by other means.

I showed that badger DNA can be successfully extracted from non-invasively collected hair and faecal samples. Faecal DNA was degraded, but by following careful protocols it was possible to obtain reliable microsatellite profiles from this source material. Even though faecal DNA typing did not prove to be very practical for estimating badger abundance, the technique was successfully employed in the investigation of defecation patterns at boundary latrines. Plucked hair DNA, on the other hand, proved to be a source of badger DNA of sufficient quality to allow straightforward genotyping. The technique could successfully be employed in the estimation of badger abundance at a local scale.

Given the relative ease with which plucked hair DNA can be genotyped, the technique can be employed in other badger research. Given the difficulty associated with catching badgers in low- and medium density populations (Rodríguez *et al.* 1996, Brøseth *et al.* 1997, Schley 2000), repeated hair capture and DNA profiling could replace conventional studies based on capture-mark-recapture of animals; including investigations into the the demography and dynamics of a population (Cheeseman *et al.* 1988, Rogers *et al.* 1997b) and movements and dispersal behaviour of badger (Woodroffe *et al.* 1993, Christian 1994, Rogers *et al.* 1998). *A priori*, hair capture is less abusive to an animal's welfare than trapping and less likely to entail abnormal behaviour on part of the animals. The disadvantage of my approach would be that it doesn't provide any information on the physical condition of the animals. Also, rather than opportunistically collecting badger carcasses, collection of remotely plucked hairs from barbed wire cattle fences, for example, could be a simple source of DNA in investigations on background genetic variability of an unstudied population. Finally, hair trapping could prove to be a good source of DNA for genetic analyses on the social structure, paternity and relatedness of badger communities. These are just three examples of the applications of a new and potentially powerful tool in the investigation of badger ecology.

6.3. Summary of Conclusions

1. I have developed an efficient protocol, named the 'comparative multiple tubes' approach, for generating reliable microsatellite profiles from badger faecal DNA by directing replication at similar, but not identical genotypes. By following this protocol, reliable profiles were obtained with fewer amplification reactions than required by other published methods. My approach would be best suited for studies with limited financial resources that included multiple faeces from each individual.
2. Systematic collection of faecal samples during a ten-day period from latrines next to three setts in a high-density badger population did not

lead to the sampling of an adequate proportion of the known population to obtain robust abundance estimates, either by simply counting the different profiles or by applying statistical techniques to the dataset.

3. DNA extracted from fresh hair samples, remotely plucked by means of hair capture devices, gave rise to reliable genetic profiles in single rounds of amplification, even if the extract was obtained from single guard hair. Furthermore, in a short period of time, a high proportion of a medium-density population was sampled, so much so that a census of the different genetic profiles gave rise to an estimate that compared well with the results from direct observations and from the complete dataset of genetic profiles generated during a 6-month period. Given its reliability and accuracy, genotyping of remotely plucked hairs can form the basis of a cost-effective approach to estimating badger abundance.
4. Faecal DNA profiling has the potential to provide detailed information about certain aspects of badger behaviour and ecology, such as territorial marking, that would be difficult to obtain by other means.

References

- Aasen E, Medrano JF (1990) Amplification of the ZFY and ZFX genes for sex identification in humans, cattle, sheep and goats. *Biotechnology* **8**, 1279-1281.
- Ackerman BB, Leban FA, Samuel MD, Garton EO (1990) *User's Manual for Program HOME RANGE. Second Edition. Technical Report 15*. Forestry, Wildlife and Range Experiment Station, University of Idaho, Moscow, Idaho, USA.
- Agriculture Committee (1999) *Fifth Report: Badgers and Bovine Tuberculosis*. Her Majesty's Stationery Office, London.
- Amos W, Whitehead H, Ferrari MJ, *et al.* (1992) Restrictable DNA from sloughed cetacean skin; its potential for use in population analysis. *Marine Mammal Science* **8**, 275-283.
- Andersen D, Rongstad O (1989) Home-range estimates of red-tailed hawks based on random and systematic relocations. *Journal of Wildlife Management* **53**, 802-807.
- Anderson DJ (1982) The home range: a new nonparametric estimation technique. *Ecology* **63**, 103-112.
- Banks SC, Horsup AB, Wilton AN, Taylor AC (2003) Genetic marker investigation of the source and impact of predation on a highly endangered species. *Molecular Ecology* **12**, 1663-1667.
- 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, Bruford MW (2000) Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. *Conservation Genetics* **1**, 173-176.
- Belkhir K (1996) *Genetix v.4.0.1., Windows™ software for population genetics*. University of Montpellier II, Laboratoire Génome et Populations, Montpellier, France.
- Bijlsma R, van de Vliet M, Pertoldi C, van Apeldoorn RC, van de Zande L (2000) Microsatellite primers from the Eurasian badger, *Meles meles*. *Molecular Ecology* **9**, 2216-2217.
- Bock WF (1986) Die Lebensraumnutzung des Dachses (*Meles meles*), erste Ergebnisse aus dem Raum Berchtesgaden. *Das Bärenseminar* **11**, 1-7.
- Boersen MR, Clark JD, King TL (2003) Estimating black bear population density and genetic diversity at Tensas River, Louisiana using microsatellite DNA markers. *Wildlife Society Bulletin* **31**, 197-207.
- Boom R, Sol CJA, Salimans MMM, *et al.* (1990) Rapid and simple method for purification of nucleic-acids. *Journal of Clinical Microbiology* **28**, 495-503.
- Bradley BJ, Boesch C, Vigilant L (2000) Identification and redesign of human microsatellite markers for genotyping wild chimpanzee (*Pan troglodytes verus*) and gorilla (*Gorilla gorilla gorilla*) DNA from faeces. *Conservation Genetics* **1**, 289-292.
- Bradley BJ, Chambers KE, Vigilant L (2001) Accurate DNA-based sex identification of apes using non-invasive samples. *Conservation Genetics* **2**, 179-181.
- Bradley BJ, Vigilant L (2002) False alleles derived from microbial DNA pose a potential source of error in microsatellite genotyping of DNA from faeces. *Molecular Ecology Notes* **2**, 602-605.
- Bricker J, Bushar LM, Reinert HK, Gelbert L (1996) Purification of high quality DNA from shed skins. *Herpetological Review* **27**, 133-134.
- Brøseth H, Knutsen B, Bevanger K (1997) Spatial organization and habitat utilization of badgers *Meles meles*: Effects of food patch dispersion in the boreal forest of central Norway. *Zeitschrift für Säugetierkunde* **62**, 12-22.

- Brown JA, Cheeseman CL, Harris S (1992) Studies on the spread of bovine tuberculosis from badgers to cattle. *Journal of Zoology, London* **227**, 694-696.
- Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multilocus and single-locus DNA fingerprinting. In: *Molecular Genetic Analysis of Populations: A Practical Approach* (ed. Hölzel AR), pp. 287-336. IRL Press, Oxford.
- Brunner H, Coman B (1974) *The Identification of Mammalian Hair*. Inkata Press, Melbourne.
- Burt WH (1943) Territoriality and home range concepts as applied to mammals. *Journal of Mammology* **24**, 346-352.
- Carpenter PJ, Dawson DA, Greig C, *et al.* (2003) Isolation of 39 polymorphic microsatellite loci and the development of a fluorescently labelled marker set for the Eurasian badger (*Meles meles*) (Carnivora: Mustelidae). *Molecular Ecology Notes* **3**, 610-615.
- Chao A, Lee S-M, Jeng S-L (1992) Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal. *Biometrics* **48**, 201-216.
- Cheeseman CL, Mallinson PJ (1979) Radio tracking in the study of bovine tuberculosis in badgers. In: *A Handbook of Biotelemetry and Radiotacking* (eds. Amlaner CJ, Macdonald DW), pp. 649-656. Pergamon Press, Oxford.
- Cheeseman CL, Wilesmith JW, Stuart FA, Mallinson PJ (1988) Dynamics of tuberculosis in a naturally infected badger population. *Mammal Review* **18**, 61-72.
- Christian SF (1993) *Behavioural ecology of the Eurasian badger (Meles meles): space use, territoriality and social behaviour*. PhD thesis, University of Sussex.
- Christian SF (1994) Dispersal and other inter-group movements in badgers, *Meles meles*. *Zeitschrift für Säugetierkunde* **59**, 218.
- Clifford SL, Jeffrey K, Bruford MW, Wickings EJ (1999) Identification of polymorphic microsatellite loci in the gorilla (*Gorilla gorilla gorilla*) using human primers: application to noninvasively collected hair samples. *Molecular Ecology* **8**, 1556-1558.
- Clutton-Brock TH, Guinness FE, Albon SD (1982) *Red Deer: Behaviour and Ecology of Two Sexes*. Edinburgh University Press, Edinburgh, Scotland.
- Colwell RK, Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **345**, 101-118.
- Constable JL, Ashley MV, Goodall J, Pusey AE (2001) Noninvasive paternity assignment in Gombe chimpanzees. *Molecular Ecology* **10**, 1279-1300.
- Coote T, Bruford MW (1996) Human microsatellites applicable for analysis of genetic variation in apes and Old World monkeys. *The Journal of Heredity* **87**, 406-410.
- Cresswell WJ, Smith GC (1992) The effects of temporally autocorrelated data on methods of home range analysis. In: *Wildlife Telemetry: Remote Monitoring and Tracking of Animals* (eds. Priede IG, Swift SM), pp. 272-284. Ellis Horwood, New York.
- da Silva J, Macdonald DW, Evans PGH (1994) Net costs of group living in a solitary forager, the Eurasian badger (*Meles meles*). *Behavioral Ecology* **5**, 151-158.
- da Silva J, Woodroffe R, Macdonald DW (1993) Habitat, food availability and group territoriality in the European badger, *Meles meles*. *Oecologia* **95**, 558-564.
- 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.
- Davies JM, Lachno DR, Roper TJ (1988) The anal gland secretion of the european badger (*Meles meles*) and its role in social communication. *Journal of Zoology, London* **216**, 455-463.

- Davison A, Birks JDS, Brookes RC, Braithwaite TC, Messenger JE (2002) On the origin of faeces: morphological versus molecular methods for surveying rare carnivores from their scats. *Journal of Zoology, London* **257**, 141-143.
- De Solla SR, Bonduriansky R, Brooks RJ (1999) Eliminating autocorrelation reduces biological relevance of home range estimates. *Journal of Animal Ecology* **68**, 221-234.
- Delahay RJ, Brown JA, Mallinson PJ, *et al.* (2000a) The use of marked bait in studies of the territorial organization of the European Badger (*Meles meles*). *Mammal Review* **30**, 73-87.
- Delahay RJ, Langton S, Smith GC, Clifton-Hadley RS, Cheeseman CL (2000b) The spatio-temporal distribution of *Mycobacterium bovis* (bovine tuberculosis) infection in a high-density badger population. *Journal of Animal Ecology* **69**, 428-441.
- Domingo-Roura X, Macdonald DW, Roy MS, *et al.* (2003) Confirmation of low genetic diversity and multiple breeding females in a social group of Eurasian badgers from microsatellite and field data. *Molecular Ecology* **12**, 533-539.
- Don RH, Cox PT, Wainwright BJ, Baker K, Mattrick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* **19**, 4008.
- Doncaster CP, Woodroffe R (1993) Den site can determine shape and size of badger territories: implications for group-living. *Oikos* **66**, 88-93.
- Donnelly P (1995) Nonindependence of matches at different loci in DNA profiles: quantifying the effect of close relatives on the match probability. *Heredity* **75**, 26-34.
- Dunn JE, Gipson PS (1977) Analysis of radio telemetry data in studies of home range. *Biometrics* **33**, 85-101.
- Dytham C (1999) *Choosing and Using Statistics: A Biologist's Guide*. Blackwell Science, Oxford.
- 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.
- Engel SR, Linn RA, Taylor JF, Davis SK (1996) Conservation of microsatellite loci across species of artiodactyls: implications for population studies. *Journal of Mammalogy* **77**, 504-518.
- Ernest HB, Penedo MC, 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. *Molecular Ecology* **9**, 433-442.
- Evans PGH, Macdonald DW, Cheeseman CL (1989) Social structure of the Eurasian badger (*Meles meles*): Genetic evidence. *Journal of Zoology, London* **218**, 587-596.
- Evett IW, Weir BS (1998) *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists*. Sinauer, Sunderland.
- Ewer RF (1968) *Ethology of Mammals*. Legos Press, London, England.
- Faber R (1971) *Climatologie du Grand-Duché de Luxembourg*. Musée National d'Histoire Naturelle & Société des Naturalists Luxembourgeois, Luxembourg.
- Farrell LE, Roman J, Sunquist ME (2000) Dietary separation of sympatric carnivores identified by molecular analysis of scats. *Molecular Ecology* **9**, 1583-1590.
- Feore S, Montgomery WI (1999) Habitat effects on the spatial ecology of the European badger (*Meles meles*). *Journal of Zoology, London* **247**, 537-549.
- Feore S, Smal CM, Montgomery WI (1991) Survey of badger setts in Northern Ireland: progress report. In: *The Badger* (ed. Hayden TJ), pp. 23-25. Royal Irish Academy, Dublin.
- Fernando P, Lande R (2000) Molecular genetic and behavioral analysis of social organization in the Asian elephant (*Elephas maximus*). *Behavioral Ecology and Sociobiology* **48**, 84-91.
- Flagstad Ø, 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.

- Foran DR, Minta SC, Heinemeyer KS (1997) DNA-based analysis of hair to identify species and individuals for population research and monitoring. *Wildlife Society Bulletin* **25**, 840-847.
- Ford G, Krumme DW (1979) The analysis of space use pattern. *Journal of Theoretical Biology* **76**, 125-155.
- Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.
- Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* **6**, 861-868.
- Garnier JN, Bruford MW, Goossens B (2001) Mating system and reproductive skew in the black rhinoceros. *Molecular Ecology* **10**, 2031-2042.
- Gerloff U, Hartung B, Fruth B, Hohmann G, Tautz D (1999) Intracommunity relationships, dispersal pattern and paternity success in a wild living community of Bonobos (*Pan paniscus*) determined from DNA analysis of faecal samples. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 1189-1195.
- 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.
- Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Science International* **112**, 17-40.
- Goldstein DB, Schlötterer C (1999) *Microsatellites: Evolution and applications*. Oxford University Press, Oxford.
- Goossens B, Chikhi L, Utami SS, de Ruiter J, Bruford MW (2000) A multi-sample, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Conservation Genetics* **1**, 157-162.
- Goossens B, Waits LP, Taberlet P (1998) Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology* **7**, 1237-1241.
- Goszczynski P (1994) Habitat use by badgers. *Proceedings of the II. North European Symposium in the Ecology of Carnivores*, 17.
- Graf M, Wandeler AI, Lüp P (1996) Die räumliche Habitatnutzung einer Dachspopulation (*Meles meles*) im schweizerischen Mittelland. *Revue Suisse de Zoologie* **103**, 835-850.
- Greenwood JJD (1996) Basic techniques. In: *Ecological Census Techniques, a Handbook* (ed. Sutherland WJ), pp. 11-110. Cambridge University Press, Cambridge.
- Griffiths HI, Thomas DH (1993) The status of the Badger *Meles meles* (L. 1758) (Carnivora, Mustelidae) in Europe. *Mammal Review* **23**, 17-58.
- Griffiths HI, Thomas DH (1997) *The conservation and management of the European badger* (*Meles meles*). Council of Europe Publishing.
- Griffiths R, Tiwari B (1993) Primers for the differential amplification of the sex-determining region Y gene in a range of mammal species. *Molecular Ecology* **2**, 405-406.
- Guo SW, Thompson EA (1992) Performing the exact tests of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 361-372.
- Halfpenny J, Biesot E (1986) *A Field Guide to Mammal Tracking in North America*. Johnson Books, Chicago.
- Hansen MM, Jacobsen L (1999) Identification of mustelid species: otter (*Lutra lutra*), American mink (*Mustela vison*) and polecat (*Mustela putorius*), by analysis of DNA from faecal samples. *Journal of Zoology, London* **247**, 177-182.

- Harris S (1982) Activity patterns and habitat utilization of badgers (*Meles meles*) in suburban Bristol: A radio tracking study. *Symposia of the Zoological Society of London* **49**, 301-323.
- Harris S, Cresswell P, Forde PG, *et al.* (1990) Home-range analysis using radio-tracking data - a review of problems and techniques particularly as applied to the study of mammals. *Mammal Review* **20**, 97-123.
- Harris S, Cresswell WJ, Jefferies D (1989) *Surveying badgers*. The Mammal Society, London.
- Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. *Genomic Research* **6**, 986-994.
- Henry C, Lafontaine L, Mouchès A (1988) Le blaireau (*Meles meles* Linnaeus 1758). In: *Encyclopédie des Carnivores de France*, p. 35. Société Française pour l'Etude et la Protection des Mammifères, Paris.
- Higuchi R, Von Beroldingen CH, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. *Nature* **332**, 543-546.
- Hofer H (1988) Variation in resource presence, utilization and reproductive success within a population of European badgers (*Meles meles*). *Mammal Review* **18**, 25-36.
- Hofmann T, Ebersbach H, Stubbe M (2000) Home range-Größe und Habitatnutzung beim Europäischen Dachs (*Meles meles* L., 1758) im nordöstlichen Harzvorland (Sachsen-Anhalt). *Beiträge zur Jagd- und Wildforschung* **25**, 199-209.
- Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of Specific Polymerase Chain-Reaction Product by Utilizing the 5' - 3' Exonuclease Activity of Thermus-Aquaticus DNA-Polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 7276-7280.
- Hooge PN, Eichenlaub B (1997) *Animal movement extension to arcview. ver. 1.1*. Alaska Biological Science Center, U.S. Geological Survey, Anchorage.
- Höss M, Kohn M, Pääbo S, Knauer F, Schröder W (1992) Excrement analysis by PCR. *Nature* **359**, 199.
- Höss M, Pääbo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Research* **21**, 3913-3914.
- Hutchings MR, Service KM, Harris S (2001) Defecation and urination patterns of badgers *Meles meles* at low density in south west England. *Acta Theriologica* **46**, 87-96.
- Hutchings MR, Service KM, Harris S (2002) Is population density correlated with faecal and urine scent marking in European badgers (*Meles meles*) in the UK? *Zeitschrift für Säugetierkunde* **67**, 286-293.
- Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299-314.
- Jansman HAH, Chanin PRF, Dallas JF (2001) Monitoring otter populations by DNA typing of spraints. *IUCN Otter Specialist Bulletin* **18**, 11-18.
- Jensen-Seaman MI, Kidd KK (2001) Mitochondrial DNA variation and biogeography of eastern gorillas. *Molecular Ecology* **10**, 2241-2247.
- Johnson DDP, Jetz W, Macdonald DW (2002) Environmental correlates of badger social spacing across Europe. *Journal of Biogeography* **29**, 411-425.
- Kaufman JH (1962) Ecology and social behaviour of the coati, *Nasua narica* on Barro Colorado Island Panama. *University of California Publications in Zoology* **60**, 95-222.
- Kendall KC, Metzgar LH, Patterson DA, Steele BM (1992) Power of sign surveys to monitor population trends. *Ecological Applications* **2**, 422-430.
- Kenward RE (1992) Quantity versus quality: programmed collection and analysis of radio-tracking data. In: *Wildlife Telemetry: Remote Monitoring and Tracking of Animals* (eds. Priede IG, Swift SM), pp. 231-246. Ellis Horwood, New York.

- Kenward RE (2001) *A Manual for Wildlife Radio Tagging*. Academic Press, London.
- Kocher TD, Thomas WK, Meyer A, *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 6196-6200.
- 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. *Trends in Ecology and Evolution* **12**, 223-227.
- Kohn MH, York EC, Kamradt DA, *et al.* (1999) Estimating population size by genotyping faeces. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 657-663.
- Kowalczyk R, Jedrzejewska B, Zalewski A (2003a) Annual and circadian activity patterns of badgers (*Meles meles*) in Białowieża Primeval Forest (eastern Poland) compared with other Palaearctic populations. *Journal of Biogeography* **30**, 463-472.
- Kowalczyk R, Zalewski A, Jedrzejewska B, Jedrzejewski W (2003b) Spatial organization and demography of badgers (*Meles meles*) in Białowieża Primeval Forest, Poland, and the influence of earthworms on badger densities in Europe. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **81**, 74-87.
- Krebs JR, Anderson RM, Clutton-Brock T, *et al.* (1997) *Bovine Tuberculosis in Cattle and Badgers*. Ministry of Agriculture, Fisheries and Food, London.
- Kruuk H (1978a) Foraging and spatial organisation of the European badger, *Meles meles* L. *Behavioral Ecology and Sociobiology* **4**, 75-89.
- Kruuk H (1978b) Spatial organisation and territorial behaviour of the European badger *Meles meles*. *Journal of Zoology, London* **184**, 1-19.
- Kruuk H (1989) *The Social Badger - Ecology and Behaviour of a group-living Carnivore (Meles meles)*. Oxford University Press, New York.
- Kruuk H, Gorman M, Leitch A (1984) Scent-marking with the subcaudal gland by the European badger, *Meles meles* L. *Animal Behaviour* **32**, 899-907.
- Kruuk H, Macdonald DW (1985) Group territories of carnivores: empires and enclaves. In: *Behavioural Ecology: Ecological Consequences of Adaptive Behaviour* (eds. Sibley RM, Smith RH). Blackwell Scientific Publications, Oxford.
- Kruuk H, Parish T (1982) Factors affecting population-density, group-size and territory size of the european badger, *Meles meles*. *Journal of Zoology, London* **196**, 31-39.
- Lee S-M, Chao A (1994) Estimating population size via sample coverage for closed capture-recapture models. *Biometrics* **50**, 88-97.
- Lie YS, Petropoulos CJ (1998) Advances in quantitative PCR technology: 5 ' nuclease assays. *Current Opinion in Biotechnology* **9**, 43-48.
- Livak KJ, Flood SJA, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods and Applications* **4**, 357-362.
- Lobert B, Lumsden L, Brunner H, Triggs B (2001) An assessment of the accuracy and reliability of hair identification of south-east Australian mammals. *Wildlife Research* **28**, 637-641.
- Long C, Killingley CA (1983) *The Badgers of the World*. Charles C Thomas, Springfield, Illinois.
- Lucchini V, Fabbri E, Marucco F, *et al.* (2002) Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. *Molecular Ecology* **11**, 857-868.

- Lüpés P (1988) Der Dachs und die Landwirtschaft. *Natur und Mensch* **30**, 136-139.
- Macdonald DW (1983) The ecology of carnivore social behaviour. *Nature* **301**, 379-384.
- Macdonald DW, Mace G, Rushton SR (1998) *Proposals for a National Mammal Monitoring Network*. Department of Environment, Transport and the Regions, London.
- Macdonald DW, Stewart PD, Johnson PJ, Porkert J, Buesching CD (2002) No evidence of social hierarchy amongst feeding badgers, *Meles meles*. *Ethology* **108**, 613-628.
- Manning T, Edge WD, Wolff JO (1995) Evaluating population-size estimators: An empirical approach. *Journal of Mammalogy* **76**, 1149-1158.
- McNay RS, Morgan JA, Bunnell FL (1994) Characterizing independence of observations in movements of Columbian black-tailed deer. *Journal of Wildlife Management* **58**, 422-429.
- Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics* **160**, 357-366.
- 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.
- Miyamoto MM, Porter CA, Goodman M (2000) *c-Myc* gene sequences and the phylogeny of bats and other eutherian mammals. *Systematic Biology* **49**, 501-514.
- Mohr CO (1947) Table of equivalent populations on North American small mammals. *American Midland Naturalist* **37**, 223-249.
- Moore N, Whiterow A, Kelly P, *et al.* (1999) Survey of badger *Meles meles* damage to agriculture in England and Wales. *Journal of Applied Ecology* **36**, 974-988.
- Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology* **10**, 1835-1844.
- Morin PA, Moore JJ, Chakraborty R, *et al.* (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* **265**, 1193-1201.
- Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* **3**, 401-411.
- 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.
- Mullis KB, Faloona F (1987) Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Methods in Enzymology* **155**, 335-350.
- 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.
- Navidi W, Arnheim N, Waterman MS (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American Journal of Human Genetics* **50**, 347-359.
- Neal E, Cheeseman C (1996) *Badgers*. T.&A.D. Poyser Natural History, London.
- O'Corry-Crowe G, Eves J, Hayden TJ (1993) Sett distribution, territory size and population density of badgers (*Meles meles* L.) in east Offaly. In: *The Badger* (ed. Hayden TJ), pp. 35-56. Royal Irish Academy, Dublin.
- Oli MK (1993) A key for the identification of hair of mammals of a snow leopard (*Panthera unica*) habitat in Nepal. *Journal of Zoology, London* **231**, 71-91.

- Ostler JR, Roper TJ (1998) Changes in size, status and distribution of badger *Meles meles* L. setts during a 20-year period. *Zeitschrift für Säugetierkunde - International Journal of Mammalian Biology* **63**, 200-209.
- Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62**, 1-135.
- Otis DL, White GC (1999) Autocorrelation of location estimates and the analysis of radiotracking data. *Journal of Wildlife Management* **63**, 1039-1044.
- Paetkau D (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology* **12**, 1375-1387.
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**, 347-354.
- Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* **3**, 489-495.
- Paetkau D, Waits LP, Clarkson PL, *et al.* (1998) Variation in genetic diversity across the range of North American Brown Bears. *Conservation Biology* **12**, 418-429.
- Palomares F, Godoy JA, Piriz A, O'Brien JO, Johnson WN (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.
- Palsbøll PJ (1999) Genetic tagging: contemporary molecular ecology. *Biological Journal Of The Linnean Society* **68**, 3-22.
- Parsons KM, Durban JW, Claridge DE, *et al.* (2003) Kinship as a basis for alliance formation between male bottlenose dolphins, *Tursiops truncatus*, in the Bahamas. *Animal Behaviour* **66**, 185-194.
- Paxinos E, McIntosh C, Rails 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.
- Pearce JM, Fields RL, Scribner KT (1997) Nest materials as a source of genetic data for avian ecological studies. *Journal of Field Ornithology* **68**, 471-481.
- Pertoldi C, Loeschcke V, Madsen AB, Randi E (2000) Allozyme variation in the Eurasian badger *Meles meles* in Denmark. *Journal of Zoology, London* **252**, 544-547.
- Piggott MP, Taylor AC (2003a) Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced species. *Australian Journal of Zoology* **51**, 341-355.
- Piggott MP, Taylor AC (2003b) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research* **30**, 1-13.
- Pigozzi G (1990) Latrine use and the function of territoriality in the European badger, *Meles meles*, in a Mediterranean coastal habitat. *Animal Behaviour* **39**, 1000-1002.
- Pires AE, Fernandes ML (2003) Last lynxes in Portugal? Molecular approaches in a pre-extinction scenario. *Conservation Genetics* **4**, 525-532.
- Poole DW, McKillop IG, Western G, Hancocks PJ, Packer JJ (2002) Effectiveness of an electric fence to reduce badger (*Meles meles*) damage to field crops. *Crop Protection* **21**, 409-417.
- Poole KG, Mowat G, Fear DA (2001) DNA-based population estimate for grizzly bears *Ursus arctos* in northeastern British Columbia, Canada. *Wildlife Biology* **7**, 105-115.
- Powell RA (1979) Mustelid spacing patterns: variations on a theme by *Mustela*. *Zeitschrift für Tierpsychologie* **50**, 153-165.

- Powell RA (2000) Animal home range and territories and home-range estimators. In: *Research Techniques in Animal Ecology: Controversies and Consequences* (eds. Boitani L, Fuller TK), pp. 65-110. Columbia University Press, New York.
- Powell RA, Zimmermann JW, Seaman DE (1997) *Ecology and Behavior of North American Black Bears: Home Ranges, Habitat and Social Organization*. Chapman and Hall, New York.
- Prigioni C (1999) *Meles meles*. In: *The Atlas of European Mammals* (eds. Mitchell-Jones AJ, Amori G, Boganowicz W, et al.). T. & A.D. Poyser, London.
- Primmer CR, Møller AP, Ellegren H (1996) A wide-ranging survey of cross-species microsatellite amplifications in birds. *Molecular Ecology* **5**, 365-378.
- Putman RJ (1984) Facts from faeces. *Mammal Review* **14**, 79-97.
- Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology and Evolution* **8**, 285-288.
- Ramekers J, Hummel S, Herrmann B (1997) How many cycles does a PCR need? Determination of cycle numbers depending of the number of targets and the reaction efficiency factor. *Naturwissenschaften* **84**, 259-262.
- Raphael MG (1994) Techniques for monitoring populations of fishers and American martens. In: *Martens, sables, and fishers: biology and conservation* (eds. Buskirk SW, Harestad AS, Raphael MG, Powell RA), pp. 224-240. Cornell University Press, Ithaca, N.Y.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 239.
- 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.
- Revilla E, Palomares F (2002) Spatial organisation, group living and ecological correlates in low-density populations of Eurasian badgers, *Meles meles*. *Journal of Animal Ecology* **71**, 497-512.
- Reynolds TD, Laundre JW (1990) Time intervals for estimating pronghorn and coyote home ranges and daily movements. *Journal of Wildlife Management* **54**, 316-322.
- Rice WR (1989) Analysing tables of statistical tests. *Evolution* **43**, 223-225.
- Rodríguez A, Martín R, Delibes M (1996) Space use and activity in a mediterranean population of badgers *Meles meles*. *Acta Theriologica* **41**, 59-72.
- Rogers LM, Cheeseman CL, Langton S (1997a) Body weight as an indication of density dependent population regulation in badgers (*Meles meles*) at Woodchester Park, Gloucestershire. *Journal of Zoology, London* **242**, 597-604.
- Rogers LM, Cheeseman CL, Mallinson PJ, Clifton-Hadley R (1997b) The demography of a high-density badger (*Meles meles*) population in the west of England. *Journal of Zoology, London* **242**, 705-728.
- Rogers LM, Delahay R, Cheeseman CL, et al. (1998) Movement of badgers (*Meles meles*) in a high-density population: individual, population and disease effects. *Proceedings of the Royal Society of London Series B-Biological Sciences* **265**, 1269-1276.
- Rogers LM, Delahay RJ, Cheeseman CL, Smith GC, Clifton-Hadley RS (1999) The increase in badger (*Meles meles*) density at Woodchester Park, south-west England: implications for disease (*Mycobacterium bovis*) prevalence. *Mammalia* **63**, 183-192.
- Rolstad J, Wegge P (1988) Spacing and habitat use of capercaillie during summer. *Canadian Journal of Zoology* **66**, 670-679.
- Roon DA, Waits LP, Kendall KC (2003) A quantitative evaluation of two methods for preserving hair samples. *Molecular Ecology Notes* **3**, 163-166.

- Rooney SM, Wolfe A, Hayden TJ (1998) Autocorrelated data in telemetry studies: time to independence and the problem of behavioural effects. *Mammal Review* **28**, 89-98.
- Roper TJ (1992) Badger *Meles meles* sett architecture, internal environment and function. *Mammal Review* **22**, 43-53.
- Roper TJ (1993) Badger setts as a limiting resource. In: *The Badger* (ed. Hayden TJ), pp. 26-34. Royal Irish Academy, Dublin.
- Roper TJ (1994) The European badger *Meles meles*: food specialist or generalist? *Journal of Zoology, London* **234**, 437-452.
- Roper TJ, Conradt L, Butler J, Christian SE (1993) Territorial marking with faeces in badgers (*Meles meles*): a comparison of boundary and hinterland latrine use. *Behaviour* **127**, 289.
- Roper TJ, Findlay SR, Lüps P, Shepherdson DJ (1995) Damage by badgers *Meles meles* to wheat *Triticum vulgare* and barley *Hordeum sativum* crops. *Journal of Applied Ecology* **32**, 720-726.
- Roper TJ, Lüps P, Lycett S (1989) Badgers as pests in English vineyards. In: *Mammals as Pests* (ed. Putman RJ), pp. 207-208. Croom Helm, London.
- Roper TJ, Shepherdson DJ, Davies JM (1986) Scent marking with faeces and anal secretions in the European badger (*Meles meles*): Seasonal and spatial characteristics of latrine use in relation to territoriality. *Behaviour* **97**, 94-117.
- Rosenberg DK, Overton WS, Anthony RG (1995) Estimation of animal abundance when capture probabilities are low and heterogeneous. *Journal of Wildlife Management* **59**, 252-261.
- Saiki RK, Gelfand DH, Stoffel S, *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.
- Saiki RK, Scharf S, Faloona F, *et al.* (1985) Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: A laboratory manual*, 2nd edn. Cold Spring Harbour Laboratory Press, Cold Spring, New York.
- Samuel MD, Pierce DJ, Garton EO (1985) Identifying Areas of Concentrated Use within the Home Range. *Journal of Animal Ecology* **54**, 711-719.
- Schley L (2000) *The Badger Meles meles and the Wild Boar Sus scrofa: Distribution and Damage to Agricultural Crops in Luxembourg*. DPhil Thesis, University of Sussex.
- Schlötterer C, Amos W, Tautz D (1991) Conservation of polymorphic simple sequence data in cetacean species. *Nature* **354**, 63-65.
- Seaman DE, Griffiths B, Powell RA (1998) KERNELHR: a program for estimating animal home ranges. *Wildlife Society Bulletin* **26**, 95-100.
- Seaman DE, Millspaugh JJ, Kernohan BJ, *et al.* (1999) Effects of sample size on kernel home range estimates. *Journal of Wildlife Management* **63**, 739-747.
- Seaman DE, Powell RA (1996) An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology* **77**, 2075-2085.
- Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. *Molecular Ecology Notes* **2**, 367-369.
- Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analysis. *Canadian Journal of Zoology* **69**, 82-90.
- Seymour SB, ed (1991) *Disinfection, Sterilisation and Preservation*. Lea & Febiger, Malvern, PA.
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

- Slate J, Coltman DW, Goodman SJ, *et al.* (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and soay sheep (*Ovis aries*). *Animal Genetics* **29**, 307-315.
- Sloane MA, Sunnucks P, Alpers D, Beheregaray LB, Taylor AC (2000) Highly reliable genetic identification of individual northern hairy-nosed wombats from single remotely collected hairs: a feasible censusing method. *Molecular Ecology* **9**, 1233-1240.
- Smal CM (1995) *The badger and habitat survey of Ireland*. National Parks and Wildlife Service and Department of Agriculture, Food and Forestry (DAFF, Republic of Ireland), Dublin.
- Smith EFG, Arctander P, Fjeldsø J, Amir OG (1992) A new species of shrike (Laniidae: *Lanius*) from Somalia, verified by DNA sequence data from the only known individual. *Ibis*, 227-235.
- Smith S, Vigilant L, Morin PA (2002) The effects of sequence length and oligonucleotide mismatches on 5' exonuclease assay efficiency. *Nucleic Acids Research* **30**, art. no.-e111.
- Solow AR (1989) A Randomization Test for Independence of Animal Locations. *Ecology* **70**, 1546-1549.
- Stewart PD, Anderson C, Macdonald DW (1997) A mechanism for Passive Range Exclusion: Evidence from the European Badger (*Meles meles*). *Journal of Theoretical Biology* **184**, 279-289.
- Stewart PD, Macdonald DW (1997) Age, sex, and condition as predictors of moult, and the efficacy of a novel fur-clip technique for individual marking of the European badger (*Meles meles*). *Journal of Zoology, London* **241**, 543-550.
- Stewart PD, Macdonald DW, Newman C, Cheeseman CL (2001) Boundary faeces and matched advertisement in the European badger (*Meles meles*): a potential role in range exclusion. *Journal of Zoology, London* **255**, 191-198.
- Stewart PD, Macdonald DW, Newman C, Tattersall FH (2002) Behavioural mechanisms of information transmission and reception by badgers, *Meles meles*, at latrines. *Animal Behaviour* **63**, 999-1007.
- Sugiyama Y, Kawamoto S, Takenaka O, Kumizaka K, Norikatsu W (1993) Paternity discrimination and inter-group relationships of chimpanzees at Bossou. *Primates* **34**, 545-552.
- Swihart RK, Slade NA (1985) Testing for independence of observations in animal movements. *Ecology* **66**, 1176-1184.
- Swihart RK, Slade NA (1997) On testing for independence of animal movements. *Journal of Agriculture, Biological, and Environmental Statistics* **2**, 48-63.
- Symondson WOC (2002) Molecular identification of prey in predator diets. *Molecular Ecology* **11**, 627-641.
- Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proceedings of the Royal Society of London Series B-Biological Sciences* **255**, 195-200.
- Taberlet P, Camarra JJ, Griffin S, *et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology* **6**, 869-876.
- Taberlet P, Fumagalli L (1996) Owl pellets as a source for genetic studies of small mammals. *Molecular Ecology* **5**, 301-305.
- 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 JL (1998) Non-invasive genetic sampling. *Trends in Ecology and Evolution* **13**, 26-27.
- Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: Look before you leap. *Trends in Ecology and Evolution* **14**, 323-327.
- Takenaka O, Takashi H, Kawamoto S, Arakawa M, Takenaka A (1993) Polymorphic microsatellite DNA amplification customised for chimpanzee paternity testing. *Primates* **34**, 27-35.
- Tallmon D, Mills LS (1994) Use of logs within home ranges of California red-backed voles on a remnant of forest. *Journal of Mammalogy* **75**, 97-101.
- Tuytens FAM, Delahay RJ, Macdonald DW, *et al.* (2000) Spatial perturbation caused by a badger (*Meles meles*) culling operation: Implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*). *Journal of Animal Ecology* **69**, 815-828.
- Tuytens FAM, Long B, Fawcett T, *et al.* (2001) Estimating group size and population density of Eurasian badgers *Meles meles* by quantifying latrine use. *Journal of Applied Ecology* **38**, 1114-1121.
- Tuytens FAM, Macdonald DW, Delahay R, *et al.* (1999) Differences in trappability of European badgers *Meles meles* in three populations in England. *Journal of Applied Ecology* **36**, 1051-1062.
- Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377-379.
- Valière N, Fumagalli L, Gielly L, *et al.* (2003) Long-distance wolf recolonization of France and Switzerland inferred from non-invasive genetic sampling over a period of 10 years. *Animal Conservation* **6**, 83-92.
- Valière N, Taberlet P (2000) Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology* **9**, 2150-2152.
- Van Winkle W (1975) Comparison of several probabilistic home-range models. *Journal of Wildlife Management* **39**, 188-190.
- Verma SK, Prasad K, Nagesh N, Sultana M, Singh L (2003) Was elusive carnivore a panther? DNA typing of faeces reveals the mystery. *Forensic Science International* **137**, 16-20.
- Verma SK, Singh L (2003) Novel universal primers establish identity of an enormous number of animal species for forensic application. *Molecular Ecology Notes* **3**, 28-31.
- Vigilant L, Hofreiter M, Siedel H, Boesch C (2001) Paternity and relatedness in wild chimpanzee communities. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 12890-12895.
- von Schantz T (1984) Spacing strategies, kin selection and population regulation in altricial vertebrates. *Oikos* **42**, 48-58.
- 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.
- Walsh PA, 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.
- Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of fecal DNA methods to field studies of Ursids. *Molecular Ecology* **6**, 1091-1098.
- Weir BS, Triggs CM, Starling L, *et al.* (1997) Interpreting DNA mixtures. *Journal of Forensic Science* **42**, 213-222.
- White GC, Anderson DR, Burnham KP, Otis DL (1982) *Capture-recapture and removal methods for sampling closed populations*. Los Alamos National Laboratory Report LA-8787-NERP, Los Alamos, New Mexico.

- White GC, Garrott RA (1990) *Analysis of Wildlife Radio-Tracking Data*. Academic Press, Inc., San Diego.
- Wilson GJ, Delahay RJ (2001) A review of methods to estimate the abundance of terrestrial carnivores using field signs and observations. *Wildlife Research* **28**, 151-164.
- Wilson GJ, Delahay RJ, de Leeuw ANS, Spyvee PD, Handoll D (2003) Quantification of badger (*Meles meles*) sett activity as a method of predicting badger numbers. *Journal of Zoology, London* **259**, 49-56.
- Woodroffe R, Macdonald DW (1992) Badger clans: demographic groups in an antisocial species. *Journal of Zoology, London* **227**, 696-698.
- Woodroffe R, Macdonald DW (1993) Badger sociality - models of spatial grouping. *Symposia of the Zoological Society of London* **65**, 145-169.
- Woodroffe R, Macdonald DW (2000) Helpers provide no detectable benefits in the European badger (*Meles meles*). *Journal of Zoology, London* **250**, 113-119.
- Woodroffe R, Macdonald DW, da Silva J (1993) Dispersal and philopatry in the European badger, *Meles meles*. *Journal of Zoology, London* **237**, 227-239.
- Woods JG, Paetkau D, Lewis D, *et al.* (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin* **27**, 616-627.
- Worton BJ (1987) A review of models of home range for animal movement. *Ecological Modelling* **38**, 277-298.
- Worton BJ (1989) Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology* **70**, 164-168.
- Worton BJ (1995) Using Monte Carlo simulation to evaluate kernel-based home range estimators. *Journal of Wildlife Management* **59**, 794-800.
- Zubaid A, Gorman ML (1993) The effect of supplementary feeding upon the sizes of the home ranges of woodmice *Apodemus sylvaticus* living on a system of maritime sand-dunes. *Journal of Zoology, London* **231**, 233-237.

Appendices

Appendices

Appendix 1	136
(a) Results of the 1009 faecal DNA amplification reactions that gave rise to 293 consensus genotypes.	136
(b) RELIOTYPE output files	145
(c) Genetic profiles obtained if consensus genotypes had been constructed in RELIOTYPE using reliability criteria alone.	151
Appendix II	152
(a) Results of the 283 faecal DNA amplification reactions that gave rise to the 40 faecal profiles used to investigate territorial marking in chapter 5.	152
(b) Sex identification of badgers by means of PCR.	156

Sample	<i>Mel</i> -105	Consensus genotype	<i>Mel</i> -102	Cons. geno.	<i>Mel</i> -117	Cons. geno.	<i>Mel</i> -106	Cons. geno.	<i>Mel</i> -111	Cons. geno.	<i>Mel</i> -109	Cons. geno.	<i>Mel</i> -113	Cons. geno.
2	<u>148/148</u> <u>148/148</u> 148/148	148/148	195/197 <u>193/197</u> 195/197 195/197	195/197	<u>187/189</u> <u>187/189</u>	187/189	220/222 <u>220/222</u> <u>220/222</u> <u>220/222</u>	132/132	<u>132/132</u> <u>132/132</u> <u>132/132</u> <u>132/132</u>	106/106	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	120/120	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>	
3	<u>148/148</u> <u>148/148</u> <u>148/148</u>	148/148	195/197 195/197 197/197	195/197	<i>187/195*</i> <u>174/187</u> <u>187/189</u> 189/189 189/189 187/189	187/189	220/222 <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u>	132/132	<u>132/132</u> <u>132/132</u> <u>132/138</u> <u>132/132</u> <u>132/132</u> <u>132/132</u>	106/106	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	120/120	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>	
7	<u>148/148</u> <u>148/148</u> 148/148	148/148	195/197 <u>195/197</u> 197/197	195/197	<i>187/191*</i> <u>187/187</u> <u>187/189</u>	187/189	220/222 <u>222/222</u> <u>222/222</u>	132/132	<u>132/132</u> <u>132/132</u> <u>132/136*</u> ¹³⁸	106/106	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	120/120	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>	

8	<u>148/148</u>	<u>197/197</u>	<u>174/187/189</u>	<u>220/222</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>197/197</u> 195/197	<u>187/189</u>	<u>220/222</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>195/195</u>			<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>	<u>195/197</u>			<u>132/132</u>		
	<u>148/148</u>	<u>195/197</u>					
16	<u>148/148</u>	<u>195/197</u>	<u>187/187</u>	<u>220/222</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>195/195</u> 195/197	<u>187/187</u>	<u>220/222</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>197/197</u>	<u>187/189</u>		<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>		<u>187/189</u>		<u>132/132</u>		
	<u>148/148</u>		<u>187/189</u>				
40	<u>148/148</u>	<u>195/197</u>	<u>189/189</u>	<u>222/224</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>197/197</u> 195/197	<u>174/189</u>	<u>220/220</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>195/195</u>	<u>187/189</u>	<u>220/222</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>	<u>197/197</u>	<u>187/189</u>	<u>220/222</u>			
	<u>148/148</u>		<u>189/189</u>				
48	<u>148/148</u>	<u>195/195</u>	<u>187/187</u>	<u>220/226</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>195/197</u> 195/197	<u>187/187</u>	<u>220/224</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>195/197</u>	<u>187/189</u>	<u>220/222</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>	<u>197/197</u>	<u>187/189</u>	<u>222/222</u>			
	<u>148/148</u>			<u>220/222</u>			
49	<u>148/148</u>	<u>197/197</u>	<u>189/189</u>	<u>220/220</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>195/195</u> 195/197	<u>187/189</u>	<u>220/220</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>195/197</u>	<u>174/189</u>	<u>220/222</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>		<u>187/193</u>	<u>220/222</u>			
			<u>185*/189</u>	<u>222/222</u>			
57	<u>148/148</u>	<u>195/197</u>	<u>187/189</u>	<u>222/224</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u> 148/148	<u>195/197</u> 195/197	<u>187/189</u>	<u>220/222</u> <u>220/222</u>	<u>132/132</u> 132/132	<u>106/106</u> 106/106	<u>120/120</u> 120/120
	<u>148/148</u>	<u>195/197</u>	<u>187/189</u>	<u>220/220</u>	<u>132/132</u>	<u>106/106</u>	<u>120/120</u>
	<u>148/148</u>	<u>195/197</u>	<u>189/189</u>		<u>132/132</u>		

Sample	<i>Mel</i> -105	Consensus genotype	<i>Mel</i> -102	Cons. geno.	<i>Mel</i> -117	Cons. geno.	<i>Mel</i> -106	Cons. geno.	<i>Mel</i> -111	Cons. geno.	<i>Mel</i> -109	Cons. geno.	<i>Mel</i> -113	Cons. geno.
1	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u> 199/199	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/193</u>	<u>173/193</u> <u>173/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
20	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/193</u> <u>174/193</u>	<u>174/193</u> <u>174/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
25	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u> 199/199	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/193</u> <u>174/193</u>	<u>174/193</u> <u>174/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
28	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/174</u> <u>174/193</u>	<u>174/193</u> <u>174/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
29	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/193</u> <u>174/193</u>	<u>174/193</u> <u>174/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
43	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>199/199</u> <u>199/199</u>	<u>174/193</u> <u>174/193</u> <u>174/193</u>	<u>174/193</u> <u>174/193</u>	<u>220/222</u> <u>220/222</u> <u>220/222</u>	<u>132/138</u> <u>132/138</u>	<u>132/138</u> <u>132/138</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/116</u> <u>106/116</u> <u>106/116</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
17	<u>138/142</u> <u>138/142</u>	138/142	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>199/199</u> <u>199/199</u>	<u>174/187</u> <u>174/174</u> <u>174/187/</u> 193	<u>174/187</u> <u>174/187</u> <u>174/187/</u> 193	<u>224/226</u> <u>224/224</u> <u>224/224</u> <u>224/224</u>	<u>130/132</u> <u>132/132</u> <u>132/132</u> <u>132/132</u>	<u>130/132</u> <u>132/132</u> <u>132/132</u> <u>130/132</u> <u>130/138</u>	<u>106/125</u> <u>106/125</u> <u>106/125</u> <u>106/125</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>106/125</u> <u>106/125</u> <u>106/125</u> <u>106/125</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>

46	<u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>174/187</u> <u>174/187</u>	<u>220/224</u> 174/187 <u>220/224</u> 220/224	<u>130/132</u> <u>130/132</u>	<u>106/125</u> <u>106/125</u> <u>106/106</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
	138/142	199/199			130/132	106/125	120/120
47	<u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>174/187</u> <u>174/187</u> <u>174/187</u>	<u>220/222</u> <u>224</u> 174/187 <u>224/224</u> 220/224 <u>220/224</u>	<u>130/132</u> <u>130/132</u> <u>130/132</u>	<u>106/125</u> <u>106/125</u> <u>106/125</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
53	<u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>174/187</u> <u>187/193</u> <u>174/187/</u> <u>193</u> <u>187/187</u> <u>187/187</u> <u>174/187</u> <u>174/187</u> <u>174/187</u>	<u>220/224</u> 174/187 <u>220/220</u> 220/224 <u>224/224</u> <u>220/224</u>	<u>130/132</u> <u>130/132</u>	<u>106/125</u> <u>106/125</u> <u>106/125</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u>
9	<u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>187/187</u> <u>187/187</u> <u>187/187</u>	<u>220/220</u> 187/187 <u>220/220</u> 220/220 <u>220/220</u>	<u>130/132/</u> <u>134*</u> <u>130/132</u> <u>130/132/</u> <u>138</u> <u>130/132</u>	<u>106/106</u> <u>106/106</u> <u>106/106</u>	<u>120/120</u> <u>118/120</u> <u>118/120</u>
41	<u>134*/</u> <u>138/142</u> <u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u> <u>191*/199</u> <u>199/199</u> <u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>187/187</u> <u>187/187</u> <u>187/193</u> <u>187/187</u> <u>187/187</u> <u>187/187</u> <u>187/187</u> <u>187/187</u>	<u>220/220</u> 187/F <u>220/220</u> 220/220 <u>220/220</u>	<u>130/132</u> <u>130/132/</u> <u>138</u> <u>132/136</u> <u>132/132</u>	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	<u>118/120</u> <u>118/120</u> <u>118/120</u> <u>118/120</u>
44	<u>138/142</u> <u>138/142</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>187/187</u> <u>187/187</u> <u>187/187</u>	<u>220/220</u> 187/187 <u>220/220</u> 220/220 <u>220/220</u>	<u>130/130</u> <u>130/130</u> <u>130/132</u>	<u>106/106</u> <u>106/106</u> <u>106/106</u>	<u>118/120</u> <u>118/120</u> <u>120/120</u>
	138/142	199/199			130/132	106/106	118/120

[illegible]

54	<u>140/144</u> <u>140/144</u>	<u>195/197</u> <u>197/197</u> <u>195/195</u> <u>195/197</u>	<u>174/187</u> <u>174/187</u> <u>187/187</u> <u>174/187</u>	<u>220/224</u> <u>220/224</u> <u>220/224</u> <u>220/224</u>	<u>174/187</u> <u>174/187</u> <u>187/187</u> <u>174/187</u>	<u>132/132</u> <u>132/132</u> <u>132/132</u> <u>132/132</u>	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	<u>126/126</u> <u>126/126</u> <u>126/126</u> <u>126/126</u>
	<u>140/144</u> <u>140/144</u>	<u>195/197</u> <u>197/197</u> <u>195/195</u> <u>195/197</u>	<u>174/187</u> <u>174/187</u> <u>187/187</u> <u>174/187</u>	<u>220/224</u> <u>220/224</u> <u>220/224</u> <u>220/224</u>	<u>174/187</u> <u>174/187</u> <u>187/187</u> <u>174/187</u>	<u>132/132</u> <u>132/132</u> <u>132/132</u> <u>132/132</u>	<u>106/106</u> <u>106/106</u> <u>106/106</u> <u>106/106</u>	<u>126/126</u> <u>126/126</u> <u>126/126</u> <u>126/126</u>
58	<u>138/138</u> <u>138/138</u> <u>138/138</u> <u>138/138</u>	<u>199/199</u> <u>199/199</u> <u>199/199</u> <u>199/199</u>	<u>174/187</u> <u>174/187</u> <u>174/187</u> <u>174/187</u>	<u>222/226</u> <u>222/226</u> <u>222/226</u> <u>222/226</u>	<u>174/187</u> <u>174/187</u> <u>174/187</u> <u>174/187</u>	<u>132/138</u> <u>132/138</u> <u>132/138</u> <u>132/138</u>	<u>106/125</u> <u>106/125</u> <u>106/125</u> <u>106/125</u>	<u>120/120</u> <u>120/120</u> <u>120/120</u> <u>120/120</u>
6	<u>142/142</u> <u>142/142</u> <u>142/142</u> <u>142/142</u>	<u>199/199</u> <u>199/199</u> <u>195/199</u> <u>199/199</u> <u>199/199</u> <u>195/199</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>130/132</u> <u>130/132</u> <u>130/132</u> <u>130/132</u> <u>130/132</u> <u>130/132</u>	<u>127/127</u> <u>127/127</u> <u>106/127</u> <u>106/127</u> <u>106/127</u> <u>106/106</u>	<u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u>
32	<u>142/142</u> <u>142/142</u> <u>142/142</u> <u>142/142</u>	<u>195/199</u> <u>195/199</u> <u>195/199</u> <u>195/199</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>222/224</u> <u>222/224</u> <u>222/224</u> <u>222/224</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>130/130</u> <u>130/132</u> <u>130/132</u> <u>130/132</u>	<u>106/127</u> <u>106/127</u> <u>106/127</u> <u>106/127</u>	<u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u>
51	<u>138/142</u> <u>142/142</u> <u>142/142</u> <u>142/142</u> <u>142/142</u> <u>142/142</u>	<u>199/199</u> <u>199/199</u> <u>195/199</u> <u>195/199</u> <u>195/199</u> <u>195/199</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u> <u>222/222</u>	<u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u> <u>174/174</u>	<u>130/130</u> <u>132/132</u> <u>130/132</u> <u>130/132</u> <u>130/132</u> <u>130/132</u>	<u>106/127</u> <u>106/127</u> <u>106/127</u> <u>106/127</u> <u>106/127</u> <u>106/127</u>	<u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u> <u>120/126</u>
4	<u>138/142</u> <u>138/142</u>	<u>195/199</u> <u>195/199</u>	<u>174/174</u> <u>174/174</u>	<u>222/224</u> <u>222/224</u>	<u>174/174</u> <u>174/174</u>	<u>130/132</u> <u>130/132</u>	<u>106/106</u> <u>106/106</u>	<u>120/126</u> <u>120/126</u>

Sample	<i>Mel</i> -105	Consensus genotype	<i>Mel</i> -102	Cons. geno.	<i>Mel</i> -117	Cons. geno.	<i>Mel</i> -106	Cons. geno.	<i>Mel</i> -111	Cons. geno.	<i>Mel</i> -109	Cons. geno.	<i>Mel</i> -113	Cons. geno.
39	<u>138/142</u>	138/142	<u>195/199</u>	195/199	174/174	174/174	<u>222/224</u>	174/174	<u>130/132</u>	130/132	<u>106/106</u>	106/106	<u>120/126</u>	120/126
	<u>138/142</u>		<u>195/199</u>		174/174		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/126</u>	
	<u>138/142</u>		<u>195/195</u>		174/174		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/126</u>	
	<u>138/142</u>		<u>195/195</u>		174/174		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/126</u>	
5	<u>138/138</u>	138/142	<u>195/199</u>	195/199	174 187	174/187	<u>222/222</u>	174/187	<u>132/132</u>	130/132	<u>106/106</u>	106/106	<u>120/126</u>	120/126
	<u>138/142</u>		<u>195/199</u>		174 174		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/120</u>	
	<u>138/138</u>		<u>195/195</u>		174 193		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/126</u>	
	<u>138/142</u>		<u>195/199</u>		174 187		<u>222/224</u>		<u>130/132</u>		<u>106/106</u>		<u>120/126</u>	
33	<u>142/142</u>	138/142	<u>195/199</u>	195/199	174/174	174/187	<u>222/224</u>	174/187	<u>130/132</u>	130/132	<u>106/106</u>	106/106	<u>120/126</u>	120/126
	<u>138/142</u>		<u>195/199</u>		174/187		<u>222/224</u>		<u>132/132</u>		<u>106/106</u>		<u>120/126</u>	
	<u>138/142</u>		<u>195/199</u>		174/187		<u>222/224</u>		<u>138/138</u>		<u>106/106</u>		<u>120/126</u>	
	<u>138/142</u>		<u>195/199</u>		174 174		<u>222/224</u>		<u>130/138</u>		<u>106/106</u>		<u>120/126</u>	
23	<u>138/138</u>	138/138	<u>197/199</u>	197/199	174/189	174/189	<u>222/226</u>	174/189	<u>132/132</u>	132/F	<u>106/106</u>	106/106	<u>120/120</u>	120/120
	<u>138/138</u>		<u>197/199</u>		174/174		<u>226/226</u>		<u>130/130</u>		<u>106/106</u>		<u>120/120</u>	
	<u>138/138</u>		<u>197/199</u>		174/189		<u>222/222</u>		<u>132/138</u>		<u>106/106</u>		<u>120/120</u>	
	<u>138/138</u>		<u>197/199</u>		174/189		<u>222/222</u>		<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	
55	<u>138/138</u>	138/138	<u>199/199</u>	199/199	174/174	174/189	<u>220/222</u>	174/189	<u>132/132</u>	132/132	<u>106/106</u>	106/106	<u>120/120</u>	120/120
	<u>138/138</u>		<u>199/199</u>		174/189		<u>222/222</u>		<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	
	<u>138/138</u>		<u>199/199</u>		174/189		<u>226/226</u>		<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	
	<u>138/138</u>		<u>197/199</u>		174/189		<u>222/222</u>		<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	

	199/199	197/199		226/226					
	199/199								
	197/199								
56	<u>136/140</u>	136/140	<u>193/193</u>	<u>174/187</u>	<u>222/222</u>	<u>132/132</u>	<u>106/127</u>	<u>120/120</u>	
	<u>136/140</u>	136/140	<u>193/193</u>	<u>174/187</u>	<u>222/222</u>	<u>132/132</u>	<u>106/127</u>	<u>120/120</u>	
	<u>136/140</u>	136/140	<u>193/193</u>	<u>174/187</u>	<u>222/222</u>	<u>132/132</u>	<u>106/127</u>	<u>120/120</u>	
					222/F	132/132	106/127	120/120	
					<u>222/222</u>	<u>132/132</u>	<u>106/127</u>	<u>120/120</u>	
					<u>222/222</u>				
					<u>222/222</u>				
21	<u>138/138</u>	138/148	<u>199/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/138</u>	138/148	<u>199/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/146*</u>		<u>193/195</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>148/148</u>		<u>199/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>136/138</u>		<u>195/195</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>136/146*</u>		<u>195/195</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/148</u>			<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
34	<u>138/148</u>	138/148	<u>195/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/148</u>	138/148	<u>195/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/148</u>	138/148	<u>195/199</u>	<u>187/187</u>	<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
					187/187	130/132	106/106	120/126	
					<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
					<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
					<u>222/222</u>	<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
37	<u>138/148</u>	138/148				<u>130/138</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/148</u>	138/148				<u>130/130</u>	<u>106/106</u>	<u>120/126</u>	
	<u>138/148</u>	138/148				<u>130/132</u>	<u>106/106</u>	<u>120/126</u>	
						<u>130/132/</u>			
						<u>138</u>			
						<u>130/138</u>			
						<u>130/130</u>			

Sample	Mel-105	Consensus genotype	Mel-102	Cons. geno.	Mel-117	Cons. geno.	Mel-106	Cons. geno.	Mel-111	Cons. geno.	Mel-109	Cons. geno.	Mel-113	Cons. geno.
42	<u>148/148</u>		<u>199/199</u>		<u>174/174</u>		<u>222/222</u>		<u>130/130</u>		<u>106/106</u>		<u>126/126</u>	
	<u>138/144</u>	144/148	<u>199/199</u>	199/199	<u>187/187</u>	174/187	<u>222/222</u>	222/222	<u>130/130</u>	130/130	<u>106/106</u>	106/106	<u>126/126</u>	126/126
	<u>144/144</u>													
	<u>146*/148</u>		<u>199/199</u>		<u>174/187</u>		<u>222/222</u>		<u>130/130</u>		<u>106/106</u>		<u>126/126</u>	
	<u>148/148</u>										<u>106/106</u>		<u>126/126</u>	
	<u>144/148</u>													
36	<u>144/144</u>								<u>130/130</u>		<u>106/106</u>		<u>126/126</u>	
	<u>144/148</u>	144/148							<u>130/130</u>	130/130	<u>106/106</u>	106/106	<u>126/126</u>	126/126
	<u>148/148</u>								<u>130/130</u>		<u>106/106</u>		<u>126/126</u>	
13	<u>136/136</u>													
	<u>142/142</u>	136/142												
	<u>134*/136</u>													
	<u>136/136</u>													
	<u>136/136</u>													
26	<u>140/148</u>										<u>106/106</u>		<u>120/126</u>	
	<u>148/148</u>	140/148									<u>106/106</u>	106/106	<u>120/126</u>	120/126
	<u>140/148</u>										<u>106/106</u>		<u>120/126</u>	
35	<u>148/148</u>								<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	
	<u>148/148</u>	148/148							<u>132/132</u>	132/132	<u>106/106</u>	106/106	<u>120/120</u>	120/120
	<u>148/148</u>								<u>132/132</u>		<u>106/106</u>		<u>120/120</u>	

(b) RELIOTYPE output files

File_1: Reliability and replication strategy, based on reliability criteria alone, suggested after including the first two genotypes in a first input file. Integers in the table are the suggested number of additional replicates at each locus. The first number in the parentheses corresponds to the expected number of additional amplification reactions necessary to obtain reliable genotypes. The second number in the parentheses corresponds to the expected number of additional amplification reactions necessary to observe each allele the required number of times (in this case 1 time). The integer before the parentheses is the greater of the two numbers in parentheses. An asterisk indicates a locus with no data. All the RELIOTYPE files in the appendix have the same structure.

Parametric Settings:
Upper_bound_on_the_dropout_rate: .75%.
Upper_bound_search_interval: 2%.
#bootstrap_rep/candidate_value_to_obtain_upper_bound: 200
Multiple_Test_Correction: to_limit_samplewide_#_of_genotype_errors_to_<_.5%_with_95%_probability.
a_98_29_%_reliability_is_required_of_each_individual_sample.
Minimum_#_of_observations/allele: 1
Single_round_of_replication_probability_of_achieving_this_minimum: .95%.

Sample replication)	p [*]	p(up) [*]	Est. Reliab.	Locus								Exp. Reliab. (after_
				Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113		
Sample 1	0	0.16	0.9761615	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9916587	
Sample 2	0	0.18	0.8777899	2(2,0)	0(0,0)	0(0,0)	0(0,0)	1(1,0)	1(1,0)	0(0,0)	0.9882988	
Sample 3	0.3301	0.6000999	0.3321932	5(5,0)	0(0,0)	0(0,0)	0(0,0)	3(3,0)	3(3,0)	3(3,0)	0.9876633	
Sample 4	0	0.18	0.9645802	0(0,0)	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9872025	
Sample 5	0.5000998	0.6300999	0.8849086	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	2(2,0)	0(0,0)	0.9872552	
Sample 6	0.1701	0.4301	0.2886649	3(3,0)	2(2,0)	3(3,0)	0(0,0)	0(0,0)	4(4,0)	0(0,0)	0.9874028	
Sample 7	0.3301	0.5800999	0.3515429	4(4,0)	0(0,0)	0(0,0)	0(0,0)	3(3,0)	3(3,0)	3(3,0)	0.9844245	
Sample 8	0	0.16	0.8520026	1(1,0)	1(1,0)	0(0,0)	0(0,0)	1(1,0)	1(1,0)	1(1,0)	0.9869143	

Sample replication	p [*]	p(up) [*]	Locus								Exp. Reliab. (after_ Accept_
			Est. Reliab.	Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113	
Sample 9	0.1701	0.3501	0.5628194	0(0,0)	2(2,0)	2(2,0)	3(3,0)	0(0,0)	1(1,0)	0(0,0)	0.9856786
Sample 11	0.5000998	0.4801	1	0(0,0)	*	*	*	*	*	1	Accept_
Genotype_w/o_further_replication											
Sample 13	1	1	1	0(0,0)	*	*	*	*	*	*	1
Accept_Genotype_w/o_further_replication											1
Sample 16	0.2501	0.5601	0.3094229	5(5,0)	0(0,0)	3(3,0)	0(0,0)	3(3,0)	3(3,0)	3(3,0)	0.9871222
Sample 17	0.3001	0.4301	0.8460803	0(0,0)	2(2,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	1(1,0)	0.9835305
Sample 19	0	0.18	0.9342042	1(1,0)	0(0,0)	0(0,0)	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0.9843011
Sample 20	0.1001	0.2001	0.9631106	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.986473
Sample 21	1	1	0.2467004	8(8,0)	6(6,0)	*	*	*	*	*	0.9841045
Sample 23	0.5000998	0.6700999	0.4739773	5(5,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	3(3,0)	3(3,0)	0.9872557
Sample 25	0	0.16	0.9761615	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9916587
Sample 26	0.2501	0.4001	0.950173	0(0,0)	*	*	*	*	1(1,0)	0(0,0)	0.9896183
Sample 28	0.1001	0.2001	0.9631106	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.986473
Sample 29	0	0.16	0.9761615	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9916587
Sample 30	0.5000998	0.7400998	0.2356377	5(5,0)	5(5,0)	5(5,0)	0(0,0)	4(4,0)	0(0,0)	0(0,0)	0.985154
Sample 31	0.6700997	0.8300997	0.2467908	0(0,0)	0(0,0)	5(5,0)	5(5,0)	0(0,0)	5(5,0)	6(6,0)	0.984312
Sample 32	0.1001	0.1801	0.9214208	1(1,0)	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9924999
Sample 33	0.2501	0.3701	0.9570561	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	1(1,0)	0(0,0)	0.991765
Sample 34	0	0.2	0.9456767	0(0,0)	0(0,0)	1(1,0)	1(1,0)	0(0,0)	1(1,0)	0(0,0)	0.9943798
Sample 35	1	1	0.1015269	10(10,0)	*	*	*	7(7,0)	6(6,0)	6(6,0)	0.9845932
Sample 36	0.5000998	0.8700997	0.2047154	0(0,0)	*	*	*	6(6,0)	5(5,0)	6(6,0)	0.9829884
Sample 37	0.1701	0.3301	0.9655345	0(0,0)	*	*	*	0(0,0)	1(1,0)	0(0,0)	0.9941429
Sample 39	0.1001	0.1801	0.964542	0(0,0)	0(0,0)	1(1,0)	0(0,0)	0(0,0)	0(0,0)	0(0,0)	0.9871872

Sample	40	0.6700997	0.8200997	0.1742674	7(7, 0)	0(0, 0)	0(0, 0)	0(0, 0)	5(5, 0)	5(5, 0)	4(4, 0)	0.9851566
Sample	41	0	0.18	0.8452437	0(0, 0)	1(1, 0)	1(1, 0)	2(2, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9847311
Sample	42	1	1	1.4189E-02	10(10, 0)	8(8, 0)	0(0, 0)	7(7, 0)	9(9, 0)	7(7, 0)	9(9, 0)	0.9833354
Sample	43	0	0.18	0.9698909	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9892472
Sample	44	0	0.18	0.8060688	0(0, 0)	1(1, 0)	1(1, 0)	2(2, 0)	1(1, 0)	1(1, 0)	0(0, 0)	0.9899032
Sample	46	0	0.16	0.9761615	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9916587
Sample	47	0.1001	0.2001	0.9631106	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.986473
Sample	48	0.5000998	0.7700998	0.146447	7(7, 0)	0(0, 0)	5(5, 0)	0(0, 0)	5(5, 0)	4(4, 0)	4(4, 0)	0.9847429
Sample	49	0.7500996	0.9200001	3.0798E-02	9(9, 0)	0(0, 0)	0(0, 0)	9(9, 0)	6(6, 0)	6(6, 0)	5(5, 0)	0.9841915
Sample	51	0.5000998	0.6300999	0.5990275	0(0, 0)	3(3, 0)	4(4, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9886063
Sample	52	0	0.48	1	0(0, 0)	*	*	*	*	*	1	Accept_
Genotype_w/o_further_replication												
Sample	53	0.2001	0.3701	0.882233	0(0, 0)	2(2, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	1(1, 0)	0.9899637
Sample	54	0.1301	0.2301	0.910115	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	1(1, 0)	1(1, 0)	1(1, 0)	0.9890229
Sample	55	0.5000998	0.7700998	0.2268747	6(6, 0)	5(5, 0)	0(0, 0)	0(0, 0)	5(5, 0)	4(4, 0)	4(4, 0)	0.9842226
Sample	56	0	0.18	0.8991274	0(0, 0)	1(1, 0)	0(0, 0)	1(1, 0)	1(1, 0)	0(0, 0)	1(1, 0)	0.9902509
Sample	57	0.1701	0.3301	0.6672911	3(3, 0)	0(0, 0)	0(0, 0)	0(0, 0)	2(2, 0)	1(1, 0)	1(1, 0)	0.9865808
Sample	58	0	0.16	0.9419208	1(1, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9887832

File 2. Reliability and replication strategy suggested after including the first two genotypes in a first input file, when, in addition to the reliability criteria, all alleles needed to be observed at least twice. The program produced seven errors, as it did not recommend further replications at seven loci where an allele had only been observed once. The errors have been underlined and the corresponding alleles are shown in Appendix 1a.

Parametric Settings:
Upper_bound_on_the_dropout_rate: 75%.
Upper_bound_search_interval: 2%.
#_bootstrap_rep/candidate_value_to_obtain_upper_bound: 200
Multiple_Test_Correction: to_limit_samplewide_#_of_genotype_errors_to_<5%_with_95%_probability.
a_98_29_%_reliability_is_required_of_each_individual_sample.
Minimum_#_of_observations/allele: 2
Single_round_of_replication_probability_of_achieving_this_minimum: 95%.

Sample replication)	p [^]	p [^] (up)	Est. Reliab.	Locus							Exp. Reliab. (after_
				Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113	
Sample 1	0	0.16	0.9761615	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9916587
Sample 2	0	0.18	0.8777899	2(2, 2)	0(0, 0)	0(0, 0)	0(0, 0)	1(1, 0)	1(1, 0)	0(0, 0)	0.9882988
Sample 3	0.3301	0.6000999	0.3321932	5(5, 4)	0(0, 0)	3(0, 3)	3(0, 3)	3(3, 3)	3(3, 3)	3(3, 2)	0.9876633
Sample 4	0	0.18	0.9645802	0(0, 0)	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9872025
Sample 5	0.5000998	0.6300999	0.8849086	3(0, 3)	3(0, 3)	3(0, 3)	3(0, 3)	3(0, 3)	3(2, 3)	3(0, 3)	0.9872552
Sample 6	0.1701	0.4301	0.2886649	3(3, 3)	2(2, 2)	3(3, 2)	2(0, 2)	0(0, 0)	4(4, 4)	0(0, 0)	0.9874028
Sample 7	0.3301	0.5800999	0.3515429	4(4, 4)	0(0, 0)	3(0, 3)	3(0, 3)	3(3, 3)	3(3, 2)	3(3, 2)	0.9844245
Sample 8	0	0.16	0.8520026	2(1, 2)	2(1, 2)	0(0, 0)	0(0, 0)	1(1, 0)	1(1, 0)	1(1, 0)	0.9869143
Sample 9	0.1701	0.3501	0.5628194	0(0, 0)	2(2, 2)	2(2, 0)	3(3, 0)	0(0, 0)	1(1, 0)	<u>0(0, 0)</u>	0.9856786

Sample 11 1	0.5000998	0.4801	1	3 (0, 3)	*	*	*	*	*	*	*	*	*	*
Sample 13 1	1	1	1	6 (0, 6)	*	*	*	*	*	*	*	*	*	*
Sample 16	0.2501	0.5601	0.3094229	5 (5, 4)	3 (3, 3)	0 (0, 0)	3 (3, 3)	3 (3, 2)	3 (3, 2)	3 (3, 2)	3 (3, 2)	3 (3, 2)	0.9871222	
Sample 17	0.3001	0.4301	0.8460803	0 (0, 0)	2 (2, 2)	2 (0, 2)	2 (0, 2)	0 (0, 0)	2 (0, 2)	0 (0, 0)	1 (1, 0)	1 (1, 0)	0.9835305	
Sample 19	0	0.18	0.9342042	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9843011	
Sample 20	0.1001	0.2001	0.9631106	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.986473	
Sample 21	1	1	0.2467004	8 (8, 7)	6 (6, 6)	*	*	*	*	*	*	*	0.9841045	
Sample 23	0.5000998	0.6700999	0.4739773	5 (5, 4)	0 (0, 0)	3 (0, 3)	3 (0, 3)	4 (0, 4)	3 (3, 3)	3 (3, 3)	3 (3, 3)	3 (3, 3)	0.9872557	
Sample 25	0	0.16	0.9761615	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9916587	
Sample 26	0.2501	0.4001	0.950173	2 (0, 2)	*	*	*	*	*	1 (1, 0)	0 (0, 0)	0 (0, 0)	0.9896183	
Sample 28	0.1001	0.2001	0.9631106	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.986473	
Sample 29	0	0.16	0.9761615	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9916587	
Sample 30	0.5000998	0.7400998	0.2356377	5 (5, 5)	5 (5, 4)	5 (5, 4)	4 (0, 4)	4 (4, 4)	4 (0, 4)	4 (0, 4)	4 (0, 4)	4 (0, 4)	0.985154	
Sample 31	0.6700997	0.8300997	0.2467908	0 (0, 0)	5 (0, 5)	5 (5, 5)	5 (5, 5)	5 (0, 5)	5 (0, 5)	5 (5, 4)	6 (6, 5)	6 (6, 5)	0.984312	
Sample 32	0.1001	0.1801	0.9214208	2 (1, 2)	0 (0, 0)	1 (1, 0)	0 (0, 0)	2 (0, 2)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9924999	
Sample 33	0.2501	0.3701	0.9570561	2 (0, 2)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0.991765	
Sample 34	0	0.2	0.9456767	0 (0, 0)	0 (0, 0)	1 (1, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0.9943798	
Sample 35	1	1	0.1015269	10 (10, 7)	*	*	*	*	7 (7, 6)	6 (6, 5)	6 (6, 5)	6 (6, 5)	0.9845932	
Sample 36	0.5000998	0.8700997	0.2047154	4 (0, 4)	*	*	*	*	6 (6, 6)	5 (5, 4)	5 (5, 4)	5 (5, 4)	0.9829884	
Sample 37	0.1701	0.3301	0.9655345	0 (0, 0)	*	*	*	0 (0, 0)	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0.9941429	
Sample 39	0.1001	0.1801	0.964542	0 (0, 0)	2 (0, 2)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9871872	
Sample 40	0.6700997	0.8200997	0.1742674	7 (7, 6)	4 (0, 4)	4 (0, 4)	5 (0, 5)	5 (5, 4)	5 (5, 4)	5 (5, 4)	4 (4, 4)	4 (4, 4)	0.9851566	
Sample 41	0	0.18	0.8452437	0 (0, 0)	1 (1, 0)	1 (1, 0)	2 (2, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9847311	
Sample 42	1	1	1.4189E-02	10 (10, 7)	8 (8, 6)	6 (0, 6)	7 (7, 6)	9 (9, 7)	9 (9, 7)	7 (7, 5)	9 (9, 7)	9 (9, 7)	0.9833354	
Sample 43	0	0.18	0.9699809	0 (0, 0)	1 (1, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.9892472	

Sample replication)	\hat{p}	$\hat{p}(\text{up})$	Est. Reliab.	Locus								Exp. Reliab. (after_
				Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113		
Sample 44	0	0.18	0.8060688	0(0, 0)	1(1, 0)	1(1, 0)	2(2, 0)	1(1, 0)	1(1, 0)	0(0, 0)	0.9899032	
Sample 46	0	0.16	0.9761615	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9916587	
Sample 47	0.1001	0.2001	0.9631106	0(0, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.986473	
Sample 48	0.5000998	0.7700998	0.146447	7(7, 5)	4(0, 4)	5(5, 4)	4(0, 4)	5(5, 4)	4(4, 4)	4(4, 3)	0.9847429	
Sample 49	0.7500996	0.9200001	3.0798E-02	9(9, 7)	5(0, 5)	4(0, 4)	9(9, 7)	6(6, 5)	6(6, 5)	5(5, 4)	0.9841915	
Sample 51	0.5000998	0.6300999	0.5990275	3(0, 3)	3(3, 3)	4(4, 4)	3(0, 3)	4(0, 4)	0(0, 0)	3(0, 3)	0.9886063	
Sample 52	0	0.48	1		0(0, 0)	*	*	*	*	*	*	
1	Accept_Genotype_w/o_further_replication											
Sample 53	0.2001	0.3701	0.882233	0(0, 0)	2(2, 2)	2(0, 2)	2(0, 2)	0(0, 0)	0(0, 0)	1(1, 0)	0.9899637	
Sample 54	0.1301	0.2301	0.910115	0(0, 0)	2(0, 2)	0(0, 0)	0(0, 0)	1(1, 0)	1(1, 0)	1(1, 0)	0.9890229	
Sample 55	0.5000998	0.7700998	0.2268747	6(6, 5)	5(5, 4)	4(0, 4)	4(0, 4)	5(5, 4)	4(4, 4)	4(4, 3)	0.9842226	
Sample 56	0	0.18	0.8991274	0(0, 0)	2(1, 2)	0(0, 0)	1(1, 0)	1(1, 0)	0(0, 0)	1(1, 0)	0.9902509	
Sample 57	0.1701	0.3301	0.6672911	3(3, 3)	0(0, 0)	0(0, 0)	<u>0(0, 0)</u>	2(2, 0)	1(1, 0)	1(1, 0)	0.9865808	
Sample 58	0	0.16	0.9419208	1(1, 0)	1(1, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0(0, 0)	0.9887832	

(c) Consensus genotypes that would have been obtained by following the replication strategy suggested by RELIOTYPE that was based on reliability criteria alone. Errors that would have been obtained are highlighted in bold. Genotypes in italic would have required further replication to be considered reliable.

		Consensus alleles at the microsatellite loci under investigation							
	Sample	Mel-105	Mel-102	Mel-117	Mel-106	Mel-111	Mel-109	Mel-113	Match
Parkmill Social Group	2	148 148	195 197	187 189	220 222	132 132	106 106	120 120	Q36
	3	148 148	195 197	187 189	220 222	132 138	106 106	120 120	
	7	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	8	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	16	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	40	148 148	195 197	174 189	220 222	132 132	106 106	120 120	
	48	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	49	148 148	195 197	187 189	220 222	132 132	106 106	120 120	
	57	148 148	195 197	187 189	220 222	132 132	106 106	120 120	X59
	1	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	20	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	25	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	28	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	29	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	43	138 142	199 199	174 193	220 222	132 138	106 116	120 120	
	17	138 142	199 199	174 187	224 226	130 132	106 125	120 120	D77
	46	138 142	199 199	174 187	220 224	130 132	106 125	120 120	
	47	138 142	199 199	174 187	220 224	130 132	106 125	120 120	
	53	138 142	199 199	174 187	220 224	130 132	106 125	120 120	
	9	138 142	199 199	187 187	220 220	130 132	106 106	118 120	U41
	41	138 142	199 199	187 193	220 220	130 132	106 106	118 120	
	44	138 142	199 199	187 187	220 220	130 132	106 106	118 120	
	31	138 140	195 197	174 187	222 222	130 132	106 106	120 126	
11	138 140	-	-	-	-	-	-	-	
19	138 138	197 199	174 187	220 222	132 132	106 106	120 126	-	
30	138 142	199 199	174 187	220 222	132 132	106 116	120 126	-	
52	136 144	-	-	-	-	-	-	Q66	
54	140 144	195 197	174 187	220 224	132 132	106 106	126 126	Q65	
58	138 138	199 199	174 187	222 226	132 138	106 125	120 120	H51	
Nettle	6	142 142	195 199	174 174	222 224	130 132	106 127	120 126	J68
	32	142 142	195 199	174 174	222 224	130 132	106 127	120 126	
	51	138 142	195 199	174 174	222 224	130 132	106 127	120 126	
	4	138 142	195 199	174 174	222 224	130 132	106 106	120 126	
	39	138 142	195 199	174 174	222 224	130 132	106 106	120 126	J56
	5	138 142	195 199	174 187	222 224	130 132	106 106	120 126	Q72
	33	138 142	195 199	174 187	222 224	130 132	106 106	120 126	
	23	138 138	197 199	174 189	222 226	130 132	106 106	120 120	
	55	138 138	197 199	174 189	220 222	132 132	106 106	120 120	
56	136 140	193 193	174 187	220 222	132 132	106 127	120 120		
Kennel	21	138 148	195 199	-	-	-	-	-	U62
	34	138 148	195 199	187 187	222 222	130 132	106 106	120 126	
	37	138 148	-	-	-	130 138	106 106	120 126	
	42	144 148	199 199	174 187	222 222	130 130	106 106	126 126	U61
	36	144 148	-	-	-	130 130	106 106	126 126	
	13	136 142	-	-	-	-	-	-	T50
	26	140 148	-	-	-	-	106 106	120 126	M58
	35	148 148	-	-	-	132 132	106 106	120 120	X30

Appendix II

(a) Results of the 283 amplification reactions that gave rise to the 40 faecal profiles used to investigate territorial marking in Chapter 5. In a first step, the samples that contained amplifiable DNA were genotyped once at all the loci and compared to reliable reference profiles (see Chapter 4). Reference profiles are in bold and the results from the initial genotyping are in normal font.

- (1) Provisional faecal profiles that were identical to reference profiles. These samples were considered correct and further amplifications were unnecessary.
- (2) Provisional faecal profiles that differed from reference profiles by one allele at one locus.
- (3) Provisional faecal profiles that differed from reference profiles by two alleles at two different loci.
- (4) Provisional faecal profiles that differed from reference profiles by two alleles at one different locus.

The amplifications performed to obtain a match with a reference profile are shown in italic. After these amplifications, the faecal profile matched with the reference profile in bold, but not with the one in italic and bold. It was not possible to amplify locus *Mel-105* from those faecal samples marked with (*).

- (5) Two special cases:
 - a. Even though faecal sample 33 differed from a reference profile EMa4 by two alleles at two loci, the two profiles had three alleles between them at the second most informative locus. Therefore, the faecal profile was compared by hand with other reference profiles. Despite an initial three-allele difference, repeated genotyping showed that faecal sample 33 originated from animal EMbH.
 - b. Faecal sample 20 did not match with any reference profile. All the loci were amplified a second time to obtain a reliable profile.

Faecal sample / Reference	Alleles at the microsatellite loci under investigation					
	<i>Mel-105</i>	<i>Mel-116</i>	<i>Mel-117</i>	<i>Mel-113</i>	<i>Mel-106</i>	<i>Mel-111</i>
(1)						
14	142 144	121 136	193 193	124 130	222 224	132 138
51	142 144	121 136	193 193	124 130	222 224	132 138
EMa1	142 144	121 136	193 193	124 130	222 224	132 138
33	144 144	123 132	187 193	124 124	222 224	132 132
35	144 144	123 132	187 193	124 124	222 224	132 132
EMa2	144 144	123 132	187 193	124 124	222 224	132 132
5	144 146	121 123	193 193	124 124	222 224	132 132
15	144 146	121 123	193 193	124 124	222 224	132 132
52	144 146	121 123	193 193	124 124	222 224	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
41	142 144	132 136	193 193	124 124	218 222	132 138
45	142 144	132 136	193 193	124 124	218 222	132 138
50	142 144	132 136	193 193	124 124	218 222	132 138
64	142 144	132 136	193 193	124 124	218 222	132 138
EMb1	142 144	132 136	193 193	124 124	218 222	132 138
26	142 142	132 136	174 193	124 124	222 222	132 138
EMb2	142 142	132 136	174 193	124 124	222 222	132 138
42	142 142	121 136	193 195	124 130	222 224	132 132
EMb3	142 142	121 136	193 195	124 130	222 224	132 132
3	142 146	121 132	174 174	124 130	222 224	132 138
12	142 146	121 132	174 174	124 130	222 224	132 138
37	142 146	121 132	174 174	124 130	222 224	132 138
38	142 146	121 132	174 174	124 130	222 224	132 138
46	142 146	121 132	174 174	124 130	222 224	132 138
47	142 146	121 132	174 174	124 130	222 224	132 138
58	142 146	121 132	174 174	124 130	222 224	132 138
EMbH	142 146	121 132	174 174	124 130	222 224	132 138
18	146 148	123 136	174 174	130 130	222 222	138 140
KH2	146 148	123 136	174 174	130 130	222 222	138 140

(2)						
6*	-	121 123	193 193	124 124	222 222	132 132
		<i>121 123</i>	<i>174 193</i>	<i>124 124</i>	<i>222 222</i>	
EMa4	142 144	121 123	174 193	124 124	222 222	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
25*	-	121 123	193 193	124 124	222 222	132 132

Faecal sample / Reference	Alleles at the microsatellite loci under investigation					
	<i>Mel-105</i>	<i>Mel-116</i>	<i>Mel-117</i>	<i>Mel-113</i>	<i>Mel-106</i>	<i>Mel-111</i>
			<i>174 193</i>			
			<i>174 193</i>			
EMa4	142 144	121 123	174 193	124 124	222 222	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
34	144 144	121 123	193 193	124 124	222 224	132 132
	<i>144 146</i>					
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
36	144 144	121 123	193 193	124 124	222 224	132 132
	<i>144 146</i>					
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
40	142 144	121 136	193 193	124 130	224 224	132 138
					<i>222 224</i>	
EMa1	142 144	121 136	193 193	124 130	222 224	132 138
43*	-	121 132	174 174	124 130	222 224	132 132
						<i>132 132</i>
						<i>132 132</i>
						<i>132 138</i>
EMbH	142 146	121 132	174 174	124 130	222 224	132 138
48*	-	132 132	193 193	124 124	218 222	132 138
		<i>132 136</i>	<i>193 193</i>	<i>124 124</i>		<i>132 138</i>
		<i>132 136</i>	<i>193 193</i>	<i>124 124</i>		
				<i>124 124</i>		
EMb1	142 144	132 136	193 193	124 124	218 222	132 138
54*	-	123 132	193 193	124 124	222 224	132 132
			193 193			
			187 187			
EMa2	144 144	123 132	187 193	124 124	222 224	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
56	142 144	132 132	193 193	124 124	218 222	132 138
		<i>132 136</i>				
EMb1	142 144	132 136	193 193	124 124	218 222	132 138
65*	-	121 121	174 174	124 130	222 224	132 138
		<i>121 132</i>				
EMbH	142 146	121 132	174 174	124 130	222 224	132 138
75*	-	132 136	193 193	124 124	222 222	132 138
		<i>132 136</i>			<i>218 222</i>	<i>132 138</i>

Faecal sample / Reference	Alleles at the microsatellite loci under investigation					
	<i>Mel-105</i>	<i>Mel-116</i>	<i>Mel-117</i>	<i>Mel-113</i>	<i>Mel-106</i>	<i>Mel-111</i>
EMb1	142 144	132 136	193 193	124 124	218 222	132 138
EMb2	142 142	132 136	174 193	124 124	222 222	132 138
77	142 146	121 132	174 174	124 130 <i>124 130</i>	222 224	132 138
EMbH	142 146	121 132	174 174	124 130	222 224	132 138
(3)						
23	142 144	121 121 <i>121 136</i>	193 193	124 130	222 224	132 132 <i>132 138</i>
EMa1	142 144	121 136	193 193	124 130	222 224	132 138
55	144 144 <i>144 146</i>	121 121 <i>123 123</i> <i>121 123</i> <i>121 123</i>	193 193 <i>193 193</i> <i>193 193</i> <i>193 193</i>	124 124	222 224	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
72	142 144	132 132 <i>132 136</i>	193 193	124 124	222 222 <i>218 218</i>	132 138
EMb1	142 144	132 136	193 193	124 124	218 222	132 138
(4)						
4	144 146	121 123	193 193	126 126 <i>124 124</i>	222 224	132 132
EMa3	144 146	121 123	193 193	124 124	222 224	132 132
19	140 140 <i>146 148</i>	123 136	174 174	130 130	222 222	138 140
KH2	146 148	123 136	174 174	130 130	222 222	138 140
(5)						
2*	-	121 132	174 174	124 124 <i>130 130</i>	222 222 <i>222 224</i>	132 132 <i>132 132</i> <i>132 138</i>
EMbH	142 146	121 132	174 174	124 130	222 224	132 138
EMa4**	142 144	121 123	174 193	124 124	222 222	132 132
20	142 142 <i>142 142</i>	121 132 <i>121 132</i>	174 195 <i>174 195</i>	124 130 <i>124 130</i>	222 224 <i>222 224</i>	132 140 <i>132 140</i>
no match						

(c) Sex identification of badgers by simultaneous amplification in the same PCR of the testis-determining factor gene (*SRY*), a Y-chromosome fragment, and a microsatellite as positive control (multiplex PCR). Samples that amplify the positive control without amplifying the *SRY* fragment are scored as females, while those that amplify both fragments are scored as males. (1) Test of the suitability of microsatellite locus *Mel-106* as positive control. (2) Test of the suitability of microsatellite locus *Mel-117* as positive control. On both gels, the PCR products in the first six lanes originate from male DNA, while the following six products originate from females and the last lane is the negative control that did not contain any DNA. (3) Multiplex PCR of *SRY* fragment / *Mel-117* used in the sex identification of DNA samples originating from all the known badgers in the study area. Lanes 1-2: EMa1; L3-4: EMa2; L5-6: EMa3; L7-8: EMa4; L9-10: EMb1; L11-12: EMb2; L13-14: EMb3; L15-16: KH1; L17-18: KH2; L19-20: KH3; L21-22: B1; L23-24: B2; L25-26: G1; L27-28: profile G; L29-30: profile C; L31-32: profile N; L33-34: profile O; L35-36: test with faecal sample 20; L37: positive control. Badgers that have not been captured are represented by their remote-DNA profile (see Table 4.3).

