Norwegian University of Life Sciences
Faculty of Environmental Science and Technology
Department of Ecology
and Natural Resource Management

Philosophiae Doctor (PhD)
Thesis 2015:72

Genetic structure and gene flow in a continuously distributed large terrestrial carnivore - the brown bear (*Ursus arctos*) in Northern Europe

Genetisk struktur og genflyt i et kontinuerlig distribuert stort rovdyr - den nordeuropeiske brunbjørn (*Ursus arctos*)

Julia Schregel

Genetic structure and gene flow in a continuously distributed large terrestrial carnivore - the brown bear (*Ursus arctos*) in Northern Europe

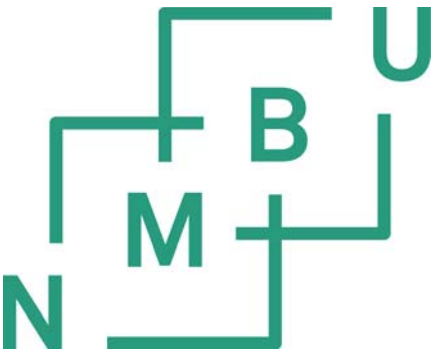
Genetisk struktur og genflyt i et kontinuerlig distribuert stort rovdyr - den nordeuropeiske brunbjørn (*Ursus arctos*)

Philosophiae Doctor (PhD) Thesis

Julia Schregel

Department of Ecology and Natural Resource Management
Faculty of Environmental Science and Technology
Norwegian University of Life Sciences

Ås 2015



Thesis number 2015:72
ISSN 1894-6402
ISBN 978-82-575-1309-2

PhD supervisors

Dr Snorre B. Hagen[1]

Dr Hans Geir Eiken[1]

Professor Jon E. Swenson[2]

[1]Norwegian Institute of Bioeconomy (NIBIO), Section of Biotechnology and Molecular Genetics, Frederik A. Dahls vei 20, N-1430 Ås, Norway

[2]NMBU, Department of Ecology and Natural Resource Management

P.O.Box 5003, N-1432 Ås, Norway

PhD evaluation committee

Professor Ettore Randi

University of Bologna

Instituto Superiore per la protezione e la Ricerca Ambientale (ISPRA)

Via Cà Fornacetta 9

I-40064 Ozzano dell'Emilia (BO), Italy

Dr Øystein Flagstad

Norwegian Institute for Nature Research (NINA)

P.O.Box 5685 Sluppen

N-7485 Trondheim, Norway

Committee coordinator

Dr Katrine Eldegard

NMBU, Department of Ecology and Natural Resource Management

P.O.Box 5003, N-1432 Ås, Norway

Table of contents

Papers I-V have separate page numbers

Summary

The aim of this thesis was to investigate large scale population structure and gene flow in the Northern European brown bear (*Ursus arctos*). Brown bears were reduced in numbers to near extinction during recent centuries due to human persecution, but have gradually recovered after protective measures were initiated. In this context, our goal was also to provide knowledge that will be of aid for current and future management.

For this, we analyzed samples obtained at high spatial density from across the Northern European distribution zone in the course of regional and national monitoring schemes during 2005-2012, which allowed statistical analysis on various spatial scales. Sample material consisted of feces and hairs sampled in the field as well as tissue samples of bears shot legally. For the genetic analysis, we employed autosomal microsatellite markers (short-tandem repeats, STRs) and male specific Y chromosomal markers (STRs and single-nucleotide polymorphisms, SNPs). We used individual-based population genetic approaches (Bayesian assignment algorithms) to unveil population structure, F-statistics to estimate genetic differentiation, and spatial autocorrelation to assess fine-scale population structure. Population size (N_c), effective population size (N_e) and the ratio between these measures may have a practical value in wildlife conservation, and we investigated the temporal stability of these measures in a focal bear group in the Pasvik Valley between Norway, Finland and Russia. We also performed clustering and cluster placement analyses to identify the number of genetic clusters across Sweden and Norway and assess their spatial arrangement. Furthermore, we used the new DResD approach to i) identify the location of gene flow barriers and corridors and ii) determine the importance of limited dispersal distance as the structuring mechanism. This was done by correcting pairwise measures of genetic distance for isolation by distance (IBD).

Initially, we validated 12 STR markers and established an analysis protocol that minimizes risks for genotyping errors and produces highly reliable genetic data valuable for increasing possibility for comparing genetic data in international brown bear studies. Next, we determined the average brown bear populations size (N_c) in the Pasvik Valley during 2005-2010 to be between 40 and 45 individuals. Also we found a high N_e/N_c ratio, indicating gene flow from other populations. In the same study, a large scale analysis of population structure based on 477 individuals from the Pasvik Valley and three other regions (Karelia in Finland and Russia, Västerbotten in Sweden and Troms in Norway), revealed four distinct genetic clusters with low migration rates between the regions and indications of limited gene flow

towards the west. Using Y chromosomal markers on 443 males from Sweden, Norway, Finland and Northwestern Russia, we studied the importance male gene flow in the recovering Northern European brown bear. We found clear differences in distribution of Y-lineages across the study area depending on the amount of gene flow from the outside during the recovery process between the eastern and the western parts of Northern Europe. In the eastern parts of Northern Europe (Northeastern Norway (Pasvik Valley and Anarjohka), Finland and Northwestern Russia), we found high Y-haplotype diversity and admixture. In contrast, in the western parts of Northern Europe (Sweden and southern Norway), we found that the large population of today (ca. 3,000 individuals) is highly structured and that it may have recovered from as few as four male lineages. The sex-specific analysis of 1531 individuals from Norway and Sweden revealed a clear difference in genetic structure between males and females that tended to decrease from south to north, indicating limitations to male gene flow in the north. In the same sample material we also evaluated the Bayesian assignment results statistically and identified four cluster core areas. We found that gene flow may be asymmetrical, with more dispersing individuals crossing from east to west than the other way. With correcting measures of individual pairwise genetic distance, we identified two barriers, one in the south and one in the north, the latter also had been suggested in our region-wide studies. Our data indicated that these barriers limit female more than male gene flow. We also showed genetic structure of males when analyzed at the small scale, but not at the large scale. This may indicate that forming assumptions about small-scale structure from results based on large-scale analyses may lead to erroneous conclusions.

Our results documented two barriers to gene flow, one localized between the western and eastern subpopulations and a second barrier in southern Scandinavia. Thus, we found at least two areas where genetic connectivity should be improved to ensure the long-term genetic health of the Scandinavian subpopulation, in the southern part of the Scandinavian Peninsula and the area north of the Bay of Bothnia. For this, transborder management coordination should be established to match the transnational nature of the brown bear population.

Sammendrag

Målsettingen med denne avhandlingen var å undersøke storskala populasjonsstruktur og genflyt hos den nordeuropeiske brunbjørnen. Brunbjørnen ble redusert i antall til nær utryddelse i løpet av de siste århundrene på grunn av menneskets forfølgelse, men har gradvis kommet tilbake etter at fredningsbestemmelser ble innført. På dette grunnlaget var også vårt mål å bidra med kunnskap som kan være til hjelp for nåværende og fremtidig forvaltning. For å oppnå dette analyserte vi prøver som igjennom regionale og nasjonale kartleggingsprogram var samlet inn med høy geografisk tetthet igjennom leveområdet I Nord-Europa i tidsrommet 2005-2012. Prøvematerialet var egnet for statistisk analyse på ulike romlige skalaer, og bestod av ekskrementer og hår som var samlet i felten samt vevsprøver fra lovlig skutte bjørner. I den genetiske analysen brukte vi autosomale mikrosatellittmarkører og Y kromosomale markører (mikrosatellitt- og enkeltnukleotidmarkører). Vi brukte individ-baserte populasjonsgenetiske tilnærminger (Bayesiske tilhørighets algoritmer) for å avdekke struktur, F-statistikk for å estimere genetisk differensiering og romlig autokorrelasjon for å bestemme finskala populasjonsstruktur. Størrelse av populasjoner og effektiv populasjonsstørrelse og forholdet mellom disse målene kan ha en praktisk verdi for artsforvaltning, og vi undersøkte den temporale stabiliteten av disse målene i gruppen av bjørner i Pasviksdalen i mellom Norge, Finland og Russland. Vi utførte også analyser av hvordan individer samlet seg i grupper og i hvilke områder, for å identifisere genetiske grupper igjennom Sverige og Norge og avdekke gruppenes romlige utbredelse. Videre brukte vi den nye DResD tilnærmingen for å i) identifisere lokaliseringen av barrierer for genflyt og ii) bestemme hvor viktig begrensa spredningsavstand er som mekanisme for dannelse av genetisk struktur. Dette ble utført ved å korrigere de parvise målene for genetisk avstand for effekten av isolasjon ved avstand (IBD). Vi startet med å validere 12 mikrosatellitt markører og etablere en analyseprotokoll som minimerer risikoen for genotypings feil og produserer svært sikre genetiske data som er verdifulle for å øke mulighetene for å sammenligne genetiske data i internasjonale studier av brunbjørn. Deretter bestemte vi gjennomsnittlig populasjonsstørrelse i Pasvikdalen i tidsrommet 2005-2010 til å være mellom 40 til 45 individer. Vi fant også en høy ratio mellom effektiv og census populasjonsstørrelse noe som tyder på genflyt inn fra andre populasjoner. I den samme studien viste en storskalaanalyse basert på 477 individer fra tre andre regioner (Karelia i Finland og Russland, Västerbotten i Sverige og Troms i Norge) fire ulike genetiske grupper med lave migrasjonsrater mellom og indikasjoner på begrensa genflyt mot vest. Ved å

bruke markører på Y kromosomet til 443 hanner fra Sverige, Norge, Finland og Nordvest-Russland, kunne vi studere betydningen av hannlig genflyt i den gjennoppståtte Nord Europeiske bjørnestammen. Vi fant klare forskjeller i fordelingen av Y-linjer igjennom studieområdet avhengig av mengden genflyt utenfra i løpet av gjenoppstandelsesprosessen mellom de østlige og vestlige delene av Nord-Europa. I de østlige delene av Nord-Europa (Nordøst i Norge (Pasvikdalen og Anarjohka), Finland, and Nordøst-Russland), fant vi høy Y-haplotypediversitet og -sammenblanding. I motsetning, i de vestlige delene av Nord-Europa (Sverige og Sør-Norge), fant vi at dagens store populasjon (ca. 3000 individer) er sterkt strukturert og at den kan ha gjenoppstått fra så få som fire hannlinjer. Kjønnsspesifikk analyse av 1531 individer fra Norge og Sverige viste en klar forskjell i genetisk struktur mellom hanner og hunner som tenderte til å avta fra sør til nord, en indikasjon på mer begrenset hannlig genflyt i nord. I det samme prøvematerialet evaluerte vi også resultatene fra de Bayesiske tilhørighets-algoritmene statistisk og identifiserte fire genetiske populasjonskjerneområder og fant at genflyten kan være asymmetrisk med flere individer som beveger seg fra øst til vest enn andre veien. Ved å korrigere målene for individuell, parvis genetisk avstand identifiserte vi to barrierer, en i sør og en i nord, den siste antydet også i våre storskala studier. Våre data indikerer at disse barrierene begrenser hunnlig genflyt mer enn hannlig genflyt. Vi viser også genetiske struktur hos hanner analysert på liten skala, men ikke på stor skala. Dette kan indikere at det å forme antagelser om småskala struktur basert på resultater fra storskalaanalyser kan føre til feilaktige konklusjoner. Våre resultater dokumenterer to ulike genflytbarrierer, en lokalisert mellom den østlige og vestlige subpopulasjonen og en annen barriere sør i Skandinavia. Dermed finner vi minst to områder der den genetisk utvekslingen bør økes for å sikre langsiktige genetisk helse hos den Skandinaviske populasjonen, den sørlige del av den skandinaviske halvøy og området nord for Bottenvika. For dette bør det etableres en koordinert, grensekryssende forvaltning for å matche den grensekryssende brunbjørnpopulasjonen.

List of papers

Paper I

Andreassen, R., Schregel, J., Kopatz, A., Tobiassen, C., Knappskog, P. M., Hagen, S. B., Kleven, O., Schneider, M., Kojola, I., Aspi, J., Rykov, A., Tirronen, K. F., Danilov, P. I., Eiken, H. G. (2012) A forensic DNA profiling system for Northern European brown bears (*Ursus arctos*). *Forensic Science International* 6(6): 798-809

Paper II

Schregel, J., Kopatz, A., Hagen, S. B., Brøseth, H., Smith, M. E., Wikan, S., Wartiainen, I., Aspholm, P. E., Aspi, J., Swenson, J. E., Makarova, O., Polikarpova, N., Schneider, M., Knappskog, P. M., Ruokonen, M., Kojola, I., Tirronen, K. F., Danilov, P. I., Eiken, H. G. (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Molecular Ecology* 21 (14): 3474-3488.

Paper III

Schregel, J., Eiken, H. G., Grøndahl, F. A., Hailer, F., Aspi, J., Kojola, I., Tirronen, K. F., Danilov, P., Rykov, A., Poroshin, E., Janke, A., Swenson, J., Hagen, S. B. Present and past Y chromosomes reveal the demographic and genetic impact of male dispersal during the recovery of the Northern European brown bear (*Ursus arctos*). *Submitted manuscript*

Paper IV

Schregel, J., Eiken, H. G., Swenson, J. E., Hagen, S. B. A multiscale analysis of sex-dependent population structure and gene-flow: The case of the Scandinavian brown bear (*Ursus arctos*). *Manuscript*

Paper V

Schregel, J., Remm, J., Eiken, H.G., Saarma, U., Swenson, J.E., Hagen, S.B. The effect of scale and isolation by distance on the analysis of population structure in large, continuous populations: The Scandinavian Brown Bear (*Ursus arctos*). *Manuscript*

Introduction

Throughout history, large carnivores have fascinated humans, being revered and feared at the same time. During the last centuries, however, humans, especially in the densely populated Europe, perceived large carnivores as competition for food sources and a threat to life and livelihood, thus persecuting them with high intensity in order to eradicate them (Woodroffe 2000; Dalerum *et al.* 2009). In the wake of a growing awareness of the importance of intact ecosystems for human survival, though, attitudes slowly began to change and policies were altered in order to maintain and protect the remaining carnivore populations (Enserink & Vogel 2006). In recent years, scientific research has revealed the importance of apex predators for a healthy ecosystem (Ripple *et al.* 2014a; Ripple *et al.* 2014b; Prowse *et al.* 2015) and large carnivores have been the subject of a large number of studies. One of the most studied large carnivore species is the brown bear (Brooke *et al.* 2014). Once almost extirpated throughout Europe, the brown bear has made a successful come-back to its Northern European distribution, enabled by a change in attitudes and policies (Swenson *et al.* 1995; Enserink & Vogel 2006; Chapron *et al.* 2014). Populations of surviving brown bears were small at the beginning of the last century (~130 individuals in Sweden and ~150 individuals in Finland (Swenson *et al.* 1995; Ermala 2003) and numbers were also low in northwestern Russia (Danilov 2005)). However, populations in Sweden, Norway, Finland and western Russia have gradually grown to substantial sizes again (Swenson *et al.* 1995; Danilov 2005; Kojola *et al.* 2006b; Kindberg *et al.* 2011). The most recent population estimates are ~2800 individuals in Sweden in 2013 (Naturvårdsverket 2014), a minimum of 136 in Norway in 2014 (Aarnes *et al.* 2015), and 1,150-1,950 in Finland in 2009 (Wikman 2010). The most recent estimates in western Russia are from 1990; ~500 individuals in Murmansk Oblast to the north and ~3 500 in Russian Karelia to the south (Chestin 1992; Danilov 1994).

The brown bear in Northern Europe, especially the Swedish population, has been studied extensively to answer questions regarding e.g. life-history traits (Bellemain *et al.* 2006; Zedrosser *et al.* 2007a; Ordiz *et al.* 2008; Zedrosser *et al.* 2009), dispersal behavior (Støen *et al.* 2005; Støen *et al.* 2006; Zedrosser *et al.* 2007b), terrain use (Nellemann *et al.* 2007) and effects of human hunting (Bischof *et al.* 2008a; Bischof *et al.* 2008b; Ordiz *et al.* 2013). Genetic studies have been conducted based on the distribution and diversity of mtDNA sequences in order to reveal female population structure and to shed light on recolonization routes after the last Ice Age (Taberlet & Bouvet 1994; Bray *et al.* 2013; Keis *et al.* 2013), and Y-chromosomal

markers have been used to investigate speciation processes, phylogeographical structure and global male gene flow patterns (Bidon *et al.* 2014; Kutschera *et al.* 2014). Studies directed more towards contemporary genetic structure have also been performed: although restricted to the Swedish population, these studies revealed a distinct subpopulation structure (Waits *et al.* 2000; Manel *et al.* 2004). Information about the genetic connectivity across the large scale and across national borders was still lacking at the start of this thesis.

Thanks to its successful recovery, the brown bear is considered to be out of immediate threat of extinction in Fennoscandia and Russia (Servheen *et al.* 1999). However, conflicts with the human population due to livestock depredation and hunting competition have to be anticipated (Graham *et al.* 2005). A recent survey has shown that the increase of bear numbers in Sweden has resulted in a more negative attitude towards the species (Eriksson *et al.* 2015). Thus, there is a need to actively manage the brown bears in Northern Europe to alleviate human-bear conflict and to ensure their long-term survival. To aid the latter, information about genetic connectivity is important, especially across a larger geographical scale. Population fragmentation caused by barriers to gene flow can lead to a reduction in genetic diversity within subpopulations, which in turn has been shown to be strongly connected to population viability (Liberg *et al.* 2005; Hogg *et al.* 2006; Hostetler *et al.* 2013). Information about the genetic structure of the Northern European brown bear population is thus important to help identify potential areas of limited gene flow. If genetic barriers exist, it is important to identify them to carry out actions to ensure or reestablish genetic connectivity among subpopulations. As brown bear distribution and population boundaries do not follow political borders, a transnational approach as performed in this thesis is especially important (Meirmans 2015; Bischof *et al.* in press).

The study of the genetic population structure of the Northern European brown bear, the genetic connectivity among subpopulations, and the identification of possible gene flow barriers is not only important in the conservation and management perspective, but may also contribute to the advancement of population genetic theory. Recent studies have highlighted issues concerned with the accuracy and applicability of standard analysis tools used to assess and describe population genetic structure. For example, Schwartz & McKelvey (2008) have shown how different sampling schemes can influence the ability of the Bayesian assignment algorithm STRUCTURE (Pritchard *et al.* 2000) to correctly identify the number of populations present in the sample. This is because of the influence of isolation by distance, i.e. a correlation

between spatial and genetic distance due to limited dispersal distances of individuals (Wright 1943), on the genetic structure of a population. The problem of inaccurate assessment of population fragmentation due to IBD is well known in population and conservation genetics (Frantz *et al.* 2009; Meirmans 2012; Landguth & Schwartz 2014). Therefore results may be inaccurate, if analytical tools use inaccurate null models, like the nonspatial island model, where migration is equal among all populations, or make assumptions that are hardly ever met in wildlife populations, like populations being in Hardy-Weinberg equilibrium (HWE), which includes e.g. the assumption of non-overlapping generations and random mating. In turn, this potentially leads to inefficient or wrong management actions. Simulation studies often are employed to test population genetic concepts to further our understanding and to improve analysis tools. For example, Landguth *et al.* (2010) tested how many generations were needed before the appearance or disappearance of gene flow barriers was detected using different analytical algorithms. The problem is that also simulation studies make assumptions when generating the data, which may not reflect conditions found most commonly in nature. For example, van Strien *et al.* (2015) showed that landscape configuration and deme topology have a strong influence on the pattern of IBD, thus simulation studies that use a simple IBD model of progressive spatial and genetic distance may not be representative of actual population structure.

Directly related to the issue of IBD and its influence on the genetic structure of a population is the issue of sex-biased dispersal. When dispersal distances and probabilities differ between the sexes, the rate and distance at which gene flow occurs is dependent on sex, which in turn may lead to differing population genetic structure of the male and female component of a population (Lawson Handley & Perrin 2007). According to this, the dispersing sex would be the genetic mediator among regions, whereas the nondispersing sex would give rise to local population structure. There is a wide range of variation on this theme in nature, because social organization, mating strategy, and social complexity interact to influence the shape and magnitude of sex-biased dispersal (Greenwood 1980; Devillard *et al.* 2004; Lawson Handley & Perrin 2007). In an attempt to understand the evolutionary causes to dispersal, empirical studies have uncovered a number of proximate conditions that influence individual dispersal decisions, such as crowding, local kinship, habitat conditions, social structure, and individual fitness (Legagneux *et al.* 2009; Solmsen *et al.* 2011; Hardouin *et al.* 2012; Vercken *et al.* 2012). From this it is easy to understand that the issue of sex-biased dispersal is not only of theoretical

relevance for the study of evolutionary processes, but may be important also to questions of species conservation: if conditions for dispersal are unfavorable, demographic and genetic connectivity among populations may suffer, thus increasing fragmentation and lowering population viability (Vilà *et al.* 2003; Long *et al.* 2005). Traditionally, sex-biased dispersal has been studied with tools such as telemetry or presence-absence data (Driscoll *et al.* 2014). However, these methods are not suitable to tackle an important problem: dispersal does not necessarily lead to gene flow (Johnson & Gaines 1990; Kitanishi & Yamamoto 2015). Therefore, nongenetic methods are not able to assess the genetic consequences of a certain type of dispersal behavior. For this reason, the inclusion of genetic methods to the study of dispersal has increased during the last years (Driscoll *et al.* 2014), promising to shed new light on the causes and consequences of sex-biased dispersal.

Combining the results of a population genetic study with information about a species' ecology and behavioral patterns may be especially fruitful to better understand population processes and how they influence each other (Lowe & Allendorf 2010). Based on this, the Northern European brown bear is a well suited model system to shed more light on the described named above. A large amount of knowledge about the brown bear's ecology and behavioral patterns exist and the species' almost continuous distribution from the western edge in Norway across Sweden, Finland and western Russia allows sampling at high spatial density across a large geographical area, and thus facilitates studies at multiple scales. In addition, the recovery history of the Northern European brown bear and its male-biased dispersal behavior enables the study of the genetic consequences of these population processes. The aim of this thesis is, thus, to study range-wide connectivity in a highly mobile species, population genetic processes on small and large scales, as well as sex-specific gene flow and structure. To this end, I will analyze genetic data obtained through the application of autosomal genetic markers (microsatellites) as well as male specific Y-chromosomal markers to samples obtained from across the entire Northern European distribution zone. In addition, I will use genetic data obtained by noninvasive sampling of hairs and feces in the course of regional and national monitoring schemes in Sweden and Norway, thus testing the usefulness of this kind of database (which typically uses a smaller amount of microsatellite markers than common in research settings due to cost efficiency restrictions) for population genetic studies. In the following, I will explain the specific objectives for this thesis, the rationale behind them, and the approach used to answer the posed questions.

Objectives and their rational

How reliable are the microsatellite markers used to identify individuals and to study the genetic structure of brown bear populations? (Paper I)

Brown bears have been monitored in Sweden and Norway with the help of noninvasive genetic identification since 2006 (Rovdata-Naturvårdsverket 2014). This is done by using a set of microsatellite markers, which have been used in most of the previous population genetic brown bear studies (e.g. Paetkau *et al.* 1997; Waits *et al.* 2000; Bellemain *et al.* 2004; Manel *et al.* 2004; Proctor *et al.* 2005). Despite this widespread use, certain characteristics of microsatellite markers, such as ambiguous interpretability of allele length, allelic drop out, and the presence of null alleles do not guarantee comparability of data, influencing the reliability of comparisons or combination of analysis results among different laboratories and studies. In addition, brown bears are often involved in human-bear conflicts, such as livestock depredation or illegal killing. In such cases, a validated set of markers would aid in the specific identification of individuals for criminal investigations and enable the construction of reference data in order to facilitate the traceability of bear products. Therefore, the aim of this paper was to formally test the validity of the microsatellite markers most commonly used in population genetic studies in brown bears by testing species specificity and sensitivity as well as providing measures of precision, stutter, and heterozygote balance.

What is the genetic structure, connectivity, and diversity of the recovering Northern European brown bear population? (Paper II, III, IV, V)

After being almost extirpated in the beginning of the 20th century, the brown bear has made a remarkable come-back after protection measures were established in the respective countries (Swenson *et al.* 1995; Danilov 2005; Kojola *et al.* 2006b; Kindberg *et al.* 2011). The viability of a population is assumed to be strongly correlated to its intrapopulation genetic variability and its interpopulation connectivity (Cegelski *et al.* 2006; Lowe & Allendorf 2010; Hedrick *et al.* 2014). A bottleneck event of such magnitude as to reduce the population size to only about 130 and 150 individuals in Sweden and Finland, respectively (Pulliainen 1990; Swenson *et al.* 1995), may have a long-lasting effect, both on genetic variability and genetic structure and connectivity among Northern European brown bears. Nowadays, brown bears are distributed more or less continuously across much of the available habitat. Despite this, most previous studies were confined by national borders. Dispersal distances show that bears are a highly

mobile species (Swenson *et al.* 1998; Støen *et al.* 2006; Zedrosser *et al.* 2007b), therefore warranting a transboundary approach to assess population structure in the recovered Northern European brown bear. In paper II, we therefore used genetic data based on the analysis of autosomal microsatellite markers from four regions of the distribution zone to assess genetic variability within the regions and the genetic connectivity among them. In this paper, we also focused on the group of bears inhabiting the transborder area Pasvik-Inari-Pechenga, which is believed to be the only place in Norway, where bears were never extirpated completely (Swenson et al. 1995; Swenson & Wikan 1996). We estimated effective (N_e) and census population size (N_c) as well as the ratio N_e/N_c to assess the stability of this group as well as gain insight into a potential connectivity towards further east. Dispersal has been shown to be male-biased in the brown bear and it is assumed that connectivity among regions is mostly provided by male dispersal (Manel *et al.* 2004). Thus, in paper III, we applied Y-chromosomal markers to assess the population structure of male bears across the entire Northern European distribution zone. Even though brown bears became functionally extinct during the demographic bottleneck in the last century in most of Norway, numbers have been increasing steadily during the last decades to at least 136 individuals in 2014 (Aarnes *et al.* 2015). Previous assessment of population genetic structure in Scandinavia was restricted to Sweden and little was known about the connectivity between the two countries (Waits *et al.* 2000; Manel *et al.* 2004). To shed light on this issue, we used a spatial autocorrelation analysis and a hierarchical approach to assess large- and fine-scale population structure of the Scandinavian bear population (paper IV). Because IBD may have a large impact on the correct assessment of population structure (Frantz *et al.* 2009), and interpretation of STRUCTURE results is especially challenging in continuously distributed populations, like the brown bear in Norway and Sweden, we applied a recently developed method that takes the effect of IBD on the population structure into account (paper V).

What is the structure, diversity, and distribution of male brown bear lineages across the Northern European distribution zone? (Paper III)

Male brown bears disperse more frequently and across farther distances from their natal home range than females (Støen *et al.* 2006; Zedrosser *et al.* 2007b). Consequently, male genes are potentially passed on more frequently than female genes (Greenwood 1980) and therefore, males may show a weaker spatial genetic structure, while contributing more to both genetic

diversity within and genetic connectivity among populations than the females (Ishibashi *et al.* 2013; Quaglietta *et al.* 2013). Recently, Bidon *et al.* (2014) identified Y-chromosomal markers for polar (*Ursus maritimus*) and brown bears and were able to document male-specific gene flow across the entire range on a phylogeographic time scale. They also showed that, consistent with the male-biased dispersal displayed by the species, the global pattern of male genetic structure displayed more admixture than that based on mtDNA. Paper II showed that there were considerable limitations to gene flow between the eastern and western part of the Fennoscandian brown bear range, whereas Tammeleht *et al.* (2010) had documented a subdivision of the Finnish brown bear population. Kopatz *et al.* (2012), on the other hand, found bidirectional gene flow between eastern Finland and Arkhangelsk, Russia. However, all these studies were based on autosomal microsatellites, which are not particularly suited for the assessment of male-specific gene flow and population structure, because they are inherited biparentally. Previous studies did not find a clear genetic signature of the demographic bottleneck in the Scandinavian (Waits *et al.* 2000) nor in the Finnish brown bears (Kopatz *et al.* 2012) using microsatellite-based data. However, Y chromosomes are more sensitive to a dramatic population decline (Greminger *et al.* 2010). In paper III, we therefore applied Y-chromosomal microsatellite (Y-STRs) and single nucleotide polymorphism markers (Y-SNPs) to samples of 491 male bears from Sweden, Norway, Finland and Western Russia to study the Y-haplotypic diversity and dispersal as well as population structure. We also used historical samples from bears from southern Norway, originating from ~1750 to ~1950, to compare prebottleneck haplotype diversity with postbottleneck haplotype diversity in Sweden and Norway.

What is the importance of male gene flow in the Northern European brown bear? (Paper III, IV, V)

Considering that natal dispersal is male-biased and females show the formation of matrilineal assemblages (Støen *et al.* 2005; Støen *et al.* 2006; Zedrosser *et al.* 2007b), males should be responsible for most of the interregional gene flow and show a much weaker population genetic structure than females. Also, because males disperse across large distances (Støen *et al.* 2006), they may have contributed considerably to the large-scale recovery of the brown bear in Fennoscandia. We tested this hypothesis in paper III by comparing the Y-haplotype diversity and distribution of the two previously identified subpopulations of the Northern European

brown bear population, the Eastern subpopulation, including Finland and Russia, and the Western subpopulation, including Sweden and Norway (Paper II; Kopatz *et al.* 2014). We also wanted to investigate the effect of male-biased dispersal on the genetic population structure using autosomal microsatellite markers. Using this biparentally inherited marker had the advantage that we could compare male vs female structure directly. For paper IV, we therefore used the genotypes of 1531 brown bears, obtained from the national monitoring programs in Sweden and Norway, to investigate the fine-scale and study area-wide structure of the subpopulation. For this, we first used a Bayesian assignment algorithm to perform a hierarchical structure analysis, followed by an analysis of spatial autocorrelation, analyzed separately by sex. Several studies have shown that the presence of IBD in the population structure can lead to an overestimation of the number of true genetic clusters when using structuring algorithms (Schwartz & McKelvey 2008; Frantz *et al.* 2009). Keis *et al.* (2013) and Hindrikson *et al.* (2013) have published two novel approaches (cluster placement analysis and DResD) that may be of great value to aid in the assessment of population structure in the presence of IBD. Thus, in paper V, we applied these methods to the same dataset used for paper IV, in order to assess the impact of male-biased gene flow on the population structure of the brown bears in Sweden and Norway, corrected for the effect of IBD.

Do genetic analyses reflect the observed male-biased dispersal and female philopatry in the brown bear? (Paper III, IV, V)

The dispersal behavior of brown bears has been studied extensively using mainly telemetric methods. However, dispersal does not necessarily translate into gene flow (Johnson & Gaines 1990; Kitanishi & Yamamoto 2015). Thus, combining previous knowledge with the results of population genetic analyses may allow the estimation of the true costs and benefits of dispersal behavior (Lawson Handley & Perrin 2007). In paper III, using Y-chromosomal markers, we therefore tested the assumption of large-scale male gene flow by assessing the structure of male lineages in the entire Northern European distribution zone. In paper IV and V, we compared male and female population structure in the Scandinavian subpopulation.

What influence do scale and isolation by distance have on the result of the population genetic analysis? (Paper IV, V)

Scale is an important issue in population and landscape genetic analyses. How an individual perceives its environment and at what scale landscape features influence habitat selection, foraging behavior, or dispersal patterns are the subjects of a growing body of studies (Mayor & Schaefer 2005; Mayor *et al.* 2009; Cushman & Landguth 2010; Galpern *et al.* 2012; Heisler *et al.* 2013). Most often the evaluation of the scale effect, though, is restricted to matching pattern to process, i.e. selecting the right grain size for the landscape model in landscape genetic analyses (Anderson *et al.* 2010; de Knegt *et al.* 2010; Galpern *et al.* 2012). However, several studies have demonstrated that the scale at which genetic data is analyzed, i.e. the spatial extent of the data considered, can have a strong effect on the outcome of a population or landscape genetic analysis (Gabrielsen *et al.* 2013; Gorospe & Karl 2013; Keller *et al.* 2013). Therefore, in papers IV and V, we took advantage of the high spatial resolution of the dataset and performed the analyses on two spatial scales to learn more about the effect of scale on the results obtained in a population genetic analysis. As for scale, the presence of IBD can present great challenges in the interpretation of results. In paper V, we used the DResD method proposed by Keis *et al.* (2013), which corrects population structure analyses for IBD, to gain insight into its effect on Bayesian assignment analysis and global estimates of population differentiation (F_{ST}).

Material and Methods

Study area

The study area for this thesis encompasses the large parts of the Northern European distribution area of the brown bear, including Sweden, Norway, Finland and western Russia, from 60°-69°N and 12°-59°E. The arctic and boreal landscapes of the study area consist variously of large, forested areas with Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*) and birches (*Betula pendula* and *B. pubescens*), mires, peat land and open tundra. Most of the study area is sparsely populated by humans, generally decreasing from south to north (Ordiz et al. 2014).

Study populations

For paper I, eight areas were sampled in Norway, Sweden, Finland, and western Russia. For paper II, samples from four locations were obtained: Västerbotten in Sweden, Troms in

Norway, the Pasvik valley at the border between Russia, Finland and Norway, and Karelia at the border between south-central Finland and Russia. For paper III, samples of male brown bears from Sweden, Norway, Finland and western Russia were obtained, so that sampling coverage was more or less continuous throughout large parts of the Northern European distribution zone, reaching as far east as Komi, bordering Arkhangelsk in Western Russia. In addition, the prebottleneck population of southern Norway was sampled by collecting material from museum specimens. Papers IV and V concentrated on the western part of the Northern European brown bear population: the study encompassed samples from all of Sweden and Norway.

Sampling

Samples from Sweden and Norway were obtained from regional and national monitoring programs, consisting mainly of noninvasively collected fecal samples. In addition we obtained hair samples and tissue samples from legally shot bears. Fecal sampling was carried out by wildlife managers and volunteers across the region and throughout the activity period of bears, i.e. spring to autumn. Hair samples were obtained with the help of systematic hair trapping in the transborder area of Norway, Finland, and Russia, as well as Eastern Finland during the summer months, following the methods described by Woods *et al.* (1999). The samples from legally shot bears accounted for the majority of samples from Finland. Samples from Russia were collected in collaboration with the research institutes in Karelia, Arkhangelsk, and Komi. For paper III, samples from various tissue types were obtained from 130 museum specimens from Southern Norway. The time of sample collection spanned 2006-2010 for papers I and II, 2006-2012 for the sampling of the extant brown bear population and 1750-1956 for the historical specimens for paper III, and 2006-2013 for papers IV and V. For the historical samples used in paper III, sampling followed the standard protocols used for the genetic sampling of historic material, using a clean room and cleaning of tools and surfaces after each individual sample to avoid cross-contamination and further DNA degradation (Casas-Marce *et al.* 2010).

DNA extraction and amplification

In the course of this thesis, three types of markers were used to obtain the genetic data to be analyzed statistically: autosomal microsatellites (STRs), Y-chromosomal microsatellites (Y-

STRs), and Y-chromosomal single nucleotide polymorphism markers (Y-SNPs). Autosomal STRs are the most commonly used markers in genetic studies and genetically based monitoring programs, due to their generally high polymorphism and expected neutrality to selection. In papers I, II, IV, and V, we applied the STRs that are most commonly used in bear genetic studies, developed for black and brown bear (Paetkau & Strobeck 1994; Paetkau *et al.* 1995; Paetkau & Strobeck 1995; Taberlet *et al.* 1997). For paper III, we used Y-STRs and Y-SNPs (Bidon *et al.* 2014). In contrast to autosomal STRs, these markers are inherited uniparentally, so they are not recombined under sexual reproduction and therefore enable the analysis of sex-specific gene flow for more than just one generation. In addition, Y-STRs possess similar characteristics as autosomal STRs in terms of diversity, mutation rate, and unambiguity of allele designation and are therefore applicable to the study of similar genetic processes (Hurles & Jobling 2001; Kayser *et al.* 2005; Roewer *et al.* 2005). Y-chromosomal markers have been used successfully to study male-specific gene flow and sex-specific population structure in humans (Rubicz *et al.* 2010; Zupan *et al.* 2013; Karmin *et al.* 2015), nonhuman primates (Langergraber *et al.* 2007; Schubert *et al.* 2011), and various other wildlife and domestic species (Meadows *et al.* 2006; Yannic *et al.* 2012; Neaves *et al.* 2013) as well as bears (Bidon *et al.* 2014; Kutschera *et al.* 2014).

The extraction of DNA from and genotyping of the various source materials of extant brown bears followed strict analysis protocols, as they were performed in a lab that is accredited according to the EN ISO/IEC 17025 standard (Norwegian Accreditation: Test 139). This means that the laboratory has a documented quality management system in place, ensuring reliability and reproducibility of the results. Genetic analyses follow the guidelines for the use of nonhuman genetic material in forensic investigations (Linacre *et al.* 2011), thus noninvasively collected samples were genotyped twice if heterozygous and thrice if homozygous for each respective marker and consensus genotyping was not performed. Size calling of the respective allele was only accepted if peak height was higher than a threshold value of 300 relative fluorescent units (RFU) for autosomal STRs and 600 RFU for Y-STRs. The extraction of DNA and pre-PCR set up of the historical samples was performed in a separate laboratory, where brown bear samples or DNA have never been stored or handled, to avoid contamination with contemporary bear DNA. Historic DNA was amplified and genotyped three times, allele size calling was confirmed at least twice.

Testing for reliability of microsatellite markers

For paper I, several tests were performed to test the reliability of the selected autosomal STRs. Sensitivity of markers was tested on two positive control samples by varying the DNA concentration in the template from 30-0.02 ng by diluting the template with sterile water. Species specificity was tested using DNA template originating from nine different wildlife and domesticated species other than brown bear. This test also included samples from humans and polar bears. Between-run precision, stutter ratio, and heterozygote balance, i.e. peak height ratio between alleles, was assessed by performing ≥30 independent runs of amplification and subsequent genotyping of two heterozygote positive controls. The tandem repeat array and the upstream and downstream sequence immediately next to it were assessed by DNA sequencing of each marker. The resulting allelic sequences were inspected manually. One marker, MU26, showed significant deviations from Hardy- Weinberg equilibrium and further tests were made to detect possible null-allele at this locus. This included the design of three new forward and two new reverse primers, aimed to amplify an extended region up- and downstream of the repeat array. We ran selected samples, i.e. samples that were typed positive for all remaining markers but showed no amplicon with MU26, as well as samples that did show positive amplification results to serve as positive controls, with these new primers in various combinations and inspected results on an agarose gel for positive or negative amplification. In addition, eight of the most commonly used markers (MU05, MU09, MU10, MU23, MU50, MU51, MU59, and G10L) were combined for an octa-plex PCR, using DNA extracted from tissue, hair, and fecal samples, which had tested positive previously in a single-plex PCR setting.

Statistical methods

Estimation of population size (N_c), density (D) and effective population size (N_e)

To estimate the annual population size (N_c) for paper II, we applied the single-session capture-mark-recapture (CMR) method based on identification of individuals, based on genetic analyses. For this we used *Capwire* (Miller *et al.* 2005) and CAPTURE (Otis *et al.* 1978), based on the closed-population heterogeneity estimator (M_h Chao). For estimating population density (D), annual effective sampling area was calculated first by creating two differently sized circular buffers around each sampling location, because home ranges sizes for the focal population were not known. In addition, we estimated the mean maximum distance (MMD) (Obbard *et al.* 2010) between resampling event for individuals with more than five resampling

events. We estimated annual effective population sizes using two different approaches: LDN_e (Waples & Do 2008), which is based on linkage disequilibrium (LD) data (Hill 1981), and ONeSAMP (Tallmon *et al.* 2008), which uses approximate Bayesian algorithms.

Assessment of basic population parameters

For paper II, we used the software GENETIX 4.05.2 (Belkhir *et al.* 1996-2004) to estimate expected and observed heterozygosity, allele numbers, inbreeding coefficient (F_{IS}), and LD. We used GENEPOP v 4.0.11 (Rousset 2008) to run an exact test for deviation from the Hardy-Weinberg equilibrium and we tested for the genetic signature of population bottlenecks using BOTTLENECK v. 1.2.202 (Cornuet & Luikart 1997; Luikart *et al.* 1998; Piry *et al.* 1999). For paper I, the software GDA (Lewis & Zaykin 2001) was used to test for Hardy-Weinberg equilibrium, linkage disequilibrium, overall inbreeding coefficient (F_{IS}), and pairwise genetic differentiation (F_{ST}). Micro-Checker (Van Oosterhout *et al.* 2004) was used to test for the presence of null alleles and to discern between erroneous allele frequency calculations caused by null alleles, allele drop-out, or stutter. For paper IV, we used GENEPOP v 4.0.11 to estimate the inbreeding coefficient (F_{IS}), but used GenAlEx 6.5 (Peakall & Smouse 2006, 2012) to calculate the number of alleles, and expected and observed heterozygosity.

Assessment of population structure with autosomal microsatellite data

Most commonly, recent studies assessing population structure use clustering or assignment methods, which have the advantage of determining the most likely number of clusters without making *a priori* assumptions about these (Latch *et al.* 2006). However, caution is needed in the interpretation of results; especially sampling design and the presence of IBD in the sample can strongly influence the reliability of the results, and it is therefore generally advised to employ more than one method (Latch *et al.* 2006; Rowe & Beebee 2007; Schwartz & McKelvey 2008; Frantz *et al.* 2009). Because this is probably true not only for Bayesian assignment methods, but for all methods used in population genetic studies, we used a combination of different methods to shed light on the most likely structure of the analyzed genetic data throughout the papers.

For papers II and IV, we used the Bayesian assignment software STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003; Hubisz *et al.* 2009) to estimate the most likely number of genetic clusters (K) in the genetic data. This software assumes that within populations or

clusters, loci are in Hardy-Weinberg and linkage equilibrium. It then assigns individual genotypes probabilistically to one or more clusters, so that each cluster fulfills the above assumptions. Whereas we used a simple approach in paper II, for paper IV, we used the software in a hierarchical manner, i.e. each identified cluster was analyzed again separately to assess possible substructure within it. In paper II, we also used GENELAND v. 3.2.4 (Guillot *et al.* 2005a) to assess population structure, which works similarly to STRUCTURE, but enables the implementation of an individually based spatially explicit model. As a third means to assess population structure in paper II, we performed a factorial correspondence analysis (FCA) with GENETIX (Belkhir *et al.* 1996-2004) to visualize relative similarity among samples.

For paper IV, we assessed the effect of sex-biased dispersal on genetic population structure. Therefore, we applied the analysis of spatial autocorrelation, which has been shown to detect differences in spatial genetic structure even when the spatial extent of gene flow differs little between the sexes (Banks & Peakall 2012). For the analysis, we used the method implemented in GenAlEx 6.5 (Peakall & Smouse 2006, 2012), which uses the multilocus approach of estimating the genetic distance among individuals developed by Smouse & Peakall (1999).

For paper V, we used two new approaches for the analysis of population structure. The first approach by Hindrikson *et al.* (2013) enables the statistical evaluation of the results returned by a STRUCTURE analysis to identify the spatial extent of the detected clusters. It uses the calculation of the inverse distance weighted average (w=1/dist) of the posterior probability (q) from all samples and for each detected cluster for each grid cell placed over the study area. With the help of bootstrap permutations, grid points or cells can then be classified as having a i) significantly higher, ii) significantly lower, or iii) no different probability of belonging to a particular cluster under the expectations for random spatial structure for the entire population. This method therefore allows the identification of cluster core areas (significantly higher probability of belonging to a cluster), areas significantly out of cluster range (significantly lower probability), as well as areas which belong to the catchment area of a cluster. The second new approach (Keis *et al.* 2013), corrects individual pairwise genetic distances for the influence of IBD. The corrected values are interpolated throughout the study area, based on the midpoints between individuals using distance weighting. Essentially, this method assesses population structure minus the structure caused by IBD and thus enables the identification of corridors, i.e. areas where the individual genetic distance among individuals is

higher than expected under IBD, and barriers, i.e. areas where the individual genetic difference is lower.

In addition to these methods, we used ARLEQUIN v. 3.5.1.2 (Excoffier & Lischer 2010) in papers II and III to calculate pairwise F_{ST} values (Weir & Cockerham 1984) among detected clusters and to conduct an analysis of molecular variance (AMOVA) to assess the distribution of genetic variance within and among groups. In paper IV, pairwise F_{ST} values were calculated using GenAlEx (Peakall & Smouse 2006, 2012), and Genepop v4.0 (Rousset 2008) to estimate the number of effective migrants. For paper II, we used BAYESASS v 1.3 (Wilson & Rannala 2003) to estimate migration among the detected clusters.

Population genetic analysis of Y-chromosomal data

Paper III is based on the application of Y-chromosomal markers to assess male gene flow and the resulting male-specific population structure. Because these data are haplotypic in nature, analysis methods to assess population structure are limited, especially in a continuously distributed population, such as the Northern European brown bear. Most commonly, haplotypic data are analyzed using *a priori* grouping of samples based on sample location (Meadows *et al.* 2006; Langergraber *et al.* 2007; Schubert *et al.* 2011), sometimes followed by a hierarchical analysis of genetic differentiation and subsequent merging of nonsignificantly differing groups (Kayser *et al.* 2005). To account for the continuous nature of the brown bear population, however, we devised a different approach and conducted a cluster analysis based on haplotype frequencies using the function 'hclust' in the program R (Rcoreteam 2013). We then used the results initially to group the data for a subsequent AMOVA, validating the results by performing several runs with different groupings in order to find the grouping that would return the least within-group variation and the largest among-group variation. In addition to this, we estimated pairwise genetic differentiation F_{ST} among sample locations, as well as among the groups resulting from the pooling of sampling locations based on the *hclust* analysis and the AMOVA, using ARLEQUIN v. 3.5.1.2 (Excoffier & Lischer 2010) and viewed the results of these estimations with the help of an analysis of principle coordinates (PCoA) with GenAlEx 6.5 (Peakall & Smouse 2006, 2012).

Results and Discussion

How reliable are the microsatellite markers used to identify individuals and to study the genetic structure of brown bear populations?

In paper I, we assessed the reliability of the microsatellite markers used most commonly in population genetic studies and management of brown bears and evaluated whether these markers are suitable for a forensic DNA profiling system. The tests for species sensitivity included species that can be regarded either as predated on by brown bears or whose hair or scats may be mistaken for those of brown bear. No amplification results were observed for three markers and only weak amplicons at different alleles than observed for brown bears for the rest of the markers for all tested species, except the polar bear. This means that is not likely to confuse genotypes of nonursid origin with those of brown bears, rendering the tested markers ideal for the application in a noninvasive sampling design. Especially in noninvasive sampling, untrained volunteers occasionally might misidentify droppings or hair as being from brown bears when they are not. Also, contamination with human DNA, which may occur during handling of samples in the field and at the laboratory, could be excluded as well. Differentiation between brown and polar bear genotypes is more challenging, as the polar bear samples displayed successful amplification with all markers and at allele size ranges overlapping with those of brown bears. Even though no other bear species were tested, we assume that they also would display positive results.

All markers showed high sensitivity with positive amplification results up to 0.6 ng DNA. Traditional DNA quantification methods are not applicable to fecal samples, as they contain DNA from food and microorganisms, in addition to the DNA of the sampled individuals. To judge the quality and quantity of DNA contained in a sample, however, the observed similarity of marker sensitivity across all tested loci showed that an initial run with two markers should be sufficient. If these positive results can be obtained at these loci, successful amplification with the remaining ones is likely. This may be of considerable importance for the efficiency of a research project or management action based on noninvasively sampled material, where the cost of genetic analysis must be kept at a minimum. In addition, a combination of eight of the most commonly used markers in octa-plex PCR showed similarly high sensitivity as in a single-plex set up, enabling a further increase of cost efficiency.

Twelve markers showed no consistent deviation from Hardy-Weinberg equilibrium in all populations sampled, however, one marker (MU26) did, most likely due to null alleles at this locus. The additional amplification runs with customized primers did not result in detectable amplicons, thus the cause of the null allele remains unknown, although a large deletion of the entire locus may be a possible explanation. MU26 was therefore dropped from the panel of markers used for this study and for brown bear monitoring. Its use in other brown bear studies should probably be evaluated. For the remaining twelve markers, we detected no obvious differences in between-run precision, and allele differing by 2 bp could be separated reliably. For seven of the tested markers, we achieved reliable separation even at a difference of 1 bd. In addition, the total average of probability of identity and sibling identity was $<1.1x10^{-9}$ and $<1.3x10^{-4}$, respectively, indicating reliable individual identification necessary for a DNA profiling system used in forensic applications. Measures of stutter ratios and heterozygote balance were elevated, as expected for dinucleotide microsatellites, and are thus not suitable to identify individual genotypes in mixtures, i.e. source material that contains DNA from more than one individual. However, taking into account the observed stutter ratios and observed peak height ratios, we optimized genotyping protocols, so that experienced analysts could achieve high reliability of genotyping results and minimize errors due to e.g. allele drop out by following the protocols developed during our study.

Sequencing showed that size variation in alleles corresponded generally with variation in repeat array length, even though three loci showed an additional influence of indels, i.e. insertion of deletion of bases in the allele length variation. Therefore, a common nomenclature of allele designation based on number of repeats should be possible with a collaborate effort of laboratories that participate in genetic brown bear research to calibrate genotyping results. This may contribute to a better comparability of different genetic studies conducted in different countries and enable more feasible collaborative, range-wide surveys.

What is the genetic structure, connectivity and diversity of the recovering Northern European brown bear population?

We assessed the population structure of Northern European brown bears at different scales and sampling resolutions, using various subsets of samples across the study area and three different types of markers (Figure 1, papers II, III, IV, V). Throughout our analyses, the results persistently showed considerable genetic structuring in the Northern European brown bear. For

the first paper, we used the brown bears in northeastern Norway living in the border area between Norway, Finland, and Russia as a focal population to investigate gene flow between

Figure 1: Location overview of the brown bear samples used in this thesis. Different subsets of samples where used for the individual studies, with a number of them being used more than once. For paper II, the subset contained samples from 477 individuals from the border area of Norway-Finland-Russia, from northern Norway, from central Sweden and from the south-central border region between Finland and Russia; for paper I, three additional areas were sampled in southern and central Norway and Western Russia and the subset consisted of 479 individual genotypes; for paper III, 443 male bears across the entire study area were sampled (dark gray circles, light gray squares), plus 14 historical genotypes from southern Norway; for paper IV and V, 1531 individuals from Sweden and Norway were sampled, here presented as gray (males) and white circles (females). For a more detailed depiction of the different dataset locations, please refer to the original articles.

eastern and western groups of the Northern European brown bear population. The Bayesian assignment analysis performed with STRUCTURE (Pritchard *et al.* 2000) using samples from this and from three other locations (paper II) clearly indicated four separate clusters (Figure 2a), corroborated by similar results using GENELAND (Guillot *et al.* 2005b) and an analysis of factorial correspondence (paper II). Pairwise F_{ST} values among locations ranged from 0.05-0.120 and estimation of migration and self-recruitment rates ranged from 0.001-0.047 and 0.963-0.990, respectively, indicating limited gene flow among locations, particularly in the east-west direction (paper II). We also estimated population size (N_c), effective population size (N_e), density (D), and N_e/N_c ratio for the focal population, the results of which indicate that it is a stable population with a relatively even sex-ratio (Frankham 1995). This result may also

have relevance for the assessment of genetic connectivity, as an elevated N_e may point to relatively high gene flow into a population, here perhaps from the east, which is also indicated by the relatively high observed heterogeneity (paper II). High gene flow may bias the N_e value upwards, causing the local N_e to approach global N_e (Pray *et al.* 1996; Palstra & Ruzzante 2008; England *et al.* 2010; Luikart *et al.* 2010) and thus enlarging the N_e/N_c ratio.

Sampling design has been shown to inflate the number of detected clusters in Bayesian assignment algorithms (Schwartz & McKelvey 2008; Tucker *et al.* 2014). To address this issue, in the subsequent papers we maximized sampling density to come as close to continuous sampling as possible. For paper IV, we ran another STRUCTURE analysis, this time with a nearly continuous sampling across Sweden and Norway. The analysis identified four clusters, showing again a clear separation between cluster 4 in the northeasternmost area and clusters 1 to 3 in the rest of the Scandinavian Peninsula in the west (Figure 2b), coinciding with those identified in previous population structure analyses (Waits *et al.* 2000; Manel *et al.* 2004).

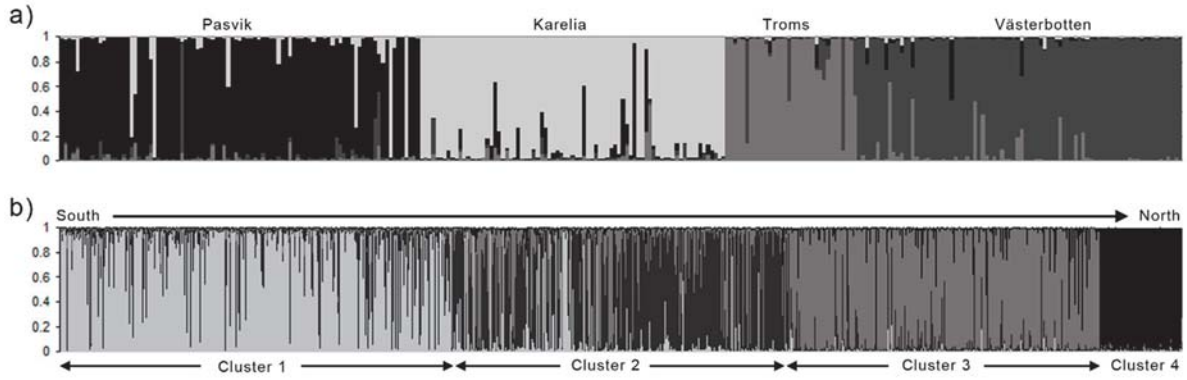


Figure 2: Results of the Bayesian assignment analysis of two different data sets of brown bear genotypes using STRUCTURE (Pritchard *et al.* 2000). a) Barplot showing the assignment probabilities for bear samples originating from four areas across the Northern European distribution zone, collected form 2005-2010 (reproduced from paper II). Each bar represents one bear partitioned into segments according to the membership value q for each of the detected clusters (represented at the y-axis). Sampling area is indicated above the barplot. b) Results of the assignment analysis of brown bear genotypes contained in the databank compiled through the genetic monitoring of bears in Norway and Sweden, originating from 2006-2013 (reproduced from paper IV and V). Samples are ordered from South to North, as indicated above the barplot. Cluster names and their extent are indicated with arrows at the x-axis.

Analysis of genetic diversity within the identified clusters showed high levels of expected and observed heterozygosity across all markers, ranging from 0.715 to 0.779 (H_e) and 0.738 to 0.810 (H_o). The derived values of the inbreeding coefficient F_{IS} were, apart from one, <0.1 and insignificant (paper IV), indicating that the demographic bottleneck had only little effect on autosomal diversity of the Scandinavian brown bear subpopulation, perhaps due to its short duration (Peery *et al.* 2012). Whereas Manel *et al.* (2004) suggested that the substantial

structure found within the Swedish bears could be an echo from the disintegration of the population in the course of the extreme reduction of bear numbers during the bottleneck, a recent study suggested instead historical ecological processes as the cause for the observed structure (Xenikoudakis *et al.* in press). However, both studies employed STRUCTURE to assess population structure, which has been shown to give biased results when IBD is present in the data (Frantz *et al.* 2009), as acknowledged by the authors of the algorithm (Pritchard *et al.* 2010). Recently, Tammeleht *et al.* (2010) proposed that the brown bears in northeastern Europe are structured mainly by IBD, and it can be assumed that this process also has a large influence on the Scandinavian subpopulation, because the distribution of individuals is more or less continuous, mostly without obvious discontinuities. Landscape barriers are difficult to discern, which makes an interpretation of the results challenging (Schwartz & McKelvey 2008). Therefore, in paper V, we used two new approaches to assess the population structure of the Scandinavian subpopulation to account for the potential effect of IBD. First, we ran a cluster placement analysis using the individual posterior probability values (q), estimated in paper IV. The result showed clear cluster core areas for each identified cluster, with signs of asymmetrical gene flow, i.e. higher gene flow from the east towards the west and southwest than the other way around (Figure 3 and paper V). This was indicated by the fact that no out-of-bounds area was identified to the west and southwest of cluster 4, whereas cluster 4 showed a significantly lower probability of belonging to the area for the other three clusters (except for male bears in the northern Swedish cluster) (paper V, appendix S1). The second approach was based on the correction of individual pairwise genetic distances for IBD to identify barriers and corridors to gene flow (Keis *et al.* 2013). We identified two barriers, one separating the southernmost cluster 1 from the cluster north of it (cluster 2), and one separating cluster 4 from the other three on the Scandinavian Peninsula (Figure 5b), thus validating the interpretation of an east-west partition of the Northern European brown bear population in paper II. For the two clusters situated in central and northern Sweden, the results of the DResD analysis did not show any barrier to gene flow between them, thus indicating that the interpretation of the STRUCTURE results of four separate subpopulations should be modified. Because a number of individuals showed very high membership values for cluster 2 and a core area could be identified with the cluster placement analysis, the existence of a separate cluster cannot be excluded. However, there seemed to be a considerable amount of migration of females from cluster 2 to cluster 3, and of

males from cluster 3 to cluster 2 (paper IV), so that, taking IBD into account, the DResD

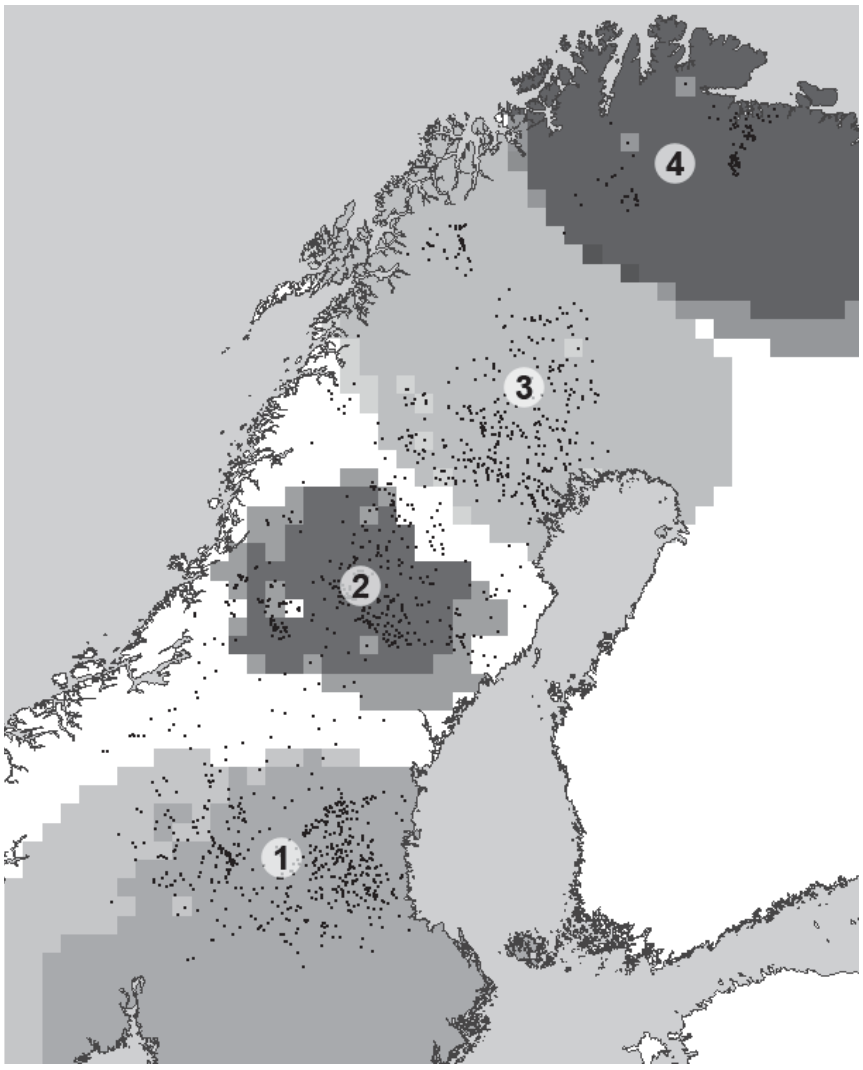


Figure 3: Geographical placement of the cluster core areas based on distance weighted interpolation of assignment probability (q) determined by 100 bootstrap permutations (reproduced from paper V). The darker grid points (25x25km) for each cluster were determined as core areas for males and females, the lighter ones for males only. The clusters are referred to in the text by the numbers given in the figure.

analysis was unable to identify a definite core for cluster 2. This may indicate a slow, ongoing merging of clusters 2 and 3. That the algorithm identified the two barriers in the south and north identified from previous analyses, supports the reliability of the results: the genetic discreetness of the southernmost group has been found by several other studies, using both mtDNA and autosomal markers from historical and contemporary samples (Taberlet & Bouvet 1994; Waits *et al.* 2000; Bray *et al.* 2013; Xenikoudakis *et al.* in press); the barrier in the north has been identified by our previous analyses (paper II, III) as well as another study estimating gene flow from Russia in the Northern European population (Kopatz *et al.* 2014). Our study showed that

a large amount of the genetic structure in the central parts of the Scandinavian population is influenced by strong IBD.

Another piece in the puzzle was provided by the results of our study based on Y-chromosomal markers (paper III). I will present and discuss the results in more detail in the next section, however, they contribute to the issue of the general genetic structuring of the brown bears in Northern Europe and should therefore be mentioned here as well. Y-haplotype diversity and distribution differed substantially between brown bears in the east and west (Figure 4). An AMOVA of the Y-STR data indicated a higher level of structure in the west,

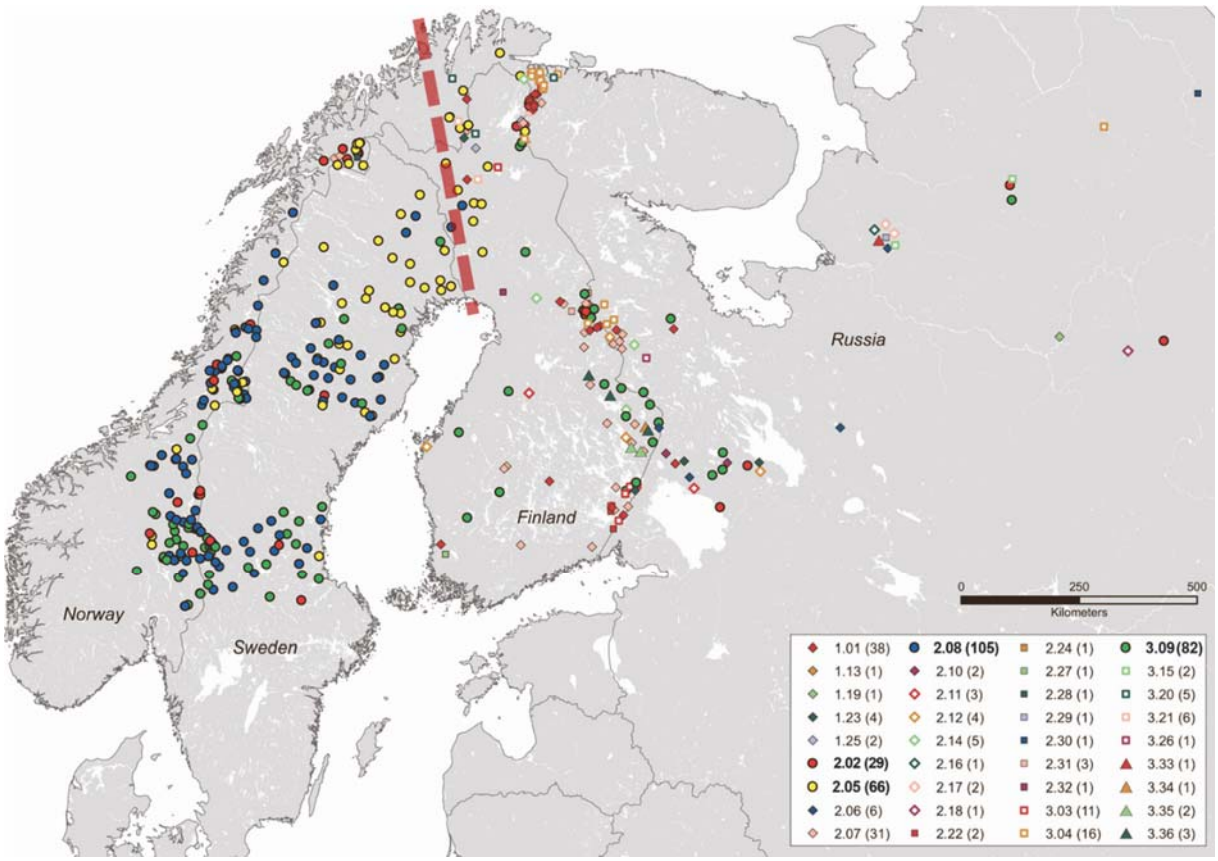


Figure 4: Geographical distribution of the detected Y-chromosome haplotypes among extant male brown bears in Northern Europe (reproduced from paper III). Each sample is depicted with the symbol used for the haplotype to which it was assigned. Haplotypes occurring in Sweden and the southern and centrals parts of Norway (Hedmark/Oppland, Sør-Trøndelag, Nord-Trøndelag and Nordland) are depicted with a circle and indicated with bold characters in the legend. The total frequency of each haplotype is in parentheses. The red dashed line indicates the boundary between the eastern and western subpopulation.

which showed an among-group percentage of variation of 20.4%, compared to 5.7% in the east (paper III). The strong male structure in the west, with Y-haplotype distribution showing comparably distinct clusters coinciding with the clusters found based on autosomal STR

assignment analysis, are not easily reconciled with the results of the sex-specific DResD analysis (paper V). Støen *et al.* (2006) proposed an impact of social mechanisms and higher-than-assumed territoriality for the brown bear, and our results corroborate this conclusion, because male dispersal among clusters has been documented in the monitoring records (81 males on the entire Scandinavian Peninsula, 11 males crossing the most distinct cluster boundary in the northern region, unpublished data). Had these males been successful in producing male offspring, cluster boundaries would not have been as distinct, as a nonlocal haplotype introduced into a new area should spread within a few generations (Ohnishi & Osawa 2014). Female brown bears disperse farther in Scandinavia than in North America (Støen *et al.* 2006) and at least 22 between-region female dispersers have been documented in the monitoring records (unpublished data). In addition, large-scale analysis of pairwise genetic distance to test for IBD showed a relatively small difference between males and females (papers IV and V). We speculate that female gene flow contributes to among-region gene flow to a larger degree than previously assumed, thereby decreasing genetic differentiation, as shown with the DResD analysis (paper V), whereas social mechanisms and/or other external factors, e.g. illegal killing of dispersing individuals, reduces male gene flow among regions, as evidenced by Y-chromosomal structure (paper III).

Based on the combined results of these studies, we suggest that the brown bears in Northern Europe are divided into an eastern and a western subpopulation. The reason for this pattern is unclear, as there are no clear landscape barriers that could explain the limited gene flow between east and west. Habitat structure may play a role, because bears seem to prefer rugged, forested terrain (Nellemann *et al.* 2007) and the open, high-elevation tundra in the boundary area may deter individuals from long-distance dispersal. Studies of polar bears and wolverines (*Gulo gulo*), while not directly comparable because of differences in life history traits and ecological preferences, have shown that unfavorable habitat can have a strong influence on long-distance dispersal (Paetkau *et al.* 1999; Kyle & Strobeck 2001, 2002; Cegelski *et al.* 2006). Another factor may be illegal killing in the reindeer husbandry area, which has been suggested to cause a barrier for wolf (*Canis lupus*) dispersal (Wabakken *et al.* 2001; Vilà *et al.* 2003; Kojola *et al.* 2006a). Several studies have shown that illegal hunting has a strong effect on the large carnivores in Scandinavia and that it is more accepted and more frequent in the north than in the south (Andrén *et al.* 2006; Persson *et al.* 2009; Liberg *et al.* 2011; Gangaas *et al.* 2013; Rauset 2013). However, both of these proposed causes to the

observed east-west division cannot account for the fact of asymmetrical dispersal, as poaching or suboptimal habitat should affect dispersal independent of direction. Also, the northeasternmost cluster of brown bears in Norway is divided again into an eastern group at the border to Russia and a western group, halfway to the population on the Scandinavian Peninsula. Despite similar distances to the brown bears in the east and the west, the brown bears at the center seem to be better connected to the east than to the west and southwest (paper V). A targeted study with the aim to increase sampling density in that particular area may be able to shed more light on this issue.

What is the structure, diversity and distribution of male brown bear lineages across the Northern European distribution zone?

In mammals, dispersal is most often male-biased, so that males, as the dispersing sex, are supposed to be responsible for genetic and demographic connectivity among regions (Lawson Handley & Perrin 2007; Greminger *et al.* 2010). Consistent with the global pattern of more Y-chromosomal admixture and less structure than found in mtDNA-based data, indicating higher migration rates in males than in females (Bidon *et al.* 2014), we expected to find no discernable structure in the distribution of Y-haplotypes across the study area (Figure 1). We detected three different haplogroups (based on the Y-SNP variation) and 5, 20 and 11 different Y-STR haplotypes, respectively, with a total of 36 Y-haplotypes. Instead of a relatively unstructured distribution of both Y haplogroups and -types, we found a division into east and west. There was a markedly different Y-haplotype diversity east and west of the boundary (indicated in figure 4), and one haplogroup occurred only in the eastern subpopulation (Figure 4 and paper III). Only four haplotypes were found in the majority of the sampled groups on the Scandinavian Peninsula, compared to the seven to thirteen haplotypes found in the various eastern locations. Two factors might help to explain this difference. First, the effective population size of sex chromosomes is only 1/4 of that of autosomes, thus making them more vulnerable to the effect of a bottleneck event (Greminger *et al.* 2010). Even though previous studies of autosomal DNA found only weak signs for a genetic bottleneck in Scandinavian brown bears (Waits *et al.* 2000; Kopatz *et al.* 2014), our results showed that the near-extinction event was more decimating for Y-haplotype diversity (paper III). It is possible that the low Y-haplotype diversity was caused mainly by historical ecological processes, e.g. during post-glacial recolonization, as has been proposed to explain the equally low mtDNA variation in this population (Bray *et al.* 2013).

However, we were able to obtain complete to nearly complete haplotypes from historic brown bear samples originating from southern Norway from 1780 to 1920, i.e., the pre- and peribottleneck period (paper III). We found six haplotypes in total, with two of them occurring also in the contemporary dataset. Four of the found haplotypes, however, were no longer present there, strongly indicating that Y-haplotype diversity was higher before the bottleneck event. Xenikoudakis *et al.* (in press) found a similar result, comparing pre- and postbottleneck mtDNA diversity. The second consideration pertains to the relatively high Y-haplotype diversity found in the Finnish groups. The Finnish bears also experienced a severe reduction in population size around the same time the Scandinavian bears did (Swenson *et al.* 1995; Ermala 2003; Danilov 2005), yet Y-haplotype diversity is much higher and comparable to that of Northwestern Russia (paper III). The reason for this may be a high amount of gene flow from the east (Kopatz *et al.* 2012; Kopatz *et al.* 2014). In addition, north-south genetic connectivity seems to be better within Finland than in the west, with the population processes that balance genetic structure acting much faster than predicted by population genetic theory (Hagen *et al.* 2015).

What is the importance of male gene flow in the Northern European brown bear?

Because brown bears show a pronounced male-biased dispersal pattern (Støen *et al.* 2005; Støen *et al.* 2006; Zedrosser *et al.* 2007b), we hypothesized that male dispersal was important in the demographic recovery of the Northern European brown bear population. To test this hypothesis, we exploited the known partition between the western and eastern subpopulation of the Northern Europe brown bear. Previous studies have shown that immigration patterns are different between east and west, with the Finnish group receiving high levels of immigration from farther east, whereas immigration into the Scandinavian subpopulation is comparably low. If it were true that among-region male gene flow was important in the recovery process, we would expect to find equally low levels of Y-chromosomal population structure, despite the differential recovery histories. However, as already discussed above, this is not what we found. Viewing pairwise F_{ST}-values with a PCoA showed the east-west partition clearly, as well as a stronger differentiation among western than eastern groups (paper III). Also the AMOVA showed that most of the among-group variation could be attributed to the substructure in the west (paper III). The different recovery histories, low versus high immigration, were thus clearly associated with different degrees of spatial genetic structuring and Y-haplotypic diversity (paper III). We infer from this that range-wide male dispersal, and thus gene flow,

was not a determinant for the successful demographic recovery, but may have had a large impact on genetic recovery.

The relative importance of male gene flow can also be evaluated using a population genetic analysis of autosomal markers. We did this by splitting the genotypic data for paper IV and V according to sex and analyzing them separately. In contrast to the Y-chromosomal markers, though, the genotype of each individual is a mixture of both parents, so that male-specific gene flow is nearly impossible to track without a thorough pedigree analysis. However, contrasting the results based on Y-chromosomal data with those obtained by the analysis of autosomal data may shed new light on the issue of male dispersal and gene flow. The result of the hierarchical structure analysis (paper IV) and the cluster placement analysis (paper V) revealed surprisingly small differences in the spatial genetic structure between males and females. Even though males have been shown to disperse much farther than females (Støen *et al.* 2006), thus making it reasonable to expect no, or at least much less, discernable population structure in males than in female bears, cluster placement analysis revealed similar spatial cohesiveness of each identified cluster between the sexes. Without any further analysis, this may be interpreted as males not having a large impact on balancing population structure. However, substracting the influence of IBD on the observed structure with the help of the DResD analysis showed that, on the large scale (150-300 km distance between pairs of analyzed individuals), males seem to breach the barriers exhibited by the females (paper V) thus contributing to gene flow among regions. It is important to remember that the DResD results do not show that there is no structure within the male brown bear population, but that the observed structure is likely to be caused by limits to dispersal, i.e. IBD. This is remarkable, as the scale on which the data were analyzed was smaller than the observed maximum dispersal distance (467 km, Støen *et al.* 2006) and close to the average dispersal distance of 119 km (Støen *et al.* 2006).

Perhaps these results reflect relatively slow processes that function towards balancing population structure. However, it is puzzling why the pace at which these processes act would be so much slower in the Scandinavian than in the Finnish subpopulation. In the latter, an increase of admixture, and thus a blurring of cluster boundaries and decreasing structure, happened at a rapid pace (Hagen *et al.* 2015). The number of effective migrants per generation between the two identified clusters in the north and south of the country was estimated to have increased from 1.60 to 3.63 over the course of 14 years, with a mean across the years of 2.21

effective migrants (Hagen et al. 2015). Between Finland and Russian Karelia, the number of effective migrants was even higher, 7.64 (Kopatz *et al.* 2014). In contrast, the estimated number of effective male migrants among clusters in the Scandinavian subpopulation ranged from 0.168-1.114 (paper IV).

Do genetic analyses reflect the observed male-biased dispersal and female philopatry in the brown bear?

The results of the hierarchical STRUCTURE analysis (paper IV), estimation of pairwise F_{ST} among detected clusters and subclusters (paper IV), cluster placement analysis (paper V), and small-scale DResD analysis (paper V) revealed only small differences between male and female population structure. Also a study-area-wide analysis of the relationship between spatial and genetic distance (paper IV and V) showed only little difference between the sexes. However, the results of the spatial autocorrelation analysis showed significant differences between male and female patterns of relatedness on the local scale, i.e., <40 km (Figure 5), reflecting the

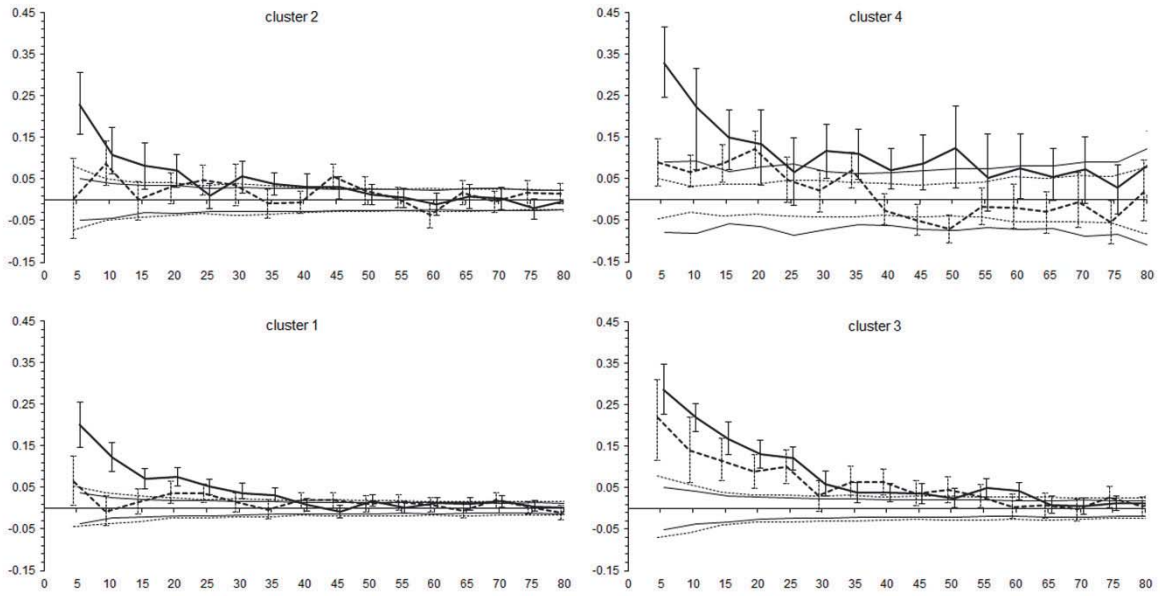


Figure 5: Spatial autocorrelation of female versus male brown bears in Scandinavia compared within each cluster, distance class of 5 km (reproduced from paper IV). The genetic correlation coefficient (r) is given as a solid line for females and a dashed line for males. The 95% confidence intervals for the null hypothesis of random distribution of genotypes, as well as bootstrap errors, are displayed in the same manner. Clusters are numbered from the south to the north, their geographic location is shown in figure 3.

observed patterns of female philopatry and male-biased natal dispersal (Støen *et al.* 2006; Zedrosser *et al.* 2007b) and coinciding with previous findings of overlapping home ranges

among related females (Støen *et al.* 2005). Such scale-dependent differences in the magnitude of bias have been observed in other systems, including humans (Wilder *et al.* 2004; Wilkins & Marlowe 2006; Heyer *et al.* 2012) and other species (Gauffre *et al.* 2009; Perez-Espona *et al.* 2010; Vangestel *et al.* 2013). Usually these differences are attributed to cost/benefit arithmetic on the small versus large scale, as e.g., the costs of dispersal may increase with distance (Perez-Espona *et al.* 2010) and it is likely that different mechanisms are responsible for the evolution of short- and long-distance dispersal (Lawson Handley & Perrin 2007; Ronce 2007). Furthermore, positive spatial autocorrelation seemed to increase towards the north for both sexes, with a surprisingly large magnitude of positive male spatial autocorrelation found in the northern Swedish cluster (cluster 4), similar to that of females in the south (cluster 1 and 2) (Figure 4). It is likely that this pattern was due to a comparably strong substructure within cluster 4, thus elevating the genetic correlation coefficient *r*, because an analysis using the multipopulation approach (Peakall *et al.* 2003; Banks & Peakall 2012) resulted in somewhat lower values (paper IV). However, because the detected subclusters overlap to a large degree, geographic distance cannot explain the relatively high levels of genetic differentiation among them (paper IV). Instead, social barriers may become increasingly strong towards the north, which would impede the successful immigration of nonlocal males into nonnatal habitat (Schulte *et al.* 2013). Another possible explanation is the impact of illegal killing. If a large percentage of dispersing males are shot, gene flow would naturally be hampered. Several studies have pointed to the fact that the rate of illegal killing is stronger in northern than southern Scandinavia and that its effect on large carnivore populations is substantial (Andrén *et al.* 2006; Persson *et al.* 2009; Liberg *et al.* 2011). Støen *et al.* (2006) documented shorter dispersal distances in the north than in the south, and attributed this to a higher rate of illegal killing in the north. The small-scale DResD analysis corroborated this interpretation, by showing areas of higher genetic similarity than expected under IBD, i.e. family clusters, as well as local patches of lower genetic similarity, i.e. barriers to gene flow on the local scale (paper V). Also Manel *et al.* (2004) and Waits *et al.* (2000) found stronger structure in the northern parts of Sweden. This was attributed to the formation of matrilineal assemblages and the existence of one dominant male. Our results do indeed show the existence of family clusters for females in the northern and southern areas (paper V). However, we find indications of this also for the males to larger extent than seems plausible under the scenario of only occasionally occurring dominant males. This may suggest a more flexible dispersal behavior, particularly for males,

than previously documented, possibly exacerbated by a strong impact of illegal killing of dispersing males.

What influence do scale and isolation-by-distance have on the results of population genetic analyses?

To assess the effect of scale on the analysis of population genetic structure, we have analyzed the Scandinavian brown bear subpopulation at two different spatial scales, using a combination of different methods. In paper IV, we compared the results of using the Loiselle kinship coefficient implemented in SPAGeDI v.1.4c (Hardy & Vekemans 2002) to infer the relationship between spatial and genetic distance on a large scale (on which it is usually employed as a test for IBD) and a small scale, i.e. only sample pairs within one cluster were included in the analysis. These cluster-wide estimates also were compared with the results of the spatial autocorrelation analysis and showed similar trends, albeit with lower values especially at the smaller distance intervals (paper IV). This difference in values was expected, of course, because the two methods use different algorithms to estimate the genetic relationship coefficient. Also the the DResD analyses showed strong differences between the two scales of analysis (Figure 6 and paper V). To explain these differences, one has to invoke the correlation between spatial and temporal scale of population genetic processes, i.e. on a small scale they may operate on a much faster rate than at the large scale. For example, natal dispersal patterns and territoriality may have a much larger impact on genetic structure on a much more immediate time frame, whereas at the large scale, rare long-distance dispersal events and long-distance gene flow operate on a slower rate, and are relatively more important than at the small scale.

Our results also show that considering the scale at which the data are analyzed potentially has a large impact on the analysis outcome and what is actually measured. This has received some attention recently in the landscape genetic community as well, with studies showing that finding a good fit between data and model was influenced by the spatial scale of the data considered (Keller *et al.* 2013; Razgour *et al.* 2014). Another issue that pertains to the importance of scale is that of up- or downscaling results. It is well known in ecology that up-scaling small-scale results to the large scale can lead to erroneous conclusions (Underwood *et al.* 2005). Our results demonstrate that also down-scaling can be problematic; on the large scale, male brown bears exhibit mostly a pattern of structure largely influenced by IBD. However, on the small scale, our analyses revealed an unexpected amount of local structure that was in contrast to their documented high level of mobility (Støen *et al.* 2006; Zedrosser *et al.* 2007).

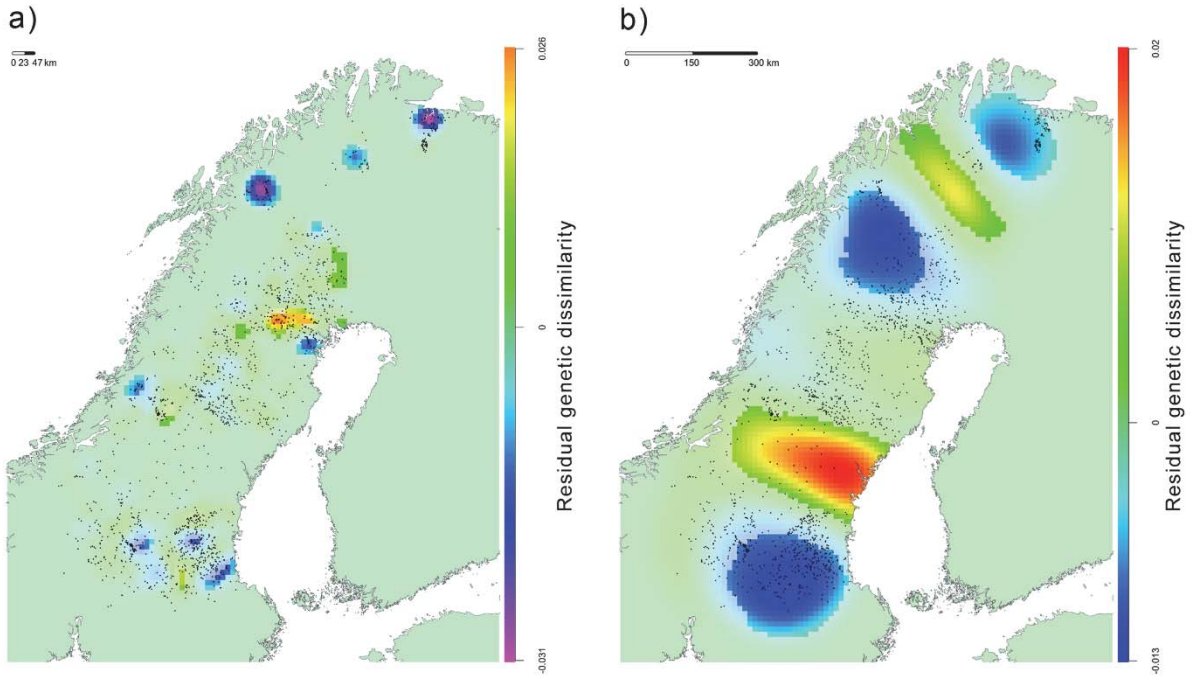


Figure 6: Areas of high and low genetic differentiation between sample pairs of brown bears in Norway and Sweden, estimated with the spatially explicit DresD method at two different scales, corrected for isolation by distance and interpolated across the study area using inverse distance weighting (reproduced from paper V). The black and white scale bar indicates the distance ranges of samples pairs included in the analyses. Saturated colored grid points indicate significant departure from IBD expectations, magenta to blue indicating lower, and green to red higher than expected genetic dissimilarity; samples are represented as black dots. a) Local scale, males and females combined, including sample pairs between 23 to 47 km apart (n=26,207); b) Population scale, males and females combined, including sample pairs between 150 to 300 km apart (n=243,373).

An important aspect of population genetic analysis is the influence of IBD on the performance of analytical tools (Schwartz & McKelvey 2008; Frantz *et al.* 2009; Pritchard *et al.* 2010). The STRUCTURE analysis we performed initially in paper IV identified four

clusters, thus a result similar to that of previous studies (Waits *et al.* 2000; Manel *et al.* 2004; Norman *et al.* 2013). The subsequent reassessment of each cluster revealed an increase in population structure towards the north, with a substantial amount of individuals in the northernmost region showing high membership values to the respective clusters (paper IV). Also the estimation of pairwise F_{ST} among clusters and subclusters showed increasing differentiation in the same direction (paper IV). These results would seem to support the interpretation of the reliability of at least the initial result of four main clusters, but stand in contrast to the DResD results (paper V), which only supported the existence of three clusters. It seems that F_{ST} values are not necessarily helpful for deciding which cluster divisions represent real barriers and which are an artefact caused by IBD. Especially the comparison of the similarity of estimated pairwise F_{ST} values between clusters 3 and 4 for males (0.048) and females (0.054) and for females between clusters 1 and 2 (0.050) (paper IV) illustrates this well. A gene flow barrier was identified with the DResD analysis in both instances for the females but not for the males, yet the F_{ST} value estimated for the males was similar to that for the females. Similarly it seems that the high F_{ST} values estimated among the subclusters identified in the reassessment of within-cluster structure were caused by strong family structure, rather than by limited gene flow among them (paper V).

Management implications and future perspective

Recently Norway and Sweden have agreed on synchronizing monitoring methodology for brown bears (Rovdata 2014). This is a laudable step towards a better comparability of the results and a right step in acknowledging the transboundary nature of the Northern European brown bear population. The next step to follow should be an agreement about the joint management of this population to ensure the establishment of a self-sustaining population (Servheen *et al.* 1999; Blanco 2012, Bischof *et al.* in press). The studies in this thesis have highlighted areas where management should be conducted in coordination between the two countries. One such area is the border zone between the cluster 1 and cluster 2 (Figure 3) in southern Scandinavia. Whereas communication between these clusters seemed severely limited on the Swedish side, the Norwegian side seemed to function as a corridor, especially for the males, enabling gene flow between clusters, albeit apparently at a low level. Another area of interest is in the northwest (cluster 3, Figure 3), along the border between Sweden and Norway. On the Norwegian side, the subpopulation seems to receive male immigrants from farther east, whereas

on the Swedish side, males seem to move east rather than the other way around, indicated by the distribution of Y-haplotypes (Figure 4). Synchronizing management decisions, such as harvest levels between Norway, Finland and Sweden, may increase connectivity between east and west.

The discovery of a limit to gene flow between the east and west of the Northern European brown bear population brings new questions, both regarding the management but also for future research projects. Even though the area between the Bay of Bothnia and the Barents Sea may contain regions where forest cover is lacking and thus may be partly suboptimal brown bear habitat, there are also large forested areas with only a few roads or other anthropogenic disturbances. Brown bears do live in this area and should be able to move around relatively freely, as human density in the north is much lower than in the south. Also, the area between the two seas in not particularly narrow, in fact it covers a similar number of kilometers from south to north as the border area between Finland and Russian Karelia. However, genetic connectivity is much higher between the latter than the former. Are the reasons for this difference in migration and gene flow related to differences in anthropogenic disturbance or can it be explained by differences in landscape and habitat suitability?

Related to this, one of the most striking results of the studies conducted for this thesis is the surprising amount of structure discovered in male bears and the contrasting differences in speed of population genetic processes between the eastern and western subpopulation of the Northern European brown bears. With males being considered to be highly mobile and the species being distributed almost continuously throughout its Northern European habitat, the structure found in our studies, using different analysis methods and genetic markers showed that one cannot easily transfer knowledge about dispersal behavior to assumptions about gene flow. Here, targeted research, combining information based on Y-chromosomal data and autosomal markers with a detailed pedigree analyses may shed more light on the conditions associated with successful male gene flow among regions and may help to resolve the contrasting results of autosomal and Y-chromosomal data for bears in Scandinavia.

Acknowledgements

Finally, the time has come that I defend my PhD. The road that led to this point was long and definitely not straight, with some rocky bits in it, and unexpected turns. But even though it took me a bit longer than anticipated (back then, when I had just had received my Diploma), I am glad the road took me these places I never planned on going to and brought me where I never thought I would end up. Along the way, many people made significant contributions, and I would like to express my gratitude to them.

Firstly, I would like to thank my supervisors, Snorre B. Hagen, Hans Geir Eiken and Jon E. Swenson, who guided and supported me. Hans Geir, besides all the fruitful discussions we had over the years and all the advice you gave, I would like to thank you especially for taking this crazy idea from Alex, that I should apply for the PhD position, seriously. The opportunity came at a time when I was struggling about what to do in the future, and I have never regretted the decision to apply. Jon, I would like to thank you for welcoming me and supporting me and for trying (and succeeding) to make the spatial distance between us seem not so relevant to my life as an external PhD student. And last but not least, I would like to thank you, Snorre, for your never waning enthusiasm to discuss our studies with me, your always open door and for your professional and personal support.

This is also a good point, to extent my gratitude to Ole Wiggo Røstad Kari Elise Moxnes, Grethe Delbeck and Kari Margrete Thue at NMBU. Although I am sure everybody who needed their help has been met with the same friendliness and efficiency as I have, but not being present at the campus, I have to come to regard their help as a lifesaver for me. Whenever I had a question, whenever I needed help with the formalities that come with being a PhD student, they responded quickly and warmly and let the many kilometers between Ås and Svanhovd disappear. Thank you for that!

Doing Science is rarely an activity one does alone and often the best studies are those done in collaboration. So here, I would like to express my thanks to my co-authors, who, just like the brown bear, were distributed over quite a large part of Northern Europe and thus embody perfectly the transboundary nature of my study species. You all accompanied me for a piece of the way, helped me over bumps, provided and analyzed data, gave valuable comments and thus took an important part in me reaching my goal. Thank you all for this collaboration and I hope to have the chance to conduct some research with you again!

Living north of the polar circle has its very special charm and I am thoroughly enjoying it, but this is not only due to the fantastic nature around me, the Northern Lights in the winter and the long, endless day in summer (although that certainly helps). A huge part of why I love to life here, are the people I work with. I would especially like to thank Siv Grethe Aarnes, who is not only a fantastic co-worker, whose professional capabilities have added significantly to my PhD, but also a good friend, who always has an open ear for me. Even though I did some of the labwork for this thesis myself, a lot of my data has been handed to me on a silver platter, thanks to the amazing work by Camilla Tobiassen, Ida Fløystad, Bård Spachmo and Beate Banken Bakke. Beate, I also would like to thank you for spending so much time with me improving my Norwegian; I am very sure that being office mate with you decidedly sped up my learning. Up here, there are only a handful of people, and the relationship one has with them is comparably crucial for making one's life an enjoyable one. Luckily, the people at Svanhovd are a special bunch. From the day I arrived, everybody made me feel welcomed and tremendously eased the transition into the new habitat and for that I would like to thank you all sincerely!

Family and friends are an important part of life, giving support, reminding one that life is about more than just work. So as for the friends, thank you, Wiebke Lammers, for being one of the most inspiring and courageous women I know and can call a friend. And thank you Eun-Joo and Thomas Park, and Ralf Steinmetz for not losing touch, even though we moved so far away from you. Also, I am very happy about the lucky fluke that Snorre and Victoria decided to move to Svanvik. Not only because Snorre turned out to be a fantastic supervisor for me, but also because we have become something like a mini-Bullerby (Bakkebygrenda for the Norwegians), with a common vegetable garden (although lately it has been mostly Victoria doing the work, which will hopefully change now and I can join in the garden work more often) and our children running in and out of our houses. Also thank you, Victoria, for introducing me to and sharing the passion for our post-apocalyptical life skill: KNITTING! What better way to spend the polar night and long winter? Plus, it is a good way to relax your brain when you have been puzzling about the latest analysis results.

As for family, I am grateful to my parents, Hans and Monika, who never expected me to be anything particular and let me explore freely what I wanted to be. Even though some of my choices must have seemed a bit weird or risky to them, there was never any doubt that when I really needed help, they would be there. That, I think has made me more courageous in

following where life seemed to take me. A huge thanks needs to go to my sister, Christina, who always had my back! Also, I would like to thank my parents-in-law, Horst and Hilkka, and my brother-in-law Andreas and his family for their support whenever needed. Finally, and most importantly, I would like to express my deepest gratitude to my dear Alex, who journeyed together with me. It is perhaps not a wise decision for both partners to start their PhD AND start a family halfway, but we made it through intact and wiser than before and that is no small feat. This thesis is dedicated to him, and to my two wonderful boys, Lasse and Hugo.

References

Aarnes SG, Fløystad I, Brøseth H*, et al.* (2015) Populasjonsovervåking av brunbjørn-DNA-analyse av prøver innsamlet i Norge i 2014. *Bioforsk rapport* **10**, 1-62.

Anderson CD, Epperson BK, Fortin MJ*, et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology* **19**, 3565-3575.

Andrén H, Linnell JDC, Liberg O*, et al.* (2006) Survival rates and causes of mortality in Eurasian lynx (*Lynx lynx*) in multi-use landscapes. *Biological Conservation* **131**, 23-32.

Banks SC, Peakall R (2012) Genetic spatial autocorrelation can readily detect sex-biased dispersal. *Molecular Ecology* **21**, 2092-2105.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) *GENETIX 4.05, logiciel sous Windows pour la génétique des populations.* Laboratoire Génome, Populations, Interactions,CNRS UMR 5171, Université de Montpellier II, Montpellier.

Bellemain E, Taberlet P (2004) Improved non-invasive genotyping method: application to brown bear (*Ursus arctos*) feaces. *Molecular Ecology Notes* **3**, 519-522.

Bellemain E, Zedrosser A, Manel S*, et al.* (2006) The dilemma of female mate selection in the brown bear, a species with sexually selected infanticide. *Proceedings of the Royal Society of London Series B-Biological Sciences* **273**, 283-291.

Bidon T, Janke A, Fain SR*, et al.* (2014) Brown and Polar Bear Y Chromosomes Reveal Extensive Male-Biased Gene Flow within Brother Lineages. *Molecular Biology and Evolution* **31**, 1353-1363.

Bischof, R., H. Brøseth and O. Gimenez. (*In press*) Large carnivores in a world divided by political borders: focus on jursidiction leads to inflated abundance estimates. *Conservation Letters*. doi: 10.1111/conl.12183

Bischof R, Fujita R, Zedrosser A, Söderberg A, Swenson JE (2008a) Hunting Patterns, Ban on Baiting, and Harvest Demographics of Brown Bears in Sweden. *Journal of Wildlife Management* **72**, 79-88.

Bischof R, Mysterud A, Swenson JE (2008b) Should hunting mortality mimic the patterns of natural mortality? *Biological Letters* **4**, 307-310.

Blanco JC (2012) Towards a population level approach for the management of large carnivores in Europe. Challenges and opportunities. European Comission.

Bray SCE, Austin JJ, Metcalf JL*, et al.* (2013) Ancient DNA identifies post-glacial recolonisation, not recent bottlenecks, as the primary driver of contemporary mtDNA phylogeography and diversity in Scandinavian brown bears. *Diversity and Distributions* **19**, 245-256.

Brooke ZM, Bielby J, Nambiar K, Carbone C (2014) Correlates of Research Effort in Carnivores: Body Size, Range Size and Diet Matter. *Plos One* **9**.

Casas-Marce M, Revilla E, Godoy JA (2010) Searching for DNA in museum specimens: a comparison of sources in a mammal species. *Molecular Ecology Resources* **10**, 502-507.

Cegelski CC, Waits LP, Anderson NJ*, et al.* (2006) Genetic diversity and population structure of wolverine (*Gulo gulo*) populations at the southern edge of their current distribution in North America with implications for genetic viability. *Conservation Genetics* **7**, 197-211.

Chapron G, Kaczensky P, Linnell JDC, *et al.* (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* 346, 1517-1519.

Chestin I (1992) The brown bear (*Ursus arctos* L.) in the USSR: numbers, hunting and systematics. *Annales Zoologici Fennici* **29**, 57-68.

Cornuet JM, Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Cushman SA, Landguth EL (2010) Scale dependent inference in landscape genetics. *Landscape Ecology* **25**, 967-979.

Dalerum F, Cameron EZ, Kunkel K, Somers MJ (2009) Diversity and depletions in continental carnivore guilds: implications for prioritizing global carnivore conservation. *Biological Letters* **5**, 35-38.

Danilov P (1994) The brown bear in Northwest Russia. *International Conference of Bear Research and Management* **9**, 199-203.

Danilov PI (2005) *Game animals of Karelia - Ecology, resources, management, protection.* Nauka, Moscow.

de Knegt HJ, van Langevelde F, Coughenour MB*, et al.* (2010) Spatial autocorrelation and the scaling of species-environment relationships. *Ecology* **91**, 2455-2465.

Devillard S, Allaine D, Gaillard JM, Pontier D (2004) Does social complexity lead to sex-biased dispersal in polygynous mammals? A test on ground-dwelling sciurids. *Behavioral Ecology* **15**, 83-87.

Driscoll DA, Banks SC, Barton PS*, et al.* (2014) The Trajectory of Dispersal Research in Conservation Biology. Systematic Review. *Plos One* **9**, 18. doi: 10.1371/journal.pone.0095053

England PR, Luikart G, Waples RS (2010) Early detection of population fragmentation using linkage disequilibrium estimation of effective population size. *Conservation Genetics* **11**, 2425-2430.

Enserink M, Vogel G (2006) Wildlife conservation - The carnivore comeback. *Science* **314**, 746-749.

Eriksson M, Sandström C, Ericsson G (2015) Direct experience and attitude change towards bears and wolves. *Wildlife Biology* **21**, 131-137.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Frankham R (1995) Effective population-size adult-population size ratios in wildlife - a review. *Genetical Research* **66**, 95-107.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology* **46**, 493-505.

Gabrielsen CG, Kovach AI, Babbitt KJ, McDowell WH (2013) Limited effects of suburbanization on the genetic structure of an abundant vernal pool-breeding amphibian. *Conservation Genetics* **14**, 1083-1097.

Galpern P, Manseau M, Wilson P (2012) Grains of connectivity: analysis at multiple spatial scales in landscape genetics. *Molecular Ecology* **21**, 3996-4009.

Gangaas KE, Kaltenborn BP, Andreassen HP (2013) Geo-spatial aspects of acceptance of illegal hunting of large carnivores in Scandinavia. *PlosOne* **8**. doi: 10.1371/journal.pone.0068849

Gauffre B, Petit E, Brodier S, Bretagnolle V, Cosson JF (2009) Sex-biased dispersal patterns depend on the spatial scale in a social rodent. *Proceedings of the Royal Society B-Biological Sciences* **276**, 3487-3494.

Gorospe KD, Karl SA (2013) Genetic relatedness does not retain spatial pattern across multiple spatial scales: dispersal and colonization in the coral, *Pocillopora damicornis*. *Molecular Ecology* **22**, 3721-3736.

Graham K, Beckerman AP, Thirgood S (2005) Human-predator-prey conflicts: ecological correlates, prey losses and patterns of management. *Biological Conservation* **122**, 159-171.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28**, 1140-1162.

Greminger MP, Krutzen M, Schelling C, Pienkowska-Schelling A, Wandeler P (2010) The quest for Y-chromosomal markers - methodological strategies for mammalian non-model organisms. *Molecular Ecology Resources* **10**, 409-420.

Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics* **170**, 1261-1280.

Guillot G, Mortier F, Estoup A (2005b) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes* **5**, 712-715.

Hagen SB, Kopatz A, Aspi J, Kojola I, Eiken HG (2015) Evidence of rapid change in genetic structure and diversity during range expansion in a large terrestrial carnivore. *Proceedings of the Royal Society B-Biological Sciences* **282**, 20150092.

Hardouin LA, Nevoux M, Robert A*, et al.* (2012) Determinants and costs of natal dispersal in a lekking species. *Oikos* **121**, 804-812.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Hedrick PW, Peterson RO, Vucetich LM, Adams JR, Vucetich JA (2014) Genetic rescue in Isle Royale wolves: genetic analysis and the collapse of the population. *Conservation Genetics* **15**, 1111-1121.

Heisler LM, Somers CM, Wellicome TI, Poulin RG (2013) Landscape-scale features affecting small mammal assemblages on the northern Great Plains of North America. *Journal of Mammalogy* **94**, 1059-1067.

Heyer E, Chaix R, Pavard S, Austerlitz F (2012) Sex-specific demographic behaviours that shape human genomic variation. *Molecular Ecology* **21**, 597-612.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetic Resources* **38**, 209-216.

Hindrikson M, Remm J, Maennil P*, et al.* (2013) Spatial Genetic Analyses Reveal Cryptic Population Structure and Migration Patterns in a Continuously Harvested Grey Wolf (*Canis lupus*) Population in North-Eastern Europe. *Plos One* **8**. doi: 10.1371/journal.pone.0075765

Hogg JT, Forbes SH, Steele BM, Luikart G (2006) Genetic rescue of an insular population of large mammals. *Proceedings of the Royal Society B-Biological Sciences* **273**, 1491-1499.

Hostetler JA, Onorato DP, Jansen D, Oli MK (2013) A cat's tale: the impact of genetic restoration on Florida panther population dynamics and persistence. *Journal of Animal Ecology* **82**, 608-620.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**, 1322-1332.

Hurles ME, Jobling MA (2001) Haploid chromosomes in molecular ecology: lessons from the human Y. *Molecular Ecology* **10**, 1599-1613.

Johnson ML, Gaines MS (1990) Evolution of dispersal - theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21**, 449-480.

Karmin M, Saag L, Vicente M*, et al.* (2015) A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research* **25**, 459-466.

Kayser M, Lao O, Anslinger K*, et al.* (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. *Human Genetics* **117**, 428-443.

Keis M, Remm J, Ho SYW*, et al.* (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *Journal of Biogeography* **40**, 915-927.

Keller D, Holderegger R, van Strien MJ (2013) Spatial scale affects landscape genetic analysis of a wetland grasshopper. *Molecular Ecology* **22**, 2467-2482.

Kindberg J, Swenson JE, Ericsson G*, et al.* (2011) Estimating population size and trends of the Swedish brown bear Ursus arctos population. *Wildlife Biology* **17**, 114-123.

Kitanishi S, Yamamoto T (2015) Comparison of genetic structure between juvenile and adult masu salmon indicates relatively low reproductive success of dispersers. *Environmental Biology of Fishes* **98**, 405-411.

Kojola I, Aspi J, Hakala A*, et al.* (2006a) Dispersal in an expanding wolf population in Finland. *Journal of Mammalogy* **87**, 281-286.

Kojola I, Hallikainen V, Nygren T, Pesonen M, Ruusila V (2006b) Recent trends and harvest in Finland's brown bear population. *Ursus* **17**, 159-164.

Kopatz A, Eiken HG, Aspi J*, et al.* (2014) Admixture and Gene Flow from Russia in the Recovering Northern European Brown Bear (*Ursus arctos*). *Plos One* **9**. doi: 10.1371/journal.pone.0097558

Kopatz A, Eiken HG, Hagen SB*, et al.* (2012) Connectivity and population subdivision at the fringe of a large brown bear (*Ursus arctos*) population in North Western Europe. *Conservation Genetics* **13**, 681-692.

Kutschera VE, Bidon T, Hailer F*, et al.* (2014) Bears in a Forest of Gene Trees: Phylogenetic Inference Is Complicated by Incomplete Lineage Sorting and Gene Flow. *Molecular Biology and Evolution* **31**, 2004-2017.

Kyle CJ, Strobeck C (2001) Genetic structure of North American wolverine (*Gulo gulo*) populations. *Molecular Ecology* **10**, 337-347.

Kyle CJ, Strobeck C (2002) Connectivity of peripheral and core populations of North American wolverines. *Journal of Mammalogy* **83**, 1141-1150.

Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**, 4179-4191.

Landguth EL, Schwartz MK (2014) Evaluating sample allocation and effort in detecting population differentiation for discrete and continuously distributed individuals. *Conservation Genetics* **15**, 981-992.

Langergraber KE, Siedel H, Mitani JC*, et al.* (2007) The Genetic Signature of Sex-Biased Migration in Patrilocal Chimpanzees and Humans. *Plos One* **2**. doi: 10.1371/journal.pone.0000973

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* **7**, 295-302.

Lawson Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology* **16**, 1559-1578.

Legagneux P, Inchausti P, Bourguemestre F, Latraube F, Bretagnolle V (2009) Effect of predation risk, body size, and habitat characteristics on emigration decisions in mallards. *Behavioral Ecology* **20**, 186-194.

Lewis PO, Zaykin D (2001) Genetic Data analysis: computer program for the analysis of allelic data.

Liberg O, Andren H, Pedersen HC*, et al.* (2005) Severe inbreeding depression in a wild wolf (*Canis lupus*) population. *Biological Letters* **1**, 17-20.

Liberg O, Chapron G, Wabakken P*, et al.* (2011) Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. *Proceedings of the Royal Society of London Series B-Biological Sciences* **279**, 910-915.

Linacre A, Gusmao L, Hecht W*, et al.* (2011) ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Science International-Genetics* **5**, 501-505.

Long ES, Diefenbach DR, Rosenberry CS, Wallingford BD, Grund MRD (2005) Forest cover influences dispersal distance of white-tailed deer. *Journal of Mammalogy* **86**, 623-629.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038-3051.

Luikart G, Allendorf FW, Cornuet J-M, Sherwin WB (1998) Distortion of Allele Frequency Distributions Provides a Test for Recent Population Bottlenecks. *Journal of Heredity* **89**, 238-247.

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

Manel S, Bellemain E, Swenson JE, Francois O (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Molecular Ecology* **13**, 1327-1331.

Mayor SJ, Schaefer JA (2005) The many faces of population density. *Oecologia* **145**, 276-281.

Mayor SJ, Schaefer JA, Schneider DC, Mahoney SP (2009) The spatial structure of habitat selection: A caribou's-eye-view. *Acta Oecologica-International Journal of Ecology* **35**, 253-260.

Meadows JRS, Hanotte O, Drogemuller C*, et al.* (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* **37**, 444-453.

Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology* **21**, 2839-2846.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*. doi: 10.1111/mec.13243

Miller CR, Joyce P, Waits LP (2005) A new method for estimating the size of small populations from genetic mark-recapture data. *Molecular Ecology* **14**, 1991-2005.

Naturvårdsverket (2014) *Björn, populationsutveckling i Sverige.* http://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-O/Bjorn-populationsutveckling-i-Sverige/. Access Date: 21.01.2015. In Swedish

Neaves LE, Zenger KR, Prince RIT, Eldridge MDB (2013) Paternally inherited genetic markers reveal new insights into genetic structuring within *Macropus fuliginosus* and hybridisation with sympatric *Macropus giganteus*. *Australian Journal of Zoology* **61**, 58-68.

Nellemann C, Støen O-G, Kindberg J*, et al.* (2007) Terrain use by an expanding brown bear population in relation to age, recreational resorts and human settlements. *Biological Conservation* **138**, 157-165.

Norman AJ, Street NR, Spong G (2013) De Novo SNP Discovery in the Scandinavian Brown Bear (*Ursus arctos*). *Plos One* **8**. doi: 10.1371/journal.pone.0081012

Obbard ME, Howe EJ, Kyle CJ (2010) Empirical comparison of density estimators for large carnivores. *Journal of Applied Ecology* **47**, 76-84.

Ohnishi N, Osawa T (2014) A difference in the genetic distribution pattern between the sexes in the Asian black bear. *Mammal Study* **39**, 11-16.

Ordiz, A., J. Kindberg, S. Sæbø, J. E. Swenson, and O.-G. Støen. 2014. Brown bear circadian behavior reveals human environmental encroachment. *Biological Conservation* **173**, 1-9.

Ordiz A, Bischof R, Swenson JE (2013) Saving large carnivores, but losing the apex predator? *Biological Conservation* **168**, 128-133.

Ordiz A, Stoen OG, Swenson JE, Kojola I, Bischof R (2008) Distance-dependent effect of the nearest neighbor: spatiotemporal patterns in brown bear reproduction. *Ecology* **89**, 3327-3335.

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62**, 7-135.

Paetkau D, Amstrup SC, Born EW*, et al.* (1999) Genetic structure of the world's polar bear populations. *Molecular Ecology* **8**, 1571-1584.

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, Strobeck C (1995) Micorsatellite analysis of genetic-variation in black-bear populations (vol 3, pg 489, 1994). *Molecular Ecology* **4**, 133-133.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (*Ursidae*) populations. *Genetics* **147**, 1943-1957.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology* **17**, 3428-3447.

Peakall R, Ruibal M, Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution* **57**, 1182-1195.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

Peery MZ, Kirby R, Reid BN*, et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology* **21**, 3403-3418.

Perez-Espona S, Perez-Barberia FJ, Jiggins CD, Gordon IJ, Pemberton JM (2010) Variable extent of sex-biased dispersal in a strongly polygynous mammal. *Molecular Ecology* **19**, 3101-3113.

Persson J, Ericsson G, Segerstrom P (2009) Human caused mortality in the endangered Scandinavian wolverine population. *Biological Conservation* **142**, 325-331.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 502-503.

Pray LA, Goodnight CJ, Stevens L, Schwartz JM, Yan GY (1996) The effect of population size on effective population size: An empirical study in the red flour beetle Tribolium castaneum. *Genetical Research* **68**, 151-155.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

Pritchard JK, Wen X, Falush D (2010) *Documentation for structure software: Version 2.3* University of Oxford, University of Chicago, Oxford.

Proctor MF, McLellan BN, Strobeck C, Barclay RM (2005) Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proceedings of the Royal Society of London Series B-Biological Sciences* **272**, 2409-2416.

Prowse TAA, Johnson CN, Cassey P, Bradshaw CJA, Brook BW (2015) Ecological and economic benefits to cattle rangelands of restoring an apex predator. *Journal of Applied Ecology* **52**, 455-466.

Pulliainen E (1990) Recolonization of Finland by the brown bear in the 1970s and 1980s. *Aquilo Seria Zoologica* **27**, 21-25.

Rauset GR (2013) Life and Death in Wolverines (Doctoral Thesis). *Swedish University of Agricultural Sciences*, Uppsala.

Razgour O, Rebelo H, Puechmaille SJ*, et al.* (2014) Scale-dependent effects of landscape variables on gene flow and population structure in bats. *Diversity and Distributions* **20**, 1173-1185.

R Core Team (2013) R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria.

Ripple WJ, Beschta RL, Fortin JK, Robbins CT (2014a) Trophic cascades from wolves to grizzly bears in Yellowstone. *Journal of Animal Ecology* **83**, 223-233.

Ripple WJ, Estes JA, Beschta RL*, et al.* (2014b) Status and Ecological Effects of the World's Largest Carnivores. *Science* **343.** doi: 10.1126/science.1241484.

Roewer L, Croucher PJP, Willuweit S*, et al.* (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Human Genetics* **116**, 279-291.

Ronce O (2007) How does it feel to be like a rolling stone? Ten questions about dispersal evolution. In: *Annual Review of Ecology Evolution and Systematics*, pp. 231-253. Annual Reviews, Palo Alto.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resour* **8**, 103-106.

Rovdata-Naturvårdsverket (2014) *Faktablad björn - Inventeringsmetodik.* http://www.rovdata.no/Portals/Rovdata/Dokumenter/Instrukser/Bj%C3%B6rn%20%C3%B6vervakningen%20i%20Skandinavien.pdf. 29.04.2015. in Swedish

Rovdata (2014) *Felles norsk-svenske metoder for bjørn og ulv er klare.* http://www.rovdata.no/Nyheter/Nyhetsartikkel/tabid/3764/ArticleId/3721/reftab/1999/Default.aspx. 20.04.2015. in Norwegian

Rowe G, Beebee TJC (2007) Defining population boundaries: use of three Bayesian approaches with microsatellite data from British natterjack toads (*Bufo calamita*). *Molecular Ecology* **16**, 785-796.

Rubicz R, Melton PE, Spitsyn V*, et al.* (2010) Genetic Structure of Native Circumpolar Populations Based on Autosomal, Mitochondrial, and Y Chromosome DNA Markers. *American Journal of Physical Anthropology* **143**, 62-74.

Schregel J, Kopatz A, Hagen SB*, et al.* (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Molecular Ecology* **21**, 3474-3488.

Schubert G, Stoneking CJ, Arandjelovic M*, et al.* (2011) Male-Mediated Gene Flow in Patrilocal Primates. *Plos One* **6**. doi: 10.1371/journal.pone.0021514

Schulte U, Veith M, Mingo V, Modica C, Hochkirch A (2013) Strong genetic differentiation due to multiple founder events during a recent range expansion of an introduced wall lizard population. *Biological Invasions* **15**, 2639-2649.

Schwartz MK, McKelvey KS (2008) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.

Servheen C, Herrero S, Peyton B (1999) Bear - Status Survey and Conservation Action Plan. IUCN, Cambridge.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Solmsen N, Johannesen J, Schradin C (2011) Highly asymmetric fine-scale genetic structure between sexes of African striped mice and indication for condition dependent alternative male dispersal tactics. *Molecular Ecology* **20**, 1624-1634.

Støen O-G, Bellemain E, Sæbø S, Swenson JE (2005) Kin-related spatial structure in brown bears (*Ursus arctos*). *Behavioral Ecology and Sociobiology* **59**, 191-197.

Støen OG, Zedrosser A, Saebo S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears (*Ursus arctos*). *Oecologia* **148**, 356-364.

Swenson JE, Sandegren F, Söderberg A (1998) Geographic expansion of an increasing brown bear population: evidence for presaturation dispersal. *Journal of Animal Ecology* **67**, 819-826.

Swenson JE, Wabakken P, Sandegren F*, et al.* (1995) The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology* **1**, 11-25.

Swenson JE & Wikan S (1996) A brown bear population estimate for Finnmark County, North Norway. *Fauna norvegica Serie A* **17**, 11-15.

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.

Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resour* **8**, 299-301.

Tammeleht E, Remm J, Korsten M*, et al.* (2010) Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Molecular Ecology* **19**, 5359-5370.

Tucker JM, Schwartz MK, Truex RL, Wisely SM, Allendorf FW (2014) Sampling affects the detection of genetic subdivision and conservation implications for fisher in the Sierra Nevada. *Conservation Genetics* **15**, 123-136.

Underwood N, Hamback P, Inouye BD (2005) Large-scale questions and small-scale data: empirical and theoretical methods for scaling up in ecology. *Oecologia* **145**, 177-178.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

van Strien MJ, Holderegger R, Van Heck HJ (2015) Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity* **114**, 27-37.

Vangestel C, Callens T, Vandomme V, Lens L (2013) Sex-Biased Dispersal at Different Geographical Scales in a Cooperative Breeder from Fragmented Rainforest. *Plos One* **8**. doi: 10.1371/journal.pone.0071624

Vercken E, Sinervo B, Clobert J (2012) The importance of a good neighborhood: dispersal decisions in juvenile common lizards are based on social environment. *Behavioral Ecology* **23**, 1059-1067.

Vilà C, Sundqvist AK, Flagstad O*, et al.* (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 91-97.

Wabakken P, Sand H, Liberg O, Bjarvall A (2001) The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978-1998. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **79**, 710-725.

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

Waples RS, Do C (2008) LDNe: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resour* **8**, 753-756.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Wikman M (2010) *Riistakannat 2010. Riistanseuranatojen tulokset.* Riista- ja kalantalouden tutkimuslaitos, Helsinki.

Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nature Genetics* **36**, 1122-1125.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* **28**, 290-300.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

Woodroffe R (2000) Predators and people: using human densities to interpret declines of large carnivores. *Animal Conservation* **3**, 165-173.

Woods JG, Paetkau D, Lewis D*, et al.* (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin* **27**, 616-627.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Xenikoudakis G, Ersmark E, Tison J-L*, et al.* (in press) Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear *Molecular Ecology*. doi: 10.1111/mec.13239

Yannic G, Basset P, Buchi L, Hausser J, Broquet T (2012) Scale -specific sex-biased dispersal in the Valais shrew unveiled by genetic variation on the Y-chromosome, autosomes, and mitochondrial DNA. *Evolution* **66**, 1737-1750.

Zedrosser A, Bellemain E, Taberlet P, Swenson JE (2007a) Genetic estimates of annual reproductive success in male brown bears: the effects of body size, age, internal relatedness and population density. *Journal of Animal Ecology* **76**, 368-375.

Zedrosser A, Dahle B, Stoen OG, Swenson JE (2009) The effects of primiparity on reproductive performance in the brown bear. *Oecologia* **160**, 847-854.

Zedrosser A, Støen O-G, Sæbø S, Swenson JE (2007b) Should I stay or should I go? Natal dispersal in the brown bear. *Animal Behaviour* **74**, 369-376.

Zupan A, Vrabec K, Glavac D (2013) The paternal perspective of the Slovenian population and its relationship with other populations. *Annals of Human Biology* **40**, 515-526.

PAPER I

Forensic Science International: Genetics 6 (2012) 798–809

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

A forensic DNA profiling system for Northern European brown bears (*Ursus arctos*)

R. Andreassen [a,*], J. Schregel [b], A. Kopatz [b,c], C. Tobiassen [b], P.M. Knappskog [b,h,i], S.B. Hagen [b], O. Kleven [b], M. Schneider [d], I. Kojola [e], J. Aspi [c], A. Rykov [f], K.F. Tirronen [g], P.I. Danilov [g], H.G. Eiken [b]

[a] *Faculty of Health Sciences, Oslo and Akershus University College, Oslo, Norway*
[b] *Bioforsk Svanhovd, Norwegian Institute for Agricultural and Environmental Research, Svanvik, Norway*
[c] *Department of Biology, University of Oulu, Oulu, Finland*
[d] *Swedish Environmental Protection Agency, SE-106 48 Stockholm, Sweden*
[e] *Finnish Game and Fisheries Research Institute, Oulu Game and Fisheries Research, Tutkijantie 2 E, FIN-90570 Oulu, Finland*
[f] *Pinezhsky Strict Nature Reserve, Pervomayskaja 123a, Pinega, Arkhangelsk Region, Russia*
[g] *Institutes of Biology, Karelian Research Centre of the Russian Academy of Science, 185910 Petrozavodsk, Russia*
[h] *Center for Medcical Genetics and Molecular Medicine, Haukeland University Hospital, Norway*
[i] *Institute of clinical medicine, University of Bergen, Norway*

ARTICLE INFO

Article history:
Received 9 November 2011
Received in revised form 10 February 2012
Accepted 11 March 2012

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

ABSTRACT

A set of 13 dinucleotide STR loci (G1A, G10B, G1D, G10L, MU05, MU09, MU10, MU15, MU23, MU26, MU50, MU51, MU59) were selected as candidate markers for a DNA forensic profiling system for Northern European brown bear (*Ursus arctos*). We present results from validation of the markers with respect to their sensitivity, species specificity and performance (precision, heterozygote balance and stutter ratios). All STRs were amplified with 0.6 ng template input, and there were no false bear genotypes in the cross-species amplification tests. The validation experiments showed that stutter ratios and heterozygote balance was more pronounced than in the tetranucleotide loci used in human forensics. The elevated ratios of stutter and heterozygote balance at the loci validated indicate that these dinucleotide STRs are not well suited for interpretation of individual genotypes in mixtures. Based on the results from the experimental validations we discuss the challenges related to genotyping dinucleotide STRs in single source samples. Sequence studies of common alleles showed that, in general, the size variation of alleles corresponded with the variation in number of repeats. The samples characterized by sequence analysis may serve as standard DNA samples for inter laboratory calibration. A total of 479 individuals from eight Northern European brown bear populations were analyzed in the 13 candidate STRs. Locus MU26 was excluded as a putative forensic marker after revealing large deviations from expected heterozygosity likely to be caused by null-alleles at this locus. The remaining STRs did not reveal significant deviations from Hardy–Weinberg equilibrium expectations except for loci G10B and MU10 that showed significant deviations in one population each, respectively. There were 9 pairwise locus comparisons that showed significant deviation from linkage equilibrium in one or two out of the eight populations. Substantial genetic differentiation was detected in some of the pairwise population comparisons and the average estimate of population substructure (F_{ST}) was 0.09. The average estimate of inbreeding (F_{IS}) was 0.005. Accounting for population substructure and inbreeding the total average probability of identity in each of the eight populations was lower than 1.1×10^{-9} and the total average probability of sibling identity was lower than 1.3×10^{-4}. The magnitude of these measurements indicates that if applying these twelve STRs in a DNA profiling system this would provide individual specific evidence.

1. Introduction

DNA markers such as microsatellites have been extensively used in conservation genetics to study population diversity, impact of genetic drift and level of inbreeding in a variety of species [1]. One of the best studied mammalian species in conservation genetics is the European brown bear (*Ursus arctos*) [2]. Most data used in recent wildlife genetic studies of brown bear are from genotyping a collection of dinucleotide STRs that were isolated from brown bear and black bear (*Ursus americanus*) [3,4]. The development of non-invasive genetic sampling techniques has allowed sampling of living populations of large carnivores like the brown bear [3,5]. Non-invasive genetic sampling techniques open for long-term genetic monitoring of threatened carnivores that

* Corresponding author. Tel.: +47 93053273.
E-mail address: rune.andreassen@hioa.no (R. Andreassen).

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2012.03.002

occur at low densities. Capture-mark-recapture (CMR) analysis allows important parameters such as abundance, survival and migration to be studied [6,7]. As part of the population management of Northern European brown bear in Norway their abundance have been monitored since 2006 by use of CMR analyses of data from non-invasive samples typed in a set of dinucleotide STRs (see e.g. [8]).

The brown bears of Northern Europe are listed as threatened, but are often involved in conflicts with humans, livestock depredations and illegal hunting. When investigating wildlife crime, genetic analyses of sample materials could provide species specific identification of bear. Furthermore, if using a set of bear specific STRs, forensic genetic analyses would have the potential to provide individual specific bear profiles from a variety of sample materials and provide a means for traceability of bear products (e.g. food, trophy objects and medicine). Experience from population management of brown bears in Norway indicates that approximately 65–70% of non-invasive sample materials may be successfully typed in more than 6 STR loci [8]. Similar success rates when analyzing non-invasive sample materials have been reported by others [9].

Genetic identification of individuals by use of a DNA profiling system would rely on relevant reference data. The lack of such reference data may be a limitation when developing DNA profiling systems for large carnivores. However, these existing dinucleotide markers commonly used in wildlife population management may be used in a forensic DNA profiling system if the necessary allele frequencies from living populations could be retrieved in co-operation with population management laboratories. Dinucleotide STRs are widely used in population monitoring and conservation genetic studies of brown bear [1]. Tetranucleotide STRs are expected to have less stutter and less difference in heterozygote balance than dinucleotide STRs, and they are thus, the preferred markers in human forensics. However, one reason why dinucleotide STRs are commonly used in conservation genetics is that these markers has proven to work well in sample materials like faeces and hair due to relatively short amplicon sizes [1,3,8–10]. European laboratories that monitor bear populations already use very similar sets of dinucleotide STRs [10–12] while tetranucleotide STRs, on the other hand, are not used at all for DNA analysis of brown bear. Thus, a selection of dinucleotide STRs seems to be the preferred forensic markers to include in a European bear DNA profiling system. The ISFG recommendations, although pointing out the benefits of using tetranucleotide STRs, support the use of dinucleotide STRs if they are already in widely use in a non-human species [15]. Several recommendations regarding validation and use of non-human DNA in forensic genetics have been suggested to justify their application as evidence in court [13–15]. Such validation studies, that demonstrate the performance of any new markers, may be particularly important if applying dinucleotide STRs instead of the tetranucleotide STRs [15].

One aim of this study were to perform the recommended validation tests on thirteen dinucleotide microsatellite markers (G1A, G10B, G1D, G10L, MU05, MU09, MU10, MU15, MU23, MU26, MU50, MU51, MU59) commonly used for bear population management and conservation genetics. The validations tests could aid in the selection of markers for a forensic DNA profiling system for the brown bear in Northern Europe. The validation tests included species specificity testing, measurements of sensitivity as well as measurements of precision, stutter and heterozygote balance. Selected common alleles from all STR loci were sequenced to explore the allelic size variation at the sequence level. Another aim of this study was to provide allele frequency distributions as well as relevant forensic genetic parameters for the selected markers from eight bear populations in Northern Europe.

2. Materials and methods

2.1. Population material from eight brown bear populations in Northern Europe

The individual profiles in the population material are from samples collected in specific areas in Northern Europe ($n = 479$). Fig. 1 shows a map of Northern Europe with the approximate location of each of the eight populations indicated (P1–P8). A total of 290 of the individuals were from Norway, of which 233 individual profiles were obtained by typing non-invasive samples (fecal scats and hair) collected in the field as part of the monitoring of bears in Norway from 2006–2009, and 57 were obtained from legally shot bears in the same period (tissue and/or blood). The individuals were from four geographically separated areas; North-eastern Norway (P1, $n = 74$), North-western Norway (P2, $n = 34$), Middle Norway (P3, $n = 81$) and South-eastern Norway (P4, $n = 101$). The individuals from Middle Sweden were collected in 2009 in the county of Västerbotten (P5, $n = 84$) and the individuals from Finland were from the Kainuu area (P7, $n = 44$) and collected during 2005–2008. The sample materials from Sweden and Finland were scats collected in the field. The individuals from Russia were from two areas; the Pinega Strict Nature Reserve in Archangelsk (P6, $n = 27$) and Karelia (P8, $n = 35$). The sample materials were scats collected in the field in Pinega during 2005–2008 and tissue samples from legally shot bears in Karelia in the period 2005–2007.

2.2. DNA extraction, PCR amplification and STR analysis

DNA was extracted from hair-roots and tissue using the Qiagen DNeasy Tissue kit (Qiagen) and from faeces using the Invitek Stool kit (Invitek). Faeces were stored in stool collection tubes with DNA stabiliser (Invitek) or in plastic bags and kept at −20 °C until DNA extraction. The hair samples were stored dry and dark in paper envelopes until DNA extraction. No quantification of DNA concentrations was performed on the extracted samples. Instead,

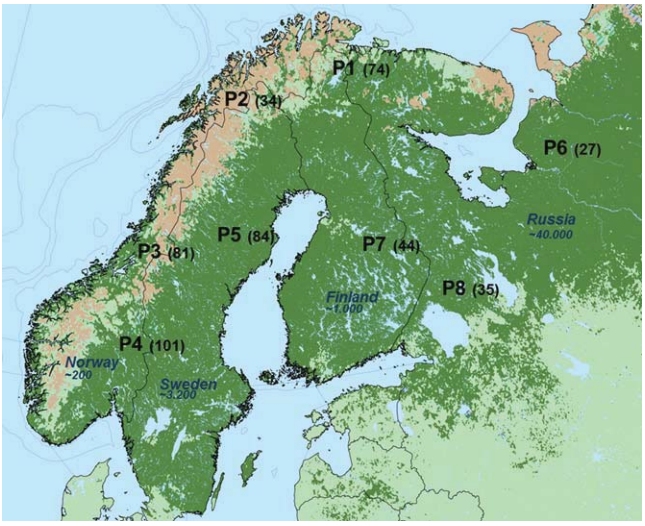


Fig. 1. The figure shows a map of northern Europe (including Norway, Sweden, Finland and Russia). The borders are shown as black lines, forest areas as green, scrub/bush and cropland as light green, tundra and mountains as brown and water as blue. The sample location of each of the eight bear populations, P1–P8 (with number of individuals in brackets) are indicated on the map. P1 (74); North eastern Norway $n = 74$, P2; North-western Norway $n = 34$, P3; Middle Norway $n = 81$, P4; South-Eastern Norway $n = 101$, P5; Middle Sweden $n = 84$, P6; Pinega in Russia $n = 27$, P7; Kainuu in Finland $n = 44$, P8; Karelia in Russia $n = 35$. Estimates of the total number of bears in each of the four countries are given in blue below the name of the country. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the results from the two first STRs analyzed (MU05/MU23 or MU09/MU10) were used to roughly judge the amount of useful template DNA in each sample. Based on these first results the samples that were negative or judged to be of too poor quality for further analysis were removed. In the samples with either very strong or weak results the amount of template DNA used were modified to optimize the PCR amplification of the following STRs. All individuals in the eight populations were analysed in the 13 STRs: G1A, G10B, G1D, G10L, MU05, MU09, MU10, MU15, MU23, MU26, MU50, MU51 and MU59. These STRs had been previously cloned and characterized by others [3,4]. The modified forward and reverse primer sequences used in this study and the reference numbers to each of the corresponding loci from Genbank are given in Table 1. A short 5′ tail [16] has been added to the reverse primer in seven of the thirteen primer pairs (see Table 1). PCRs were performed in 10 μl reaction volumes containing 1× PCR Gold buffer (ABI), 200 μM dNTP (Eurogentec), 1,5 mM $MgCl_2$ (ABI), 0,5 μM of each primer (MedProbe Inc.), 1 U AmpliTaqGold DNA polymerase (ABI), 1× BSA (NEB) and 1 μl template DNA. PCR conditions for loci G1A, MU10, MU05, MU09, MU23, MU50, MU51, MU59 and G10L were 10 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 58 °C, and 1 min at 72 °C, and final extension for 15 min at 72 °C on an ABI 2720. PCR conditions for loci G1D, G10B, MU15 and MU26 were 10 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 60 °C, and 1 min at 72 °C, and final extension for 5 min at 72 °C. Capillary electrophoresis was carried out on an ABI 3730 and the PCR fragments were analyzed in GeneMapper 4.0 (ABI). Allele sizes were measured using Genescan 500LIZ standard.

Duplicates of samples from four different individuals that together represented the common alleles at each marker were included as positive controls in all runs. Negative controls were included for every 7th sample. The positive controls were used to adjust for between-run-variation, and manual inspection of chromatograms was included as part of the final allele designations. All samples were typed at least two times by independent PCR runs followed by analysis on ABI 3730. Any sample typed as a homozygote genotype at any locus was confirmed by a minimum of three replicates (peak height threshold values >300RFU). The laboratory routinely types bear STRs for population management purposes. The extraction of samples and the analysis of the STRs are accredited according to the EN ISO/IEC 17025 standard (Norwegian Accreditation: Test 139).

A test of a multiplex PCR combining eight of the validated loci into one single PCR was carried out using Qiagen multiplex PCR-kit as described by the manufacturer's protocol except using 2 μl template and a total reaction volume of 10 μl. The concentration of each primer was 0.2 μM, except MU10 (0.4 μM) and MU59 (0.8 μM), and the PCR conditions were as in the singleplex PCRs. Loci MU09, MU10, MU23 and MU59 were labeled with FAM, MU05 with NED, MU50 with PET, MU51 and G10L with VIC (see Table 1). The multiplex was tested in sample extracts from tissue (1–2 ng template), hair (2–4 ng template) and faeces (unknown concentration of bear DNA) that previously had been successfully genotyped using singleplex PCR and shown to be single source samples originating from different individuals.

2.3. Sensitivity, species specificity, precision, stutter ratios and heterozygote balance

Two DNA samples (positive control samples) were used to test the effect of different template concentrations when amplifying the thirteen STR loci in singleplex PCRs. The following template input was tested in 10 μl reaction volumes: 30 ng, 20 ng, 3 ng, 2 ng, 0.6 ng, 0.4 ng, 0.3 ng, 0.2 ng, 0.03 ng and 0.02 ng.

All 13 markers were tested for cross-species amplification against two DNA samples from each of nine wild and domesticated

Table 1
Primer sequences for 13 STR markers.

Locus	Forward primer	Reverse primer	Structure[b]	Allele size range[c]	Ta[d]	Octaplex-PCR conditions[e]	GenBank acc #	Reference
G1D	5′ TCTCTTTTCCTTTAGGGGACTC	5′ CTAGCACCCAGCAAGGTATAATA	Simple	123–139 bp/12–20R	60 °C		U22094.1	Paetkau and Strobeck [4]
G10B	5′ ATTTTCTTGAGGACTTTTGCATATA	5′ *GTTTCTT*AACCTCCATCCATACAATACAAC[a]	Simple	94–122 bp/10–24R	60 °C		U22084.1	Paetkau and Strobeck [4]
G1A	5′ ACCCTGCATACTCTCCTCTGATG	5′ GCACTGTCCTTGCGTAGAAGTGAC	Simple	177–195 bp/15–24R	58 °C		U22095.1	Paetkau and Strobeck [4]
G10L	5′ CAGGACAGGATATTGACATTGA	5′ GATACAGAAACCTACCCATGCG	Complex	166–194 bp/21–35R	58 °C	0.2 μM, VIC	U22088.1	Paetkau and Strobeck [4]
MU10	5′ TTCAGATTTCATCAGTTTGAC	5′ TTTGTATCTTGGTTGTCAGC	Complex	133–153 bp/18–28R	58 °C	0.4 μM, FAM	Y09642.1	Taberlet et al. [3]
MU23	5′ GCCTGTGTGCTATTTTATCC	5′ *GTTTCTTT*GCTTGCCTAGACCACC[a]	Complex	164–180 bp/18–26R	58 °C	0.2 μM, FAM	Y09645.1	Taberlet et al. [3]
MU51	5′ GCCAGAATCCTAAGACACCT	5′ *GTTTCTT*GAAAGGTTAGATGGAAGAGATG[a]	Simple	131–153 bp/12–23R	58 °C	0.2 μM, VIC	Y09648.1	Taberlet et al. [3]
MU59	5′ GCTGCTTTGGGACATTGTAA	5′ *GTTTCTT*CAATCAGGCATGGGGAAGAA[a]	Simple	224–256 bp/11–27R	58 °C	0.8 μM, FAM	Y09649.1	Taberlet et al. [3]
MU50	5′ GTCTCTGTCATTTCCCCATC	5′ GAGCAGGAAACATGTAAGATG	Complex	106–136 bp/15–30R	58 °C	0.2 μM, PET	Y09647.1	Taberlet et al. [3]
MU05	5′ ATGTGGATACACTGGAATAGACC	5′ *GTTTCTT*GTCACATCAACTCAAACTTGTTAT[a]	Simple	109–133 bp/10–22R	58 °C	0.2 μM, NED	Y09640.1	Taberlet et al. [3]
MU09	5′ GCCAGCATGTGGGTATATGTGT	5′ *GTTTCTT*AGCAGCATATTTTTGGCTTTGAAT[a]	Simple	98–128 bp/11–26R	58 °C	0.2 μM, FAM	Y09641.1	Taberlet et al. [3]
MU15	5′ CATCTGAATTATGCAATTAAACAGC	5′ *GTTTCTT*GTTTTTGTTTAGCAGGTTTGTCTC[a]	Simple	104–120 bp/16–24R	60 °C		Y09644.1	Taberlet et al. [3]
MU26	5′ GCTGAGAATTCCATTGTATGTAT	5′ ATCTTCTTTACACATAACATACCAT	Complex	82–100 bp/11–20R	60 °C		Y09646.1	Taberlet et al. [3]

[a] A 5′ tail shown in italics has been added to reverse primer.
[b] Repeat structure in each locus.
[c] Absolute allel size range in basepairs/predicted size range in repeats.
[d] Annealing temperatures.
[e] Loci that may be combined in a single octaplex PCR with the current PCR-protocol if adjusting primer concentrations and dye labels as shown for each of the loci.

animals as well as humans. The following species were included; elk (*Alces alces*), reindeer (*Rangifer tarandus*), wolverine (*Gulo gulo*), eurasian lynx (*Lynx lynx*), wolf (*Canis lupus*), hare (*Lepus timidus*), red deer (*Cervus elaphus atlanticus*), domesticated cat (*Felis catus*), red fox (*Vulpes vulpes*), and human (one male and one female). Another bear species, the polar bear (*Ursus maritimus*), was also included in the species tests. Two brown bear DNA samples (DNA extracted from tissue) was used as positive controls. PCR and STR analysis was performed as described in Section 2.2.

Measurements of between-run precision were obtained from all candidate STRs by ≥30 independent amplifications and subsequent runs of two heterozygote positive controls. Measurements of stutter ratio and heterozygote balance were also obtained from these runs. Within-run precision was measured in a sample from an individual with a heterozygous genotype at locus G10L and a homozygous genotype at locus G10B. Stutter ratio was calculated by dividing the peak height (RFU) of the stutter peak in position-1R (one repeat less than the true allele) by the peak height of the true allele. Paired students *t*-test was used to tests for significant differences in stutter ratios between alleles within a locus. Heterozygote balance was calculated by dividing the peak height (RFU) of the short allele by the peak height of the longer allele. Heterozygote balance was calculated in this manner to give information about the direction of the imbalance.

2.4. DNA sequencing of tandem repeat arrays and immediate flanking sequences

The tandem repeat array and the immediate upstream and downstream sequences at each of the thirteen loci was analysed by DNA sequencing. PCR products amplified from individuals that were homozygous for common alleles were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (ABI) as recommended by the manufacturer. Forward and reverse PCR-primers were used as sequencing primers in forward and reverse sequencing reactions, respectively. Forward and reverse sequences from each sample were aligned by use of Sequencher 4.7 software. The allelic sequences from a locus were aligned and, finally, the sequence and size variation within each locus was determined by manual inspection.

2.5. Additional PCR tests of locus MU26

New pairs of primers were applied to amplify an extended region upstream and downstream of the repeat array at locus MU26. The primers used were as follows; two different forward primers, each tested in combination with the same reverse primer (F1a = 5′ CTGCCTCAAATGACAAGA, F1b = 5′ TTAAGGAGGGCACTTGAT and R1 = 5′ GGCCTTTTTACATTAGTTGAT) as well as a second forward and reverse primer set (F2 = GCCTCAAATGACAAGATTTC, R2 = TCAATTAAAATAGGAAGCAGC). Samples were amplified using reaction conditions as described in Section 2.2. Four samples that had previously been successfully amplified at locus MU26 using the standard primers were used as positive controls. Samples from ten individuals that were previously successfully amplified in all loci but MU26 (putative null-allele homozygote individuals) were amplified in three different PCR tests using primers F1a and R1, F1b and R1, F2 and R2. PCR products were visualized on agarose gels and manually typed as positive (a clear PCR band of expected size) or negative (no PCR product).

2.6. Analysis of data

The GDA software v1.0 [17] was used for evaluation of Hardy–Weinberg equilibrium (HWE), linkage disequilibrium and population structure (F_{IS}, overall and pairwise F_{ST}). The estimation of population pairwise F_{ST} values was also performed with Arlequin version 3.5.1.2 [18]. Bonferroni corrected significance levels were applied when testing HWE and linkage disequilibrium. Thus, when testing HWE the significance level was 0.0005 while when testing linkage disequilibrium the significance levels was 0.00009. PowerStats (the gene count method) was used to obtain the observed allele frequencies and for the calculation of the forensic efficiency parameters [19]. API-Calc was used to estimate average probability of identity [20]. Micro-checker was used to test for presence of null-alleles and to discriminate between errors in allele frequency estimates caused by null-alleles, allele drop-out or stutter in a locus (MU26) with homozygous excess [21].

3. Results

3.1. Species specificity and sensitivity of the STR markers

The specificity tests of the 13 loci showed that PCR tests using non-bear templates resulted in no PCR-products at loci G1D, MU15 and MU51 while at other loci there were weak amplicons observed in some of the loci and species combinations. These weak amplicons had fragment sizes that were outside of the allele size range of the markers and could not be mistaken as bear genotype results. Template DNA from polar bear was successfully amplified with allele size ranges that partially overlapped those in brown bear profiles.

The sensitivity tests using different template input concentrations showed that all markers were successfully amplified when template input was in the range 30–0.6 ng. When decreasing the template input to 0.2 ng seven markers were successfully amplified while six of the loci showed drop-out of alleles (MU05, MU50, MU15, G10B, G1D, and MU09). Drop out was observed in all markers with template input less than 0.2 ng. All loci were successfully amplified and analyzed using template DNA extracted from tissue (frozen or dried), hair, blood and faeces.

3.2. Measurements of precision, stutter and heterozygote balance

The results from measurements of precision, stutter ratio and heterozygote balance are summarized in Table 2. The between-run measurements of precision revealed a standard deviation (S.D.) that were 0.16 bp or less in seven markers and from 0.16 to 0.20 bp in five markers while locus G10L revealed a standard deviation of 0.30 bp. There was no obvious difference in precision between alleles with different size within a locus. The within-run measurements of precision at loci G10L and G10B showed standard deviations of 0.16 bp and 0.08 bp, respectively. Fig. 2 a shows a chromatogram with results from locus MU23 and demonstrates separation of alleles differing by 1 bp (see also Section 3.3).

Stutter was observed in all markers as peaks at positions that were from one to several repeats less than the true allele (-1R, -2R, -3R etc.) and in some cases one repeat larger than the true allele (+1R). Fig. 2b and c shows the chromatograms from two individuals that are heterozygous and homozygous at loci MU59 and MU50, respectively. The figures illustrate the characteristic patterns of stutter observed in the dinucleotide markers tested. The stutters revealed decreasing peak heights with increasing size difference to the true allele in all markers. The stutter ratios (peak height of stutter/peak height of true allele) for peaks in position-1R (largest stutter peak) were recorded for short allele and large allele in all loci except MU05 and MU10. At these loci the stutter ratios were recorded for the short allele only since true alleles were separated by one repeat only in the control sample used for performance measurements. The median stutter ratios and upper 95 percentiles are given in Table 2. There was a significant difference in stutter ratio between the alleles within a locus with

Table 2
Measurements of precision, heterozygote balance and stutter ratio.

Locus	Alleles/genotype[a]	Mean (bp)[b]	S.D. (bp)[c]	Het. balance[d]	Stutter ratio[e]
MU05 allele A	125	125.39	0.18		0.56 (0.62)
MU05 allele B	127	127.51	0.17	1.77 (1.39–2.72)	–
MU09 allele A	110	109.74	0.20		0.49 (0.63)
MU09 allele B	116	116.01	0.20	1.48 (0.76–2.64)	0.60 (0.80)
MU10 allele A	149	149.92	0.12		0.53 (0.61)
MU10 allele B	151	152.11	0.10	1.73 (1.28–2.33)	–
MU23 allelle A	166	166.32	0.18		0.54 (0.64)
MU23 allelle B	173	173.65	0.20	1.36 (0.98–2.06)	0.67 (0.69)
MU50 allele A	120	120.59	0.19		0.54 (0.62)
MU50 allele B	124	124.79	0.19	1.61 (0.99–2.35)	0.74 (0.90)
MU51 allele A	139	138.94	0.17		0.32 (0.39)
MU51 allele B	149	149.78	0.16	2.01 (1.22–3.15)	0.54 (0.61)
MU59 allele A	240	240.29	0.10		0.52 (0.64)
MU59 allele B	256	256.54	0.09	1.53 (1.03–2.59)	0.75 (0.87)
G10L allele A	174	174.17	0.30		0.38 (0.46)
G10L allele B	182	182.50	0.30	1.36 (0.67–2.86)	0.50 (0.58)
G10L allele A[f]	174	173.9	0.15		0.33 (0.38)
G10L allele B[f]	182	182.25	0.16	1.55 (0.89–2.17)	0.44 (0.49)
G1D allele A	129	129.50	0.11		0.40 (0.45)
G1D allele B	133	133.51	0.10	1.13 (0.96–1.79)	0.45 (0.47)
G10B allele A	98	97.34	0.13		0.37 (0.40)
G10B allele B	110	109.82	0.15	2.18 (1.77–2.54)	0.62 (0.67)
G10B hom B[g]	110	–	–	–	–
G10B hom B[g]	110	109.69	0.08	–	0.62 (0.68)
G1A allele A	181	181.06	0.16		0.57 (0.65)
G1A allele B	189	189.38	0.13	1.47 (1.08–2.09)	0.70 (0.73)
MU15 allele A	110	109.86	0.15		0.58 (0.63)
MU15 allele B	116	116.13	0.14	1.55 (1.29–2.02)	0.74 (0.79)
MU26 allele A	84	82.98	0.07		0.14 (0.18)
MU26 allele B	96	96.09	0.09	1.79 (1.34–2.30)	0.37 (0.39)

[a] Alleles as designated with a size based nomenclature.
[b] Mean value of allele sizes when measured with pop7 on ABI3730.
[c] Standard deviations (S.D.) from between-run measurements of ≥30 runs of a control sample.
[d] Median heterozygote balance ratio with upper and lower 95% percentile in paranthesis.
[e] Median stutter ratios of alleles with upper 95% percentile in parenthesis.
[f] Within-run measurements of precision (S.D. = 0.16 bp), heterozygote balance and stutter ratio.
[g] Within-run measurements of a sample with a homozygous genotype (S.D. = 0.08 bp).

smaller stutter ratios for the shorter allele. The median stutter ratios at the different loci ranged from 0.14 to 0.58 for the short allele and 0.37 to 0.75 for the larger alleles. The single largest stutter observed in any marker was 0.91 (MU59). A very small stutter peak was observed in position +1R in some cases and the average stutter ratio of these peaks were 0.09 (data not shown).

The variation in heterozygote balance was recorded in all loci. Heterozygote balance was calculated by dividing the peak height (RFU) of the short allele by the peak height of the larger allele to provide values that gives information about the direction of the imbalance. A heterozygote balance ratio less than one was rare. Thus, the shorter alleles were usually the ones with the largest peak heights, but ratios less than one were recorded in a few cases at loci G10L, MU09, MU23, MU50 and G1D (data not shown). The median values for heterozygote balance ratios for the thirteen STRs ranged from 1.13 at G1D to 2.18 at G10B. The median value from each locus as well as upper and lower 95 percentile is given in Table 2. The upper 95 percentiles of stutter and the lower 95 percentiles of heterozygote balance ratios did not overlap at any locus but MU09. This locus performed slightly less with no overlap between upper and lower 92.5 percentiles of stutter and heterozygote balance, respectively. We did not observe any heterozygote balance ratio larger than 3.3 (i.e. large allele approximately 30% of the peak height of the short allele).

3.3. Repeat numbers and repeat array structure in common alleles

Fifty alleles representing the most common alleles from all loci were sequenced. The alleles were aligned with the sequence data from Genbank and the repeats designated in accordance with the original characterization of the dinucleotide tandem repeat arrays at each locus [3,4]. The results from the sequence analysis with allele size, repeat numbers and repeat structure of the alleles from each of the loci are summarized in Supplementary File 1. Eight of the loci showed a simple tandem repeat array structure while loci MU50, MU10, MU23 and MU26 revealed compound repeat structures with $(CT)_n(GT)_n$, $(TG)_n(TA)_1(TG)_n$, $(GT)_n(AT)_n(GT)_n$ and $(TG)_n(TA)_n(TG)_n$, respectively. At locus G10L there was an insertion of TTGTT within the repeat array that was not present in the Genbank sequence from black bear (*Ursus americanus*).

For ten loci we measured size differences that corresponded to the variation in repeats. However, for three loci (MU23, MU59 and G1A) additional differences affecting the allele sizes were revealed outside of the tandem repeat array. The sequence analysis of MU23

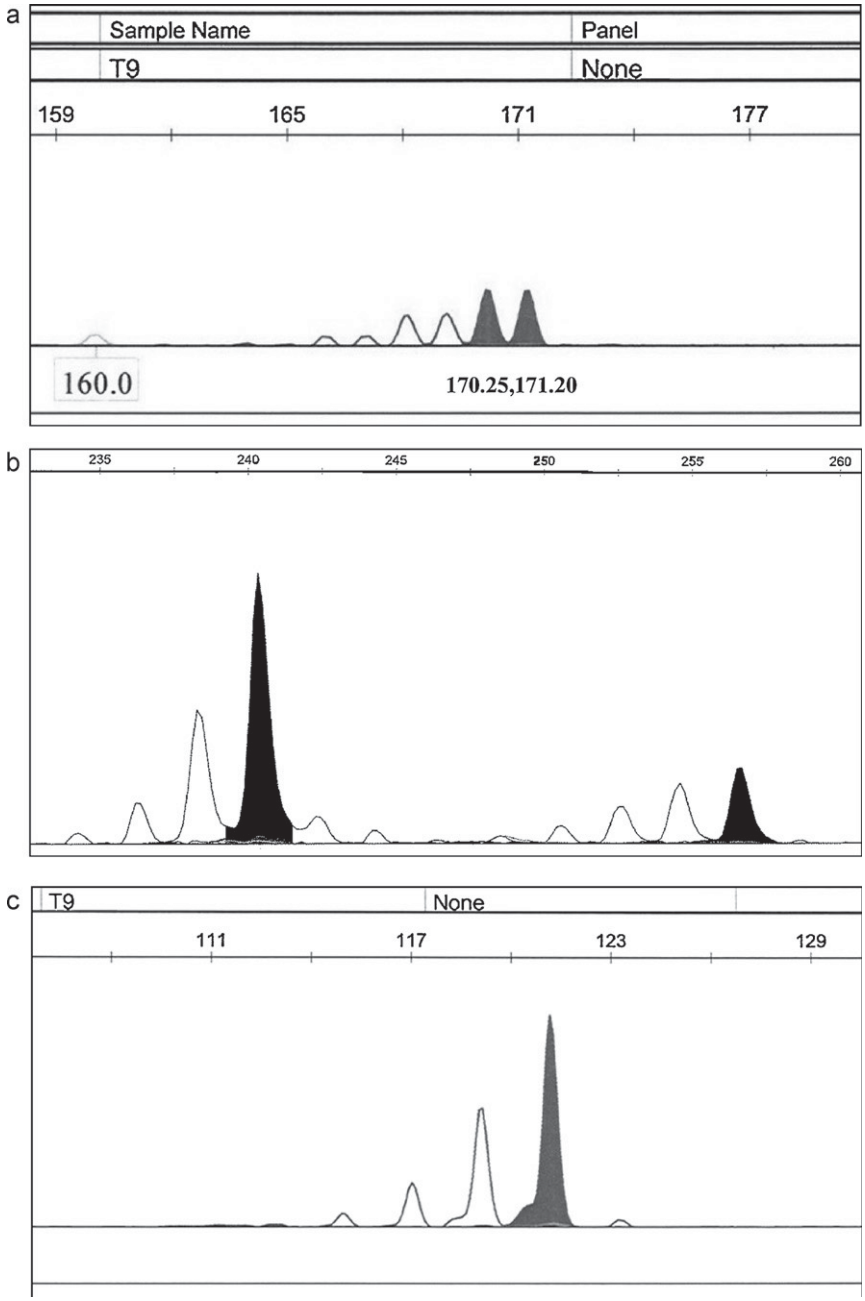


Fig. 2. The figure shows chromatograms with genotype results from three individuals. (a) A heterozygous genotype with alleles sized as 170.25 bp and 171.20 bp in locus MU23. (b) A heterozygous genotype with typical heterozygote balance ratio and stutter peaks in locus MU59. (c) A homozygous genotype with typical stutter peaks in locus MU50.

revealed an indel of a G downstream of the repeat array (position 211, Genebank Acc. No. Y09645) that produced alleles with identical number of repeats that differ by 1 bp. This finding was in agreement with the size measurements that showed micro-variation among alleles at this locus (Fig. 2a). Sequence analysis at locus MU59 showed that there was a 4 bp indel located in position 121 (Genbank Acc. No Y09649) while at locus G1A there was a 4 bp indel located in position 164 (Genbank Acc. No U22095) downstream of the repeat array. Both these indels contribute allele size differences that cannot be distinguished from size variation caused by variation in repeat numbers.

The Genbank sequences from the four ''G-named loci'' (see Genbank Acc. No in Table 1) were isolated from black bear. Our novel genomic sequences in the present study were from brown bear. Comparing Genbank sequences from the black bear with our sequences from brown bears revealed additional sequence differences in the DNA flanking the tandem repeat arrays at some of the loci. The brown bear sequences at locus G1A showed an insertion of an A in position 185. In the brown bear sequences at locus G10B we revealed a substitution (A/T) with a T at position 79 upstream of the repeat array. At the G1D locus, we detected an insertion of a C at position 148 downstream of the repeat array in the brown bear sequence.

Putative SNPs were detected at two loci (G10B, G1D), and in both cases the less frequent variant was observed in two or more individuals. The one observed at locus G10B was a transversion (A/C) in position 93 while the one observed in locus G1D was a transition (A/G) in position 134.

Table 3
Population data from eight Northern European brown bear populations.

Pop[a]		G1D	G10B	MU05	MU09	MU15	G1A	G10L	MU10	MU23	MU50	MU51	MU59	MU26	Av. Het[e] Tot. Av. PI[f]
P1	H_O	0.905	0.676	0.889	0.989	0.811	0.838	0.630	0.757	0.726	0.753	0.824	0.833	0.384	0.80
	H_E	0.844	0.739	0.831	0.861	0.746	0.829	0.609	0.784	0.686	0.833	0.805	0.849	0.550	1.1×10^{-12}
	p	0.2075	0.0015	0.2771	0.7918	0.0278	0.4465	0.2015	**0.0000**	0.7350	0.0393	0.0981	0.1828	**0.0004**	
P2	H_O	0.706	0.765	0.705	0.765	0.382	0.764	0.823	0.735	0.764	0.852	0.823	0.757	0.333	0.74
	H_E	0.708	0.689	0.713	0.689	0.340	0.665	0.808	0.674	0.740	0.739	0.782	0.708	0.294	3.7×10^{-10}
	p	0.1610	0.2203	0.5515	0.9381	0.8325	0.7596	0.1437	0.8968	0.6500	0.3315	0.4081	0.8475	1.0000	
P3	H_O	0.741	0.630	0.802	0.876	0.667	0.689	0.759	0.703	0.778	0.728	0.753	0.814	0.550	0.75
	H_E	0.665	0.598	0.707	0.824	0.640	0.710	0.741	0.762	0.796	0.749	0.751	0.829	0.607	3.9×10^{-10}
	p	0.0287	0.4912	0.1921	0.0850	0.5162	0.7525	0.2512	0.0053	0.1587	0.4340	0.5068	0.0278	0.1309	
P4	H_O	0.564	0.683	0.673	0.841	0.536	0.645	0.811	0.752	0.633	0.772	0.812	0.750	0.154	0.71
	H_E	0.643	0.689	0.683	0.795	0.579	0.649	0.743	0.785	0.661	0.789	0.808	0.739	0.371	1.1×10^{-9}
	p	0.1471	0.1237	0.8887	0.2259	0.6987	0.6440	0.2218	0.4093	0.9000	0.5393	0.9781	0.9525	**0.0000**	
P5	H_O	0.643	0.654	0.619	0.773	0.643	0.631	0.595	0.666	0.773	0.798	0.798	0.750	0.423	0.69
	H_E	0.676	0.629	0.643	0.829	0.619	0.699	0.640	0.747	0.819	0.813	0.761	0.829	0.610	4.0×10^{-10}
	p	0.1521	0.1487	0.5853	0.5018	0.0162	0.2109	0.0346	0.5365	0.7368	0.8662	0.6400	0.0068	0.0015	
P6	H_O	0.885	0.846	0.807	0.923	0.808	0.960	0.807	0.800	0.577	0.884	0.808	0.920	0.538	0.83
	H_E	0.845	0.841	0.853	0.839	0.711	0.857	0.840	0.826	0.669	0.804	0.785	0.904	0.748	2.9×10^{-12}
	p	0.7487	0.8250	0.0265	0.1112	0.9634	0.9521	0.8159	0.1153	0.0221	0.1625	0.1965	0.8743	0.0556	
P7	H_O	0.818	0.591	0.714	0.818	0.909	0.840	0.809	0.750	0.780	0.650	0.772	0.953	0.500	0.78
	H_E	0.809	0.870	0.760	0.864	0.822	0.816	0.775	0.781	0.833	0.740	0.846	0.897	0.703	2.1×10^{-12}
	p	0.0571	**0.0000**	0.3425	0.1596	0.5734	0.4662	0.0940	0.5003	0.0265	0.4431	0.6140	0.2356	**0.0003**	
P8	H_O	0.771	0.742	0.857	0.882	0.800	0.771	0.686	0.771	0.839	0.771	0.771	0.828	0.657	0.79
	H_E	0.826	0.828	0.831	0.891	0.780	0.803	0.775	0.828	0.884	0.749	0.787	0.915	0.659	1.2×10^{-12}
	p	0.0446	0.0843	0.6359	0.4453	0.6734	0.5681	0.2750	0.3462	0.6540	0.2053	0.7959	0.2190	0.9231	
No. Alleles[b]		9	13	11	13	9	10	12	10	13	10	10	16	–	
Mean H_O[c]		0.73	0.68	0.75	0.85	0.68	0.74	0.73	0.73	0.73	0.77	0.80	0.81	–	
Av. PI[d]		0.12	0.14	0.12	0.07	0.19	0.10	0.14	0.09	0.09	0.09	0.09	0.07	–	

[a] Expected (H_E) and observed (H_O) heterozygosities in eight brown bear populations. P1, North-eastern Norway ($n = 74$); P2, North-western Norway ($n = 34$); P3, Middle Norway ($n = 81$); P4, Southern Norway ($n = 101$); P5, Middle Sweden ($n = 84$); P6, Pinega-Russia ($n = 26$); P7, Kainuu-Finland ($n = 44$); P8, Karelia-Russia ($n = 35$). Significant deviations from Hardy–Weinberg equilibrium are marked in bold (MU26; P1, P4 and P7, G10B; P7, MU10; P1).
[b] Number of alleles observed at each of the loci in the total population material (MU26 not included).
[c] Mean heterozygosity at each of the loci in the total population material (MU26 not included).
[d] Average probability of identity at each of the loci from allele frequencies in the total population ($F_{ST} = 0.09$).
[e] The average heterozygosity in each of the populations.
[f] The total average probability of identity when applying twelve STRs (MU26 not included) and population specific allele frequencies (see Supplementary File 1).

3.4. Population data

A total of 479 individuals from eight Northern European brown bear populations were typed in 13 STR loci (G1A, G10B, G1D, G10L, MU05, MU09, MU10, MU15, MU23, MU26, MU50, MU51, MU59). Allele frequencies from each of the populations are given in Supplementary File 2. The expected and observed heterozygosity frequencies for all loci in each of the eight populations are given in Table 3. Locus MU26 showed a significant deficiency of heterozygotes in three of the populations (P1, P4 and P7), and additional studies of MU26 (see Section 3.5) suggested that null-alleles were present at this locus. Consequently, the MU26 locus was excluded as a putative forensic marker and allele frequency estimates and other forensic parameters are not presented for this locus.

The heterozygosity frequencies ranged from 0.54 (MU15, P4) to 0.95 (MU59) among the remaining twelve loci, although with a single exception for locus MU15 that revealed a heterozygosity frequency of 0.38 in P2. The mean heterozygosity in the eight populations ranged from 0.69 (P5) to 0.83 (P6). Deviations ($p > 0.05$) from Hardy–Weinberg equilibrium (HWE) were observed in 19 out of 104 tests. However, after Bonferroni-correction there were significant deviations in five out of the 104 tests ($p < 0.0005$, marked in bold in Table 3). The five tests with HWE deviations were at loci G10B and MU10 in populations from Kainuu (P7) and North-eastern Norway (P1), respectively, as well as the above mentioned three populations with significant deficiency of heterozygotes at MU26.

Tests for deviation from linkage equilibrium by comparing locus pairs across populations showed that sixteen percent of the 12 loci showed linkage disequilibrium (LD) at a significance level of $p < 0.05$ (data not shown). None of these locus combinations were consistent in linkage disequilibrium in all populations. After Bonferroni-correction of significance levels (528 tests, $p < 0.00009$) there were 9 pairwise locus combinations that remained significant (data not shown). Six of these significant deviations from LD were in pairwise comparisons of locus G10B to another locus in the Kainuu population (P7). However, locus G10B showed also a significant deviation from HWE in this population (see Table 3), and when controlling the HWE effect on the LD test ("the preserving genotypes" option in GDA) none of these deviations remained significant. The observed heterozygosity at MU59 was high and at same time stutter ratios were elevated (see Table 2). A possible genotyping error (a homozygous genotype with a large stutters designated as heterozygous genotype) could result in such a high heterozygosity. Although our genotyping protocol should be robust, we compared the observed number of heterozygous genotypes consisting of alleles that differed by only one repeat to the expected number of such genotypes. The results showed only sparse deviation between the numbers of observed versus expected genotypes. We also re-checked all the genotype results from the two most common heterozygous genotypes with alleles that differed by one repeat (246/248 and 248/250), and in all cases the shorter sized alleles showed the largest peak heights in the two independent analyses. Altogether, these results indicate that the high heterozygosity at MU59 is not caused by genotyping errors.

The overall F_{ST}-value was 0.09 (bootstrap confidence interval 0.07–0.10). The population pairwise F_{ST}'s were significant for all comparisons ($p < 0.0001$) and ranged from 0.008 (P7–P8) to 0.16 (P2–P4) (Supplementary File 3). Estimates of the inbreeding factor (F_{IS}) as well as confidence intervals for F_{IS} are given for each of the populations in footnotes in Supplementary File 2 (MU26 not included). The inbreeding factor ranged from −0.064 (P2) to 0.044 (P8). Significant levels of inbreeding (lower confidence interval above 0) were observed in two of the populations (P5 and P8). The overall value of F_{IS} was 0.005 (not including locus MU26). The overall value of F_{IS} across populations at MU26 was 0.22. However, this large value is likely caused by the presence of null-alleles rather than inbreeding (see also Section 3.5).

The number of alleles and mean heterozygosity at each of the loci in the total population as well as the average probability of identity at each locus (Av. PI) is given in Table 3. The total average probabilities of identity for each population (Tot. Av. PI in Table 3) refers to the probability of observing two copies of any profile in that population when applying the twelve STRs [20]. When estimating the total average PI we used the population specific allele frequencies and F_{IS} – values and the average F_{ST} – value. The largest total average PI in any population was 1.1×10^{-9}. The largest total average probability of sibling identity in any population was 1.3×10^{-4}. Some additional common forensic efficiency parameters (power of discrimination, power of exclusion, matching probability and typical paternity index) are given for each of the populations in Supplementary File 2.

3.5. Heterozygote deficiency at locus MU26

We used Micro-checker 1.0 [21] to further explore the significant heterozygote deficit in MU26 observed in the three populations P1, P4 and P7 by comparing the expected and observed frequency of homozygotes within different allele size classes. The comparison revealed a larger than expected frequencies of homozygotes in most allele size classes in these three populations (data not shown). Furthermore, samples from 17 out of the 219 individuals in P1, P4 and P7 did not produce results in MU26 while they could be genotyped in all other loci. One reasonable explanation for the lack of results in these samples could be that they are from individuals that are homozygous for the null-allele.

New primers located 5′ and 3′ of the original forward and reverse primers were applied to amplify locus MU26 by use of three different PCR tests. Template DNA from ten individuals that were previously successfully amplified in all loci but MU26 (putative null-allele homozygote individuals) were used in an attempt to amplify locus MU26 by use of these new primers. Applying the three new PCR tests locus MU26 was successfully amplified in four positive controls while it could not be amplified in any of the ten individuals.

3.6. Octa-plex PCR for eight common STRs tested on brown bear samples

Tests of a multiplex PCR amplification showed that eight of the validated STRs may be combined and successfully amplified at sensitivity levels as low as 1 ng in a single reaction using primers identical to the ones used in singleplex PCRs. The genotypes from individuals amplified by use of template extracts from tissue, hair and faeces were identical to the genotypes at the eight loci when amplified in singleplex PCRs (Fig. 3).

4. Discussion

4.1. Species specificity and sensitivity of the STR markers

The test species represented animals that bears predate on or animals from where hair and scats may falsely be collected as bear samples. Humans were included to assure that contamination by human DNA, e.g. in the laboratory, would not produce false results. The results demonstrate that presence of DNA from the test species would not lead to any false positive results.

Although even more species may be tested, the results indicate that all the markers have high species specificity. Assignment of individual identity would depend on results from many STRs, but if testing only for presence of bear (a species specific test), positive

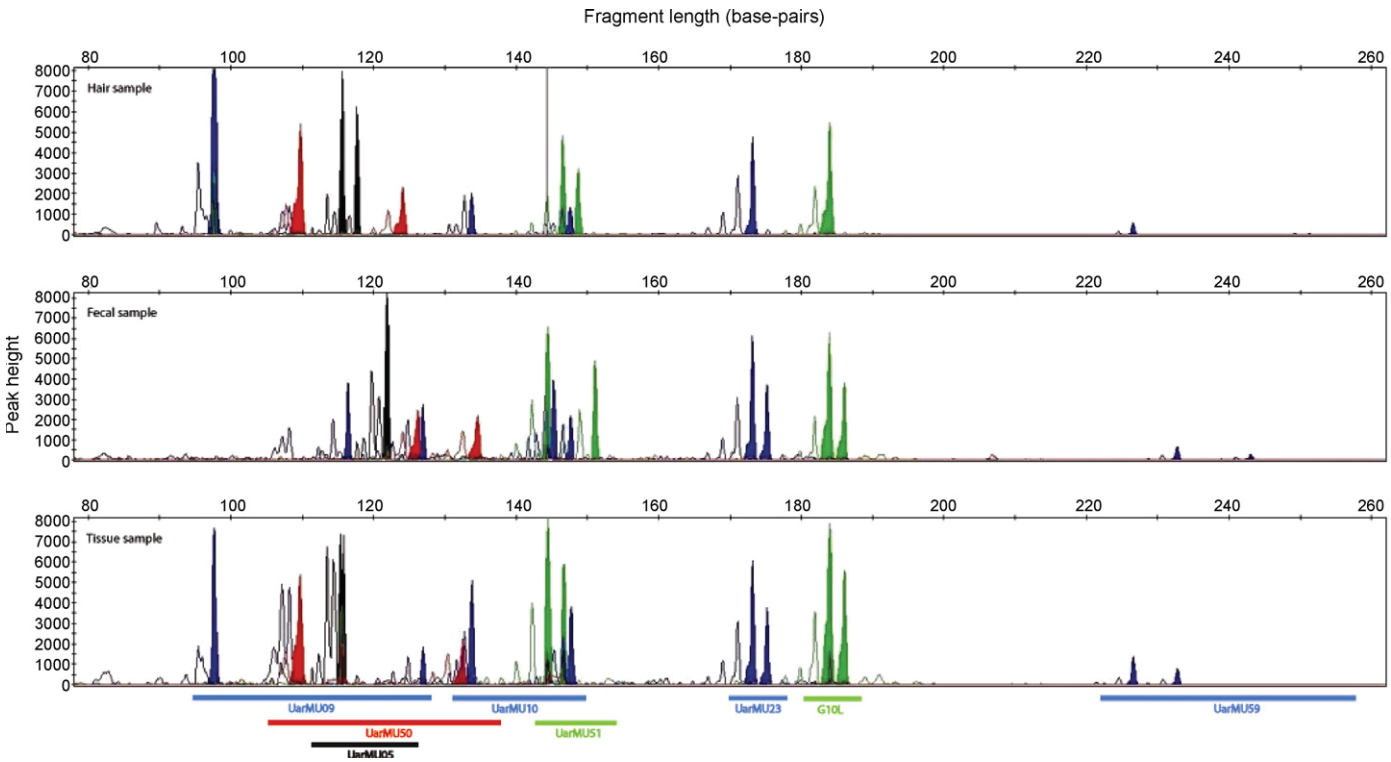


Fig. 3. The chromatograms in the figure show results from a multi-plex PCR consisting of eight of the validated and commonly used STRs (octa-plex). Template DNA was from extracts of hair (2.4 ng), faces (amount of bear DNA not determined) and muscle tissue (1.2 ng). The results in the upper panel are from hair, the middle panel from faeces and the lower panel from tissue, and the results shown in the three panels are from different individuals. The octa-plex PCR consisted of the following STRs (dye); MU05 (NED), MU09 (FAM), MU10 (FAM), MU23 (FAM), MU59 (FAM), MU50 (PET), MU51 (VIC). The fragment range for each STR-locus is indicated below the lower panel (all loci ranges except G10L starting with the abbreviation Uar (*Ursus arctos*)).

results (genotype results) from a smaller number of loci would provide strong bear specific evidence.

Several of the markers have been shown to produce genotype results in other bear species [4,22]. The loci were successfully amplified in polar bear when amplified with the primers from this study, and the sequence comparisons to black bear from Genbank at four loci (see Section 3.3) revealed few sequence differences. Thus, although the primers have not been tested in many bear species, it is likely that they would be successfully amplified in other bear species and should not be regarded as brown bear specific.

Traditional DNA quantification methods (spectrophotometric estimation) would fail to measure the amount of bear specific DNA in sample materials like faeces which also contain DNA from food and microorganisms. Neither would the amount of DNA degradation and presence of PCR inhibitors be detected by use of such DNA quantification methods. Development of a species specific qPCR assay could provide better estimates of template concentrations, thus, be a good addition to a bear DNA profiling system, but in absence of any qPCR method we have simply used the results from the two first STR loci typed to roughly judge the quality of the DNA extract (DNA-concentration, degree of degradation) in each sample. The measurements of sensitivity showed that there were no large differences among the markers validated. They all worked well in a large range of template concentrations. The use of two initial STRs to roughly judge the quality of the DNA extract would be unwise if there were any large differences in sensitivity among the markers. However, the similar sensitivity range of the validated STRs indicates, that if failing in the two initial STRs due to a low sample DNA concentration, it is likely that this sample will also fail in many of the other markers, and, thus, supports the use of results from the initial STRs as an indicator of template success rate in the following STRs.

The measurements of sensitivity as well as the other performance measurements (see Section 4.2) are from results using a singleplex PCR amplification as this is the current procedure used when analyzing the validated loci at our laboratory. While the aim of this study was not to develop new multiplex assays, a multiplex PCR that combine several of our validated and also widely used bear STRs into single PCR reactions would be valuable to both population monitoring and forensic use of these markers (less time consuming and more cost effective). A recent study has reported the performance of a multiplex PCR that included eight of the validated STRs in this study [9]. Likewise, we have shown that an octa-plex PCR consisting of eight of the most commonly used markers in conservation genetics and population monitoring may be successfully amplified in DNA extracts from tissue, hair and faeces, and at approximately same sensitivity levels as in the singleplex PCRs. This demonstrates that subsets of the validated loci may be combined into multiplex assays. Preferably, such multiplex assays should not perform much less than the singleplex methods, and the measured performance levels from the respective validated singleplex assays in this study should serve as reference values when evaluating performance and loci-balance in multiplex assays of the validated STRs.

4.2. Precision, stutter and heterozygote balance

There is a correlation between the size of the standard deviation and the power to discriminate between alleles of different size. If the bin range is equal to three standard deviations and the bin ranges do not overlap, 99.7% of identical alleles are sized within the same bin [23]. The measurements of precision in this study indicate that alleles with 2 bp size differences can be discriminated with a confidence of more than 99.7% in all loci. Seven of the loci also show a between-run precision that allows a similar high discrimination of alleles with only 1 bp size differences ($3 \times$ S.D. <0.50 bp). Within-run measurements of the least performing

locus, G10L, revealed that the precision improved considerably when eliminating the between-run factors (from 0.30 bp to 0.16 bp). A similar improvement was demonstrated at locus G10B (from 0.15 bp to 0.08 bp).

Allelic ladders are usually applied to control the between-run factors from affecting precision. In our study we have applied four positive controls as allelic ladders and the bin windows were adjusted according to the size measurements of the alleles in these control samples. Given the demonstrated precision and the typing procedure applied in our study, it is likely that any common alleles with size differences of 1 bp would be detected. Such size differences were observed at locus MU23, only. Thus, micro-variation (1 bp size variation) was not common in our North European brown bear populations at the other loci tested. The micro-variation at MU23 contributes to the size diversity at this locus, and we believe that if applying a similiar analyzing system as in this study that has the power to discriminate between 1 bp size differences MU23 could still be included in a brown bear DNA profiling system.

Stutter is a common artifact inherent in STR amplification. The height of stutter peaks depend on the type of repeat units (di, tri, tetra, penta) and is expected to be much more pronounced in dinucleotides than in tetranucleotides. The heterozygote balance at dinucleotide loci is also expected to differ substantially from the ones observed in tetranucleotide loci with larger differences in peak heights between alleles [24–27]. In agreement with these expectations the stutter ratios observed in the thirteen dinucleotide STRs were considerably larger than those reported in tetranucleotide loci and they seem to increase with number of repeats within each locus. The imbalance between alleles in a heterozygous genotype was also substantially larger than in tetranucleotide loci and there was a direction of the imbalance with shorter alleles being the ones with largest peak heights.

The presence of three or more alleles in any of the markers would be required to distinguish between a mixture of two (or more) individuals and a single source sample (one individual). Although not tested in particular, the heterozygosity level of the validated markers supports that in a mixture of two or more unrelated individuals there would usually be three alleles in at least one of the loci. As in other DNA profiling systems consisting of a combination of STRs, the minimal number of individuals in a mixture may be estimated on basis of maximum number of alleles observed in any of the markers. However, given the elevated ratios of stutter and heterozygote balance in the dinucleotide loci, we believe that interpretation of the individual genotypes in a mixture would be difficult if not impossible.

There are two common types of genotyping errors in single source samples associated with large stutter ratios and large variation in heterozygote balance. One such error would be failure of detecting a drop-out of a larger allele. This would lead to mistyping a heterozygous genotype as a homozygous genotype. Such errors could occur if the interval between the homozygote peak height threshold (threshold for accepting a true homozygous genotype) and the analytical threshold (detection limit, usually 50 RFU) does not reflect the heterozygote balance ratios. The larger alleles are usually above 60% peak heights of the shorter alleles in tetranucleotide loci and the homozygote peak height threshold used are about 150–200 RFU [25]. The measurements of heterozygote balance ratio in the dinucleotide STRs showed that peak heights of the larger alleles were usually in the range 40–70% of the shorter allele, but never less than 30% (a heterozygote balance ratio of 3.3). The more pronounced differences in allelic imbalance indicate that larger homozygote peak height threshold values should be applied when genotyping the dinucleotide loci. In our study we used 300 RFU to assure that a too narrow interval between homozygote peak height and analytical peak height thresholds did not lead to this kind of genotyping errors.

The other common type of genotyping error would be failure to distinguish between a heterozygous genotype where the smaller allele has the lesser peak height and a homozygous genotype with a large stutter. Such errors could occur at a locus if the values of heterozygote balance ratios and stutter ratios overlap, e.g. if the heterozygote balance ratio are less than one (shorter allele has lesser peak height than larger alleles in a true heterozygous genotype) and the stutter ratios are high (close to one). However, heterozygote balance ratios were usually well above one (Table 2), and the upper 95 percentiles of stutter and the lower 95 percentiles of heterozygote ratios did not overlap at twelve of the loci while one locus performed slightly less (MU09). Thus, usually the random variation in stutter and heterozygote ratios does not result in overlapping values. Genotyping should, however, be carried out manually by expert analysts with detailed knowledge of the expected heterozygote balance and stutter ratios in the STRs used. Improved confidence in allele designation may also be achieved by independent analysis of duplicates from each sample. A procedure, as applied in our study, where all samples are typed in duplicate and all homozygote genotypes are typed in triplicate would further limit the chance that imprecision, random variation of stutter ratios or heterozygote balance in a single typing would lead to errors in allele designation.

MU09 showed a small overlap between stutter and heterozygote balance values. Thus at this locus there are particular challenges when genotyping heterozygous individuals with alleles that differ by one repeat. If using markers like MU09 the expert analyst should, in our opinion, pay particular attention to allele designation of such genotypes. The heterozygote balance ratios at MU09 were in some cases less than one, but most true heterozygous samples would show a value larger than one in at least one of the two individual analyses of the sample, thus, recognized as a true heterozygous genotype (or if in doubt even typed a third time). In rare occasions a true a heterozygous genotype would show heterozygote balance values less than one in two independent analyses. If so, it would be regarded as a potential homozygote and typed independently a third time. The chance that a true heterozygous genotype would show heterozygote balance ratios that could be interpreted as stutter by chance in three independent analyses would be very rare. However, if this should occur, that all three analyses resulted in values that are in the overlap between stutter and heterozygote balance, the sample should simply be scored as an ambiguous genotype and not included in the DNA-profile given for this sample. We believe that such ''challenging'' loci may be used if the protocols and standard operating procedures documents the individual performance of the loci used and provide threshold guidelines to assure that if values of stutter and heterozygote ratio are at levels that results in ambiguous genotypes they are not included in the DNA-profile.

4.3. Variation of repeat numbers and repeat array structure

Allele designations used in the population genetic studies of brown bear are usually based on allele size measurements. The designation to a certain allele size depends on the primers and electrophoretic systems (as well as size standards and polymere) used at the laboratory. Thus, identical alleles are usually designated with different sizes at different laboratories. If a laboratory applies an in-house allelic ladder, any change in nomenclature would not itself improve the confidence in allele designation. However, it makes inter laboratory co-operation and sharing of population frequency data less complicated. Thus, a common nomenclature and inter laboratory calibration based on sharing of standard DNA samples would be a benefit for the conservation genetic community as well as for the use of these markers for forensic purposes.

The size variation in the alleles from ten of the STR loci validated in our study seemed to depend on variation of number of repeats, only. A nomenclature where allele designation is based on number of repeats in a given allele could therefore, as suggested by the forensic community [13,15], be a unique and better way to designate these alleles. Three loci showed indels in addition to the variation in number of repeats, and more alleles (allelic ladders) from all loci should be sequenced before details of the system for allele nomenclature at single loci are suggested. To make such a change to a consensus nomenclature valuable it should be a collaborative task in which laboratories that routinely genotype European brown bear participate and agrees on implementing a common nomenclature. Thus, our samples with the sequenced alleles are potential standards that may be used in an inter laboratory calibration which may represent the first steps towards a change to a common nomenclature.

The species differences revealed by comparison to the Genbank sequence from black bear as well as the polymorphisms discovered at some loci must be taken into account if designing new primers to amplify these microsatellites in brown bear. The species differences, if confirmed as fixed differences by typing more individuals from both species, have the potential to be utilized for species specific amplification of the markers.

4.4. Heterozygote deficiency at locus MU26

Null-alleles as well as genotyping errors caused by stutter (interpreting a heterozygous genotype as a homozygous genotype with large stutter) or drop-out of large alleles could lead to higher than expected presence of false homozygous genotypes. However, when testing the expected versus observed distribution of homozygous genotypes in P1, P4 and P7 within allele size classes at locus MU26 the frequencies of homozygous genotypes were larger than expected in most of the allele sizes. If the homozygote excess was caused by scoring error due to stutter or large allele drop-out one would expect a deficiency and excess of particular genotypes [21]. Thus, there was no evidence that the heterozygote deficiency was caused by genotyping errors.

Heterozygote deficiency in all size classes could also result from inbreeding. However, inbreeding is expected to affect the whole genome, and there was a substantial difference in the inbreeding coefficient observed at MU26 (0.22) and the average inbreeding coefficient from the other 12 microsatellite markers (0.005). Thus, although there is a certain degree of inbreeding in some populations, we believe that null-alleles are the main cause of the extreme heterozygote deficit observed at MU26.

A common cause of null-alleles is polymorphisms located at the primer sites [28]. However, there was no amplification of locus MU26 when re-amplifying this locus by use of new primer pairs and template DNA from putative null-allele homozygous individuals. Thus, the cause of null-alleles remains unknown. A large deletion of the entire locus may, however, be a possible explanation.

4.5. Population data

The locus heterozygosity and number of alleles observed at the twelve loci in the total population (see Table 3) indicates that all loci are highly polymorphic. In wildlife species like the brown bear the ideal conditions that must be assumed for Hardy–Weinberg equilibrium (e.g. infinite population size, random mating) does not exist. Nevertheless, the tests for HWE demonstrated that there were few significant deviations from expected heterozygosity at twelve of the loci, while it lead to the identification of null-alleles at locus MU26. None of the remaining twelve loci were in linkage disequilibirium in all populations. Thus, in lack of any mapping information, the assumption that the loci were not closely linked was not invalidated. There were moderate levels of population substructure (average F_{ST} 0.09) and significant levels of inbreeding were revealed in only two populations. This suggests that a general theta adjustment of 0.09 could be used in match probability estimates in the Northern European brown bear meta-population. An inclusion of an estimate of inbreeding in the match probability could, as suggested in Dawney et al. [29], be included at homozygote loci. However, if the alternative hypothesis implies that the alternative matching individual is from the same geographical location we believe that a sibling match probability, representing the most conservative match estimate, could also be reported since home range overlap is positively correlated with relatedness in brown bears [2].

In general, the level of population structure increased with distance between the populations, although this was not consistent for all single comparisons. The population pairwise F_{ST}'s (Supplementary File 3) indicated close relationship between the populations of Kainuu and Russian Karelia (P7 and P8, $F_{ST} = 0.0086$). In contrast, there was substantially larger differentiation between North-eastern Norway and North-western Norway (P1 and P2, $F_{ST} = 0.1033$) that may not have been caused by distance alone. The estimates of population substructure and inbreeding were included when estimating the total average probability of identity and total average probability of sibling identity [20]. The magnitude of these estimates indicates that a DNA profiling system applying the twelve STRs (G1D, G10B, G10L, G1A, MU05, MU09, MU10, MU15, MU23, MU50, MU51 and MU59) would provide individual specific DNA profiles.

Conflict of interest

None.

Acknowledgements

We would like to thank Siv Grete Aarnes, Beate Banken Bakke Ingrid Helle Danielsen and Mari Bergsvåg for skilful technical assistance in the laboratory. Also, we thank the State Nature Inspectorate in Norway, the Västerbotten County Administration in Sweden and the Finnish Hunters' Association in Finland, and all others that have contributed to collecting samples. Financial support for this project has been provided by the Norwegian Ministry of Environment, the Directorate for Nature Management and the Finnmark County Governor's Office in Norway.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2012.03.002.

References

[1] N.J. Ouborg, C. Pertoldi, V. Loeschcke, R.K. Bijlsma, P.W. Hedrick, Conservation genetics in transition to conservation genomics, Trends Genet. 26 (2010) 177–187.
[2] J.E. Swenson, P. Taberlet, E. Bellemain, Genetics and conservation of European brown bears *Ursus arctos*, Mammal Rev. 41 (2011) 87–98.
[3] P. Taberlet, J.J. Camarra, S. Griffin, E. Uhrés, O. Hanotte, L.P. Waits, C. Dubois-Paganon, T. Burke, J. Bouvet, Noninvasive genetic tracking of the endangered Pyrenean brown bear population, Mol. Ecol. 6 (1997) 869–876.
[4] D. Paetkau, C. Strobeck, Microsatellite analysis of genetic variation in black bear populations, Mol. Ecol. 3 (1994) 489–495.
[5] L.P. Waits, D. Paetkau, Noninvasive genetic sampling tools for wildlife biologist: a review of applications and recommendations for accurate data collection, J. Wildlife Manage. 69 (2005) 1419–1433.

[6] P.M. Lukacs, K.P. Burnham, Review of capture–recapture methods applicable to noninvasive genetic sampling, Mol. Ecol. 14 (2005) 3909–3919.
[7] M.K. Schwartz, G. Luikart, R.S. Waples, Genetic monitoring as a promising tool for conservation and management, Trends Ecol. Evol. 22 (2006) 25–33.
[8] M.E. Smith, H.G. Eiken, L.E. Ollila, C. Tobiassen, S.G. Bjervamoen, P.E. Aspholm, I. Wartiainen, Hair snares applied to detect brown bears in the vicinity of farms in the Pasvik Valley 2008, Bioforsk Report 169 (2008), http://www.barentswatch.com/innhold/fauna/bjorn/Bioforsk_Rapport_169_2008_hair_snare_EN.pdf.
[9] T. Skribinsek, M. Jelenicic, L. Waits, I. Kos, P. Trontelj, Highly efficient multiplex PCR of noninvasive DNA does not require pre-amplification, Mol. Ecol. Res. 10 (2010) 495–501.
[10] A. Kopatz, H.G. Eiken, S.B. Hagen, M. Ruokonen, R. Esparza-Salas, J. Schregel, I. Kojola, M.E. Smith, I. Wartiainen, P.E. Aspholm, et al., Connectivity and population subdivision at the fringe of a large brown bear (*Ursus arctos*) population in North Western Europe, Conserv. Genet. (2011), doi:10.1007/s10592-012-0317-2.
[11] E. Tammeleht, J. Remm, M. Korsten, J. Davidson, I. Tumanov, A. Saveljev, P. Männil, I. Kojola, U. Saarma, Genetic structure in large, continuous mammal populations: the example of brown bears in North-western Eurasia, Mol. Ecol. 19 (2010) 5359–5370.
[12] E. Bellemain, J.E. Swenson, O. Tallmon, S. Brunberg, P. Taberlet, Estimating population size of elusive animals with DNA from hunter-collected feces: four methods for brown bears, Conserv. Biol. 19 (2005) 150–161.
[13] B. Budowle, P. Garofano, A. Hellman, M. Ketchum, S. Kanthaswamy, W. Parson, W. vanHaeringen, S. Fain, T. Broad, Reccomendations for animal DNA forensic and identity testing, Int. J. Legal Med. 119 (2005) 295–302.
[14] R. Ogden, N. Dawnay, R. McEwing, Wildlife DNA forensics-bridging the gap between conservation genetics and law enforcement, Endang. Species Res. 9 (2009) 179–195.
[15] A. Linacre, L. Gusmao, W. Hech, A.P. Hellmann, W.R. Mayr, W. Parson, M. Prinz, P.M. Schneider, N. Morling, ISFG: reccomendations regarding the use of non-human (animal) DNA in forensic genetic investigations, Forensic Sci. Int. Genet. 5 (2011) 501–505.
[16] M.J. Brownstein, J.D. Carpten, J.R. Smith, Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping, Biotechniques 20 (1996) 1004–1010.
[17] P.O. Lewis, D. Zaykin, Genetic data analysis: computer program for the analysis of allelic data, Version 1.0 (2001) http://lewis.eeb.uconn.edu/lewishome/software.html.
[18] L. Excoffier, H.E.L. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (2010) 564–567.
[19] PowerStats, Promega; http://www.promega.com/genetictools/powerstats/Default.htm.
[20] K.L. Ayres, A.D.J. Overall, API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives, Mol. Ecol. Notes 4 (2004) 315–318.
[21] C. VanOosterhout, W.F. Hutchinson, D.P.M. Wills, P.P. Shipley, MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data, Mol. Ecol. Notes 4 (2004) 535–538.
[22] D. Paetkau, S.C. Amstrup, E.W. Born, W. Calvert, A.E. Derocher, G.W. Garner, F. Messier, I. Stirling, M.K. Taylor, O. Wiig, C. Strobeck, Genetic structure of the world's polar bear populations, Mol. Ecol. 8 (1999) 1571–1584.
[23] H.M. Wenz, J.M. Robertson, S. Menchen, F. Oaks, D.M. Demorest, D. Scheibler, B.B. Rosenblum, C. Wike, D.A. Gilbert, J.W. Efcavitch, High-precision genotyping by denaturing capillary electrophoresis, Genome Res. 8 (1998) 69–80.
[24] A. Debernardi, E. Suzanne, A. Formant, L. Pene, A.B. Dufour, J.R. Lobry, One year variability of peak heights, heterozygous balance and inter-locus balance for the DNA positive control of AmpF*l*STR© Identifiler© STR kit, Forensic Sci. Int. Genet. 5 (2010) 43–49.
[25] P. Gill, R. Sparkes, C. Kimpton, Development of guidelines to designate alleles using an STR multiplex system, Forensic Sci. Int. 89 (1997) 185–197.
[26] D. Kube, D. Schmidt, M. Mormann, A.C. Uhlemann, J. Tomiuk, H. Tesch, P.G. Kremsner, Semiautomated and simultaneous analysis of the interleukin-10 gene microsatellites IL-10G and IL-10R by fluorescence-based polymerase chain reaction reveals significant differences in allele distributions between Caucasians (Germany) and Africans (Gabon), Eur. Cytokine Netw. 12 (2001) 537–544.
[27] R.J. Slebos, D.M. Umbach, C.A. Sommer, G.A. Horner, J.Y. Choi, J.A. Taylor, Analytical and statistical methods to evaluate microsatellite allelic imbalance in small amounts of DNA, Lab Invest. 84 (2004) 649–657.
[28] M. Delamoye, C. Duverneuil, K. Riva, M. Leterreux, S. Taieb, P. De Mazancourt, False homozygosities at various loci revealed by discrepancies between commercial kits: implications for genetic databases, Forensic Sci. Int. 143 (2004) 47–52.
[29] N. Dawnay, R. Ogden, R.S. Thorpe, L.C. Pope, D.A. Dawson, R. McEwing, A forensic STR profiling system for the Eurasian badger: a framework for developing profiling systems for wildlife species, Forensic Sci. Int. Genet. 2 (2008) 47–53.

PAPER II

MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 3474–3488 doi: 10.1111/j.1365-294X.2012.05631.x

Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east–west border population in the Pasvik Valley

JULIA SCHREGEL,*† ALEXANDER KOPATZ,*‡ SNORRE B. HAGEN,* HENRIK BRØSETH,§ MARTIN E. SMITH,* STEINAR WIKAN,* INGVILD WARTIAINEN,* PAUL E. ASPHOLM,* JOUNI ASPI,‡ JON E. SWENSON,†§ OLGA MAKAROVA,¶ NATALIA POLIKARPOVA,¶ MICHAEL SCHNEIDER,** PER M. KNAPPSKOG,* MINNA RUOKONEN,‡ ILPO KOJOLA,†† KONSTANTIN F. TIRRONEN,‡‡ PJOTR I. DANILOV‡‡ and HANS GEIR EIKEN*

**Bioforsk Soil and Environment, Svanhovd, Norwegian Institute for Agricultural and Environmental Research, NO-9925 Svanvik, Norway, †Department of Ecology and Natural Resources Management, University of Life Sciences, NO-1432 Ås, Norway, ‡Department of Biology, University of Oulu, PO Box 3000, FIN-90014 Oulu, Finland, §Norwegian Institute for Nature Research, NO-7485 Trondheim, Norway, ¶Pasvik Strict Nature Reserve, 184424 Rajakoski, Murmansk Region, Russia, **Swedish Environmental Protection Agency, SE-106 48 Stockholm, Sweden, ††Finnish Game and Fisheries Research Institute, Oulu Game and Fisheries Research, Rakentajantie 33, FIN-90014 Oulu, Finland, ‡‡Institute of Biology, Karelian Research Centre of the Russian Academy of Science, 185910 Petrozavodsk, Russia*

Abstract

Noninvasively collected genetic data can be used to analyse large-scale connectivity patterns among populations of large predators without disturbing them, which may contribute to unravel the species' roles in natural ecosystems and their requirements for long-term survival. The demographic history of brown bears (*Ursus arctos*) in Northern Europe indicates several extinction and recolonization events, but little is known about present gene flow between populations of the east and west. We used 12 validated microsatellite markers to analyse 1580 hair and faecal samples collected during six consecutive years (2005–2010) in the Pasvik Valley at 70°N on the border of Norway, Finland and Russia. Our results showed an overall high correlation between the annual estimates of population size (N_c), density (D), effective size (N_e) and N_e/N_c ratio. Furthermore, we observed a genetic heterogeneity of ~0.8 and high N_e/N_c ratios of ~0.6, which suggests gene flow from the east. Thus, we expanded the population genetic study to include Karelia (Russia, Finland), Västerbotten (Sweden) and Troms (Norway) (477 individuals in total) and detected four distinct genetic clusters with low migration rates among the regions. More specifically, we found that differentiation was relatively low from the Pasvik Valley towards the south and east, whereas, in contrast, moderately high pairwise F_{ST} values (0.91–0.12) were detected between the east and the west. Our results indicate ongoing limits to gene flow towards the west, and the existence of barriers to migration between eastern and western brown bear populations in Northern Europe.

Keywords: capture–mark–recapture, DNA, effective population size, microsatellites, migration rates, N_e/N_c ratio, noninvasive genetic sampling, population structure

Received 11 October 2011; revision received 26 March 2012; accepted 6 April 2012

Correspondence: Julia Schregel, Fax: +47 78 99 56 00; E-mail: julia.schregel@bioforsk.no

Introduction

Noninvasive genetic methods, based on the analysis of hair and faecal samples, are increasingly used in wildlife biology as a feasible and cost-effective tool to monitor large carnivore populations (Bellemain *et al.* 2005; Proctor *et al.* 2010). The approach has great potential value in the conservation and the management of large carnivores, as it might enable the estimation of important population parameters from genetic data, that is, census and effective population sizes, population density, genetic diversity, degree of inbreeding and gene flow among populations (Quéméré *et al.* 2010; Roberts *et al.* 2011; Wang *et al.* 2011). In particular, knowledge about the connectivity among populations assessed by estimating the degree of genetic differentiation and gene flow among populations is important as it is believed to counteract the effects of genetic drift (Mills *et al.* 2003) and be strongly linked to the long-term viability of populations (Schwartz *et al.* 2002; Long *et al.* 2005).

There is little knowledge about the diversity and connectivity of large carnivore populations, especially across national borders and on larger scales (Dalerum *et al.* 2009). One reason is that large predators were almost extirpated in Western Europe and much of North America (Enserink & Vogel 2006; Dalerum *et al.* 2009). Conflicts with humans and the resulting persecution and habitat destruction, combined with life history traits, such as large home ranges, long dispersal distances and long generation times, make large predators vulnerable (Crooks 2002) as has been shown recently for brown bears (*Ursus arctos*) (Miller & Waits 2003; Proctor *et al.* 2005; Kendall *et al.* 2009). Large predators are now recovered in many places. Elucidating their functions in natural ecosystems and requirements for long-term survival has become a major research interest (Smith *et al.* 2003; Estes *et al.* 2011). In this context, noninvasive genetic data on large-scale connectivity patterns among populations of large predators may contribute to the conservation and management of these species without disturbing them. This is especially important as invasive methods, like capturing and equipping animals with GPS-collars, have several drawbacks. It has been shown for brown and black bears (*Ursus americanus*) that trapping may have long-term negative effects, such as reduced body condition (Cattet *et al.* 2008). GPS-tagging is also expensive and cannot reveal large-scale biological patterns involving numerous individuals and populations.

The demographic history of brown bears in Northern Europe indicates several extinction and recolonization events (Swenson *et al.* 1995; Danilov 2005). In Norway and Sweden, the brown bear population nearly went extinct during the 19th and 20th centuries due primarily to state-financed persecution. The species was functionally extirpated in Norway, whereas three to four small and separate relict populations survived in Sweden (Swenson *et al.* 1995). This historical population fragmentation also is evident in the current genetic population structure (Waits *et al.* 2000; Manel *et al.* 2004). In Finland and northwestern Russia, similar bottlenecks have been recorded for brown bears from observations and hunting statistics (Pulliainen 1990; Ermala 2003; Danilov 2005). The genetic connectivity among these and other brown bear populations in Northern Europe is not clear. In particular, we lack information about the gene flow between the westernmost brown bear populations of Norway and Sweden and the eastern ones of Russia and Finland. A recent genetic study of brown bear populations from six different geographical areas in Finland, Estonia and Russia suggested large-scale gene flow from Finland far into southeastern European Russia, whereas the more southern populations formed three distinct genetic clusters (Tammeleht *et al.* 2010). Moreover, a phylogenetic study of mitochondrial DNA determined a common maternal lineage among four different brown bear haplotypes in northern Eurasia, indicating the historical existence of a large, genetically uniform group throughout the area (Korsten *et al.* 2009). In a recent study, we found a more restricted pattern of effective migration and gene flow among the populations in the region (Kopatz *et al.* 2012). However, the gene flow between the western and eastern parts of the Northern European brown bear populations still remains to be understood.

In this study, we have used noninvasively obtained genetic data from the brown bear population in the Pasvik Valley at the border of Norway, Finland and Russia to investigate the degree of genetic connectivity between western and eastern brown bear populations in Northern Europe. To address this issue, we have studied the Pasvik bear population's genetic structure, connectivity and variability in relation to a regional area, including the bear populations of Karelia (Russia, Finland), Västerbotten (Sweden) and Troms (Norway). Thus, our study includes individuals of both the westernmost brown bear populations in Northern Europe as well as the eastern brown bear populations.

Population size (N_c), effective population size (N_e) and the ratio between them are important indicators of population viability (Luikart *et al.* 2010). We used data from the Pasvik Valley during 6 years (2005–2010) to determine the magnitude and between-year variation in the N_e/N_c ratio. The N_e/N_c ratio might allow us to infer N_e from N_c (and vice versa) and be useful for planning management actions to increase N_e (Ficetola *et al.* 2010; Brekke *et al.* 2011).

Material and methods

Study areas

Samples were collected at four different locations in Northern Europe (Fig. 1). The focus population was located in the Pasvik Valley at the border between Norway, Finland and Russia (~70°N, 30°E) and the study area encompassed ~5000 km^2. The three other sampling areas were located to the west and south of Pasvik Valley: (i) Troms County Norway, ~420 km to the west, ~70°N, 20°E, encompassing ~5000 km^2; (ii) Västerbotten County, Sweden, ~725 km to the southwest, ~65°N, 17°E, encompassing ~45 000 km^2; and (iii) the transboundary area in Karelia (Finland and Russia), ~600 km to the south, ~64–60°N, 30–37°E and encompassing ~130 000 km^2. The airline distances between the study areas Troms, Västerbotten and Karelia are as follows: Troms-Västerbotten: ~460 km, Troms—Karelia: ~830 km, Västerbotten—Karelia ~680 km.

Sampling

Hair and faecal samples were collected opportunistically in the field (Pasvik from 2005 to 2010, Troms in 2006, 2008 and 2009; Västerbotten in 2009; Karelia in 2005 and 2007, Table 1). In 2007 and 2008, additional hair samples were obtained from the Pasvik population, using hair snares placed systematically in geographical grids, with trap design, collection protocol and lure composition adapted from previous studies (Kendall 1999; Woods *et al.* 1999; Romain-Bondi *et al.* 2004). In 2007, we used 56 traps for 2 months in a 5 km × 5 km grid, and in 2008, we used 20 traps for 1 month in a 2.5 km × 2.5 km grid. Additionally, to further increase the coverage, we included tissue samples from legally harvested bears (Table 1). Brown bear monitoring in the Pasvik Valley is included in both the Norwegian Large Predator Monitoring Program and as a part of the management of a certified transboundary park (Europarc Federation). In Västerbotten, a county-wide brown bear faecal collection programme was conducted for population estimation, during which a large number of individuals (N = 270) was detected. This population was recently estimated to consist of around 300 individuals (Kindberg *et al.* 2011), and thus, we sampled ~90% of the population. To minimize the risk of family over-representation, which can bias the results of the algorithms used for the population structure analyses (Anderson & Dunham 2008), we used a subset of 84 individuals for statistical testing. To avoid large families and at the same time ensure sufficient geographical and gender distribution, we selected randomly three males and three females from each municipality in Västerbotten in this subset.

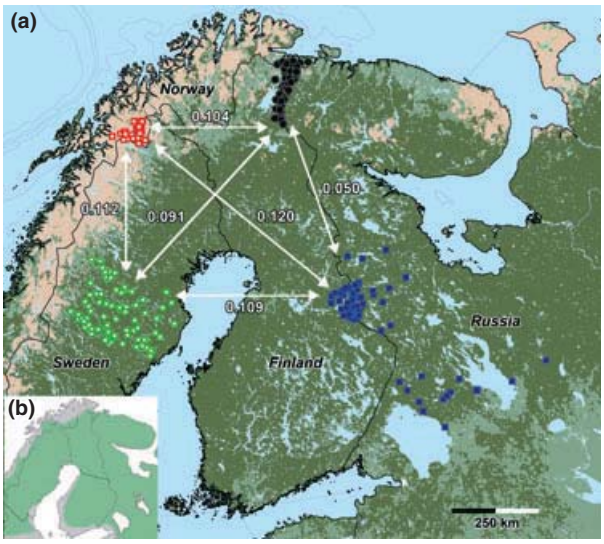


Fig. 1 (a) The four sampling locations in Northern Europe and pairwise F_{ST} values among them. Each mark represents the average position of a genotyped brown bear. Black filled circles: Pasvik (n = 94), red open squares: Troms (n = 34), green open circles: Västerbotten (n = 84) and blue filled squares: Karelia (n = 79). The map legend is as follows: blue = water bodies; dark green = forest cover; light green = brush/scrub/grassland; light brown = tundra. All F_{ST} values are significant, the arrows indicate the pairs of populations compared. (b) Map showing brown bear distribution across Northern Europe. Green = area with possible brown bear occurrence (see also http://www.lcie.org), dashed line = southern border of the reindeer husbandry area in the three Nordic countries.

Molecular analysis

Faecal samples were stored in stool collection tubes with DNA stabilizer (Invitek) or in plastic bags and kept at minus 20 °C until DNA extraction. The hair samples were stored dry and dark in paper envelopes until DNA extraction. To extract DNA, we used the PSP Spin Stool DNA Plus Kit (Invitek) for the faecal samples and the DNeasy Tissue Kit (Qiagen) for the hair and tissue samples, following the manufacturers' instructions. We used the following 12 dinucleotide markers (short tandem repeats, STRs) to genotype the DNA samples: G1A, G1D, G10B, G10L (developed for the black bear; Paetkau & Strobeck 1994; Paetkau *et al.* 1995; Paetkau & Strobeck 1995); Mu05, Mu09, Mu10, Mu15, Mu23, Mu50, Mu51 and Mu59 (developed for the brown bear; Taberlet *et al.* 1997). All of the STRs used here have been validated with respect to species specificity, sensitivity, accuracy and probability of identity (Eiken *et al.* 2009; Andreassen *et al.* 2012). Sex was determined as described by Kopatz *et al.* (2012).

A detailed description of PCR protocols and the fragment analysis as well as protocols for individual

Table 1 Brown bear sample collection and genetic analyses* from four locations in Northern Europe

	Pasvik (2005–2010)	Troms (2006, 2008–2009)	Västerbotten (2009)	Karelia (2005–2007)	Total
No. of samples	1580	307	1355	123	3365
Faeces	1180	239	1346	89	2854
Hair	92	67	3	0	162
Hair from hair traps†	281	0	0	0	281
Tissue	27	1	6	34	68
No. of samples genotyped*	901	178	914	113	2106
No. of males	54	19	138	49	260
No. of females	37	15	131	29	212
n.d.	3	0	1	1	5
No. of bears	94	34	270‡	79	477

n.d., not determined.
*Genotyping was performed using 12 different STRs and an amelogenin gene XY-assay (see Materials and methods).
†Only for 2007 and 2008.
‡From Västerbotten, only a subset of 84 individuals was used in the population genetic analyses (see Materials and methods), while the remaining individuals were typed for only 8 STRs and gender in this study.

identification can be found in Andreassen *et al.* (2012). In this study, the genetic analysis was performed as follows. PCR mixes were set up with 10 μl reaction volumes and contained 1× PCR Gold buffer (ABI), 200 μM dNTP (Eurogentec), 1.5 mM MgCl2 (ABI), 0.5 μM of each primer (MedProbe Inc.), 1 U AmpliTaq-Gold DNA polymerase (ABI), 1× BSA (NEB) and 1 μl template DNA. The conditions for PCRs for the loci G1A, MU10, MU05, MU09, MU23, MU50, MU51, MU59 and G10L were 10 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 58 °C and 1 min at 72 °C. A final extension phase was set for 15 min at 72 °C on an ABI 2720. PCR conditions for loci G1D, G10B and MU15 were similar, except for a higher annealing temperature of 60 °C, and a shorter final extension of 5 min. PCR products were run on an ABI 3730, and the PCR fragments were analysed with GeneMapper 4.0 (Applied Biosystems).

The first and the last four samples on every 96-well plate were positive controls, and every eighth sample was a negative control. The positive controls functioned also as a control for between-run variation; all genotypes were assigned manually. The samples were genotyped independently twice if allele designation showed a heterozygote and three times if it showed a homozygous genotype for the specific markers (peak height threshold values >300 RFU). A sample was only assigned an identity if all runs across all markers were consistent. If not, an identity was not assigned and the sample was discarded from further analyses, and, accordingly, we did not construct consensus DNA profiles. We only accepted a single negative result for STRs if the sample showed consistent results for the overall DNA profile. PCRs for sex determination were run twice with positive controls. Our procedures followed the strict guidelines for forensic examination of animal DNA material, which are in accordance with the requirements published by Linacre *et al.* (2011). The laboratory procedures, that is, the extraction of samples and the analysis of the STRs, were accredited according to the EN ISO/IEC 17025 standard. The uniqueness of the DNA profiles was verified by calculating the probability of identity of each sample using the software Gimlet version 1.3.3 (Valiere 2002). Tests for allelic dropout, presence of null alleles and scoring errors caused by stutter peaks were performed with Micro-Checker version 2.2.3 (Van Oosterhout *et al.* 2004).

Statistical analysis

Genetic diversity, inbreeding and linkage disequilibrium (LD). We used the software Genetix 4.05.2 (Belkhir *et al.* 1996–2004) to calculate observed and expected heterozygosities, allele numbers, inbreeding coefficients and LD for all sampled locations. As implemented by Genetix, we tested for LD between pairs of loci for all areas using the method of Black & Kraftsur (1985).

We used Genepop version 4.0.11 (Rousset 2008) to run the exact test for deviations from Hardy–Weinberg equilibrium (HWE) for all loci and geographical locations. All combinations of locations were tested with unbiased *P* values by a Markov chain method of 1000

burn-in iterations, 500 batches and 1000 iterations per batch.

Population bottlenecks. We used the software BOTTLENECK v. 1.2.02 (Cornuet & Luikart 1997; Luikart *et al.* 1998; Piry *et al.* 1999) to test for genetic signatures of a demographic bottleneck, that is, whether the heterozygosity in the studied populations was larger than the heterozygosity expected from the number of alleles found in the sample if the population were at mutation drift equilibrium. We applied the two-phase mutation model using 95% single-step mutations to estimate the expected heterozygosities (20 000 iterations). Significance of the differences between observed and expected heterozygosities was tested using the Wilcoxon test.

Population structure. We analysed population structure using both population- and individual-based approaches. First, we utilized the Bayesian approach to detect the number of genetic clusters (K) using the software STRUCTURE version 2.3.3 (Pritchard *et al.* 2000; Falush *et al.* 2003; Hubisz *et al.* 2009). For this analysis, we assumed population admixture and correlated allele frequencies within the population. To achieve consistency of results, we performed ten independent runs for each K value (number of genetic clusters) between one and ten. For each run, we set a burn-in period of 100 000 Markov Chain Monte Carlo (MCMC) iterations, followed by sampling of 1 000 000 iterations. Because the log-likelihood estimated with the STRUCTURE software often displays higher variance between runs for the higher K values, we calculated the rate of change in the log probability of data between successive K values (ΔK) to determine the most likely number of clusters (Evanno *et al.* 2005).

In a second step, we used an individually based spatially explicit model implemented in the software GENELAND version 3.2.4 (Guillot *et al.* 2005). We ran five independent runs, where the parameters for possible populations were K = 1–10, and the number of MCMC iterations was 10 000 000, with a thinning of 100. The maximum rate of Poisson process was set to 100, and the maximum number of nuclei was 300.

Finally, to visualize the extent of regional differentiation, we ran a factorial correspondence analysis (FCA) with GENETIX 4.05.2 (Belkhir *et al.* 1996–2004). We also used the software ARLEQUIN version 3.5.1.2 (Excoffier & Lischer 2010) to calculate pairwise F_{ST} values (Weir & Cockerham 1984) among detected populations with 10 000 burn-in iterations, 100 batches and 500 iterations per batch. We also ran an analysis of molecular variance (AMOVA) to identify genetic structure among and within populations, using 10 000 permutations.

Migration rates among populations. To estimate migration rates among the four populations, we used a Bayesian approach implemented in the software BAYESASS 1.3 (Wilson & Rannala 2003). Contrary to the classical methods (Paetkau *et al.* 1995; Rannala & Mountain 1997; Cornuet *et al.* 1999), this approach may provide rates of recent migration among populations. The number of burn-in iterations was set to 6 000 000 followed by 3 000 000 iterations and a thinning of 2000. Initial input parameters of allele frequencies, migration and inbreeding coefficient were set at 0.15 for each, respectively. As recommended, we adjusted the delta values to 0.07 (allele frequency), 0.05 (inbreeding coefficient) and 0.15 (migration), so that acceptance rates for changes in these parameters would be between 40 and 60% (Faubet *et al.* 2007). We carried out three independent runs to confirm the consistency of results. To examine differences between the sexes, the same analysis with the same settings was run with data sets split according to sex. Individual membership values q_i, estimated in the population structure analysis with the program Structure, can indicate possible migrants. Therefore, individuals with a q_i value >0.7 for a different population than the one it was sampled in was recorded to identify possible migrants.

Annual estimates of population size (N_c), density (D) and effective population size (N_e) for Pasvik 2005–2010. We used the DNA-based single session capture–mark–recapture (CMR) method to estimate N_c and N_e, as it has been shown to work well with capture heterogeneity and small population sizes (Miller *et al.* 2005) and has also been compared with and found more efficient than other field-based methods (Solberg *et al.* 2006). To avoid biased estimates and to maximize both the detection and sampling frequencies of individuals, we used the combined data of the opportunistic and systematic sampling approaches to estimate N_c and N_e (Boulanger *et al.* 2008; Gervasi *et al.* 2008; De Barba *et al.* 2010). The annual estimates of N_c were made using both *Capwire* (Miller *et al.* 2005), based on the two innate rates model (TIRM) and using ordered samples (Miller *et al.* 2005, 2007; Bromaghin 2007), and CAPTURE (Otis *et al.* 1978), based on the M_h Chao (a closed-population heterogeneity estimator). To estimate population density (D), we first estimated annual effective sampling areas to correct for geographical closure violation by creating a concave buffer around each sample location. As no home-range estimates were available for bears in the Pasvik population, we applied both an upper and a lower buffer: (i) a wide buffer of 15 km around the samples, equivalent to a circular home-range size of 707 km^2; and (ii) a narrow buffer of 7.5 km, equivalent to a circular home-range

size of 177 km^2. The upper and lower buffers were based on home-range sizes estimated from telemetry data of males and females, respectively, from neighbouring populations in Sweden (Dahle & Swenson 2003; Støen *et al.* 2006). In addition, the mean maximum distance (MMD) between resampling events (Obbard *et al.* 2010) and the equivalent circular home-range sizes of individual bears were determined for individuals with at least five resampling events during a year in Pasvik.

The effective population size N_e is an indicator of the factors affecting the strength of inbreeding and genetic drift processes (Wright 1931, 1938). N_e was estimated annually with the software LDN$_e$ (Waples & Do 2008), which is based on LD data. The method uses the principle that, with declining N_e, LD is generated by genetic drift and thus LD can be used to calculate N_e (Hill 1981). We also calculated N_e with the online software OneSamp (Tallmon *et al.* 2008), which utilizes approximate Bayesian computation and allows user-specified priors. We tested for consistency using differing priors (minimum and maximum effective population size) in the analysis settings (Tallmon *et al.* 2008).

To determine the magnitude and stability of the N_e/N_c ratio across years, we calculated the N_e/N_c ratio for all 6 years for the Pasvik population. In this context, we also tested for a correlation between N_e and N_c across years, to test the hypothesis that N_e may be estimated from N_c (and vice versa), using the Pearson's product-moment correlation implemented in the software R (R Development Core Team 2011). We also used the same function to test for correlation between the different estimators used for the estimation of N_c and N_e across the years.

Results

Sampling and genetic analysis

In total, 3365 samples were collected for genetic analyses in the four regions (Fig. 1 and Table 1). In each region, the vast majority of samples were faecal samples collected opportunistically, followed by hair and tissue samples. Systematic hair trapping was performed only in the Pasvik Valley. Successful genotyping with 12 different STRs was obtained for 2106 samples from 477 different bears: Pasvik, n = 94; Troms, n = 34; Västerbotten, n = 270; Karelia, n = 79 (Fig. 1a and Table 1). The number of bears identified annually in Pasvik in 2005–2010 ranged from 27 in 2005 to 44 in 2007 (Table 2). In 2007 and 2008 in Pasvik, several individuals (2007: 19 and 2008: 3) were detected with the hair traps only. The effect was most pronounced in 2007, when the hair trap area was largest (1400 km^2).

Genetic diversity, inbreeding and LD

We determined the expected and observed heterozygosities, the number of different alleles and the inbreeding coefficient F_{IS} for all 12 STRs for 290 individuals (Table 3). Deviations from HWE ($P < 0.05$) were observed in 8 of 48 tests, although after Bonferroni correction, only one marker (G10B) in the Karelia population deviated significantly from HWE (Table 3). Mean H_{exp} ranged from 0.68 (Troms) to 0.82 (Karelia), and mean H_{obs} from 0.69 (Västerbotten) to 0.80 (Pasvik). Mean F_{IS} values ranged from −0.02 in Pasvik to 0.04 in Västerbotten, whereas the only significant F_{IS} value was

Table 2 Annual estimates of census population size (N_c) and density of brown bears in the Pasvik Valley (2005–2010). Estimates of census population size (N_c) using both the two innate rates model (TIRM) and the M_h Chao estimator are shown. For population density (D) estimates, the N_c estimates were corrected for geographical closure by first estimating the effective sampling area with two different buffer widths around each sample: Buff7.5 ≈ 177 km^2; Buff15 ≈ 707 km^2 (see Materials and Methods) Obs./ind. = - mean number of observations per individual bear; No. of ind. = number of individuals

				Census populations size N_c		Population Density Ind./1000 km^2			
				TIRM	M_h Chao	TIRM		M_h Chao	
Year	No. of samples	Obs./ind.	No. of ind.	N_c (CI 95%)	N_c (CI 95%)	Buff7.5	Buff15	Buff7.5	Buff15
2005	68	2.52	27	36 (27–49)	39 (31–70)	11.1	4.8	12.1	5.2
2006	50	2.08	24	39 (25–57)	41 (29–86)	12.3	4.5	12.9	4.8
2007	141	3.20	44	56 (46–66)	67 (52–112)	12.3	6.2	14.7	7.4
2008	144	3.89	37	46 (37–53)	53 (43–80)	14.5	7.2	16.7	8.3
2009	137	4.42	31	33 (31–36)	43 (35–79)	9.9	4.7	12.8	6.1
2010	80	3.48	23*	27 (23–33)	29 (25–54)	8.6	4.9	9.3	5.2
Mean	103	3.27	31	39.5	45.3	11.5	5.4	13.1	6.2

*One individual only represented by a tissue sample was deleted from the data set.

Table 3 Expected, (H_{exp}) and observed (H_{obs}) heterozygosities, number of different alleles (A) and inbreeding values (F_{IS}) calculated for the 12 short tandem repeats in four Northern European brown bear populations

	Pasvik 2005–2010 (n = 93)					Troms 2006, 2008–2009 (n = 34)					Västerbotten 2009 (n = 84)					Karelia 2005–2007(n = 79)				
Marker	A	H_{exp}	H_{obs}	P	F_{IS}	A	H_{exp}	H_{obs}	P	F_{IS}	A	H_{exp}	H_{obs}	P	F_{IS}	A	H_{exp}	H_{obs}	P	F_{IS}
G1D	9	0.83	0.87	0.411	−0.04	5	0.70	0.71	0.166	0.00	6	0.67	0.64	0.152	0.05	9	0.82	0.81	0.069	0.01
G10B	10	0.76	0.76	**0.022**	0.00	6	0.68	0.76	0.244	−0.10	7	0.62	0.64	0.141	−0.03	11	0.85	0.65	**0.001***	**0.25****
Mu05	8	0.82	0.89	**0.035**	−0.08	6	0.70	0.71	0.560	0.01	6	0.64	0.62	0.589	0.04	8	0.79	0.77	0.128	0.04
Mu09	13	0.84	0.85	0.427	0.00	6	0.68	0.76	0.942	−0.11	9	0.82	0.77	0.509	0.07	9	0.87	0.85	0.453	0.04
Mu15	6	0.76	0.82	0.492	−0.07	4	0.39	0.38	0.828	0.04	5	0.62	0.64	**0.016**	−0.04	9	0.80	0.85	0.486	−0.05
G1A	8	0.80	0.80	0.665	0.01	5	0.66	0.76	0.769	−0.15	5	0.70	0.63	0.206	0.10	9	0.80	0.80	0.094	0.01
G10L	9	0.64	0.63	0.284	0.02	6	0.80	0.82	0.154	−0.02	7	0.64	0.60	**0.033**	0.07	10	0.78	0.75	0.339	0.05
Mu10	8	0.73	0.75	0.416	−0.01	4	0.67	0.74	0.900	−0.09	7	0.74	0.67	0.535	0.11	10	0.79	0.76	0.279	0.05
Mu23	9	0.69	0.73	**0.028**	−0.05	5	0.73	0.76	0.654	−0.03	8	0.81	0.77	0.737	0.06	9	0.84	0.68	**0.014**	0.19
Mu50	8	0.84	0.86	0.484	−0.02	6	0.73	0.85	0.333	−0.16	7	0.81	0.80	0.861	0.02	8	0.74	0.71	0.583	0.05
Mu51	8	0.82	0.84	0.066	−0.01	6	0.77	0.82	0.405	−0.05	7	0.76	0.80	0.632	−0.05	10	0.83	0.77	0.510	0.07
Mu59	9	0.82	0.80	0.116	0.03	5	0.70	0.76	0.844	−0.07	9	0.82	0.75	**0.006**	0.10	13	0.90	0.92	0.465	−0.01
Mean	8.75	0.78	0.80		−0.02	5.33	0.68	0.74		−0.06	6.92	0.72	0.69		0.04	9.58	0.82	0.78		0.02

Significant deviations from Hardy–Weinberg equilibrium ($P < 0.05$) are marked in bold.
*The only significant deviation after Bonferroni correction.
**The only significant F_{IS} value. $P < 0.05$.

detected for the marker G10B in the Karelian population (Table 3). The highest number of alleles for a single STR was 13 (MU09 in Pasvik and MU59 in Karelia), and the mean numbers of alleles for all STRs were highest in Karelia (9.6) and lowest in Troms (5.3). After sequential Bonferroni correction, significant LD was found in 52 of 66 marker pairs, with 37 of these observed in Pasvik. None of the remaining 15 marker pairs showed significant LD in more than two populations and were not the same in all of the sampled populations.

Population bottlenecks

Allele frequencies showed no signs of a genetic bottleneck in any of the tested populations. All tests for heterozygote excess were negative (Wilcoxon test; $P > 0.190$ for all populations).

Population structure

The four methods we used to test for genetic differentiation resulted in the same four genetic clusters. First, the Bayesian approach in the program Structure found the highest mean likelihood [lnP(D)] for $K = 4$ (Fig. 2a,b), after correction using Evanno's ad-hoc approach (Evanno *et al.* 2005), as did the software GENELAND with geographical coordinates and a priori correlated allele frequencies (results not shown). Similarly, the visualization of the extent of regional differentiation with FCA suggested four clusters, with the first axis explaining 5.7% and the second axis explaining 3.8% of the variation (Fig. 3). Pairwise F_{ST} values between populations ranged from 0.050 (between Pasvik and Karelia) and 0.120 (between Karelia and Troms), and the overall average substructuring was 0.1 (Table 4). All these comparisons were statistically significant ($P \leq 0.01$) (Table 4). AMOVA analysis revealed that 9.18% of the genetic variation was among, and 90.82% was within, the populations ($P < 0.001$).

Migration

We estimated high rates of self-recruitment in each population and low migration rates among the four locations using the Bayesian method (Table 5). The estimated rates of self-recruitment were high, ranging from 94.1% to 98.9%. Our results indicated that 96.3% of the bears sampled in Pasvik originated from the same population, and only 3.7% of the individuals originated from the other three populations. The highest estimated migration rates were found from Västerbotten to Troms (4.7%), from Karelia to Pasvik (2.3%), and from Västerbotten to Pasvik (1.2%). The lowest migration rates were found from Karelia to Västerbotten (0.1%), from Troms to Pasvik (0.2%) and Karelia (0.2%) and from Västerbotten to Karelia (0.2%). However, these differences in migration rates were not significant. Running these same analyses for males and females separately showed no signs of differences in migration rates between the sexes (data not shown). In total, eight individuals were identified as

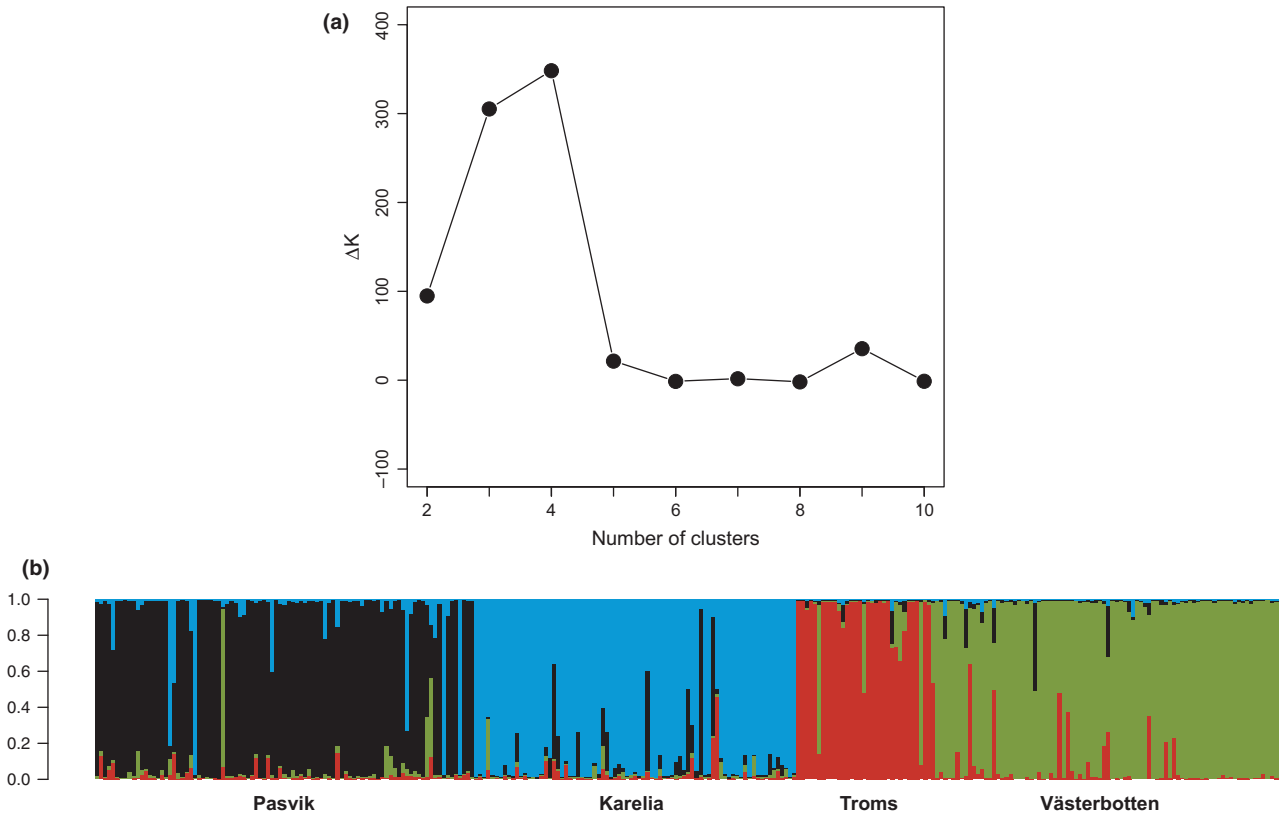


Fig. 2 (a) Population structure analysis of individual brown bear genotypes (12 STRs) from four locations in Northern Europe using the program Structure. Results were processed with the Evanno approach, *x*-axis: No. of clusters, *y*-axis: Δ*K*. (b) Population structure analysis of individual brown bear genotypes (12 STRs) from four locations in Northern Europe using the program Structure, individual admixture for $K = 4$, each bar represents one individual partitioned into segments, the length of each segment corresponds to the individual membership value (q_i).

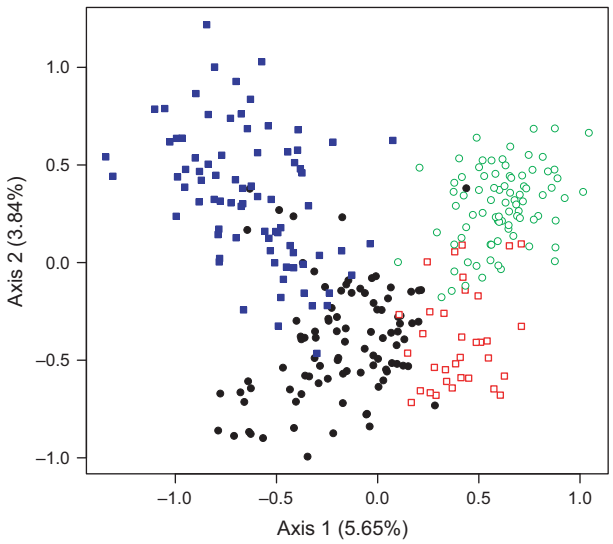


Fig. 3 Factorial correspondence analysis of individual brown bear genotypes (12 short tandem repeats) from four different geographical locations in Northern Europe; blue filled squares = Karelia, black filled circles = Pasvik, red open square = Troms, green open circle = Västerbotten.

possible migrants by having an estimated membership value $q_i > 0.7$ in Structure. In Pasvik, six individuals had $q_i > 0.7$. Five of these were assigned to Karelia (three males, one female and one of unknown sex). The last one was assigned to Västerbotten (a male). In addition, a male from Troms was assigned to Västerbotten, and a male from Karelia was assigned to Pasvik.

Annual estimated population size (N_c), density and effective population size (N_e) for Pasvik 2005–2010

The CMR estimates of N_c from the TIRM model and the M_h Chao were similar, with a correlation value of $r = 0.969$ ($P = 0.001$) between the two estimators across the years (Table 2). Both methods showed the largest N_c value in 2007 and the lowest value in 2010. Mean estimated population size in Pasvik was 39.5 (TIRM) and 45.3 (M_h Chao), ranging from 27 to 56 (TIRM) and from 29 to 67 (M_h Chao) between years (Table 2). The mean number of observations per individual ranged from 2.1 in 2006 to 4.4 in 2009.

Table 4 Pairwise F_{ST} values (ARLEQUIN 3.11) for brown bears from four locations in Northern Europe

	Pasvik	Troms	Västerbotten
Troms	0.104*		
Västerbotten	0.091*	0.112*	
Karelia	0.050*	0.120*	0.109*

*$P < 0.01$.

The estimated mean maximum distance (MMD) between resampling events was 21.7 km ($n = 46$), which corresponds with a circular home-range size of 370 km^2. Thus, this was slightly lower than the mean of the upper and lower buffers (442 km^2) that refer to previous telemetry-determined home-range sizes from Sweden.

The estimated effective sampling areas in the Pasvik Valley were used to estimate densities (Table 2). We found overall mean population densities between 5.4 and 13.1 individuals/1000 km^2. Annual population density estimates (D) using the N_c TIRM results ranged from 4.5 to 7.2 (Buff7.5) and from 8.6 to 14.5 (Buff15) individuals/1000 km^2, whereas densities ranged from 4.8 to 8.3 (Buff15) and from 9.3 to 16.7 (Buff7.5) individuals/1000 km^2, using the N_c M_h Chao results.

The results obtained with differing priors were consistent in the ONESAMP model, so we proceeded using only the results obtained with the a priori information of N_e min = 2 and N_e max = 100. The LDN_e estimates of N_e ranged from 9.1 (in 2006) to 21.1 (in 2007), with a mean of 13.5 individuals, and the ONESAMP estimates of N_e were from 18.1 (in 2009) to 36.9 (in 2007), with a mean of 25.1 individuals (Table 6). The correlation value between the LDN_e and the ONESAMP methods across all years was $r = 0.520$ ($P = 0.290$).

We found a significant correlation across years between the N_c estimates and the ONESAMP estimates of N_e ($r = 0.858$; $P = 0.029$ (TIRM) and 0.815; $P = 0.048$ [M_h Chao)], but there was no significant correlation between N_c and LDN_e estimates ($r = 0.618$; $P = 0.191$ (TIRM) and 0.561; $P = 0.247$ [M_h Chao)]. However, we have detected high LD in this population sample, and the LDN_e method is not recommended when closely related individuals are sampled. Therefore, annual N_e/N_c ratios were calculated only with the N_e estimate based on the ONESAMP method. The annual N_e/N_c ratios ranged from 0.53 to 0.82, with a mean of 0.64 (N_c from TIRM), and from 0.42 to 0.76, with a mean of 0.57 (N_c from M_h Chao) (Table 6). Note, however, that the corresponding estimates of the N_e/N_c ratio based on LDN_e, rather than the ONESAMP, would have been considerably smaller, as the LDN_e estimates of N_e were always considerably smaller than the ONESAMP estimates of N_e (Table 6).

Discussion

Using the Pasvik and three surrounding populations, we investigated the genetic diversity and the gene flow among western and eastern brown bear populations in Northern Europe using mainly noninvasive genetic sampling. We found four distinct genetic clusters with low migration rates among the populations. The overall results of the study indicate present limitations to gene flow between the eastern and western populations. The high genetic heterogeneity we found among bears in the Pasvik Valley, on the border between east and west, is comparable with results from Kirov in central Russia (Tammeleht *et al.* 2010), Pinega in Archangelsk in northwestern Russia (Kopatz *et al.* 2012), and Karelia in southwestern Russia (this study). Thus, we suggest that the Pasvik population has genetic contact with other bear populations to the east. In contrast, we detected substantial substructuring among our study populations with moderately high F_{ST} values and separate genetic clusters between east and west. F_{ST} values have been found to be substantially lower between populations that show some degree of bidirectional migration (Waits *et al.* 2000; Proctor *et al.* 2005; Kendall *et al.* 2009) than between subpopulations separated by direct barriers to migration (Proctor *et al.* 2005). Our results are somewhere between these two extremes, and both, the geographical distances and the degree of genetic differentiation, may be comparable to studies on the two most distant subpopulations in Sweden more than 10 years ago (Waits *et al.* 2000). At that time, the authors proposed that the observed population substructuring was

Table 5 Bayesian analysis (BAYESASS 1.3) of migration rates and self-recruitment among four brown bear populations in Northern Europe

	Pasvik	Troms	Västerbotten	Karelia
From Pasvik to	0.963 (0.93–0.98)	0.006 (0.00–0.03)	0.005 (0.00–0.01)	0.006 (0.00–0.02)
From Troms to	0.002 (0.00–0.01)	0.941 (0.87–0.98)	0.005 (0.00–0.02)	0.002 (0.00–0.01)
From Västerbotten to	0.012 (0.00–0.03)	0.047 (0.01–0.10)	0.989 (0.96–1.00)	0.002 (0.00–0.01)
From Karelia to	0.023 (0.01–0.05)	0.006 (0.00–0.02)	0.001 (0.00–0.01)	0.990 (0.96–1.00)

The 95% CIs are given in brackets.

Table 6 Annual estimates of effective population size (N_e) and the N_e/N_c ratio in brown bears in the Pasvik Valley, 2005–2010. The estimates for N_e using the ONESAMP method were applied to calculate N_e/N_c ratios using the two innate rates model (N_{c1}) and the M_h Chao (N_{c2}) N_c estimates

Year	n	LDN_e (P_{crit} = 0.05) N_e (CI 95%)	ONESAMP N_e (CI 95%)	N_e/N_{c1}	N_e/N_{c2}
2005	27	9.1 (7.2–11.6)	25.3 (21.5–33.9)	0.703	0.649
2006	24	18 (13.1–26.1)	20.5 (18.5–24.4)	0.526	0.500
2007	44	21.1 (17.4–26.0)	36.9 (33.5–42.0)	0.659	0.551
2008	36*	10.2 (8.3–12.3)	27.4 (24.4–33.4)	0.596	0.517
2009	31	10.7 (8.1–14.4)	18.1 (16.1–22.8)	0.548	0.421
2010	24	12 (9.2–15.9)	22.1 (19.1–28.3)	0.819	0.762
Mean	31	13.5	25.1	0.642	0.567

n, number of detected individuals by DNA analysis.
*One individual with a partial genotype was deleted from the data set.

because of a pattern of isolation by distance (IBD) and residual genetic differentiation caused by the 19th and 20th century bottleneck event and subsequent population fragmentation. Recently, Tammeleht *et al.* (2010) proposed that brown bears in northeastern Europe also are structured by IBD. Thus, the observed substructuring in this study may have resulted from a combination of IBD and the demographic history of the northern European bear populations. However, our results do not indicate any bottlenecks and are clearly suggestive of limited gene flow in the region, especially towards the west. A high rate of self-recruitment in all of our study areas suggests that barriers hinder migration. The geographical distances between the Pasvik bears and the other bear populations are similar, but substructuring is definitely more pronounced towards the west than towards the south. In comparison, we detected a relatively low degree of substructuring with the bears in Pinega, which are twice as far to the east (see Kopatz *et al.* 2012). Thus, apparently, additional mechanisms than merely spatial distance are necessary to fully explain the genetic differentiation among the bear populations in Northern Europe.

Migration between Karelia and Pasvik might be aided by the relatively undisturbed area along the Russian border. In this area, the 'Fennoscandian Green Belt' (Karivalo & Butorin 2006), transborder movements of bears have been recorded previously (Pulliainen 1990; Swenson & Wikan 1996; Kojola *et al.* 2003), but more precise, recent information about these movements is not available. In comparison, to move between Pasvik and Västerbotten, migrating bears would have to cross an area with reindeer husbandry. The reindeer husbandry area (Fig. 1b) has been suggested to constitute a migration barrier for the northern European wolf population (*Canis lupus*), because of illegal hunting (Wabakken *et al.* 2001; Vilà *et al.* 2003; Kojola *et al.* 2006), which also may be the case for brown bears. Recent studies have shown that illegal killing has a substantial effect on the large carnivore populations in the region and occurs more frequently in northern than in southern Scandinavia (Andrén *et al.* 2006; Persson *et al.* 2009; Liberg *et al.* 2012). Also, in some areas, migration to the east may be hindered by border fences from Soviet times. They are located all along the Russian border and are believed to act as barriers to large carnivore migration (e.g. for wolves, Aspi *et al.* 2009).

The limited availability of suitable habitat also might reduce gene flow among the populations. Generally, brown bears are adaptable to a wide range of habitat types, although in northern Europe, they seem to prefer rugged, forested terrain (Nellemann *et al.* 2007) at lower elevations (May *et al.* 2008). Thus, the relatively open areas of high-elevation tundra, scrub and brush west of the Pasvik Valley might impede migration between western and eastern populations. Also, human disturbance seems to be a major factor influencing home-range selection (Nellemann *et al.* 2007), and although the density of human settlements is generally low in the study area, it is possible that other human activities, such as forest industry (logging, forest roads, etc.), may have a negative impact on migration in some areas. However, the impact of these factors is not known, and we suggest that such possible barriers to migration and gene flow may be investigated in future studies.

Results from studies on polar bears (*Ursus maritimus*) and wolverine (*Gulo gulo*), equally able to disperse across large distances, have shown similar patterns of reduced gene flow because of barriers to migration (Paetkau *et al.* 1999; Kyle & Strobeck 2001, 2002; Cegelski *et al.* 2006). Although the results are not directly comparable, because of differences in, for example, life history traits and ecological requirements, these

similarities suggest that long-distance dispersal is apparently strongly influenced by the quality of the habitat to be crossed and can be easily disturbed by unfavourable circumstances.

In the Pasvik Valley, we detected bears in densities that were comparable to that found in northern Sweden (~10 bears/1000 km^2), but lower than in southern Sweden (Støen *et al.* 2006). These densities seem to be lower than in North America (Mowat *et al.* 2005; Kendall *et al.* 2008), although these differences may be due to differences in methodology. Danilov (2005) estimated substantially higher densities in Russian Karelia, but these estimates are based only on hunting bag and bear observations. However, densities may vary among the areas. Brown bears are not distributed evenly across Northern Europe, and core and peripheral areas are identifiable (Swenson *et al.* 1998; Kojola & Laitala 2000; Kojola *et al.* 2003; Kojola & Heikkinen 2006; Kindberg *et al.* 2011).

Population size (N_c), effective population size (N_e) and the ratio between these measures may have a practical value in conservation, because they have been suggested to be important indicators of population viability (Ficetola *et al.* 2010; Luikart *et al.* 2010; Brekke *et al.* 2011). Estimates of these parameters in the Pasvik Valley with two different methods during six consecutive years showed little annual variation, but substantial methodological variation. We detected almost a threefold difference between the LDN_e and the OneSamp estimates for N_e. However, the LDN_e estimate may be biased low in small and extensively sampled populations because of family over-representation (Luikart *et al.* 2010) and the high number of significant linkage disequilibria found in the Pasvik population may indicate the sampling of closely related individuals (Slate & Pemberton 2007). Accordingly, we used the OneSamp method to calculate the N_e/N_c ratios. Noninvasive genetic sampling data yield no information about the age of individuals and the sampling of overlapping generations may generate a biased N_e estimate (Luikart *et al.* 2010). The result of such an N_e estimate may be somewhere between the number of breeding pairs and the effective population size (Waples 2005). The N_e results achieved by the OneSamp analysis may be the more accurate, as it uses multiple summary statistics and therefore more information from the data (Luikart *et al.* 2010). Keeping these methodological uncertainties in mind, we found substantial correlations between the census and effective population size estimates and our results indicate that the population's N_e may be estimated directly from N_c (and vice versa) as previously suggested by Brekke *et al.* (2011) and Ficetola *et al.* (2010). This must be tested using other data sets that incorporate geographical variation and potentially confounding factors, because such estimates have important practical implications.

The mean annual N_e/N_c ratio of approximately 0.6 is very high compared with other published ratios for brown bear. For North American grizzly bears, the N_e/N_c ratio ranged from 0.20–0.38 in a simulated population using a demographic estimate (Harris & Allendorf 1989) and from 0.04–0.19 using a genetic estimate (Paetkau *et al.* 1998). In southern Sweden, the ratio ranged from 0.06 to 0.14, also using a genetic estimate (Tallmon *et al.* 2004). The two latter studies were conducted in populations in which isolation and/or bottleneck events led to low heterozygosity, which may explain the very low ratio compared with our estimates. However, the demographic estimate also is lower than in our study in Pasvik. Nunney (2000) and Storz *et al.* (2002) have shown that differences in methods for estimating N_e may produce results that are not necessarily comparable, which could also cause differences in the N_e/N_c ratio. Nevertheless, our estimate does not seem to represent genetic stochastic effects, as it was relatively stable for all 6 years. As it seems that fluctuations in population size, variance in family size and unequal sex ratio have a negative impact on the N_e/N_c ratio (Frankham 1995), the relatively high ratio may indicate that Pasvik is a stable population. In addition, the N_e/N_c ratio has been shown to be generally higher in smaller populations (Palstra & Ruzzante 2008), which might apply to the Pasvik population as well. Pasvik may receive continuous migration from the east, which is indicated by the relatively high genetic heterogeneity, and this may bias the N_e upwards by causing the local N_e to approach the global or metapopulation N_e (Pray *et al.* 1996; Palstra & Ruzzante 2008; England *et al.* 2010; Luikart *et al.* 2010), and thus resulting in a higher N_e/N_c ratio. Although research regarding the usefulness of the N_e/N_c ratio is still ongoing, our results show that in an apparently stable bear population, the ratio seems to be relatively stable and may have the potential to be used in management and conservation actions (Luikart *et al.* 2010).

Noninvasive sampling schemes have been applied to different large carnivore species mostly for monitoring purposes (e.g. wolf, Marucco *et al.* 2009; tiger (*Panthera tigris*), Mondol *et al.* 2009; wolverine, Brøseth *et al.* 2010); however, our study demonstrates that noninvasively obtained genetic data may be used to investigate population genetic structure on a large spatial and temporal scale. To our knowledge, we have also been one of the first to apply this kind of data to study the relationship of N_e and N_c in large carnivores. Thus, there is only a limited amount of comparable studies, although it would be desirable to be able to compare our findings with those from similar species in the future.

Both the genetic substructuring and the N_e/N_c ratios may support the same conclusion of higher genetic variation and gene flow towards the east than the west and the apparent existence of barriers to migration between those areas. In this context, the small population of the Pasvik Valley may represent a genetic border, as the gene flow decreases towards the west and, to some degree, also to the south. Other populations, especially from Northern Norway and Sweden, as well as from the areas between Pasvik and Karelia, should be included in future studies to improve our understanding of migration routes and population structure in Northern European brown bear populations. If the reasons for the lower gene flow are poaching and fragmentation, our findings raise concerns about the future conservation of brown bear populations in Northern Europe.

Acknowledgements

We would like to thank Siv Grete Aarnes, Camilla Tobiassen, Mari Bergsvåg, Leif Ollila, Pekka Sulkava, Tuomo Ollila, Tanja Kyykkä, Bjørn Mentyjärvi, Espen Aarnes, Sari Magga, Matti Heino, Jari Ylönen, Alexander Mershiev, Vladimir V. Belkin, Nikolai L. Rendakov for excellent assistance in the laboratory and the field. Also, we thank The State Nature Inspectorate in Norway, the Finnish Hunters'Association and all others that have contributed to collecting samples and for their collaboration. Financial support for this project has been provided by the Norwegian Ministry of Environment, the Directorate for Nature Management in Norway and the Finnmark County Governor's Office in Norway.

References

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources*, **8**, 1219–1229.

Andreassen R, Schregel J, Kopatz A *et al.* (2012) A forensic DNA profiling system for Northern European brown bears (*Ursus arctos*). *Forensic Science International Genetics*, DOI: 10.1007/s12686-012-9634-5.

Andrén H, Linnell JDC, Liberg O *et al.* (2006) Survival rates and causes of mortality in Eurasian lynx (*Lynx lynx*) in multi-use landscapes. *Biological Conservation*, **131**, 23–32.

Aspi J, Roininen E, Kiiskila J *et al.* (2009) Genetic structure of the northwestern Russian wolf populations and gene flow between Russia and Finland. *Conservation Genetics*, **10**, 815–826.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier.

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

Black WC, Kraftsur ES (1985) A FORTRAN software for the calculation and analysis of two-locus linkage disequilibrium coefficients. *Theoretical and Applied Genetics*, **70**, 491–496.

Boulanger J, Kendall KC, Stetz JB *et al.* (2008) Multiple data sources improve DNA-based mark-recapture population estimates of grizzly bears. *Ecological Applications*, **18**, 577–589.

Brekke P, Bennett PM, Santure AW, Ewen JG (2011) High genetic diversity in the remnant island population of hihi and the genetic consequences of re-introduction. *Molecular Ecology*, **20**, 29–45.

Bromaghin JF (2007) The genetic mark-recapture likelihood function of capwire. *Molecular Ecology*, **16**, 4883–4884.

Brøseth H, Flagstad Ø, Wärdig C *et al.* (2010) Large-scale non-invasive genetic monitoring of wolverines using scats reveals density dependent adult survival. *Biological Conservation*, **143**, 113–120.

Cattet M, Boulanger J, Stenhouse G, Powell RA, Reynolds-Hogland ML (2008) An evaluation of long-term capture effects in ursids: implications for wildlife welfare and research. *Journal of Mammalogy*, **89**, 973–990.

Cegelski CC, Waits LP, Anderson NJ *et al.* (2006) Genetic diversity and population structure of wolverine (*Gulo gulo*) populations at the southern edge of their current distribution in North America with implications for genetic viability. *Conservation Genetics*, **7**, 197–211.

Cornuet JM, Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Crooks KR (2002) Relative sensitivities of mammalian carnivores to habitat fragmentation. *Conservation Biology*, **16**, 488–502.

Dahle B, Swenson JE (2003) Seasonal range size in relation to reproductive strategies in brown bears *Ursus arctos*. *Journal of Applied Ecology*, **72**, 660–667.

Dalerum F, Cameron EZ, Kunkel K, Somers MJ (2009) Diversity and depletions in continental carnivore guilds: implications for prioritizing global carnivore conservation. *Biology Letters*, **5**, 35–38.

Danilov PI (2005) *Game Animals of Karelia—Ecology, Resources, Management, Protection*. Nauka, Moscow.

De Barba M, Waits LP, Genovesi P *et al.* (2010) Comparing opportunistic and systematic sampling methods for non-invasive genetic monitoring of a small translocated brown bear population. *Journal of Applied Ecology*, **47**, 172–181.

Eiken HG, Andreassen RJ, Kopatz A *et al.* (2009) Population data for 12 STR loci in Northern European brown bear (*Ursus arctos*) and application of DNA profiles for forensic casework. *Forensic Science International: Genetics Supplement Series*, **2**, 273–274.

England PR, Luikart G, Waples RS (2010) Early detection of population fragmentation using linkage disequilibrium estimation of effective population size. *Conservation Genetics*, **11**, 2425–2430.

Enserink M, Vogel G (2006) Wildlife conservation—The carnivore comeback. *Science*, **314**, 746–749.

Ermala A (2003) A survey of large predators in Finland during the 19th–20th centuries. *Acta Zoologica Lituanica*, **13**, 15–20.

Estes JA, Terborgh J, Brashares JS *et al.* (2011) Trophic downgrading of planet earth. *Science*, **333**, 301–306.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res*, **10**, 564–567.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology*, **16**, 1149–1166.

Ficetola GF, Padoa-Schioppa E, Wang J, Garner TWJ (2010) Polygyny, census and effective population size in the threatened frog, *Rana latastei*. *Animal Conservation*, **13**, 82–89.

Frankham R (1995) Effective population-size adult-population size ratios in wildlife—a review. *Genetical Research*, **66**, 95–107.

Gervasi V, Ciucci P, Boulanger J *et al.* (2008) A preliminary estimate of the apennine brown bear population size based on hair-snag sampling and multiple data source mark–recapture huggins models. *Ursus*, **19**, 105–121.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Harris RB, Allendorf FW (1989) Genetically effective population-size of large mammals—an assessment of estimators. *Conservation Biology*, **3**, 181–191.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetic Resources*, **38**, 209–216.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Karivalo L, Butorin A (2006) The fennoscandian green belt. In: *The Green Belt of Europe-from Vision to Reality* (eds Terry A, Ullrich K and Riecken U), pp. 37–45. IUCN, Gland.

Kendall KC (1999) Sampling grizzlies with noninvasive techniques. *Natural Park Service Resource Year in Review*, **1999**,20–22.

Kendall KC, Stetz JB, Roon DA *et al.* (2008) Grizzly Bear Density in Glacier National Park, Montana. *Journal of Wildlife Management*, **72**, 1693–1705.

Kendall KC, Stetz JB, Boulanger J *et al.* (2009) Demography and genetic structure of a recovering grizzly bear Population. *Journal of Wildlife Management*, **73**, 3–17.

Kindberg J, Swenson JE, Ericsson G *et al.* (2011) Estimating population size and trends of the Swedish brown bear *Ursus arctos* population. *Wildlife Biology*, **17**, 114–123.

Kojola I, Heikkinen S (2006) The structure of the expanded brown bear population at the edge of the Finnish range. *Annales Zoologici Fennici*, **43**, 258–262.

Kojola I, Laitala HM (2000) Changes in the structure of an increasing brown bear population with distance from core areas: another example of presaturation female dispersal? *Annales Zoologici Fennici*, **37**, 59–64.

Kojola I, Danilov PI, Laitala HM, Belkin V, Yakimov A (2003) Brown bear population structure in core and periphery: analysis of hunting statistics from Russian Karelia and Finland. *Ursus*, **14**, 17–20.

Kojola I, Aspi J, Hakala A *et al.* (2006) Dispersal in an expanding wolf population in Finland. *Journal of Mammalogy*, **87**, 281–286.

Kopatz A, Eiken HG, Hagen S *et al.* (2012) Connectivity and population subdivision at the fringe of a large brown bear (*Ursus arctos*) population in North Western Europe. *Conservation Genetics.*, **13**, 681–692.

Korsten M, Ho SYW, Davison J *et al.* (2009) Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Molecular Ecology*, **18**, 1963–1979.

Kyle CJ, Strobeck C (2001) Genetic structure of North American wolverine (*Gulo gulo*) populations. *Molecular Ecology*, **10**, 337–347.

Kyle CJ, Strobeck C (2002) Connectivity of peripheral and core populations of North American wolverines. *Journal of Mammalogy*, **83**, 1141–1150.

Liberg O, Chapron G, Wabakken P *et al.* (2012) Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. *Proceedings of the Biological Science*, **279**, 910–915.

Linacre A, Gusmão L, Hecht W *et al.* (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Science International-Genetics*, **5**, 501–505.

Long ES, Diefenbach DR, Rosenberry CS, Wallingford BD, Grund MRD (2005) Forest cover influences dispersal distance of white-tailed deer. *Journal of Mammalogy*, **86**, 623–629.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.

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

Manel S, Bellemain E, Swenson JE, Francois O (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Molecular Ecology*, **13**, 1327–1331.

Marucco F, Pletcher DH, Boitani L *et al.* (2009) Wolf survival and population trend using non-invasive capture-recapture techniques in the Western Alps. *Journal of Applied Ecology*, **46**, 1003–1010.

May R, van Dijk J, Wabakken P *et al.* (2008) Habitat differentiation within the large-carnivore community of Norway's multiple-use landscapes. *Journal of Applied Ecology*, **45**, 1382–1391.

Miller CR, Waits LP (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation. *Proceedings of the National Academy of Sciences USA*, **100**, 4334–4339.

Miller CR, Joyce P, Waits LP (2005) A new method for estimating the size of small populations from genetic mark-recapture data. *Molecular Ecology*, **14**, 1991–2005.

Miller CR, Joyce P, Waits LP (2007) Ordered vs. unordered samples: response to Bromaghin. *Molecular Ecology*, **16**, 4885.

Mills LS, Schwartz MK, Tallmon DA, Lair KP (2003) Measuring and interpreting connectivity for mammals in coniferous forests. In: *Mammal Community Dynamics: Management and Conservation in the Coniferous Forests of Western North America* (eds Zabel CJ and Anthony RG), pp. 587–613. Cambridge University Press, Cambridge.

Mondol S, Karanth KU, Kumar NS *et al.* (2009) Evaluation of non-invasive genetic sampling methods for estimating tiger population size. *Biological Conservation*, **142**, 2350–2360.

Mowat G, Heard DC, Seip DR *et al.* (2005) Grizzly *Ursus arctos* and black bear *U. americanus* densities in the interior mountains of North America. *Wildlife Biology*, **11**, 31–48.

Nellemann C, Støen O-G, Kindberg J *et al.* (2007) Terrain use by an expanding brown bear population in relation to age, recreational resorts and human settlements. *Biological Conservation*, **138**, 157–165.

Nunney L (2000) The limits to knowledge in conservation genetics—The value of effective population size. *Evolutionary Biology*, **32**, 179–194.

Obbard ME, Howe EJ, Kyle CJ (2010) Empirical comparison of density estimators for large carnivores. *Journal of Applied Ecology*, **47**, 76–84.

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, **62**, 7–135.

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

Paetkau D, Strobeck C (1995) Microsatellite analysis of genetic-variation in black-bear populations. *Molecular Ecology*, **4**, 133.

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, Shields GF, Strobeck C (1998) Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Molecular Ecology*, **7**, 1283–1292.

Paetkau D, Amstrup SC, Born EW *et al.* (1999) Genetic structure of the world's polar bear populations. *Molecular Ecology*, **8**, 1571–1584.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428–3447.

Persson J, Ericsson G, Segerstrom P (2009) Human caused mortality in the endangered Scandinavian wolverine population. *Biological Conservation*, **142**, 325–331.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pray LA, Goodnight CJ, Stevens L, Schwartz JM, Yan GY (1996) The effect of population size on effective population size: an empirical study in the red flour beetle *Tribolium castaneum*. *Genetical Research*, **68**, 151–155.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Proctor MF, McLellan BN, Strobeck C, Barclay RM (2005) Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proceedings of the Biological Science*, **272**, 2409–2416.

Proctor M, McLellan B, Boulanger J *et al.* (2010) Ecological investigations of grizzly bears in Canada using DNA from hair, 1995–2005: a review of methods and progress. *Ursus*, **21**, 169–188.

Pulliainen E (1990) Recolonization of Finland by the brown bear in the 1970s and 1980s. *Aquilo Seria Zoologica*, **27**, 21–25.

Quéméré E, Crouau-Roy B, Rabarivola C, Louis EE Jr, Chikhi L (2010) Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. *Molecular Ecology*, **19**, 1606–1621.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Available from: http://www.R-project.org/

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences USA*, **94**, 9197–9201.

Roberts DG, Baker J, Perrin C (2011) Population genetic structure of the endangered Eastern Bristlebird *Dasyornis brachypterus*; implications for conservation. *Conservation Genetics*, **12**, 1075–1085.

Romain-Bondi KA, Wielgus RB, Waits L *et al.* (2004) Density and population size estimates for North Cascade grizzly bears using DNA hair-sampling techniques. *Biological Conservation*, **117**, 417–428.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Schwartz MK, Mills LS, McKelvey KS, Ruggiero LF, Allendorf FW (2002) DNA reveals high dispersal synchronizing the population dynamics of Canada lynx. *Nature*, **415**, 520–522.

Slate J, Pemberton JM (2007) Admixture and patterns of linkage disequilibrium in a free-living vertebrate population. *Journal of Evolutionary Biology*, **20**, 1415–1427.

Smith DW, Peterson RO, Houston DB (2003) Yellowstone after wolves. *BioScience*, **53**, 330–340.

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

Støen OG, Zedrosser A, Saebo S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears *Ursus arctos*. *Oecologia*, **148**, 356–364.

Storz JF, Ramakrishnan U, Alberts SC (2002) Genetic effective size of a wild primate population: influence of current and historical demography. *Evolution*, **56**, 817–829.

Swenson JE, Wikan S (1996) A brown bear population estimate for Finnmark County, North Norway. *Fauna norvegica Ser. A*, **17**, 11–15.

Swenson JE, Wabakken P, Sandegren F *et al.* (1995) The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology*, **1**, 11–25.

Swenson JE, Sandegren F, Söderberg A (1998) Geographic expansion of an increasing brown bear population: evidence for presaturation dispersal. *Journal of Animal Ecology*, **67**, 819–826.

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

Tallmon DA, Bellemain E, Swenson JE, Taberlet P (2004) Genetic monitoring of scandinavian brown bear effective population size and immigration. *Journal of Wildlife Management*, **68**, 960–965.

Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, **8**, 299–301.

Tammeleht E, Remm J, Korsten M *et al.* (2010) Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Molecular Ecology*, **19**, 5359–5370.

Valiere N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377–379.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Vilà C, Sundqvist AK, Flagstad O *et al.* (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 91–97.

Wabakken P, Sand H, Liberg O, Bjarvall A (2001) The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978–1998. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **79**, 710–725.

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

Wang IJ, Johnson JR, Johnson BB, Shaffer HB (2011) Effective population size is strongly correlated with breeding pond size in the endangered Californian tiger salamander, *Ambystoma californiense*. *Conservation Genetics*, **12**, 911–920.

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, **14**, 3335–3352.

Waples RS, Do C (2008) LDNe: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population-structure. *Evolution*, **38**, 1358–1370.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.

Woods JG, Paetkau D, Lewis D *et al.* (1999) Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, **27**, 616–627.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright S (1938) Size of populations and breeding structure in relation to evolution. *Science*, **87**, 430–431.

The study is part of a larger study on brown bear genetics and ecology carried out at the research section at Bioforsk Svanhovd in Norway, lead by S.B.H and H.G.E. The research work is also part of the PhD studies of J.S. and A.K. J.A., P.M.K. and M.R. are senior geneticists with interests in conservation of animal species. The research work of P.E.A., O.M., N.P. and K.F.T. is focused on northern forest ecosystems. H.B., M.E.S., M.S., S.W., I.W. and I.K. work on monitoring and conservation genetics of large carnivores in Northern Europe. J.E.S. is a Professor of Ecology and Natural Resources Management and P.I.D. is a Professor of Biology.

Data accessibility

For microsatellite genotype data and sampling location per identified individual, please see Table S1 in the supporting materials provided with the online version of this article.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary of the microsatellite genotype data and sampling location per identified individual.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

PAPER III

Present and past Y chromosomes reveal the demographic and genetic impact of male dispersal during the recovery of the Northern European brown bear (*Ursus arctos*)

Julia Schregel[1,2], Hans Geir Eiken[1], Finn Audun Grøndahl[3], Frank Hailer[4,10], Jouni Aspi[4], Ilpo Kojola[6], Konstantin Tirronen[7], Piotr Danilov[7], Alexander Rykov[8], Eugene Poroshin[9], Axel Janke[10,11], Jon E. Swenson[2,12], Snorre B. Hagen[1]

[1] Norwegian Institute for Bioeconomics, NIBIO - Svanhovd, 9925 Svanvik, Norway
[2] Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, 1432 Ås, Norway
[3] Randsfjordmuseene Ltd., Kongevegen 92, 2770 Jaren, Norway
[4] School of Biosciences, Cardiff University, Cardiff CF10 3AX, Wales, United Kingdom
[5] Department of Genetics and Physiology, University of Oulu, P.O. Box 3000, 90014 Oulu, Finland
[6] Natural Resources Institute, P. O. Box 16, 96301 Rovaniemi, Finland
[7] Institute of Biology, Karelian Research Centre of the Russian Academy of Science, 185910 Petrozavodsk, Russian Federation
[8] Pinezhsky Strict Nature Reserve, Pervomayskaja 123a, Pinega, Arkhangelsk Region, Russian Federation
[9] Institute of Biology, Komi Research Centre of the Russian Academy of Science, 016761 Syktvkar, Russian Federation
[10] Biodiversity and Climate Research Centre (BiK-F), Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, 60325 Frankfurt am Main, Germany
[11] Goethe University Frankfurt, Institute for Ecology, Evolution & Diversity, 60438 Frankfurt am Main, Germany
[12] Norwegian Institute for Nature Research, 7485 Trondheim, Norway

Corresponding authors:
Julia Schregel, e-mail: julia.schregel@bioforsk.no;
Phone: +47 92 06 58 64; fax: +47 78 99 56 00
Snorre B. Hagen, e-mail: snorre.hagen@bioforsk.no;
Phone: +47 93 24 01 97; fax: +47 78 99 56 00

Key-words:

Y-SNP, Y-STR, haplogroups, haplotypes, gene flow, male dispersal

Abstract

High-resolution, male-inherited Y-chromosomal markers are a useful tool for population genetic analyses of wildlife species, but to date have only been applied in this context to relatively few species besides humans. Using nine Y-chromosomal microsatellite and three Y-chromosomal single nucleotide polymorphism markers (Y-SNPs), we studied the impact of male gene flow in populations of the brown bear (*Ursus arctos*) in Northern Europe, where the species declined dramatically in numbers and geographic distribution during the last centuries but is expanding now. We found 36 haplotypes in 443 male extant brown bears from Sweden, Norway, Finland and Northwestern Russia. In 14 individuals from southern Norway from 1780 to 1920, we found two Y haplotypes present in the extant population as well as four Y haplotypes not present among the modern samples. Our results suggested major differences in genetic connectivity, diversity, and structure between the eastern and the western populations in Northern Europe. In the west, our results indicated that the recovered population originated from only four male lines, displaying pronounced spatial structuring suggestive of strong limits to male gene flow. In the east, we found a contrasting pattern, with high haplotype diversity and admixture. This first population genetic analysis of male brown bears shows conclusively that male gene flow had little impact on demographic recovery but a large impact on genetic recovery.

Introduction

The genetic structure and diversity of a population are greatly influenced by gene flow. It has been shown in a large number of studies that lack of immigration, and thus lack of gene flow from outside the population, can lead to a significant reduction of population viability (e.g. Dunn *et al.* 2011; Hedrick *et al.* 2014). The concept of "genetic rescue" is based on the prevention or reversal of this situation and an increase in gene flow has been shown to lead to higher fitness and demographic growth (La Haye *et al.* 2012; Heber *et al.* 2013). Most mammals display male-biased dispersal, i.e. males disperse farther and more frequently from the natal home range than females. Therefore males may display higher levels of gene flow among regions than do females (Greenwood 1980), resulting in males generally exhibiting a weaker spatial genetic structure and a stronger contribution to both genetic diversity within and genetic connectivity among populations than the females (Ishibashi *et al.* 2013; Quaglietta *et al.* 2013). Thus, it can be hypothesized that in species with male-biased dispersal, males are responsible for counteracting the effects of genetic drift and population fragmentation, and therefore may play an important role in population recovery and demographic growth. However, this hypothesis has not been sufficiently evaluated, partly because the male contribution to population genetic connectivity, diversity, and structure is poorly understood in most wildlife species.

Previous studies have mostly assessed male gene flow either indirectly, by contrasting data on autosomal microsatellites (Short-Tandem-Repeats, STRs) and mtDNA (e.g. Baker *et al.* 2013; Hua *et al.* 2013; Ishibashi *et al.* 2013) or by comparing male versus female genetic spatial autocorrelation (Banks & Peakall 2012). However, due to differences in marker systems (in the case of comparisons between mtDNA and autosomal markers) and recombination of bi-parental markers, these approaches only give an imprecise measure of the occurrence and

influence of male migration and gene flow. In contrast, the analysis of male-inherited, Y-chromosomal variation enables the direct assessment of male gene flow independently of mtDNA and autosomal DNA variation (Hurles & Jobling 2001; Kayser *et al.* 2005; Roewer *et al.* 2005). In addition, being similar to autosomal STRs in terms of diversity, mutation rates, and methodological unambiguity of allele designation (Hurles & Jobling 2001; Kayser *et al.* 2005; Roewer *et al.* 2005), Y-STRs can be applied to answer similar questions regarding population structure and connectivity at the same temporal and spatial resolution as autosomal STRs. Y-STRs have therefore been increasingly used to investigate population genetic processes both in humans (e.g. Rubicz *et al.* 2010; Salazar-Flores *et al.* 2010; Zupan *et al.* 2013; Tateno *et al.* 2014; Karmin *et al*. 2015) and other primates (Langergraber *et al.* 2007; Schubert *et al.* 2011; Nietlisbach *et al.* 2012; Inoue *et al.* 2013). In wildlife species, population genetic studies using Y-STRs are still rare, although a growing number of phylogeographic and phylogenetic studies exist, such as for brown hares (*Lepus europaeus*) in Europe and Anatolia (Mamuris *et al.* 2010), sheep (*Ovis ssp.*) (Meadows *et al.* 2006), colobine monkeys (*Colobinae*) (Roos *et al.* 2011), or snow voles (genus: *Chionomys*) (Yannic *et al.* 2012b).

Here, we apply high-resolution Y-chromosomal markers to assess the importance of male gene flow during the recovery of a large terrestrial carnivore from near extinction across its distribution range in Northern Europe. Large terrestrial carnivores declined dramatically in abundance and geographic distribution during the last centuries, but are recovering now and expanding into areas of previous extirpation (Chapron *et al.* 2014). This "carnivore comeback" provides an opportunity to determine the role of genetic connectivity as population recovery proceeds (Hagen et al. 2015). For the study species, we used the brown bear (*Ursus arctos*) in Northern Europe, owing to the pronounced male-biased dispersal and female philopatry that has been documented in the species (McLellan & Hovey 2001; Støen *et al.* 2005; Støen *et al.*

2006; Zedrosser *et al.* 2007), which suggests that connectivity among populations occurs primarily via male dispersal (Manel *et al.* 2004). Recently, using autosomal STRs, Kopatz *et al.* (2014) documented that the demographic recovery of the brown bears in Finland was supported by immigration from Russia, whereas in Sweden and Norway, brown bear numbers increased from near extinction without significant immigration of eastern bears. In addition, genetic connectivity between the eastern (Finland, Russia) and western (Norway, Sweden) subpopulations of the Northern European brown bear has been found to be low (Schregel *et al.* 2012; Kopatz *et al.* 2014). These differential recovery histories and the low connectivity between the eastern and western parts of the species' range offer an opportunity for applying Y-chromosomal markers to study the impact of male gene flow in the recovering brown bear populations in Northern Europe.

Our expectation was that the different recovery histories, i.e. high vs. low immigration during the recovery process, might explain differences in the Y-haplotype diversity of the post-bottleneck brown bears between the eastern and western subpopulations in Northern Europe. Specifically, we expected that high versus low immigration during the recovery process was associated with high versus low Y haplotype diversity in the post-bottleneck subpopulations, respectively. Given the evidence that the connectivity among the brown bear subpopulations in Northern Europe occurs primarily by male dispersal (Manel *et al.* 2004; Støen *et al.* 2005; Støen *et al.* 2006; Zedrosser *et al.* 2007), we also expected that male gene flow may be a potentially important driver of the range-wide demographic recovery process. A demographic impact of male gene flow relevant to explain the range-wide recovery would assume a high degree of Y haplotype admixture within the two subpopulations, independent of recovery history. This would then be consistent with the global pattern of male-biased gene flow across continents and in phylogeographic time scales, as recently reported based on the same Y-

chromosomal markers we have used here (Bidon *et al.* 2014). In contrast, a pronounced spatial structuring of Y haplotypes within the two subpopulations would indicate that large-scale demographic population recovery has occurred relatively independently of male gene flow. We also included historical samples from southern Norway between ~1750 and ~1950, i.e. in an area where brown bears became functionally extinct around the 1920s (Swenson *et al.* 1995). This was done to assess any temporal changes in the genetic composition of the post-bottleneck population. Our prediction was that recovery from near extinction without supporting immigration would be associated with a long-lasting reduction in Y-haplotype diversity.

Thus, we tested whether or not male dispersal was a key driver of large-scale genetic connectivity and the recent population recovery among Northern European brown bear subpopulations. In addition, by comparing historic and modern haplotypes and haplotype diversity, we could obtain empirical data about the genetic consequences of the historic population bottleneck for Y-haplotypic diversity.

Material and methods

Demographic history of the brown bears in Northern Europe

Once abundant in Northern Europe, brown bears were persecuted to near extinction during recent centuries (Servheen *et al.* 1999). In Norway, in the west of Northern Europe, the brown bear became functionally extinct around the 1920s, whereas in Sweden, east of Norway, ~130 individuals survived in three refugee areas (Swenson *et al.* 1995). In Finland, in central Northern Europe, ~150 individuals survived in the northern and eastern parts of the country (Ermala 2003). To the east of Finland, in Northwestern Russia, a similar population decline occurred during the 1930-1940s in the southwestern part of Russian Karelia, as well as on the Karelian Isthmus of the Leningrad region (Danilov 2005). After protective measures had been

initiated during the 1960s and 1970s, simultaneously in all four countries, these populations have gradually recovered demographically (Swenson *et al.* 1995; Danilov 2005; Kojola *et al.* 2006; Kindberg *et al.* 2011).

Sampling

Extant individuals from 2006-2012

We used DNA samples of verified male individuals, analyzed and stored in the course of regional and national monitoring programs conducted in Sweden and Norway, as well as during previous studies conducted in Finland (Schregel *et al.* 2012, Kopatz *et al.* 2012, 2014, Hagen *et al.* 2015). We only used DNA samples that had been identified positively with no less than six autosomal STRs plus one sex identification marker. The laboratory protocols of DNA extraction and analysis are accredited according to the EN ISO/IEC 17025 standard (Norwegian accreditation: test 139). According to these protocols, to verify an identification, the genotyping must have been performed independently three times if a sample was homozygous for the respective marker and twice if heterozygous. Samples that did not meet these requirements were not given a positive identification and hence were not used for this study. Further details of the DNA extraction, PCR amplification, and genotyping protocols are described in (Andreassen *et al.* 2012). Following the criteria outlined above, we selected a total of 491 DNA samples of male brown bears, which had been collected in Norway, Sweden, Finland, and northwest Russia during 2006-2012. Of these, 236 were extracted from fecal and 93 from hair samples, obtained noninvasively during monitoring programs (see also Schregel *et al.* 2012). In addition, 162 were extracted from tissue samples obtained from legally shot bears. The distribution of samples is illustrated in Figures 1, 2 and 4. Sampling procedures have been previously described elsewhere (Eiken *et al.* 2009; Andreassen *et al.* 2012; Kopatz *et al.* 2012; Schregel *et al.* 2012; Kopatz *et al.* 2014).

Historical samples from Norway 1750-1950

We sampled a total of 130 historical specimens from museums and private collections in southern Norway. These samples originated from approximately 1750 to 1956, and represented bears that existed in Norway before the total extinction of the resident brown bear population around 1970 (Swenson *et al.* 1995). Whenever possible, multiple samples, including various tissues (muscle, brain, nose, and mouth palate), skin, hairs, and tooth powder, were obtained from each specimen. We collected each sample separately to minimize the danger of contamination with other historical bear samples, wrapped the sample in aluminum foil, placed it in a self-sealing plastic bag, and stored it at room temperature.

We removed the tooth powder samples with a drill after scraping the tooth surface with a scalpel and discarding the powder from the beginning of drilling at a speed of around 60 rpm to avoid overheating. We drilled in a clean room and cleaned the tools, as well as surfaces, after each sample, first with 10% chlorine solution, followed by a rinse with water, and then dried with sterile tissue to minimize cross-contamination. The powder was immediately placed in a tube containing 300 µl of extraction buffer from the Qiagen Investigator© kit (Qiagen) and stored at -20°C until DNA extraction.

Genetic analysis

Analysis of Y-chromosomal STRs

We genotyped all samples with 9 Y-STR markers developed specifically for bears (Bidon *et al.* 2014) (Table S1) and validated for brown bears (Aarnes *et al.* in revision). Prior to amplification, we divided the STRs into two multiplex sets, which were optimized so that the fragments did not overlap in length and had even peak sizes (Aarnes *et al.* in revision). The PCR reaction was set up in a 10 µl reaction volume containing 5 µl 2x Multiplex PCR

MasterMix (Qiagen), 1xBSA (NEB), and 1 µl template DNA (~1 ng). The primer concentrations we used are listed in Table S1. We used a touchdown PCR program with an initial heating of 95°C for 10 min, followed by 10 cycles of 30 s at 94°C, 30 s of 69°C (reduced by 1°C for each consecutive cycle), and 1 min at 72°C, followed by 20 cycles of 30 s at 94°C, 30 s at 59°C, and 1 min. at 72°C, and then an elongation step of 45 min at 72°C. For the historical DNA samples, the second cycling step was extended from 20 cycles to 30 cycles. The first and the last four samples on every 96-well plate were positive controls, consisting of three reference males and one reference female. The female sample was included to control for primer contamination, as well as male specificity of the markers. Negative controls (without template DNA) were included for every seventh sample.

We analyzed PCR products on an ABI 3730 genetic analyzer and scored the PCR fragments with GeneMapper 4.0 (Applied Biosystems). Fragment lengths of the positive control samples were determined independently prior to the analysis of the sample material to serve as size calling control in the main analyses. Fragment length was identified automatically, but each result was also examined manually. We accepted size calling when the peak height of a fragment was higher than a threshold value of 600 relative fluorescent units (RFU). All samples were typed and confirmed at least two times. Haplotypes were assigned according to the numbering system of Y-STR profiles shown in Table S2. A new haplotype definition was accepted only if it was found more than once or if the discerning allele/s were typed and confirmed at least twice. Samples that failed to amplify at one or more markers were only assigned a haplotype if the PCR amplification results of the remaining markers allowed this to be done unambiguously.

Analysis of Y-chromosomal SNPs

We chose three different bear-specific Y-linked SNPs (UAY579.1B545, UAY318.2C713, UAY318.2C839) for haplogroup analysis, based on the previously confirmed sequence variation among Northern European brown bears for only these 3 SNPs (see Bidon et al. 2014). The SNP primers were designed using the Custom TaqMan® Assay Design Tool (Applied Biosystems) and labeled with FAM and VIC dye (Table S2). PCRs were set up in a 9 µl reaction volume, containing 5 µl TaqMan Genotyping MasterMix (2X), 0.25 µl TaqMan assay mix (40X), 0.045 µl BSA(100X) and 1 µl DNA (~1ug), and then amplified on the ABI 7300 Real-Time PCR System (Applied Biosystems) with a protocol of 2 min at 60°C, 10 min at 95°C, followed by 40 cycles of 15 s at 95°C, and 1 min at 60°C. To control for amplification success and between-run variability, four positive controls (three males and one female) were included in the first and last row of each plate. To control for contamination, a blank was included for every seventh sample. Generally, analysis was carried out only once per sample. Haplogroups were assigned according to the numbering of Y-SNP genotypes in Table S2.

Genetic analysis of historical brown bear samples

To avoid contamination with DNA from the extant brown bears, the historical samples were stored and handled, i.e. DNA extraction and PCR setup, in a separate laboratory building, in which brown bear samples had never been handled or stored. We extracted DNA using the Qiagen Investigator kit following the manufacturer's instructions for tissue and tooth/bone powder, respectively. To control for contamination during extraction, a blank sample was included for every 11th sample. Extracted DNA was eluted in 40 µl elution buffer and stored at -20°C.

First, all samples were PCR amplified with three sex-specific markers, as previously described, two located on the Y and one on the X chromosome (Bidon *et al.* 2013). Based on the analysis results, only samples that showed all three peaks, with the two Y-specific markers displaying a peak height of ~1500 RFU, were selected for further analysis. Earlier test runs (results not shown) showed that historical samples displaying a lower peak height had failed or shown only partial PCR amplification on the nine Y-STRs. In case more than one sample per individual displayed eligible amplification results, we selected the one with the highest peak heights for the subsequent analyses. We then amplified the selected samples independently three times with the nine Y-STR multiplex set-up, using the same touchdown PCR program as for extant samples, but with a total of 40 PCR cycles (see above). No positive controls were included in the PCR to avoid contamination, but a negative control was included for every fifth sample. To control for between-run variation during the analysis with the genetic analyzer, we included four independently amplified positive controls with known allele sizes. Allele size calling was confirmed at least two times. The Y-SNP analysis for the historical DNA samples followed the same procedures as for the modern DNA samples.

Statistical analysis

To assess the degree of haplotype admixture and diversity across the study area, we calculated haplotype frequencies, mean number of pairwise difference (MPD), and haplotype diversity for each of the 18 sampling locations, using the program Arlequin 3.5.1.2 (Excoffier & Lischer 2010). To illustrate similarity of haplotypes and possible correlation between geographic proximity and haplotype groups, we constructed a median-joining network using the software Network v.4.6.1.2. (http://www.fluxus-engineering.com/sharenet.htm).

Next, we wanted to assess and compare the population genetic structure of male lineages within and between the two known subpopulations in the west (i.e., the Scandinavian Peninsula), and the east (i.e., northeastern Norway, Finland and western Russia). The haplotypic, nonrecombinant nature of Y-STR data renders most frequently used approaches to identify genetic clusters or groups in a continuous population inapplicable. A common way to group samples is therefore to use the sampling location as the group-defining variable. This works well for large-scaled studies, where the sampling area stretches across several continents (e.g. Meadows *et al.* 2006; Bidon *et al.* 2014) and for studies in group-living species (e.g. Langergraber *et al.* 2007; Schubert *et al.* 2011). This approach is sometimes followed by hierarchical analysis of population differentiation, i.e. subsequent pooling or grouping sampling locations which display nonsignificant F_{ST} values among each other (e.g. Kayser *et al.* 2005). Like many other species, however, the brown bear is distributed more-or-less continuously across the study area, with no easily discernable delineation between groups. Therefore, we devised a different strategy to find genetic clusters in Y-STR data from continuous populations. This approach consisted of applying a cluster analysis based on haplotype frequencies from all of the sampling locations using the function 'hclust' in the program R (R coreteam 2013). We used the results of this analysis to initially group the data for the subsequent AMOVA analysis, performed with Arlequin (Excoffier & Lischer 2010). To validate the results of the clustering analysis, we performed several runs with different groupings. Our aim with this was to find the grouping that would minimize within-group variation and maximize among-group variation, while taking into account geographic location of the sampling area. To investigate whether there was a difference in the magnitude of population structure in the two independently recovering Northern European brown bear subpopulations, we also performed an AMOVA for the Scandinavian and Finland/-Russian sampling areas separately. As another means to test our

grouping, we used the number of different alleles between haplotypes to estimate genetic distance among sampling locations, which is similar to a weighted Fst (Excoffier & Lischer 2011), with the software Arlequin (Excoffier & Lischer 2010).

To estimate the genetic differentiation among the groups, we then pooled the data based on the results of the AMOVA analyses, as well as estimated pairwise F_{ST} among sampling location and estimated pairwise F_{ST} among these groups, again using number of different alleles as a measure of distance. To view the results of the two pairwise F_{ST} estimations, we performed an analysis of principal coordinates (PCoA) using the program GenAlEx 6.5.01 (Peakall & Smouse 2006, 2012).

Results

Contemporary Y chromosomal diversity and distribution

Y haplogroups and Y haplotypes were determined for 443 male brown bears sampled during 2006 to 2012. The remaining 49 samples could not be assigned to any specific Y haplotype, due to failed PCR amplification at one or several loci, and were discarded from the study. Y-SNPs analysis showed that two Y SNPs (UAY579.1B545 and UAY318.2C713) were variable among the 443 male bears. For Y-SNP UAY318.2C713 only 146 samples out of 443 could be determined. However, the detected SNP variation confirmed the haplogroup/haplotype combinations shown in Table S3. Although the third Y-SNP in this study (UAY318.2C839) had previously been found to be variable among bears in populations further east in Europe (Bidon et al. 2014), it was monomorphic in our samples. Accordingly, we found only three different Y-SNP haplogroups (Figure 1, Table S3), containing five, 20, and 11 Y-STR haplotypes, respectively. The resulting 36 different Y haplotypes and their total and local frequencies are shown in Table 1. An overview of Y haplotype distribution in Norway, Sweden,

Finland and northwest Russia is shown in Figure 2, with close-up maps in the supplementary files (Figures S1, S2 and S3).

We detected an east-west division according to recovery history both in the number and diversity of Y haplotypes. All Y haplotypes in Y haplogroup 1 were found exclusively in the eastern subpopulation, i.e. in northeast Norway, Finland and northwest Russia, but were absent from the west, i.e. Sweden and the western and southern parts of Norway. Consequently, we found only four Y haplotypes in the west (among 192 males), compared with 7 to 13 Y haplotypes at the various sampling locations in the east (Table 2). Accordingly, Y haplotype diversity (hd) per sampling location ranged from 0.439 to 0.971 (mean = 0.727), with higher values to the east (Table 2). Specifically, the diversity decreased relatively modestly from AOKM in Russia to PA in northeastern Norway, before it showed an abrupt decline from FLL and AN towards NB, TR and further south, indicating the geographic delineation, between the two independently recovering subpopulations. A similar geographic pattern was observed for the mean number of pairwise differences (MPD) per sampling location, which ranged from 0.909 to 4.200 (mean = 3.061) (Table 2).

Only two Y haplotypes, 2.02 and 3.09, were distributed across the entire study area (Figures 1-3, Table 1), with 2.02 previously found to be present also in the Ural region (Bidon *et al.* 2014), consistent with an old origin of these Y haplotypes (e.g. Morral *et al.* 1994). In contrast, more than half of the haplotypes (19 of 36) were found in only one sampling location (Table 1), which may suggest a relatively recent origin. However, eight of these were found in Northern Russia, which, despite its low number of samples, displayed the highest haplotype diversity (Table 2). Hence, more extensive sampling efforts are needed in Russia to investigate the full haplotype diversity there.

The hypothesis that male gene flow underlies the range-wide demographic recovery process was not supported by our Y haplotype data. Whereas the brown bears in the eastern subpopulation (Northeast Norway, Finland and Northwest Russia) displayed a high degree of Y haplotype admixture, the western subpopulation (Central and Southern Norway as well as Sweden) displayed Y haplotype frequencies that were clearly regionally distinct, providing evidence of large-scale demographic recovery under limited male gene flow within the western subpopulation (Figure 1 and 2). A more detailed presentation of location-specific Y haplotype frequencies and distribution can be found in the Figures S1-S3 in the appendix. The constructed median-joining network is shown in Figure 3. Qualitative assessment showed that genetically close haplotypes were not restricted to locations that were geographically close to each other.

Contemporary population genetic structure of male lineages

We investigated the apparent link between the spatial scale of genetic structure and the recovery history of brown bears further by using cluster (hclust) analysis to group sampling locations according to haplotype frequencies, followed by AMOVA to estimate the amount of genetic variation among and within these groups. The cluster analysis recovered six Y haplotype clusters separated into two main groups (Figure 4). The first main branch contained the southern Swedish and southern Norwegian sampling locations, which clustered in an east-west direction across the Swedish-Norwegian border, with a southern and a northern group (groups 1 and 2, light red and dark red colors, respectively; Figure 4a and b). The second main branch, which contained all of the other sampling locations, split again into two subbranches. The first subbranch clustered in a north-south direction, with TR, NB, and FLL clustering together in the north (group 3, yellow color; Figure 4a and b) and AN, PA, and KA in the South (group 4, green color; Figure 4a and b). The second subbranch contained the most eastern sampling

locations, with FSK, FNK and RNK grouped together into one cluster (group 5, light blue color; Figure 4a and b) and RSK and AOKM grouped together into a second cluster (group 6, dark blue color; Figure 4a and b). The few samples constituting the latter cluster were distributed across a large area. Haplotype frequencies indicated that groups 1, 2, and 3 were different from the other groups (Figure 5).

Overall, the AMOVA showed that among-group genetic variation was high (20.4 %) compared to within-group genetic variation (1.1 %). This was expected, as groupings were chosen to maximize the former and minimize the latter. However, the magnitude of among-group variation differed geographically according to recovery history. The AMOVA run for the western groups 1-3 only (i.e. Sweden and Norway without AN and PA) showed a similar high among-group variation (18 %) as for the analysis for the total dataset, whereas AMOVA on the eastern groups 4-6 only (i.e. AN, PA with the Finnish and Russian groups) displayed an among-group variation of only 5.7 % (Table 3). Also the within-group variation indicated a stronger spatial genetic structuring in the western subpopulation than in the rest of Northern Europe: when all groups in the western subpopulations were combined, within-group variation increased by 6.4 %, whereas the same treatment of all eastern groups resulted in an increase of only 2.5% (alternative grouping resulted generally in an increase of within-group variation as shown in the appendix). Additionally, pairwise F_{ST} values among sampling locations, which ranged from -0.057 (NT-ST) to 0.517 (RNK - NO), largely supported the grouping based on allele frequencies, with both lower absolute values and a lower degree of clustering in PCoA among hclust/AMOVA-group members than among nonmembers (Figure S4, Table S5). Exceptions were AN, which was displayed at an equal distance between group 3 and group 4, and FSK, which clustered closely with KA and PA, as opposed to the clustering analysis (Figure S4). Our analysis strategy of using clustering analysis based on haplotype frequencies offered

a simple alternative to hierarchical pooling of predetermined groups based on population differentiation estimates, as our results showed a reasonable concordance between the two approaches.

Using both approaches may increase the robustness of grouping decisions. Therefore, we further evaluated the validity of our findings using the results of the clustering analysis, the AMOVA, and the pairwise F_{ST} values to pool sampling locations and perform another estimation of genetic differentiation. The resulting pairwise F_{ST} values among these groups ranged from 0.039 to 0.354, with all of the values being significant (Table 4). Again, the PCoA analysis of pairwise F_{ST}-values displayed a divide in terms of genetic differentiation between the western groups 1-3 and eastern groups 4-6 (Figure 6 and S4), indicating less gene flow and/or a higher impact of genetic drift within Scandinavia than in the east.

Pre-bottleneck Y haplotype diversity of brown bears samples from Norway (1750-1950)

We also investigated the genetic impact of the demographic bottleneck by assessing temporal changes to the genetic composition of the post-bottleneck population. We collected 215 historic (1750-1950) samples from 130 individuals from the extinct brown bear population in southern Norway. We were able to amplify 62 bears successfully, 20 females and 42 males, with the gender test (see Methods). The 42 males were subsequently analyzed for the three Y-SNPs and nine Y-STRs to compare with the present bear population (see above). An overview of the locations of the individuals in Norway and their Y haplotypes and are shown in Figure 7 and the year of origin and Y haplotype and group profile are given in Table S3.

We obtained a nearly complete Y haplogroup and Y haplotype profile (i.e. one to two Y-SNPs and seven to nine Y-STRs) for 14 of the 42 historical samples (Table S3). We detected four haplotypes, each in only one individual, that were not present among the 443 bears sampled

during 2006-2012, nor were they present in the data on brown bears from across the global distribution zone (Bidon *et al.* 2014). Furthermore, seven individuals were assigned to Y haplotype 2.05, which is currently most frequent in Northern Scandinavia, indicating a possible distributional shift of this Y haplotype. The remaining three individuals were assigned to Y haplotype 2.08, which occurs at a relatively high frequency across most of Scandinavia today.

Discussion

Most large terrestrial carnivores experienced extensive bottlenecks during the last centuries (Woodruff 2001). Gene flow may be important for the successful recovery of these previously extirpated or isolated populations (Vilà *et al.* 2003; Fredrickson *et al.* 2007; Hedrick *et al.* 2014; Whiteley *et al.* 2015). Using Y-chromosomal markers identified using genomic sequences of the brown and the polar bear (Bidon *et al.* 2014), we have documented a clear pattern of Y haplotypes across the reexpanding post-bottleneck populations of brown bears in northern Europe. This suggested little or no impact of male gene flow on the demographic recovery process. Furthermore, we showed that a lack of male gene flow during the recovery process has left a dramatically reduced Y haplotype diversity in the post-bottleneck brown bear subpopulation in central and southern Scandinavia, which potentially may have a negative effect on long-term viability.

Regional differences in Y haplotype diversity

Our previous data on autosomal STRs showed extensive gene flow between Finnish and Russian brown bears (Kopatz *et al.* 2012, 2014), but low gene flow farther westwards (Schregel *et al.* 2012; Kopatz *et al.* 2014). These different gene flow histories, i.e. low versus high immigration, between the western and eastern populations (Kopatz *et al.* 2014) may have led

to a permanent difference in their Y haplotype diversity following the demographic bottleneck. In the populations in the northern and eastern parts of Northern Europe, connected by a substantial amount of gene flow (Kopatz *et al.* 2012, 2014), we found 35 different Y haplotypes among 207 males. Using mtDNA analyses, Eastern European Russia has been identified as the main source population for the postglacial brown bear colonization of Northern Europe (Taberlet & Bouvet 1994; Korsten *et al.* 2009; Keis *et al.* 2013) and autosomal STRs (Kopatz *et al.* 2014; Hagen *et al.* 2015). Thus, a decrease in haplotype diversity from east to west can be expected (Excoffier *et al.* 2009; DeGiorgio *et al.* 2011), which is consistent with the gradual decrease in Y haplotype diversity from Russia towards Finland and Northern Norway that we observed in our study. In contrast to this, we found exceptionally low diversity of male lineages in central and southern Scandinavia, with only four Y haplotypes among 192 extant males. We sampled all areas that have a substantial number of brown bears today, except Jämtland in Sweden, and assume that we have discovered all common haplotypes in the area. The demographic bottleneck in Finland was similar in extent to that in Sweden and Norway (Swenson *et al.* 1995; Ermala 2003; Danilov 2005), yet Y haplotype diversity observed in Finland was comparable to that of Northwest Russia. Thus, although the demographic bottleneck seems to have had a severe and, in the absence of sufficient immigration, long-lasting effect on Y haplotype diversity in Sweden and Norway, it seems that extensive immigration from Russia has had an alleviating effect on Y haplotype diversity in Finland.

Male gene flow as a driver in the population recovery process?

Given the pronounced male-biased dispersal behavior in brown bears (Støen *et al.* 2006; Zedrosser *et al.* 2007), we hypothesized that male gene flow would play an important role in the ongoing range-wide demographic recovery process. Instead, we found that the different

immigration histories in the eastern and western brown bear populations in Northern Europe were associated with different degrees of spatial genetic structuring of the post-bottleneck populations. Whereas the eastern brown bears displayed a high degree of Y haplotype admixture, the western brown bears displayed regionally distinct Y haplotype frequencies, suggesting the existence of barriers to male gene flow. The Scandinavian brown bear population in the west has increased from ~130 to over 3000 individuals between the 1970s and 2010 (Kindberg et al. 2011). Further east, in Finland, a demographic increase of a similar magnitude has been documented (Ermala 2003; Kojola & Heikkinen 2006). In spite of this substantial population growth in both areas, the PCoA of pairwise F_{ST}-values displayed a clear difference in Y haplotype structure between the east and west. The AMOVA showed that the overall among-group genetic variation of ~20% was mostly due to the local substructure in the west. Thus, our results suggested a large variability in the degree of male gene flow across the study area, rather than homogeneous gene flow and admixture. Hence, we believe that other mechanisms than male gene flow underlie the large-scale demographic recovery of the Northern European brown bear population. In humans, sex-biased genetic structure decreases with geographical distance, suggesting that long-distance gene flow is relatively independent of sex (Wilder *et al.* 2004a,b, Heyer *et al.* 2012). Also female brown bears have been shown to disperse more frequently and across farther distances in Sweden than previously assumed (Støen *et al.* 2006). Thus, both males and females may contribute to large-scale genetic connectivity, which should be investigated further.

History as a determinant of genetic structuring in Scandinavia

It has been proposed that the spatial genetic structure of the post-bottleneck brown bear population in Scandinavia reflects the recovery from refuge areas (Waits *et al.* 2000; Manel *et*

al. 2004), combined with the impact of colonization history (Bray *et al.* 2013), resulting in three genetic clusters in the contemporary subpopulation (Manel *et al.* 2004). These three clusters correspond roughly to the AMOVA groups 1 to 3 in our study. In the absence of Y-linked data, it was assumed that the observed clusters were connected via male gene flow (Manel *et al.* 2004). In contrast, our data suggested that contemporary restrictions to male gene flow prevented or slowed a dissolving of population structure and therefor that the genetic structure observed today is not merely an echo of the fragmentation caused by the human persecution. A similar structuring of the pre- and post-bottleneck populations, suggesting historic, ecological causes for the observed spatial genetic structure rather than anthropogenic ones was also recently proposed by Xenikoudakis et al. (2015). Y-STRs are nonrecombinant and directly inherited from father to son. A nonlocal haplotype introduced by a successfully reproducing immigrant male should therefore spread within a few generations (Ohnishi & Osawa 2014). With a generation time of ca. 10 years (Tallmon *et al.* 2004), at least three to four generations have passed since brown bears were placed under protection (Swenson *et al.* 1995). This seems to be enough time for Y haplotypes to spread into new areas, if dispersing males are able to produce viable, male offspring. Based on this, the clear difference in Y haplotype distribution (and hence limited male gene flow) that we observed between Västerbotten and Norrbotten was not expected. These two adjacent Swedish counties have no obvious geographical barriers between them, which suggests that other mechanisms are behind the presumed lack of Y haplotype exchange. Støen *et al.* (2006) showed inversely density-dependent natal (offspring) dispersal and that indicated both potential social restraints on dispersal and a stronger-than-previously-assumed territoriality in brown bears. Our noninvasive genetic capture-recapture monitoring records of brown bears regularly identify migrants among the different areas in Sweden (22 females vs. 81 males; unpublished data). Although at least two females and 11

males have dispersed from Västerbotten into Norrbotten (unpublished data), such dispersal has apparently not reduced genetic dissimilarities, at least in regard to Y-STRs. This may support the conclusions of Støen *et al*. (2006) regarding an impact of social mechanisms and higher-than-previously-assumed territoriality on population genetic structuring in brown bears, making it difficult for male immigrants to establish their own territory and/or mate successfully with a resident female. However, if a similar behavioral pattern can be assumed for brown bears in the eastern subpopulation, it is unclear why it does not result in a similar effect on the population structure there. One possible difference may be the occurrence of illegal hunting, which seems to be generally more accepted and occurring at a higher rate in northern Sweden compared to southern Scandinavia (Gangaas et al. 2013; Rauset 2013). It has been shown that this has a strong influence on the mortality rates of large carnivore populations (Andrén et al. 2006; Persson et al. 2009; Liberg et al. 2011). If disproportionately more dispersing males are killed, this would limit male gene flow among regions considerably and might contribute to the occurrence of the observed spatial patterns. However, we are not aware of any data available for the rate of occurrence and/or acceptance of illegal hunting in the area of the eastern subpopulation, so that a targeted study should be conducted to test this hypothesis.

The effect of the bottleneck on Y haplotype diversity and population genetic structure

The observed low diversity in the western subpopulation could be expected, because of the very low estimated number of bottleneck survivors (Swenson et al. 1995). However, our results suggest that there were proportionally fewer males than females, potentially only four reproducing males, among the ~130 surviving bears. This was corroborated by the data obtained from the historical samples. Among 14 genotyped historical samples from southern Norway (1780-1920), we found two haplotypes also present in today's population, but also four

haplotypes not present among the modern samples. The high coverage of sampling of extant males in Scandinavia strongly suggests that these four Y haplotypes were not present in the post-bottleneck population, and thus that the diversity of male lines and haplotypes in Scandinavia (i.e. Norway and Sweden) was much higher historically than it is today in the recovered population. This is supported by a recent study comparing historic and modern Scandinavian brown bears, which showed both a strong reduction of mtDNA diversity and a significant loss of autosomal allelic richness in the same area (Xenikoudakis *et al.* 2015). In contrast, Waits *et al.* (2000) who tested statistically for a genetic bottleneck using only modern samples, found weak evidence for this in Sweden and only in the southern parts of the country using autosomal STRs. Similarly, using autosomal STRs and modern samples from Norway, Sweden, Finland and Western Russia, also Kopatz et al. (2014) found only weak signs of a genetic bottleneck in Sweden and Norway. Moreover, several studies have found that autosomal STR diversity is relatively high and similar across the Northern European distribution zone (Waits *et al.* 2000, Tammenleht *et al.* 2010; Schregel *et al.* 2012, Kopatz *et al.* 2014), despite of the dramatic demographic decline (Swenson *et al.* 1995). Our results, using Y-chromosomal diversity, differ greatly from these studies, probably because the Y chromosome has an effective population size (N_e) equivalent to only 1/4 of that of autosomes. Consequently, Y-linked genetic diversity is more sensitive to demographic bottlenecks than autosomal genetic diversity (Greminger *et al.* 2010).

The Y haplotypes 2.08 and 2.05 were found in three and seven historic bears, respectively, indicating that they may have been as frequent before the bottleneck as they are in the present population. However, although 2.08 currently occurs at relatively high frequency across most of Scandinavia, 2.05 is more common in the northern parts of Scandinavia, being the dominant haplotype in Norrbotten. The other abundant haplotype in the extant Scandinavian

population, 3.09, was not found among the historical samples. This may suggest a shift in Y haplotype distribution and diversity following the demographic bottleneck and subsequent demographic changes, although the low number and limited geographical distribution of our historical samples may not give an accurate representation of the pre-bottleneck population. Thus, if possible, both the area of sampling and the number of samples should be extended to investigate this issue further. MtDNA studies indicate that the Scandinavian Peninsula was colonized from two directions in the postglacial period, from the northeast into northern Scandinavia and from the south (Taberlet & Bouvet 1994; Bray *et al.* 2013). A contact zone between the southern and northern mtDNA clades may have been located in the area of Sør- and Nord-Trøndelag in Norway and Jämtland and Gävleborg in Sweden (Bray *et al.* 2013). All of our historical samples were obtained from the area of the southern mtDNA clade. The four haplotypes we found only in the historical samples may thus be "southern" haplotypes, which may have reached Scandinavia in the postglacial recolonization from the south. This may also be the case for the haplotypes 2.08 and 2.05. Due to the bottleneck in the 19th and 20th centuries, haplotype distribution apparently shifted and several male lineages were eradicated, resulting in a population with a possibly very different genetic composition than the original population, thus indicating a founder effect (Hundertmark & Van Daele 2010; Andersen *et al.* 2014).

Perspectives on conservation

Demographic population recovery from only a very small number of surviving individuals has been shown to occur in many populations with very low autosomal genetic variability, providing exceptions to the widely-held assumption that high genetic diversity is vital for population viability (e.g. Visscher *et al.* 2001; Hoelzel *et al.* 2002; Reed 2010; Chan *et al.* 2011;

Baldursdottir *et al.* 2012; Taft & Roff 2012). Nevertheless, a large number of studies show that a significant reduction in genetic variability generally is correlated with reduced viability, mostly due to increased levels of inbreeding depression (e.g. Liberg *et al.* 2005; Hogg *et al.* 2006; O'Grady *et al.* 2006; Heber *et al.* 2013; Hostetler *et al.* 2013). However, to date this effect has been studied only with autosomal genetic markers, and we are unaware of a study with a similar objective employing sex-chromosomal markers. We have shown that Y haplotype diversity was probably reduced drastically during the demographic bottleneck in Sweden and Norway. The functionality of the Y chromosome is still poorly understood (Sayres *et al.* 2014), but recent research has suggested that the Y chromosome is more than just a determinate of the sex of its bearer, and is essential for male survival (Bellott *et al.* 2014). Regulatory functions on the chromosome may influence gene expression across the entire genome and hence biological functions throughout the lifetime (Bellott *et al.* 2014; Clark 2014; Cortez *et al.* 2014). Based on previous studies, estimates of genetic bottlenecks and diversity in autosomal markers suggested widespread genetic recovery across the entire Northern European brown bear population (Waits *et al.* 2000, Schregel *et al.* 2012, Kopatz *et al.* 2014). However, our results showed that this process is incomplete among Swedish and most Norwegian bears and management actions that are aimed at increasing gene flow among regions may be needed to ensure long-term viability of this population.

Acknowledgements

This work received support from Bioforsk, Norwegian Institute for Agricultural an Environmental Research and by LOEWE (Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz). Sampling in Russia was supported by the budget theme of RAS № 0221-2014-0006. We would like to thank S.K. Khohanov and A.S. Khozlova

for helping with the collection of samples from the Komi region in Russia. Also, we would like to thank Bård Spachmo, Camilla Tobiassen and Siv Grethe Aarnes for technical assistance, as well as L. P. Waits and three anonymous reviewers for valuable comments to the manuscript.

Author contributions

JS, HGE and SBH designed and conceived the study; JS, HGE, FAG, FH, JA, AJ and SBH developed the methods; HGE, FAG, IK KT, PD, AR, EP and SBH collected, compiled and quality controlled the samples. Main analyses were performed by JS, HGE and SBH. The manuscript was written by JS, HGE and SBH with input from all other authors.

Data Accessibility

The raw data (i.e. the genetic profiles) have been deposited in the corresponding author's institutional repository (i.e. NIBIO - Norwegian Institute for Bioeconomy) and can be made available by contacting the corresponding author.

References

Aarnes SG, Hagen SB, Andreassen R, *et al.* (in press) Y-chromosomal testing of brown bears (*Ursus arctos*): Validation of a multiplex PCR-approach for nine STRs suitable for fecal and hair samples.

Andersen A, Simcox DJ, Thomas JA, Nash DR (2014) Assessing reintroduction schemes by comparing genetic diversity of reintroduced and source populations: A case study of the globally threatened large blue butterfly (Maculinea anion). *Biological Conservation* **175**, 34-41.

Andreassen R, Schregel J, Kopatz A, *et al.* (2012) A forensic DNA profiling system for Northern European brown bears (Ursus arctos). *Forensic Science International-Genetics* **6**, 798-809.

Andrén H, Linnell JDC, Liberg O, et al. (2006) Survival rates and causes of mortality in Eurasian lynx (Lynx lynx) in multi-use landscapes. *Biological Conservation* **131**, 23-32.

Baker CS, Steel D, Calambokidis J, *et al.* (2013) Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series* **494**, 291-306.

Baldursdottir BK, Kristjansson T, Hallsson JH (2012) Diversity of the Icelandic goat breed assessed using population data. *Acta Agriculturae Scandinavica Section a-Animal Science* **62**, 53-65.

Banks SC, Peakall R (2012) Genetic spatial autocorrelation can readily detect sex-biased dispersal. *Molecular Ecology* **21**, 2092-2105.

Bellott DW, Hughes JF, Skaletsky H*, et al.* (2014) Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**, 494-499.

Bidon T, Frosch C, Eiken HG*, et al.* (2013) A sensitive and specific multiplex PCR approach for sex identification of ursine and tremarctine bears suitable for non-invasive samples. *Molecular Ecology Resources* **13**, 362-368.

Bidon T, Janke A, Fain SR*, et al.* (2014) Brown and Polar Bear Y Chromosomes Reveal Extensive Male-Biased Gene Flow within Brother Lineages. *Molecular Biology and Evolution* **31**, 1353-1363.

Bray SCE, Austin JJ, Metcalf JL*, et al.* (2013) Ancient DNA identifies post-glacial recolonisation, not recent bottlenecks, as the primary driver of contemporary mtDNA phylogeography and diversity in Scandinavian brown bears. *Diversity and Distributions* **19**, 245-256.

Chan C-H, Robertson HA, Saul EK*, et al.* (2011) Genetic variation in the kakerori (Pomarea dimidiata), an endangered endemic bird successfully recovering in the Cook Islands. *Conservation Genetics* **12**, 441-447.

Chapron G, Kaczensky P, Linnell JDC, et al. (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* **346**, 1517-1519.

Clark AG (2014) The vital Y chromosome. *Nature* **508**, 463-465.

Colonna V, Pagani L, Xue Y, Tyler-Smith C (2011) A world in a grain of sand: human history from genetic data. *Genome Biology* **12,** 234

Cortez D, Marin R, Toledo-Flores D*, et al.* (2014) Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488-493.

Cronin MA, Amstrup SC, Garner GW, Vyse ER (1991) Interspecific and intraspecific mitochondrial-DNA variation in North American Bears (Ursus). *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **69**, 2985-2992.

Danilov PI (2005) *Game animals of Karelia - Ecology, resources, management, protection.* Nauka, Moscow.

DeGiorgio M, Degnan JH, Rosenberg NA (2011) Coalescence-Time Distributions in a Serial Founder Model of Human Evolutionary History. *Genetics* **189**, 579-593.

Eiken HG, Andreassen RJ, Kopatz A*, et al.* (2009) Population data for 12 STR loci in Northern European brown bear (Ursus arctos) and application of DNA profiles for forensic casework. *Forensic Science International: Genetics Supplement Series* **2**, 273-274.

Enserink M, Vogel G (2006) Wildlife conservation - The carnivore comeback. *Science* **314**, 746-749.

Ermala A (2003) A survey of large predators in Finland during the 19th – 20th centuries. *Acta Zoologica Lituanica* **13**, 15-20.

Excoffier L, Foll M, Petit RJ (2009) Genetic Consequences of Range Expansions. In: *Annual Review of Ecology Evolution and Systematics*, pp. 481-501.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Excoffier L, Lischer HEL (2011) *Arlequin ver 3.5.1.3 user manual: Arlequin ver 3.5: An Integrated Software Package for Population Genetics Data Analysis.* http://cmpg.unibe.ch/software/arelquin3. 26.05.2014

Fredrickson RJ, Siminski P, Woolf M, Hedrick PW (2007) Genetic rescue and inbreeding depression in Mexican wolves. *Proceedings of the Royal Society B-Biological Sciences* **274**, 2365-2371.

Gangaas KE, Kaltenborn BP, Andreassen HP (2013) Geo-spatial aspects of acceptance of illegal hunting of large carnivores in Scandinavia. *PlosOne* **8**. doi: 10.1371/journal.pone.0068849

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28**, 1140-1162.

Greminger MP, Krutzen M, Schelling C, Pienkowska-Schelling A, Wandeler P (2010) The quest for Y-chromosomal markers - methodological strategies for mammalian non-model organisms. *Molecular Ecology Resources* **10**, 409-420.

Hagen SB, Kopatz A, Aspi J, Kojola I, Eiken HG (2015) Evidence of rapid change in genetic structure and diversity during range expansion in a large terrestrial carnivore. *Proceedings of the Royal Society B-Biological Sciences.* doi: 10.1098/rspb.2015.0092.

Heber S, Varsani A, Kuhn S, *et al.* (2013) The genetic rescue of two bottlenecked South Island robin populations using translocations of inbred donors. *Proceedings of the Royal Society B-Biological Sciences* **280**, 020122228. doi: 10.1098/rspb.2012.2228

Hedrick PW, Peterson RO, Vucetich LM, Adams JR, Vucetich JA (2014) Genetic rescue in Isle Royale wolves: genetic analysis and the collapse of the population. *Conservation Genetics* **15**, 1111-1121.

Heyer E, Chaix R, Pavard S, Austerlitz F (2012) Sex-specific demographic behaviours that shape human genomic variation. *Molecular Ecology* **21**, 597-612.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf BJ, Alvord G (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology* **15**, 567-575.

Hogg JT, Forbes SH, Steele BM, Luikart G (2006) Genetic rescue of an insular population of large mammals. *Proceedings of the Royal Society B-Biological Sciences* **273**, 1491-1499.

Hostetler JA, Onorato DP, Jansen D, Oli MK (2013) A cat's tale: the impact of genetic restoration on Florida panther population dynamics and persistence. *Journal of Animal Ecology* **82**, 608-620.

Hua PY, Zhang LB, Guo TT, Flanders J, Zhang SY (2013) Dispersal, Mating Events and Fine-Scale Genetic Structure in the Lesser Flat-Headed Bats. *Plos One* **8**, doi: 10.1371/journal.pone.0054428

Hundertmark KJ, Van Daele LJ (2010) Founder effect and bottleneck signatures in an introduced, insular population of elk. *Conservation Genetics* **11**, 139-147.

Hurles ME, Jobling MA (2001) Haploid chromosomes in molecular ecology: lessons from the human Y. *Molecular Ecology* **10**, 1599-1613.

Inoue E, Akomo-Okoue EF, Ando C, *et al.* (2013) Male genetic structure and paternity in western lowland gorillas (Gorilla gorilla gorilla). *American Journal of Physical Anthropology* **151**, 583-588.

Ishibashi Y, Zenitani J, Saitoh T (2013) Male-biased Dispersal Causes Intersexual Differences in the Subpopulation Structure of the Gray-sided Vole. *Journal of Heredity* **104**, 718-724.

Karmin M, Saag L, Vicente M, *et al*. (2015) A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research* 2015; doi: 10.1101/gr.186684.114

Kayser M, Lao O, Anslinger K, *et al.* (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. *Human Genetics* **117**, 428-443.
Keis M, Remm J, Ho SYW, *et al.* (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *Journal of Biogeography* **40**, 915-927.
Kindberg J, Swenson JE, Ericsson G, *et al.* (2011) Estimating population size and trends of the Swedish brown bear Ursus arctos population. *Wildlife Biology* **17**, 114-123.
Kojola I, Hallikainen V, Nygren T, Pesonen M, Ruusila V (2006) Recent trends and harvest in Finland's brown bear population. *Ursus* **17**, 159-164.
Kojola I, Heikkinen S (2006) The structure of the expanded brown bear population at the edge of the Finnish range. *Annales Zoologici Fennici* **43**, 258-262.
Kopatz A, Eiken HG, Aspi J, *et al.* (2014) Admixture and Gene Flow from Russia in the Recovering Northern European Brown Bear (Ursus arctos). *Plos One* **9**, doi: 10.1371/journal.pone.0097558.
Kopatz A, Eiken HG, Hagen SB, *et al.* (2012) Connectivity and population subdivision at the fringe of a large brown bear (Ursus arctos) population in North Western Europe. *Conservation Genetics* **13**, 681-692.
Korsten M, Ho SYW, Davison J, *et al.* (2009) Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Molecular Ecology* **18**, 1963-1979.
Kutschera VE, Bidon T, Hailer F, *et al.* (2014) Bears in a Forest of Gene Trees: Phylogenetic Inference Is Complicated by Incomplete Lineage Sorting and Gene Flow. *Molecular Biology and Evolution* **31**, 2004-2017.
Langergraber KE, Siedel H, Mitani JC, *et al.* (2007) The Genetic Signature of Sex-Biased Migration in Patrilocal Chimpanzees and Humans. *Plos One* **2**, doi: 10.1371/journal.pone.0000973
Liberg O, Andren H, Pedersen HC, *et al.* (2005) Severe inbreeding depression in a wild wolf (Canis lupus) population. *Biology Letters* **1**, 17-20.
Liberg O, Chapron G, Wabakken P, et al. (2011) Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. *Proceedings of the Royal Society of London Series B-Biological Sciences* **279**, 910-915.
MacDonald AJ, FitzSimmons NN, Chambers B, Renfree MB, Sarre SD (2014) Sex-linked and autosomal microsatellites provide new insights into island populations of the tammar wallaby. *Heredity* **112**, 333-342.
Mamuris Z, Moutou KA, Stamatis C, Sarafidou T, Suchentrunk F (2010) Y DNA and mitochondrial lineages in European and Asian populations of the brown hare (Lepus europaeus). *Mammalian Biology* **75**, 233-242.
Manel S, Bellemain E, Swenson JE, Francois O (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Molecular Ecology* **13**, 1327-1331.
McLellan BN, Hovey FW (2001) Natal dispersal of grizzly bears. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **79**, 838-844.
Meadows JRS, Hanotte O, Drogemuller C, *et al.* (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* **37**, 444-453.
Morral N, Bertranpetit J, Estivill X, *et al.* (1994) The origin of the major cystic-fibrosis mutation (DELTA-F508) in European populations. *Nature Genetics* **7**, 169-175.

Napolitano C, Johnson WE, Sanderson J, *et al.* (2014) Phylogeography and population history of Leopardus guigna, the smallest American felid. *Conservation Genetics* **15**, 631-653.
Neaves LE, Zenger KR, Prince RIT, Eldridge MDB (2013) Paternally inherited genetic markers reveal new insights into genetic structuring within Macropus fuliginosus and hybridisation with sympatric Macropus giganteus. *Australian Journal of Zoology* **61**, 58-68.
Nietlisbach P, Arora N, Nater A, *et al.* (2012) Heavily male-biased long-distance dispersal of orang-utans (genus: Pongo), as revealed by Y-chromosomal and mitochondrial genetic markers. *Molecular Ecology* **21**, 3173-3186.
O'Grady JJ, Brook BW, Reed DH, *et al.* (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation* **133**, 42-51.
Ohnishi N, Osawa T (2014) A difference in the genetic distribution pattern between the sexes in the Asian black bear. *Mammal Study* **39**, 11-16.
Ou W, Takekawa S, Yamada T, *et al.* (2014) Temporal change in the spatial genetic structure of a sika deer population with an expanding distribution range over a 15-year period. *Population Ecology* **56**, 311-325.
Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.
Persson J, Ericsson G, Segerstrom P (2009) Human caused mortality in the endangered Scandinavian wolverine population. *Biological Conservation* **142**, 325-331.
Quaglietta L, Fonseca VC, Hajkova P, Mira A, Boitani L (2013) Fine-scale population genetic structure and short-range sex-biased dispersal in a solitary carnivore, Lutra lutra. *Journal of Mammalogy* **94**, 561-571.
Ramachandran S, Deshpande O, Roseman CC, *et al.* (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15942-15947.
Rauset GR (2013) Life and Death in Wolverines (Doctoral Thesis). *Swedish University of Agricultural Sciences*, Uppsala.
R Core Team (2013) R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria.
Reed DH (2010) Albatrosses, eagles and newts, Oh My!: exceptions to the prevailing paradigm concerning genetic diversity and population viability? *Animal Conservation* **13**, 448-457.
Roewer L, Croucher PJP, Willuweit S, *et al.* (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Human Genetics* **116**, 279-291.
Roos C, Zinner D, Kubatko LS, *et al.* (2011) Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. *Bmc Evolutionary Biology* **11**, 77. doi:10.1186/1471-2148-11-77
Rubicz R, Melton PE, Spitsyn V, *et al.* (2010) Genetic Structure of Native Circumpolar Populations Based on Autosomal, Mitochondrial, and Y Chromosome DNA Markers. *American Journal of Physical Anthropology* **143**, 62-74.
Saarma U, Kojola I (2007) Matrilineal genetic structure of the brown bear population in Finland. *Ursus* **18**, 30-37.

Salazar-Flores J, Dondiego-Aldape R, Rubi-Castellanos R*, et al.* (2010) Population Structure and Paternal Admixture Landscape on Present-Day Mexican-Mestizos Revealed by Y-STR Haplotypes. *American Journal of Human Biology* **22**, 401-409.

Sayres MAW, Lohmueller KE, Nielsen R (2014) Natural Selection Reduced Diversity on Human Y Chromosomes. *PLoS genetics* **10**. doi: 10.1371/journal.pgen.1004064

Schregel J, Kopatz A, Hagen SB*, et al.* (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Molecular Ecology* **21**, 3474-3488.

Schubert G, Stoneking CJ, Arandjelovic M*, et al.* (2011) Male-Mediated Gene Flow in Patrilocal Primates. *Plos One* **6**. doi: 10.1371/journal.pone.0021514

Servheen C, Herrero S, Peyton B (1999) Bear - Status Survey and Conservation Action Plan. IUCN, Cambridge.

Stevenson CD, Ramsey AD, Nevin OT, Sinclair W (2013) Assessing Grey Squirrel Dispersal Patterns within the Landscape Using Sequence Variation. *Landscape Research* **38**, 553-559.

Støen O-G, Bellemain E, Sæbø S, Swenson JE (2005) Kin-related spatial structure in brown bears Ursus arctos. *Behavioral Ecology and Sociobiology* **59**, 191-197.

Støen OG, Zedrosser A, Saebo S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears Ursus arctos. *Oecologia* **148**, 356-364.

Swenson JE, Wabakken P, Sandegren F*, et al.* (1995) The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology* **1**, 11-25.

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, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.

Taft HR, Roff DA (2012) Do bottlenecks increase additive genetic variance? *Conservation Genetics* **13**, 333-342.

Tallmon DA, Luikart G, Beaumont MA (2004) Comparative evaluation of a new effective population size estimator based on approximate bayesian computation. *Genetics* **167**, 977-988.

Tammeleht E, Remm J, Korsten M, *et al.* (2010) Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Molecular Ecology* **19**: 5359-5370.

Tateno Y, Komiyama T, Katoh T*, et al.* (2014) Divergence of East Asians and Europeans Estimated Using Male- and Female-Specific Genetic Markers. *Genome Biology and Evolution* **6**, 466-473.

Vilà C, Sundqvist AK, Flagstad O*, et al.* (2003) Rescue of a severely bottlenecked wolf (Canis lupus) population by a single immigrant. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 91-97.

Visscher PM, Smith D, Hall SJG, Williams JA (2001) A viable herd of genetically uniform cattle - Deleterious alleles seem to have been purged in a feral strain of inbred cows. *Nature* **409**, 303-303.

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

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends in Ecology & Evolution* **30**, 42-49.

Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004a) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nature Genetics* **36**, 1122-1125.

Wilder JA, Mobasher Z, Hammer MF (2004b) Genetic evidence for unequal effective population sizes of human females and males. *Molecular Biology and Evolution* **21**, 2047-2057.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* **28**, 290-300.

Woodruff DS (2001) Declines of biomes and biotas and the future of evolution. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5471-5476.

Xenikoudakis G, Ersmark E, Tison JL, Waits L, Kindberg J, Swenson JE, Dalén L (2015) Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear. *Molecular Ecology*. doi: 10.1111/mec.13239

Yannic G, Basset P, Buchi L, Hausser J, Broquet T (2012a) Scale -specific sex-biased dispersal in the Valais shrew unveiled by genetic variation on the Y-chromosome, autosomes, and mitochondrial DNA. *Evolution* **66**, 1737-1750.

Yannic G, Burri R, Malikov VG, Vogel P (2012b) Systematics of snow voles (Chionomys, Arvicolinae) revisited. *Molecular Phylogenetics and Evolution* **62**, 806-815.

Zedrosser A, Støen O-G, Sæbø S, Swenson JE (2007) Should I stay or should I go? Natal dispersal in the brown bear. *Animal Behaviour* **74**, 369-376.

Zerjal T, Xue YL, Bertorelle G*, et al.* (2003) The genetic legacy of the mongols. *American Journal of Human Genetics* **72**, 717-721.

Zupan A, Vrabec K, Glavac D (2013) The paternal perspective of the Slovenian population and its relationship with other populations. *Annals of Human Biology* **40**, 515-526.

Table 1: Y-chromosomal haplotype frequencies of extant male brown bears by sampling location in Northern Europe, Scandinavian locations listed from South to North: **HO**= Hedmark/Oppland, **DV**= Dalarna/Värmland, **GA**= Gävleborg, **ST**= Sør-Trøndelag, **NT**= Nord-Trøndelag, **NO**= Nordland, **VB**= Västerbotten, **TR**= Troms, **NB**= Norrbotten; Finnmark/Lappland and Finland/Russia listed from West to East: **AN**= Anarjohka, **FLL**= Lappland, **PA**= Pasvik , **KA**= Kainuu, **RNK**= Russian Northern Karelia, **FNK**= Finnish Northern Karelia , **FSK**= Finnish Southern Karelia, **RSK**= Russian Southern Karelia, **AO**= Arkhangelsk, **KM**= Kemi. Haplotype names: the first digit of the respective haplotype names indicates the haplogroup (Y-SNP alleles); the last two digits were assigned according to the order in which the haplotypes were registered during microsatellite genotyping.

	Sampling Location																	Total
	Scandinavia									Finnmark/Lappland			Finland /Russia					
hpt	HO (n=58)	DV (n=16)	GA (n=14)	ST (n=12)	NT (n=40)	NO (n=12)	VB (n=40)	TR (n=18)	NB (n=26)	AN (n=15)	FLL (n=14)	PA (n=54)	KA (n=30)	RNK (n=13)	FNK (n=26)	FSK (n=26)	RSK (n=14)	AOKM (n=15)	(n=443)
1.01	-	-	-	-	-	-	-	-	-	0.200	0.071	0.352	0.267	-	0.039	0.154	-	-	0.086
1.13	-	-	-	-	-	-	-	-	-	-	-	-	0.033	-	-	-	-	-	0.002
1.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
1.23	-	-	-	-	-	-	-	-	-	0.067	-	-	-	-	-	0.039	-	-	0.009
1.25	-	-	-	-	-	-	-	-	-	0.067	-	0.019	-	-	-	-	-	-	0.005
2.02	0.103	0.063	0.071	0.083	0.075	0.083	0.025	0.389	-	-	-	0.037	0.067	-	-	-	0.143	0.134	0.065
2.05	0.034	-	0.071	0.167	0.175	0.167	0.250	0.444	0.731	0.333	0.500	0.056	-	-	-	-	-	-	0.149
2.06	-	-	-	-	-	-	-	-	-	-	-	0.019	0.033	-	0.039	-	0.143	0.067	0.014
2.07	-	-	-	-	-	-	-	0.111	-	0.133	-	0.093	0.200	0.308	0.192	0.269	-	-	0.070
2.08	0.414	0.563	0.429	0.583	0.575	0.750	0.575	-	0.115	-	0.143	-	-	-	-	-	-	-	0.239
2.10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.143	-	0.005
2.11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.039	0.039	0.071	-	0.007
2.12	-	-	-	-	-	-	-	-	-	-	-	-	-	0.077	0.077	-	0.071	-	0.009
2.14	-	-	-	-	-	-	-	-	0.154	-	0.071	0.019	0.033	0.077	0.039	-	-	-	0.011
2.16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
2.17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.134	0.005
2.18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
2.22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.077	-	-	0.005
2.24	-	-	-	-	-	-	-	-	-	-	-	0.019	-	-	-	-	-	-	0.002
2.27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.039	-	-	0.002
2.28	-	-	-	-	-	-	-	0.056	-	-	-	-	-	-	-	-	-	-	0.002
2.29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
2.30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
2.31	-	-	-	-	-	-	-	-	-	-	-	-	0.100	-	-	-	-	-	0.007
2.32	-	-	-	-	-	-	-	-	-	-	0.071	-	-	-	-	-	-	-	0.002
3.03	-	-	-	-	-	-	-	-	-	-	0.071	0.056	0.033	-	-	0.231	-	-	0.025
3.04	-	-	-	-	-	-	-	-	-	-	-	0.185	0.067	0.231	-	-	-	0.067	0.036
3.09	0.448	0.375	0.429	0.167	0.175	-	0.150	-	-	-	-	0.037	0.133	0.154	0.346	0.154	0.214	0.067	0.185
3.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.134	0.005
3.20	-	-	-	-	-	-	-	-	-	0.133	-	0.037	0.033	-	-	-	-	-	0.011
3.21	-	-	-	-	-	-	-	-	-	0.067	0.071	0.074	-	-	-	-	-	-	0.014
3.26	-	-	-	-	-	-	-	-	-	-	-	-	-	0.077	-	-	-	-	0.002
3.33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.067	0.002
3.34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.039	-	-	-	0.002
3.35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.077	-	-	-	0.005
3.36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.115	-	-	-	0.007

Table 2: Y-STR diversity of extant male brown bears per sampling location in Northern Europe calculated with *Arlequin 3.5.1.2*; n=sample size. **no. hpt.**= number of haplotypes found, **MPD** = Mean number of pairwise differences, **hd** = haplotype diversity, both listed with their respective standard deviation (± s.d.).

Region	Sampling location	Country	n	no. hpt.	MPD	± s.d.	hd	± s.d.
Southern Scandinavia	**HO** Hedmark. Oppland	Norway	58	4	3.180	1.668	0.627	0.034
	DV Dalarna. Värmland	Sweden	16	3	3.100	1.698	0.575	0.080
	GA Gävleborg	Sweden	14	4	3.396	1.848	0.670	0.082
Central Scandinavia	**ST** South Trøndelag	Norway	12	4	2.515	1.455	0.652	0.133
	NT North Trøndelag	Norway	40	4	2.421	1.341	0.618	0.067
	NO Nordland	Norway	12	3	0.909	0.676	0.439	0.158
	VB Västerbotten	Sweden	40	4	2.249	1.264	0.600	0.062
Troms	**TR** Troms	Norway	18	4	2.092	1.225	0.673	0.069
Norbotten	**NB** Norrbotten	Sweden	26	3	1.975	1.154	0.446	0.105
Lappland	**FLL** Lappland	Finland	14	7	2.945	1.640	0.758	0.116
Anarjohka	**AN** Anarjohka/Lemmenjoki	Norway/Finland	15	7	3.600	1.934	0.857	0.065
Pasvik	**PA** Pasvik/Inari/Pechenga	Norway/Finland /Russia	54	13	3.708	1.903	0.832	0.038
Karelia	**KA** Kainuu	Finland	30	11	4.200	2.146	0.876	0.036
	RNK Russian Northern Karelia	Russia	13	7	3.603	1.953	0.872	0.067
	FNK Finnish Northern Karelia	Finland	26	10	3.317	1.760	0.843	0.053
	FSK Finnish Southern Karelia	Finland	26	8	3.649	1.909	0.849	0.037
	RSK Russian Southern Karelia	Russia	14	8	4.044	2.147	0.923	0.044
Northern Russia	**AOKM** Arkhangelsk/Komi Oblast	Russia	15	12	4.190	2.204	0.971	0.033

Table 3: Results of the AMOVA based on Y-STR data, performed on male brown bears in Northern Europe; the percentages of variation displayed where calculated for the entire dataset (AMOVA groups 1-6), and separately for the western (AMOVA groups 1-3) and eastern part of the population (AMOVA groups 4-6). See Figure 3 for the sampling locations per group.

	Percentage of variation		
	AMOVA groups 1 - 6	AMOVA groups 1 - 3	AMOVA groups 4 - 6
among groups	20.40**	17.95**	5.68*
within groups among sampling locations	1.14^{ns}	-0.21^{ns}	2.82*
within populations	78.45**	82.26**	91.50**

* $p<0.05$; ** $p<0.001$; ns=non significant

Table 4: Pairwise F_{ST} values among groups of sampling locations of male brown bears in Northern Europe, based on Y-STR data, using number of different alleles as measure of genetic distance. The groups are as follows: **group 1**=HO, DV, GA; **group 2**=ST, NT, NO, VB; **group 3**=TR, NB, FLL; **group 4**=AN, PA, KA; **group 5**=RNK, FNK, FSK; **group 6**=RSK, AOKM.

	Group 1	Group 2	Group 3	Group 4	Group 5
Group 2	0.148**				
Group 3	0.242**	0.158**			
Group 4	0.175**	0.315**	0.250**		
Group 5	0.139**	0.354**	0.290**	0.071**	
Group 6	0.087*	0.293**	0.270**	0.080**	0.039*

* $p < 0.05$; ** $p < 0.001$

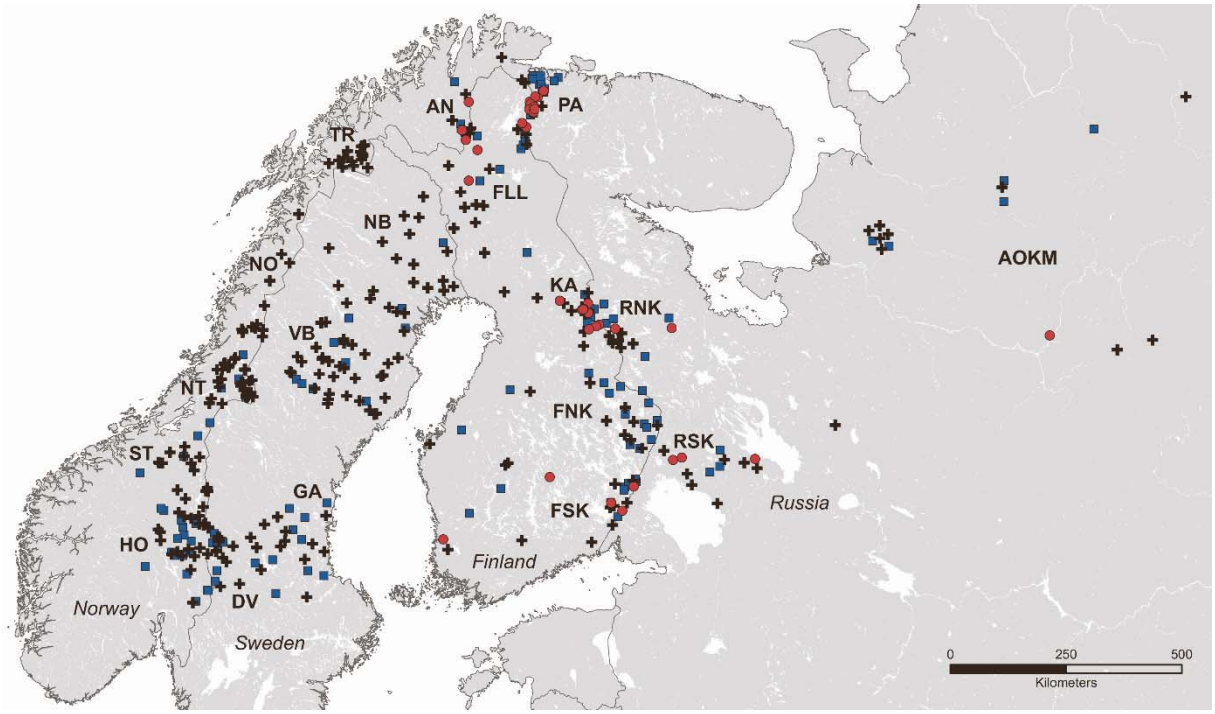


Figure 1: Y haplogroup distribution of extant male brown bears in Northern Europe. Historical samples are shown separately in Figure 7. Symbols correspond to the major (SNP-based) Y chromosome haplogroups: red circles = haplogroup 1, black cross = haplogroup 2, blue squares = haplogroup 3. The location of sampling areas are indicated as follows: **HO**= Hedmark/Oppland, **DV**= Dalarna/Värmland, **GA**= Gävleborg, **ST**= Sør-Trøndelag, **NT**= Nord-Trøndelag, **NO**= Nordland, **VB**= Västerbotten, **TR**= Troms, **NB**= Norrbotten; **AN**= Anarjohka, **FLL**= Finnish Lappland, **PA**= Pasvik Valley, **KA**= Kainuu, **RNK**= Russian Northern Karelia, **RSK**= Russian Southern Karelia, **FNK**= Finnish northern Karelia, **FSK**=Finnish southern Karelia, **AOKM**= Arkhangelsk, Kemi.

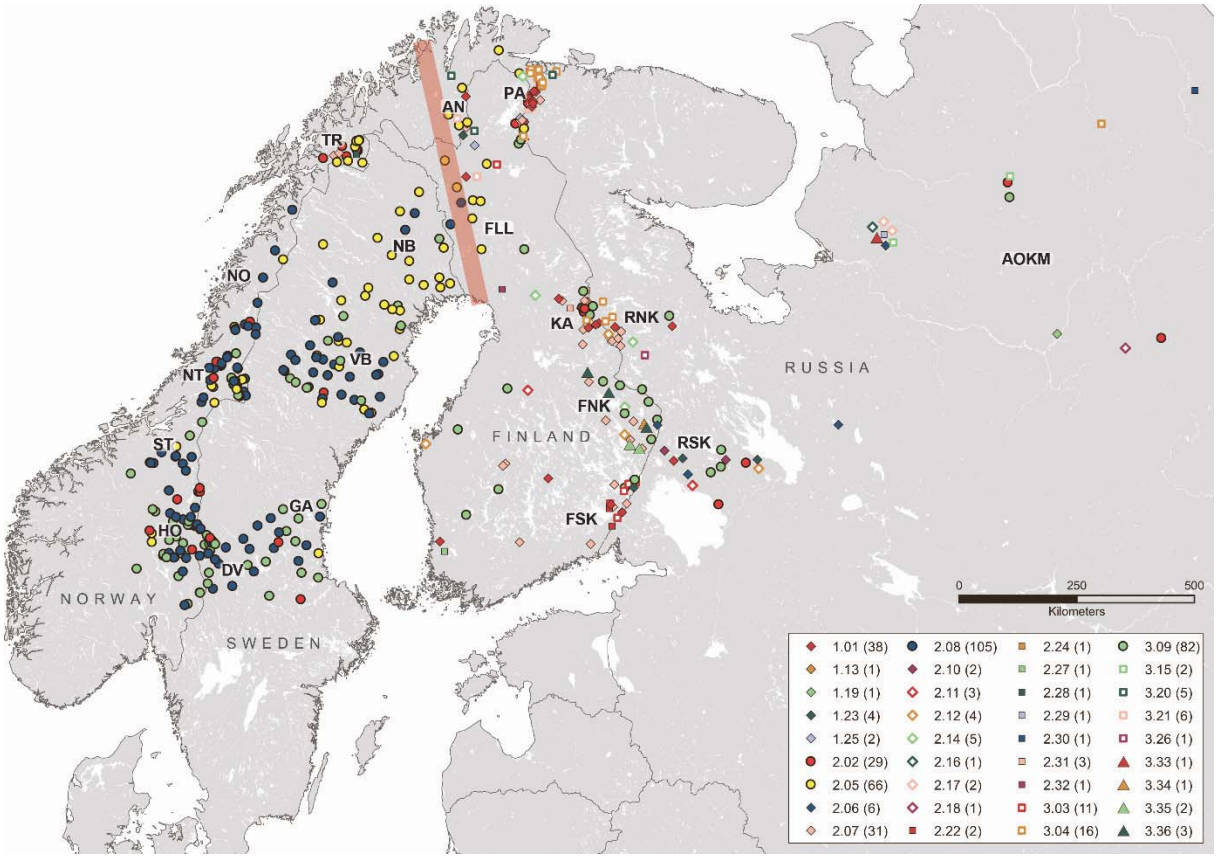


Figure 2: Geographical distribution of the detected Y chromosome haplotypes among extant male brown bears in Northern Europe. Each sample is depicted with the symbol used for the haplotype to which it was assigned. Brackets in the legend: total frequency of each haplotype. Haplotypes occurring in Sweden and the southern and centrals parts of Norway (HO, ST, NT and NO) are depicted with a circle and indicated with bold characters in the legend. The geographical location of the proposed division between the western and eastern subpopulations is indicated by the transparent red line.

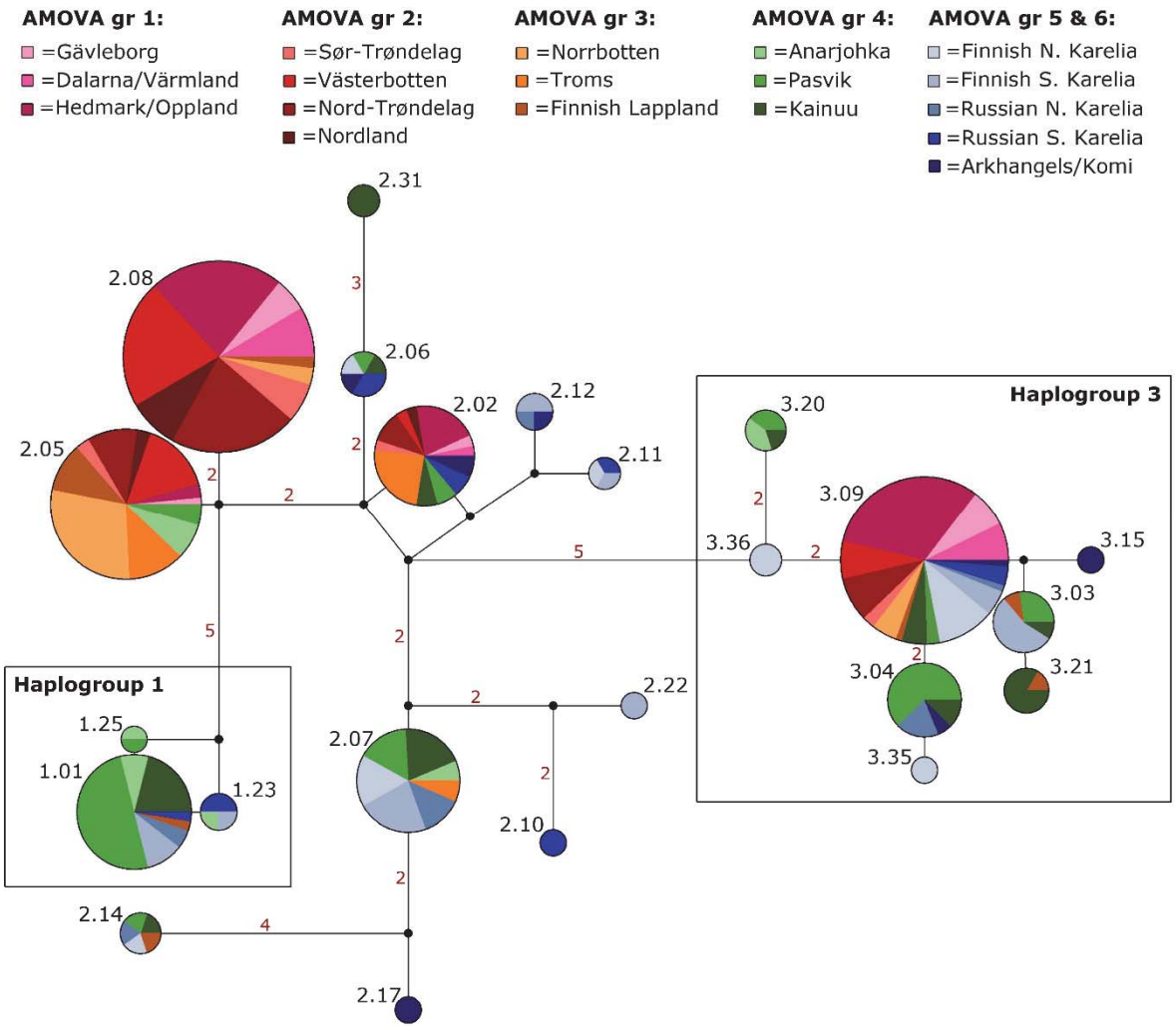


Figure 3: Phylogenetic relationship among Y chromosome haplotypes of extant male brown bears in Northern Europe, constructed with the help of the program Network v4.6.1.2. (Fluxus Technology) by using the RM-MJ method, based on the Y-STR and Y-SNP data. Only haplotypes with an occurrence of >1 were used for the construction. The area of each pie chart reflects the frequency of the respective haplotype in the extant study population, colors correspond to the sampling location of the haplotypes. The names of the haplotypes are written in black; the number of mutations between each node of the tree is given in red for each branch (if >1). Haplotypes of haplogroups 1 and 3 are indicated by boxes, all remaining haplotypes belong to haplogroup 2.

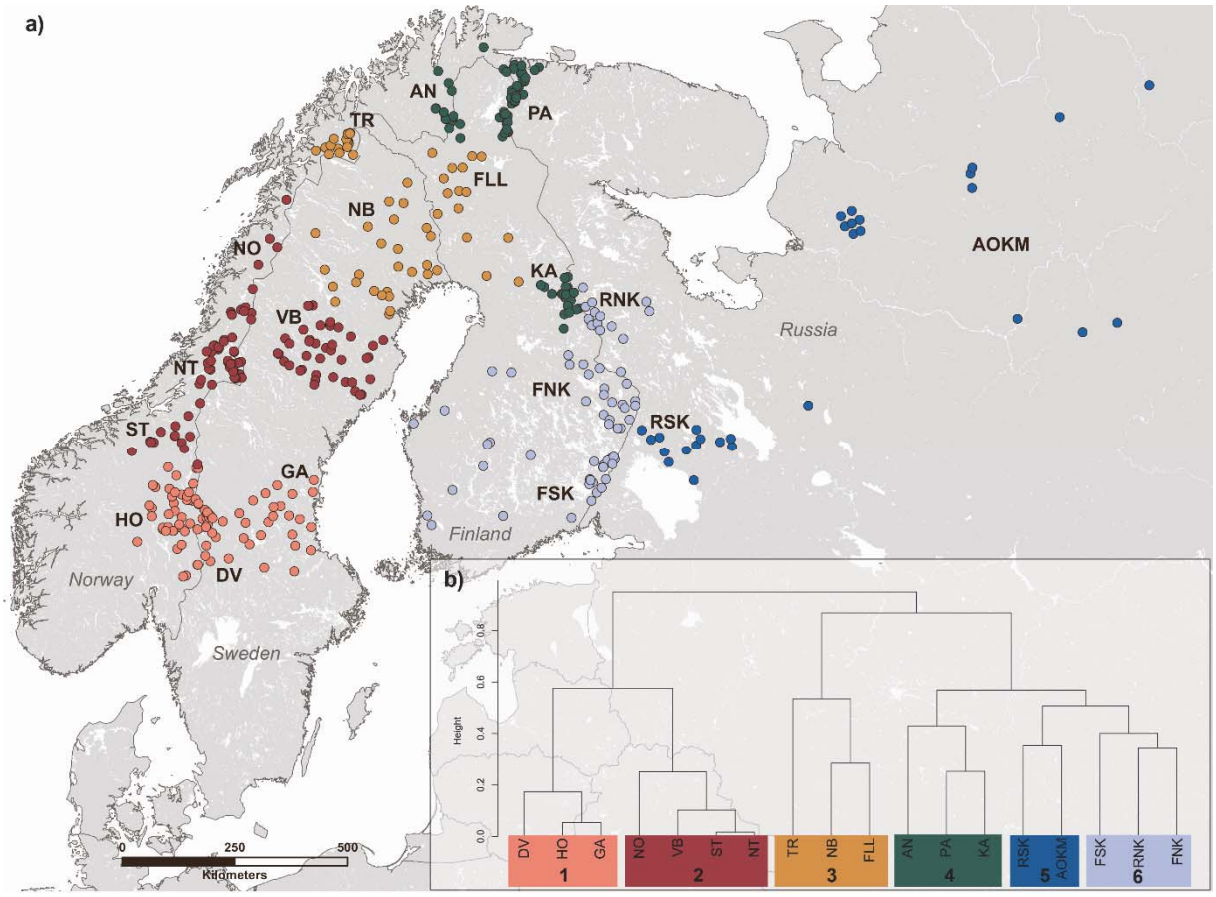


Figure 4: Geographic distribution of the AMOVA groups. a) Overview of the geographic location of the different haplotype groups of male brown bears in Northern Europe formed from the *hclust* analysis and the subsequent AMOVA. The color of each circle represents the group to which the individuals were assigned. The colors correspond to those given in inset b). The location of sampling areas are indicated as follows: **HO**= Hedmark/Oppland, **DV**= Dalarna/Värmland, **GA**= Gävleborg, **ST**= Sør-Trøndelag, **NT**= Nord-Trøndelag, **NO**= Nordland, **VB**= Västerbotten, **TR**= Troms, **NB**= Norrbotten; **AN**= Anarjohka, **FLL**= Finnish Lappland, **PA**= Pasvik Valley, **KA**= Kainuu, **RNK**= Russian Northern Karelia, **RSK**= Russian Southern Karelia, **FNK**= Finnish Northern Karelia, **FSK**=Finnish Southern Karelia, **AOKM**= Arkhangelsk, Komi. b) Result of the clustering analysis performed with the function *hclust* in R.; the colors correspond to those used in the map. Groupings for the subsequent AMOVA are indicated by the colored rectangles, the number in each rectangle indicates the name of the respective AMOVA group.

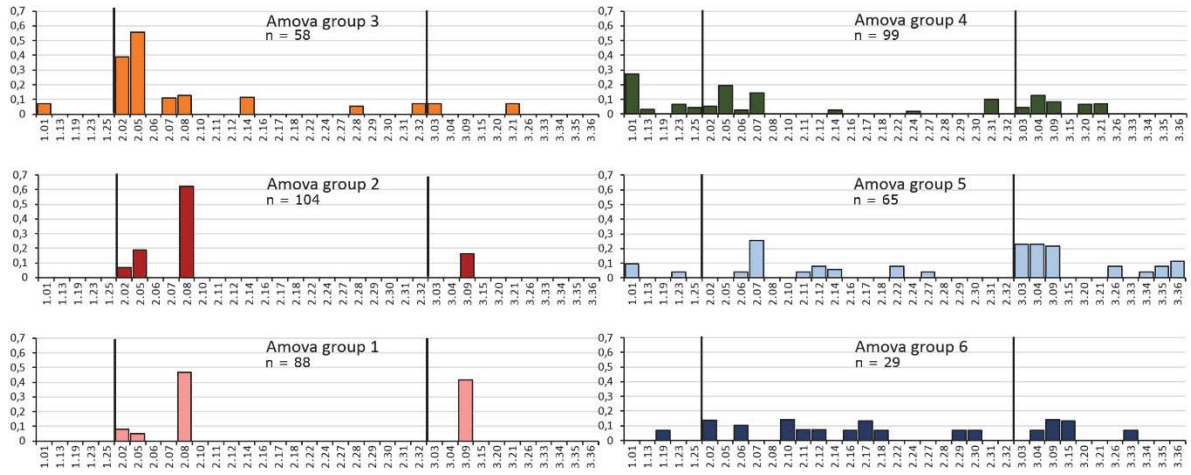


Figure 5: Y haplotype frequencies of extant male brown bears in Northern Europe calculated for each AMOVA group, the colors correspond to the ones used in Figure 3. All x-axes are identical, each bar representing the respective haplotype frequencies in each AMOVA group. Haplotypes are listed in consecutive order, beginning with haplotypes of haplogroup 1. The black vertical lines in each plot separate the haplogroups from each other.

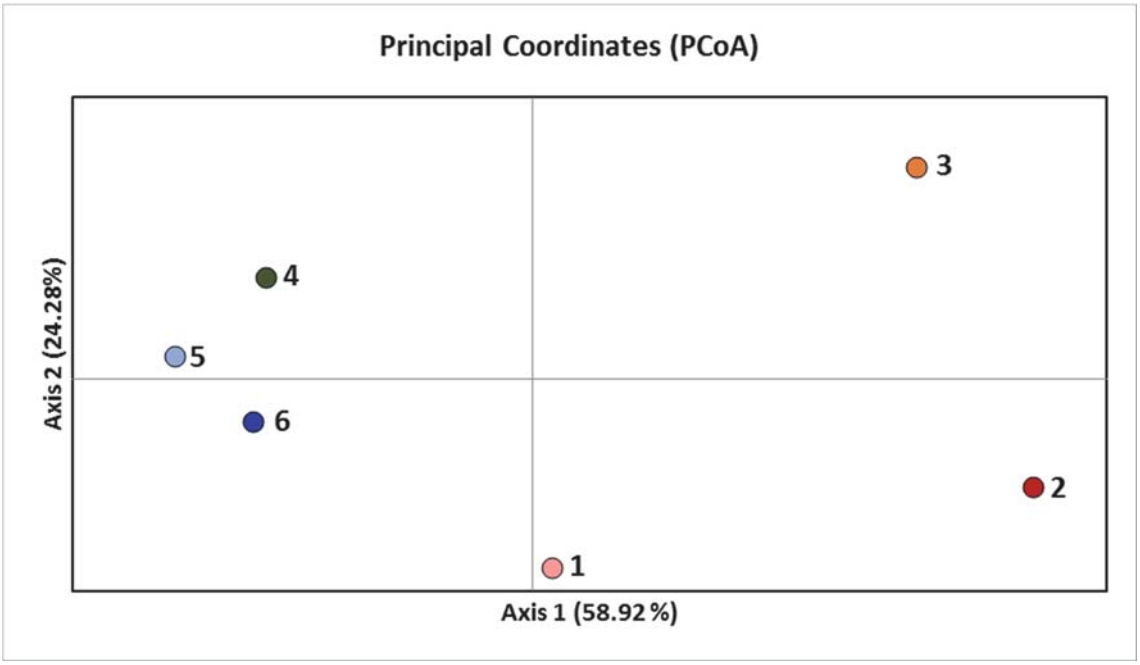


Figure 6: PCA ordination graph of the genetic differentiation based on matrices of pairwise F_{ST} between AMOVA groups. The circles are colored corresponding to Figure 3. The percentage values on the axes indicate how much variation is explained by the respective axis.

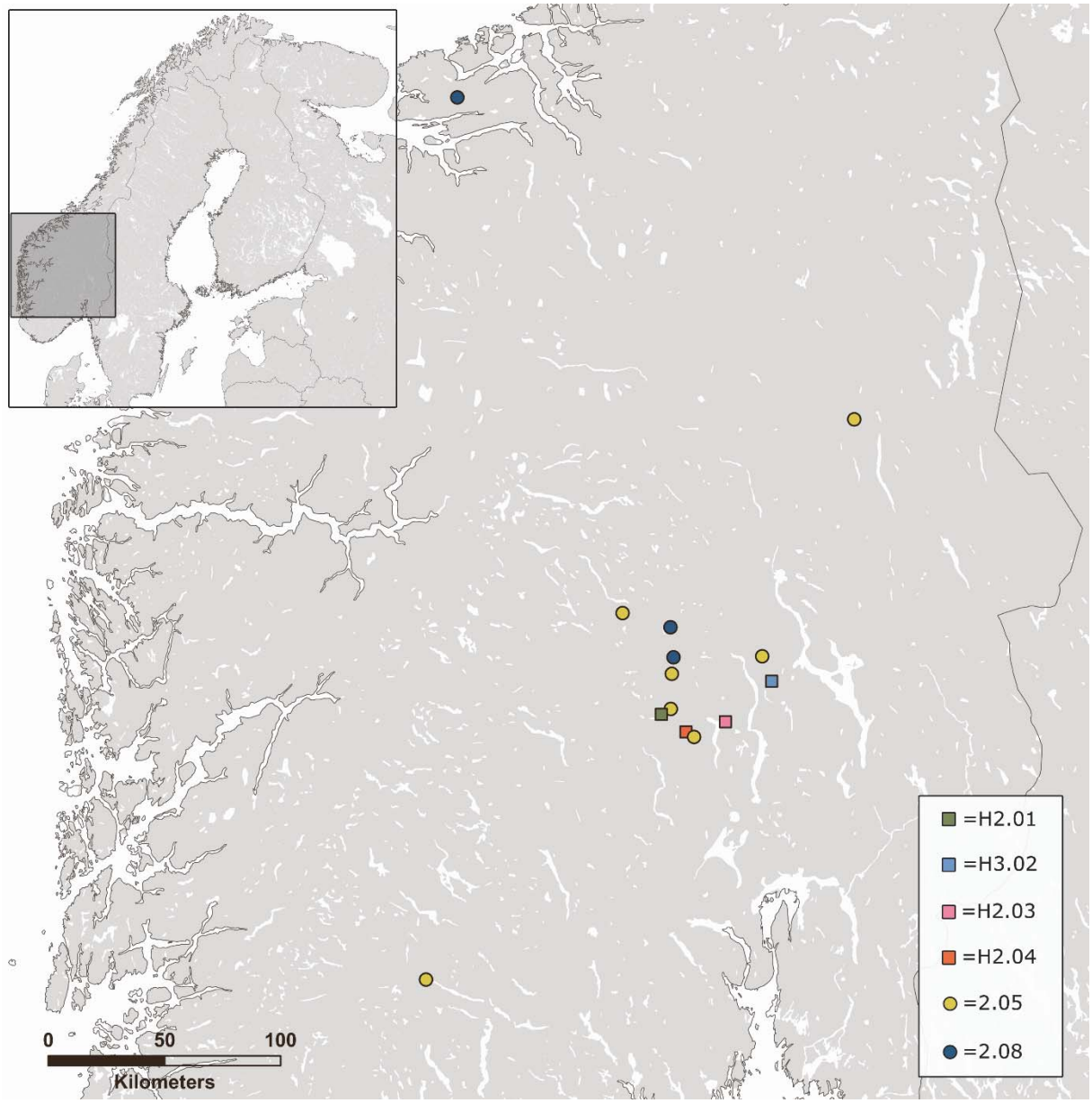


Figure 7: Location and assigned haplotype of 14 historical male brown bear samples in Norway (1780-1920). Circles symbolize the haplotypes that also were found in the extant brown bear population and squares indicate the four haplotypes exclusive for the historical samples.

Appendix 1

Table S1: Primer information, including primer sequence, observed allele size range and concentrations in the two multiplex sets A and B of the brown bear in Northern Europe. The concentration given is the final concentration of the forward and reverse primer, respectively, of each primer pair in the 10 µl reaction mix.

Set	Marker	Primer sequence (5'-3')[a]	Size range (bp)	Final conc.
A	UarY369.1	F:TCCCTGAATGAGCAGTAGCC R:GGGGTATTGCGTTGCATTGG	259-271	0.2 µM
A	UarY15020.1	F:TGCAATTTCTCTCAAACAACTTCCT R:GCGATGAAGGTCAGAGCAGT	187-189	0.1 µM
A	UarY318.1	F:GGGATCAAGCCCCACATCAA R:ACTTGTAGATGCACATCTGTGGT	281-289	0.2 µM
A	UarY69217.1	F:CTCCACCTTGTCTGCCACTC R:TTCCCCTCCCTTTCTGTCCT	243	0.1 µM
A	UarY318.6	F:GCTGGCTGTCTCTCTCTCTGA R:AAATTCCTTTGGAAACGTCCT	400-408	0.3 µM
B	UarY318.4	F:TACCTGGCTGGCTTTCTTGG R:CACTGTTGGTTTTGGCTCCG	213-215	0.1 µM
B	UarY318.2	F:CAGGCTGACACTGGGGATTT R:AAGAGGGAGTCATCTGGGGT	235	0.3 µM
B	UarY369.4	F:AGGCATCCATTCTATCACCAC R:TGTGGATGTATCTGCCCAAC	186-200	0.1 µM
B	UarY318.9	F:CACTCAGGCACCCCTCTATC R:TGGCCAGGATACAGAAACAAC	127-131	0.1 µM

[a] F = forward primer, R=reverse primer, sequence from Bidon et al. (2014)

Table S2: Primer and reporter sequences including the used dye for the three Y-SNP markers applied in this study of the brown bear in Northern Europe.

Marker	Primer sequence (5'-3')	Reporter Sequence (5'-3') - dye
UAY579.1B545	F: GTGCCATTTAAATCTTGGATTCTGCAT R: GTTTAGTAGTAGTCAAGATCCGAATAAAATGGT	CCTGACCC**C**ACTTAAA - VIC CCTGACCC**A**ACTTAAA - FAM
UAY318.2C713	F: AGGAGTGGCAATGCTGATTAAAAAA R: TGGCAAAGTTGGATGACATGATCA	ATAGGTGAGTTGT**T**TTGATTT - VIC ATAGGTGAGTTGT**C**TTGATTT - FAM
UAY318.2C839	F: ATTAAATACTCCAGACTTAAGTTTCAATGACA R: TTTTAAAGTATAATGTTTTCACTAAACATATGTTGATACT	ATTGGTAAATGTT**G**CATGTAT - VIC AATTGGTAAATGTT**A**CATGTAT - FAM

F = forward primer, R = reverse primer

Table S3: Full allelic profile for all haplotypes found in the extant brown bear study population in Northern Europe. The first digit of the respective haplotype names indicates the haplogroup, defined by the specific combination of Y-SNP alleles, the latter two digits indicate the haplotype profile and were assigned sequentially.

Haplotype	Y-STR									Y-SNP			Haplogroup
	UarY318.4	UarY369.1	UarY15020.1	UarY318.2	UarY318.1	UarY369.4	UarY69217.1	UarY318.6	UarY318.9	UAY579.1B545	UAY318.2C713	UAY318.2C839	
1.01	215	265	187	235	285	192	243	400	131	A	C	G	1
1.13	215	265	187	235	285	192	243	402	131	A	C	G	1
1.19	215	267	187	235	285	198	243	400	131	A	C	G	1
1.23	215	265	187	235	281	192	243	400	131	A	C	G	1
1.25	213	265	187	235	285	192	243	400	131	A	C	G	1
2.02	213	263	189	235	281	192	243	406	127	A	T	G	2
2.05	213	265	189	235	281	190	243	406	131	A	T	G	2
2.06	213	269	189	235	281	192	243	406	127	A	T	G	2
2.07	213	265	189	235	281	196	243	402	127	A	T	G	2
2.08	213	261	189	235	281	192	243	406	131	A	T	G	2
2.10	213	263	187	235	281	198	243	404	129	A	T	G	2
2.11	213	261	189	235	281	192	243	402	127	A	T	G	2
2.12	213	259	189	235	281	192	243	404	127	A	T	G	2
2.14	213	271	189	235	281	200	243	402	129	A	T	G	2
2.16	213	269	187	235	281	194	243	408	129	A	T	G	2
2.17	213	267	189	235	285	196	243	402	129	A	T	G	2
2.18	213	265	189	235	281	198	243	402	127	A	T	G	2
2.22	213	263	189	235	281	198	243	406	127	A	T	G	2
2.24	213	263	189	235	281	192	243	406	129	A	T	G	2
2.27	213	265	189	235	281	196	243	406	131	A	T	G	2
2.28	213	265	189	235	285	196	243	402	127	A	T	G	2
2.29	213	271	189	235	281	198	243	402	129	A	T	G	2
2.30	213	269	187	235	281	194	243	402	129	A	T	G	2
2.31	213	269	187	235	281	194	243	406	129	A	T	G	2
2.32	213	265	189	235	281	198	243	400	127	A	T	G	2
3.03	213	265	187	235	285	186	243	406	127	C	T	G	3
3.04	213	271	187	235	285	186	243	404	127	C	T	G	3
3.09	213	267	187	235	285	186	243	404	127	C	T	G	3
3.15	213	263	187	235	285	186	243	404	127	C	T	G	3
3.20	213	267	187	235	285	190	243	400	127	C	T	G	3
3.21	213	265	187	235	285	186	243	408	127	C	T	G	3
3.26	213	267	187	235	281	186	243	404	127	C	T	G	3
3.33	213	269	187	235	285	186	243	404	127	C	T	G	3
3.34	213	271	187	235	285	186	243	406	127	C	T	G	3
3.35	213	271	187	235	289	186	243	404	127	C	T	G	3
3.36	213	267	187	235	285	190	243	404	127	C	T	G	3

Table S4: Complete haplotype profile for the historical brown bears in Norway with at least seven successfully amplified markers. The haplotypes H2.01-04 and H3.02 were found only in the historical samples and not in the extant study population.

			Y-STR									Y-SNP		
Individual	**Province**	**Year**	**UarY318.4**	**UarY369.1**	**UarY15020.1**	**UarY318.2**	**UarY318.1**	**UarY69217.1**	**UarY69217.1**	**UarY318.9**	**UarY318.9**	**UAY579.1B545**	**UAY318.2C713**	**Haplotype**
HP-FE-01	Oppland	~1920	213	265	189	235	281	190	243	404	131	A	T	**H2.01**
HP-GR-01	Oppland	-	213	265	187	235	289	186	243	-	127	C	T	**H3.02**
HP-WA-03	Oppland	1840-70	213	271	189	235	-	190	243	-	131	A	T	**H2.03**
HP-AS-01	Oppland	-	215	263	189	235	281	190	243	-	131	A	T	**H2.04**
HP-2215	-	-	213	265	189	235	281	190	243	-	131	A	T	**2.05**
HP-7378	Telemark	1907	213	265	189	235	281	190	243	-	131	A	T	**2.05**
HP-AT-10	Hedmark	1780-1820	213	-	189	235	281	190	243	-	131	A	-	**2.05**
HP-BR-01	Oppland	~1870	213	265	189	-	281	190	243	-	131	A	T	**2.05**
HP-GA-01	Oppland	1880-1900	213	265	189	235	281	190	243	-	131	A	T	**2.05**
HP-SR-01	Oppland	1910	213	265	189	235	281	190	243	-	131	A	T	**2.05**
HP-SU-01	Oppland	~1850	213	265	189	235	281	190	243	-	131	A	T	**2.05**
HP-M-01	Buskerud	1899	213	261	189	235	281	192	243	406	131	A	T	**2.08**
HP-HA-02	Oppland	-	213	261	189	235	281	192	243	-	131	A	T	**2.08**
HP-R-04	Møre og Romsdal	-	213	261	189	235	281	192	243	-	131	A	T	**2.08**

Table S5: Pairwise F_{ST} values among sampling locations of brown bears in Northern Europe. Calculations performed in *Arlequin 3.5.1.2* based on Y-STR data, using number of different alleles as measure of genetic distance; the following sampling locations were grouped together for the AMOVA analysis: **group 1** = HO, GA, DV; **group 2** = ST, NT, VB, NO; **group 3**= TR, NB, FLL; **group 4**= AN, PA, KA; **group 5**= RNK, FNK, FSK; **group 6**= RSK, AOKM. The F_{ST} values among members of these respective groups are marked with gray. *) $p < 0.05$; **) $p < 0.001$

	HO	DV	GA	ST	NT	NO	VB	TR	NB	FLL	AN	PA	KA	RNK	FNK	FSK	RSK
DV	-0.024																
GA	-0.044	-0.059															
ST	0.088	0.010	0.045														
NT	0.114*	0.041	0.078	-0.057													
NO	0.263*	0.215*	0.274*	0.014	0.032												
VB	0.153*	0.083	0.119*	-0.049	-0.018	0.022											
TR	0.255**	0.258**	0.244*	0.163*	0.177**	0.306**	0.181**										
NB	0.270**	0.278*	0.256*	0.178*	0.182**	0.321**	0.153*	0.141*									
FLL	0.210*	0.187*	0.165*	0.101	0.130*	0.251*	0.111*	0.072	-0.005								
AN	0.219**	0.213*	0.173*	0.216*	0.256**	0.386**	0.260**	0.216**	0.180	0.048							
PA	0.193**	0.210**	0.165**	0.286**	0.319**	0.421**	0.338**	0.326**	0.331**	0.216**	0.061*						
KA	0.138**	0.145*	0.102*	0.216**	0.261**	0.366**	0.285**	0.246**	0.287**	0.160*	0.041	0.007					
RNK	0.145*	0.186*	0.122*	0.308**	0.349**	0.517**	0.384**	0.330**	0.395**	0.262**	0.180*	0.107*	0.037				
FNK	0.126*	0.182*	0.113*	0.325**	0.357**	0.508**	0.393**	0.363**	0.406**	0.311**	0.229**	0.140**	0.080*	-0.026			
FSK	0.152**	0.174*	0.123*	0.248**	0.286**	0.415**	0.312**	0.236**	0.296**	0.166*	0.070*	0.049*	0.000	0.027	0.078*		
RSK	0.062	0.066	0.033	0.152*	0.199**	0.348**	0.242**	0.224**	0.317**	0.183**	0.120*	0.096*	0.021	0.057	0.082*	0.065	
AOKM	0.104*	0.126	0.075	0.234*	0.285**	0.427**	0.324**	0.278**	0.356**	0.226**	0.172**	0.122*	0.033	0.007	0.031	0.063*	0.011

Location-specific haplotype distribution and frequencies

The Y haplotype 2.08 was clearly dominant among brown bears in Sweden and southern Norway, with a frequency of more than 0.5 in five of eight sampling locations, whereas in Northern Norway, Finland, and Russia, no single Y haplotype occurred in a frequency higher than 0.5 (Table 1). Y haplotype 2.05 was mainly distributed in northern Sweden and Finnish Lapland (FLL), and occurred in over 70 % of the males in Norrbotten (NB), which has high bear densities (Kindberg *et al.* 2011) (Table 1, figure S1). However, it occurred at only low frequencies in adjacent areas. Moreover, Y haplotype 3.09 was common in southern Norway (Hedmark/Oppland (HO), *f=0.448*) and southern Sweden (Dalarna/Värmland (DV), *f=0.375;* Gävleborg (GA), *f=0.429*), but also occurred in most other sampling locations in Northern Europe, albeit at lower frequencies (*f=0.037 - 0.214*), with the exception of Finnish Northern Karelia (FNK), where it showed frequencies similar to the western locations (*f=0.346)*. In contrast, Y haplotype 2.08 was frequent in the southern parts of Sweden and Norway (*f=0.414-0.750)*, was absent from Troms (TR) in northern Norway, occurred only at low frequencies in NB (*f=0.115)* and FLL (*f=0.143)*, and was absent further east (Table 1, Figure 2). Similarly, two haplotypes frequently occurring in the eastern locations, 2.07 (*f=0.093 - 0.308*) and 2.14 (*f=0.019 - 0-077*), were each found in only one location in the west (2.07: TR, *f=0.111*; 2.14: NB, *f=0.154*). These results suggested asymmetrical dispersal, with more males dispersing from west to east than vice versa, which is in agreement with previous studies (Kopatz *et al.* 2014).

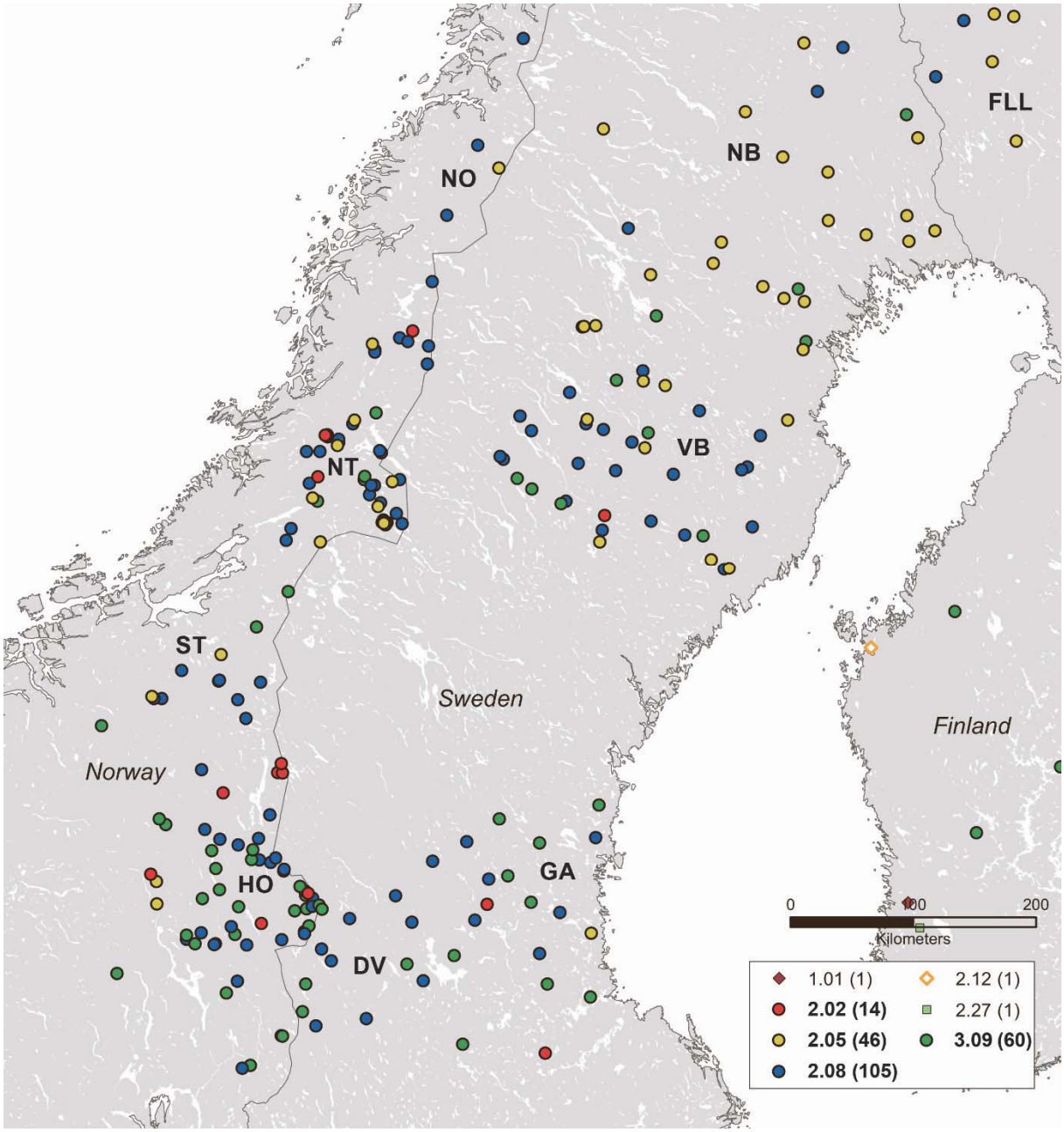


Figure S1: Close up map of the distribution of Y haplotypes among brown bears in Central and Southern Norway and Sweden, including the sampling locations Dalarna/Värmland *DV*, Hedmark/Oppland *HO*, Gävleborg *GA,* Sør Trøndelag *ST*, Nord Trøndelag *NT*, Nordland *NO* Västerbotten *VB* and Norrbotten *NB*. The symbols used for the respective haplotypes correspond to those used in the overview map, Figure 2, in the main article.

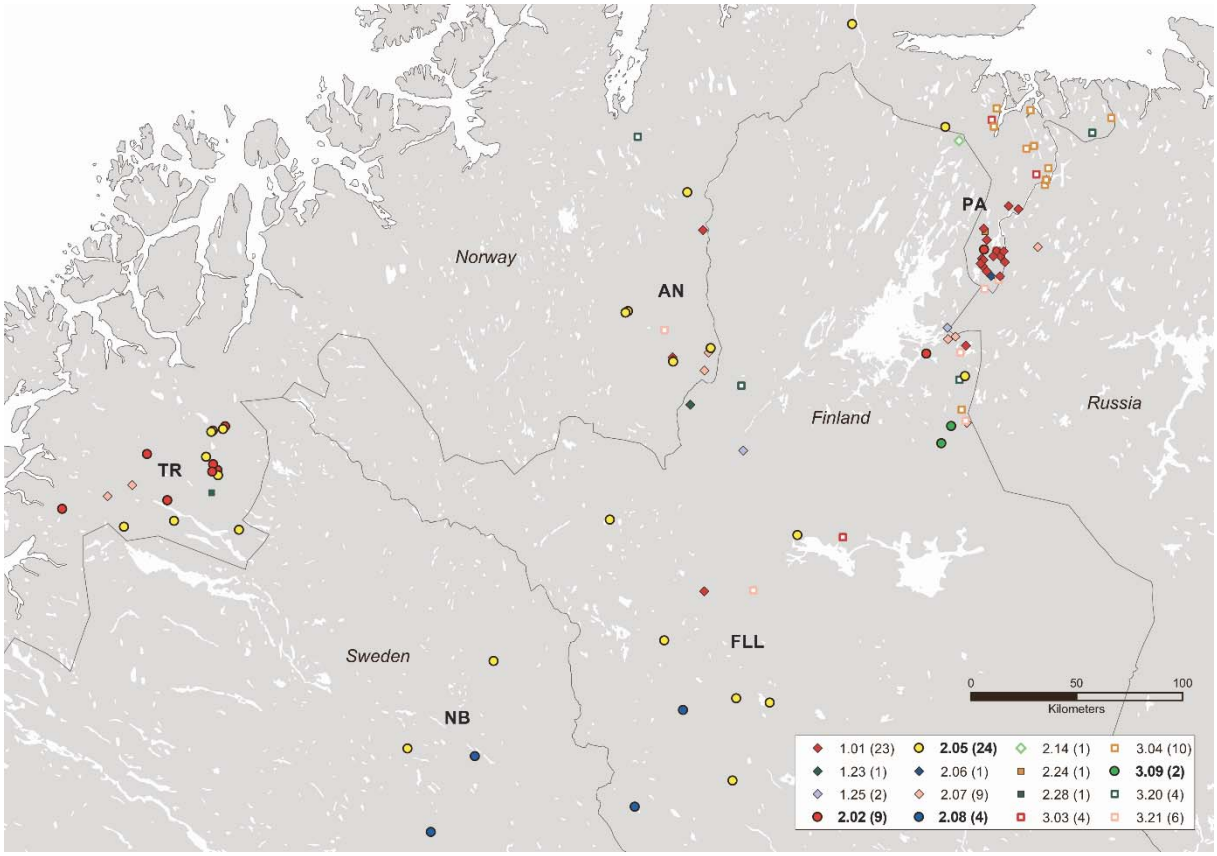


Figure S2: Close up map of the distribution of haplotypes among brown bears in Northern Norway and Finland, including the sampling locations Troms *TR*, Anarjohka *AN*, Finnish Lapland *FLL* and Pasvik *PA*. The symbols used for the respective haplotypes correspond to those used in the overview map, Figure 2, in the main article.

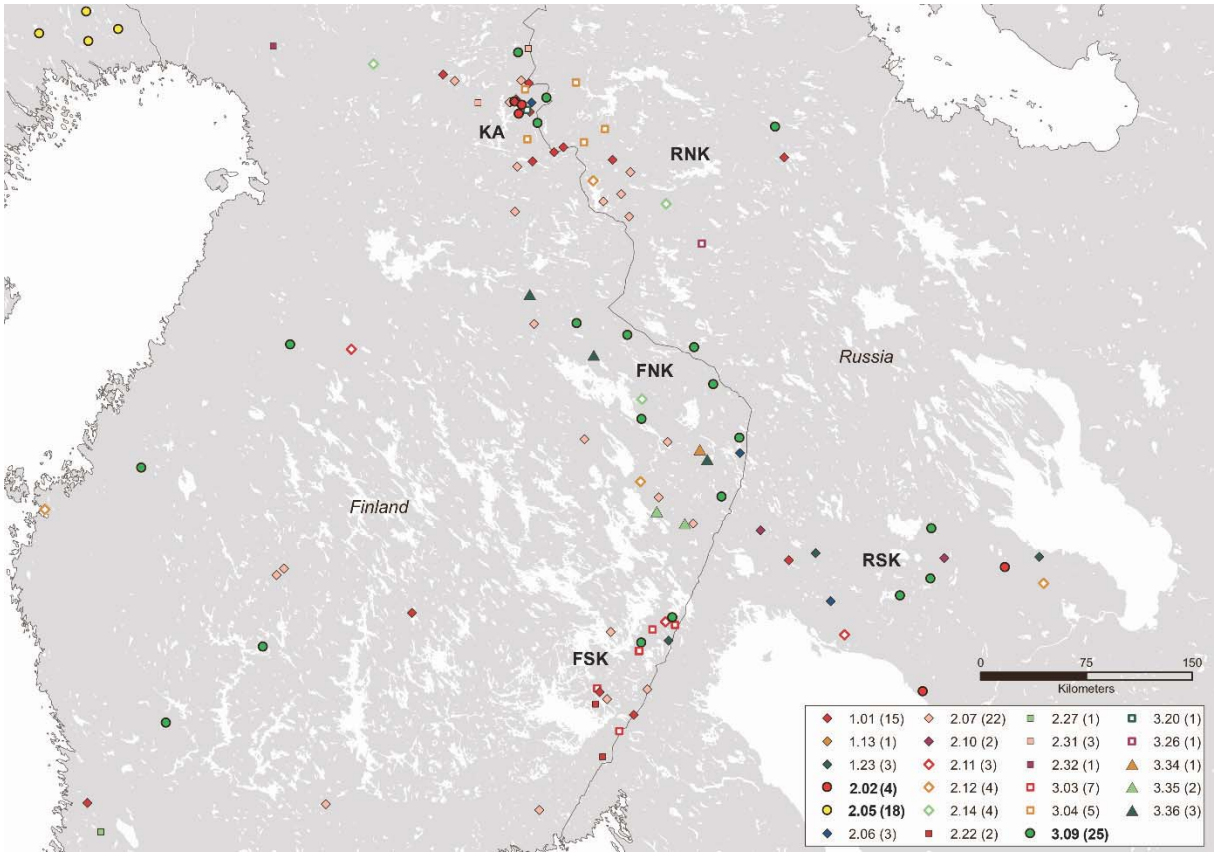


Figure S3: Close up map of the distribution of haplotypes among brown bears in in Central and Southern Finland, including the sampling locations of Kainuu *KA*, Finnish Northern Karelia *FNK*, Finnish Southern Karelia *FSK*, Russian Northern Karelia *RNK* and Russian Southern Karelia *RSK*. The symbols used for the respective haplotypes correspond to those used in the overview map, Figure 2, in the main article.

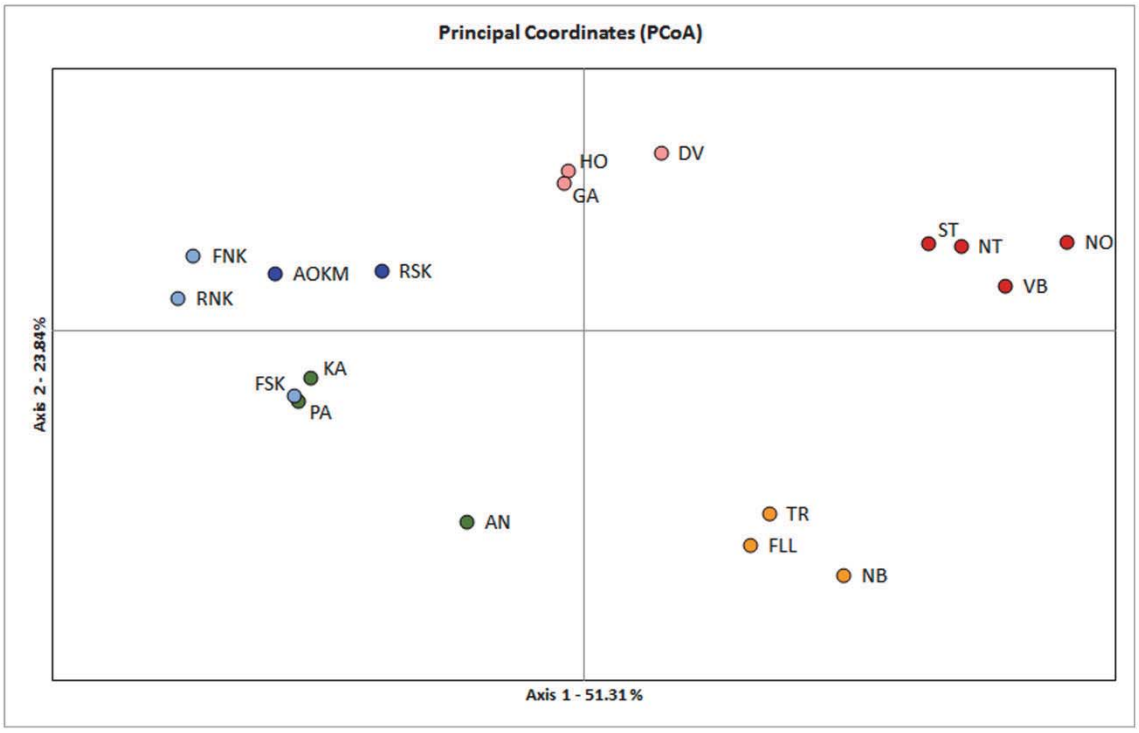


Figure S4: Results of the PCoA Analysis performed with GenAlEx 6.5.01, based on the results of the estimation of pairwise F_{ST} values among sampling locations of brown bears in Northern Europe performed with Arlequin 3.5.1.2. The circles are colored according to the colors chosen for the different AMOVA groups and correspond to those shown in Figure 4 in the main article. The percentage values on the axes indicate how much variation is explained by the respective axis.

Appendix 2

Results of the AMOVA, testing alternative groupings

The results are presented in three sections: i) entire data set used, ii) only the "western" groups on the Scandinavian peninsula, iii) only the "eastern" groups of northwestern Norway, Finland and Western Russia. The grouping presented in the article with the estimated values is given always first in the separate sections, using fat font. Significant values are indicated with *.

Section I: entire dataset

Genetic structure to test :

No. of Groups = 6
#Group1:"HO","DV","GA"
#Group2:"ST","NT","NO","VB"
#Group3:"NB","FLL","TR"
#Group4:"AN","PA","KA"
#Group5:"FSK","RNK","FNK"
#Group6:"RSK","AOKM"

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	**5**	**155.450**	**0.40144 Va**	**20.40***
Among populations within groups	**12**	**24.529**	**0.02251 Vb**	**1.14**
Within populations	**425**	**656.109**	**1.54379 Vc**	**78.45***

Genetic structure to test :

No. of Groups = 2
#Group1:"HO","DV","GA","ST","NT","NO","VB"
#Group2:"NB","TR","FLL","KA","PA","AN","RNK","FNK","FSK","RSK","AOKM"

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	63.445	0.24643 Va	12.11*
Among populations within groups	16	116.534	0.24401 Vb	12.00*
Within populations	425	656.109	1.54379 Vc	75.89*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 2
#Group1:"HO","DV","GA","ST","NT","NO","VB","NB","TR"
#Group2:"FLL","KA","PA","AN","RNK","FNK","FSK","RSK","AOKM"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	81.008	0.33148 Va	16.00*
Among populations within groups	16	98.971	0.19650 Vb	9.48*
Within populations	425	656.109	1.54379 Vc	74.52*

```
---------------------------
Genetic structure to test :
---------------------------

No. of Groups = 2
#Group1:"HO","DV","GA","ST","NT","NO","VB","NB","TR","FLL"
#Group2:"KA","PA","AN","RNK","FNK","FSK","RSK","AOKM"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	87.617	0.36858 Va	17.62*
Among populations within groups	16	92.362	0.17900 Vb	8.56*
Within populations	425	656.109	1.54379 Vc	73.82*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 3
#Group1:"HO","DV","GA","VB","NO","NT","ST"
#Group2:"FLL","TR","NB"
#Group3:"AOKM","RSK","FSK","FNK","RNK","KA","PA","AN"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	112.389	0.38096 Va	18.58*
Among populations within groups	15	67.590	0.12575 Vb	6.13*
Within populations	425	656.109	1.54379 Vc	75.29*

```
--------------------------
Genetic structure to test :
--------------------------
No. of Groups = 3
#Group1:"HO","DV","GA","VB","NO","NT","ST","NB","FLL","TR"
#Group2:"AOKM","RSK","FSK","FNK","RNK"
#Group3:"KA","PA","AN"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	102.796	0.34616 Va	16.92*
Among populations within groups	15	77.183	0.15546 Vb	7.60*
Within populations	425	656.109	1.54379 Vc	75.48*

Section II: Only "Western" groups

```
--------------------------
Genetic structure to test:
--------------------------
No. of Groups = 3
#Group1:"HO","DV","GA"
#Group2:"NT","ST","NO","VB"
#Group3:"FLL","TR","NB"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	**2**	**47.716**	**0.27946 Va**	**17.95***
Among populations within groups	**7**	**8.477**	**-0.00321 Vb**	**-0.21**
Within populations	**240**	**307.438**	**1.28099 Vc**	**82.26***

```
--------------------------
Genetic structure to test :
--------------------------
No. of Groups = 2
#Group1:"HO","GA","DV","ST","NT","NO","VB"
#Group1:"NB","TR","FLL"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	24.773	0.23306 Va	14.34*
Among populations within groups	8	31.421	0.11066 Vb	6.81*
Within populations	240	307.438	1.28099 Vc	78.84*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 2
#Group1:"HO","GA","DV"
#Group2:"ST","NT","NO","VB"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	22.943	0.24029 Va	15.65*
Among populations within groups	5	2.814	-0.03327 Vb	-2.17
Within populations	185	245.825	1.32879 Vc	86.52*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 3
#Group1:"HO","GA","DV", "ST", "NT"
#Group2:"NO","VB","NB"
#Group3:"FLL","TR"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	39.317	0.22993 Va	14.73*
Among populations within groups	7	16.877	0.04995 Vb	3.20*
Within populations	240	307.438	1.28099 Vc	82.07*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 3
#Group1:"HO","GA","DV", "ST", "NT"
#Group2:"NO","VB"
#Group3:"FLL","TR","NB"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	23.480	0.14685 Va	10.63*
Among populations within groups	6	12.640	0.04981 Vb	3.61
Within populations	183	216.818	1.18480 Vc	85.76*

Section III: Only "Eastern" groups

Genetic structure to test :
No. of Groups = 3
#Group1:"FSK","FNK","RNK"
#Group2:"KA","PA","AN"
#Group3:"AOKM","RSK"

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	**2**	**20.117**	**0.11709 Va**	**5.68***
Among populations within groups	**5**	**16.052**	**0.05801 Vb**	**2.82***
Within populations	**185**	**348.671**	**1.88471 Vc**	**91.50***

Genetic structure to test :
No. of Groups = 3
#Group1:"FSK","FNK","RNK","KA"
#Group2:"PA","AN"
#Group3:"AOKM","RSK"

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	18.025	0.08388 Va	4.09*
Among populations within groups	5	18.144	0.08114 Vb	3.96*
Within populations	185	348.671	1.88471 Vc	91.95*

Genetic structure to test :
No. of Groups = 3
#Group1:"FSK","FNK"
#Group2:"KA","PA","AN"
#Group3:"AOKM","RSK","RNK"

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	18.254	0.08994 Va	4.39*
Among populations within* groups	5	17.915	0.07572 Vb	3.69*
Within populations	185	348.671	1.88471 Vc	91.92*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 3
#Group1:"KA","PA","AN","RNK","FNK"
#Group2:"FSK","RSK"
#Group3:"AOKM"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	2	7.904	-0.02702 Va	-1.34
Among populations within groups	5	28.265	0.15598 Vb	7.75*
Within populations	185	348.671	1.88471 Vc	93.60*

```
---------------------------
Genetic structure to test :
---------------------------
No. of Groups = 2
#Group1:"KA","PA","AN","RNK","FNK"
#Group2:"AOKM","FSK","RSK"
```

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	4.539	-0.01196 Va	-0.59
Among populations within groups	6	31.630	0.14763 Vb	7.31*
Within populations	185	348.671	1.88471 Vc	93.28*

PAPER IV

A multiscale analysis of sex-dependent population structure and gene flow: the case of the Scandinavian brown bear (*Ursus arctos*)

Julia Schregel[1,2], Hans Geir Eiken[1], Jon E. Swenson[2,3], Snorre B. Hagen[1]

[1] Norwegian Institute of Bioeconomy, NIBIO - Svanhovd, 9925 Svanvik, Norway
[2] Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, 1432 Ås, Norway
[3] Norwegian Institute for Nature Research, 7485 Trondheim, Norway

Keywords: spatial autocorrelation, sex-biased dispersal, fine-scale genetic structure, female philopatry

Abstract:

Dispersal, i.e. the permanent movement of an individual out of its natal home range, has a large impact on the demographic and genetic structure of a population and may show considerable variation between the sexes. The inclusion of genetic methods to the study of dispersal behavior may help identify the true costs and benefits of different dispersal patterns and to reveal the genetic impact of these patterns on the population structure. Here, we analyze the fine-scale genetic structure of the Scandinavian brown bear, a species with male-biased dispersal and female philopatry. Our data consists of 1531 individual genotypes based on eight microsatellite markers, obtained from 2006-2013 from Sweden and Norway in the course of monitoring schemes. We first analyzed the genetic structure of the population using the Bayesian clustering algorithm STRUCTURE in a hierarchical manner, followed by the analysis of spatial autocorrelation and estimation of genetic differentiation among clusters and subclusters for each sex separately. To reveal scale-dependency of sex-specific population structure we also analyze the data at the large scale. Our results show a clear sex-bias in the spatial genetic

structure, with females displaying a positive correlation between spatial and genetic distance. However, there is also a clear area affect, with a general increase of population structure and genetic differentiation towards the north. In particular, males displayed signs of positive spatial autocorrelation in the more northern areas, pointing to a larger variability of sex-specific gene flow patterns than could be assumed based on previous, telemetric data. Contrasting the clear, fine-scale sex-bias, on the large scale, the magnitude of bias between males and females was less pronounced, suggestive of scale-dependent cost/benefit arithmetic for dispersal.

Introduction

Dispersal, the permanent movement of an individual out of its natal home range, is a major driver of change in a population, impacting its dynamics, demography, and genetics across a wide range of spatial and temporal scales (Ronce 2007). Documenting the variation in dispersal patterns, both within and between species, is important to understand the ecological and evolutionary significance of dispersal (Murrell *et al.* 2002; Rousset & Gandon 2002). Knowledge about dispersal patterns is also important for wildlife conservation and management, e.g., to outline rescue plans for threatened species, improve inter-population connectivity, and predict impacts of climatic change, biological invasions, and/or anthropogenic disturbances. Because dispersal does not necessarily translate into gene flow (Johnson & Gaines 1990; Kitanishi & Yamamoto 2015), the application of traditional approaches to the study of dispersal, e.g., radio telemetry, may sometimes not be able to fully assess the costs and benefits of a certain dispersal pattern. Particularly over large geographic areas, the application of genetic methods offers additional knowledge that may help to determine the impact, mechanisms, and patterns of dispersal (Lawson Handley & Perrin 2007). Consequently, the application of genetic methods to the study of dispersal, though relatively new, is quickly increasing (Driscoll *et al.* 2014).

Numerous mechanisms have been proposed to explain the adaptive significance of dispersal (reviewed in Clutton-Brock & Lukas 2012; Dobson 2013; Lawson Handley & Perrin 2007), ranging from colonization of new territories and escape from high population densities to reduction of inbreeding

and kin competition (Perrin & Goudet 2001). Consistent with the assumption that dispersal is a plastic trait, empirical studies have shown that the timing, rate, and distance of dispersal may be influenced by a number of proximate conditions (Ronce 2001), such as crowding, local kinship, habitat conditions, social structure, and individual fitness (Legagneux *et al.* 2009; Solmsen *et al.* 2011; Hardouin *et al.* 2012; Vercken *et al.* 2012). The explanation for dispersal probably differs between long- and short-distance dispersal (Ronce 2001); whereas long-distance dispersal may be needed to colonize a new territory or escape crowding, short-distance dispersal may be sufficient to reduce inbreeding or kin competition (Perrin & Goudet 2001). Moreover, there is considerable variation in the timing of dispersal, with many species undergoing secondary dispersal later in life and not all individuals of a species undergo natal dispersal as juveniles (e.g, Dobson 1982; Smale *et al*. 1997; Schülke 2003). Studying the variation within and between species in the distribution of dispersal times, rates, and distances may therefore help us understand the underlying mechanisms of dispersal.

Many species display sex-biased dispersal, i.e., one sex disperses farther and more frequently from the natal home range than the other, thus potentially passing on their genes more frequently over larger areas than the less or nondispersing sex (Greenwood 1980; Lawson Handley & Perrin 2007; Greminger *et al.* 2010). Accordingly, the philopatric sex should shape genetic population structure locally, whereas the dispersing sex should act as genetic mediator among and within regions. A wide number of variations of this pattern has been documented in wild populations, probably because social organization and mating strategy are tightly linked to the shape of dispersal behavior (reviewed in Lawson Handley & Perrin 2007). For example, the magnitude of the sex bias of dispersal behavior seems to correlate positively with the degree of social complexity (Greenwood 1980; Devillard *et al.* 2004). Within a species, the ratio of the male to female dispersal distance also may vary depending on whether the population is at a stable state or expanding (Swenson *et al.* 1998; Kojola *et al.* 2003; Kojola & Heikkinen 2006).

Patterns of dispersal and gene flow also may vary depending on the scale at which they are studied, which indicates that the role of explanatory processes varies in a scale-dependent manner (Fontanillas *et al.* 2004; Vuilleumier & Fontanillas 2007; Yannic *et al.* 2012). Hence, estimates derived

from small-scale studies, i.e., on the scale at which individuals are able to interact with each other (Waples & Gaggiotti 2006), might not be transferred easily to larger scales (Underwood *et al.* 2005; de Knegt *et al.* 2010). Despite this, there have been only a few analyses of genetic data on different scales. Examples are Gorospe & Karl (2013), who showed that in coral reefs, isolation by distance (IBD) is detectable on the intrareef scale, but disappears when enlarging the spatial frame to the interreef scale; and Gabrielsen *et al.* (2013), who reported a similar phenomenon in the wood frog, *Lithobates sylvaticus*, which showed isolation by distance (IBD) and genetic structuring locally, but not regionally. Moreover, scale dependency of the analysis of genetic data also has been shown in a few landscape genetic studies, e.g., Keller *et al.* (2013) reported that the fit of models improved greatly when considering individuals/populations only at a certain (here the lowest) spatial scale, whereas fit was generally worse when analyzing all populations/individuals at the same time, i.e. at the largest spatial scale. The magnitude of the sex bias in dispersal also may depend on scale; whereas local (short-distance) dispersal may be strongly biased towards one sex, regional (long-distance) dispersal may be more balanced between the sexes (Wilder *et al.* 2004; Gauffre *et al.* 2009; Heyer *et al.* 2012; Vangestel *et al.* 2013).

Brown bears (*Ursus arctos*) have a circumpolar distribution, exhibit both long- and short-distance dispersal (Støen *et al.* 2006), and have been the subject of several phylogenetic and population genetic studies using mitochondrial, autosomal, and Y-chomosomal markers (Cronin *et al.* 1991; Taberlet & Bouvet 1994; Taberlet *et al.* 1998; Saarma & Kojola 2007; Korsten *et al.* 2009; Keis *et al.* 2013). Because of a successful comeback after near extinction (Swenson *et al.* 1995; Kojola & Määtä 2004), the brown bear in Northern Europe has functioned as model species in several large-scale genetic studies of population connectivity, recovery and range expansion (Kopatz et al. 2014; Hagen et al. 2015). Radio-telemetry studies have shown that the brown bear exhibits male-biased dispersal, with male subadult dispersers emigrating at higher rates and across larger distances than female dispersers (McLellan & Hovey 2001; Støen *et al.* 2006; Zedrosser *et al.* 2007). Natal dispersal distances are inversely correlated with the local density in the vicinity of an individual (Støen *et al.* 2006). In terms of dispersal probability, males are not significantly affected by population density or population sex-

ratio (Zedrosser *et al.* 2007). Female dispersal probability, on the other hand, seems inversely correlated to body weight and the age of the mother, indicating female competition for philopatry (Zedrosser *et al.* 2007). The influence of the mother's age may indicate that the competition among siblings for philopatry decreases with the mother's age (Zedrosser *et al.* 2007). This may also indicate that the presence or absence of female kin in the vicinity of the female offspring has an influence on the dispersal probability of subadult females (Zedrosser *et al.* 2007). In addition, a genetic analysis has shown that females tend to form matrilineal assemblages, with a positive correlation between relatedness and home range overlap (Støen *et al.* 2005).

Knowledge of the dispersal behavior of brown bears seems thus to be well understood. However, it is based for the most part on the physical movement of individuals in space and there is no knowledge about how much of this dispersal actually translates into gene flow. In this study, we wanted to expand the knowledge about dispersal behavior in this continuously distributed large carnivore and investigate what influence the observed dispersal behavior has on population genetic structure. Because this system is so well understood in terms of physical movement, the application of genetic methods may contribute to the understanding of the evolutionary causes of dispersal by pinpointing where dispersal and gene flow concur and where they do not. This way, proximate conditions associated with dispersal and the possibly resulting gene flow may be identified more easily, thus furthering the understanding of the underlying processes. To this aim, we performed a sex-specific analysis of genetic population structure, using the genotypes of 1531 brown bears, obtained in the course of the national monitoring programs in Sweden and Norway. We tested whether male and female population structure differed and whether fine-scale population structure reflected the pattern of male-biased dispersal and female philopatry. The spatial extent of sampling covered the entire distribution zone of the Scandinavian brown bear population with a high sampling density, which allowed us to analyze the data at different spatial scales, thereby gaining new knowledge about the genetic effect of male-biased dispersal and female philopatry on the small as well as on the large scale.

Material and Methods

Study species

The Scandinavian brown bear experienced a severe reduction in population size during the last century, caused mainly by persecution, but numbers have increased considerably in recent decades (Swenson et al. 1994, 1995; Swenson et al. 2000). More precisely, in Sweden ~130 individuals survived in three refuge areas, whereas Norway was virtually devoid of bears by the 1920s (Swenson *et al.* 1995). After the implementation of protection measures, brown bears have recovered and since the 1960s and 1970s, the population has steadily increased to an estimated number 3,300 in 2008 (Swenson *et al.* 1995; Kindberg *et al.* 2011; Naturvårdsverket 2014). Since then, there has been a slight negative trend, with the latest population size estimate of 2,800 bears in Sweden in 2013 (Naturvårdsverket 2014). During the course of population expansion in Sweden, Norway was repopulated and there is now a permanent population of a minimum of 147 individuals, though mostly in vicinity of the Norwegian-Swedish border, (Aarnes *et al.* 2014). The genetic structure of the Swedish population seems to consist of three genetic clusters, which correspond roughly to the position of the three refuge areas, which may also be characterized as female core areas (Swenson *et al.* 1998; Waits *et al.* 2000; Manel *et al.* 2004). Gene flow towards the brown bear population in Finland and Russia is limited (Schregel *et al.* 2012; Kopatz *et al.* 2014).

Sampling and genetic analysis

The Scandinavian brown bear population is monitored by identifying individuals by genotyping noninvasively obtained samples, mostly fecal, but also hair samples. The tissue of legally shot bears is also analyzed. Every sample is recorded in the monitoring database with the location and identity, if a genotype had been assigned unambiguously. For this study we utilized the genotypes based on the eight microsatellite markers (Short Tandem Repeats, STRs) used in the course of the monitoring, resulting in a total of 1461 brown bears (707 females, 754 males). Genetic monitoring was established in 2006, with new individuals identified and added every

year since, so that the origin of individual genotypes ranged from 2006 to 2013. To ensure unambiguous identification of individuals, the laboratory follows strict analysis protocols, which are accredited according to the EN ISO/IEC 17025 standard and have been described in detail previously (Eiken *et al.* 2009; Andreassen *et al.* 2012; Kopatz *et al.* 2012; Schregel *et al.* 2012). In addition, we included 70 samples (35 males and 35 females) from legally shot bears from the county of Jämtland in Sweden. The Swedish National Veterinary Institute stores tissue samples of shot bears. These samples were chosen to achieve a fairly even sample distribution for each sex across the county. Most of the selected samples originated from 2005-2013. Three females were from 1995, 1998 and 1999, and two males from 2003 and 2004.

Statistical Analysis

We analyzed population structure in a hierarchical manner to examine the effects of sex-biased dispersal and spatial scale of genetic analysis. First, we analyzed the entire dataset using STRUCTURE v.2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003), assuming population admixture and correlated allele frequencies. We set the maximum number of populations to K=10 and performed ten independent runs for each K, with a burn-in period of 100,000 Markov-Chain-Monte-Carlo (MCMC) iterations and a subsequent sampling of 1,000,000 MCMC iterations. We processed the results using Structure Harvester (Earl & Vonholdt 2012), which implements the *ad hoc* approach of (Evanno *et al.* 2005). In a second step, we ran an analysis separately for males and females, using the same parameters as above. Based on the results of this second-step analysis, we assigned individuals to one of the inferred clusters, using a membership value of $q \geq 0.7$ as a threshold value. Even though a membership value of $q \geq 0.6$ has been proposed as sufficient to assign an individual to a genetic cluster (Coulon *et al.* 2008), we applied the slightly higher threshold value, which has been used in previous bear studies

(Pelletier *et al.* 2012; Kopatz *et al.* 2014). In a last step, we reanalyzed each inferred cluster separately for each sex. In this analysis, we set the maximum number of inferred populations to K=5; all other parameters were as described above.

After the assessment of genetic structure, we analyzed the genetic differentiation between pairs of individuals at different spatial scales. First, we estimated the scale of isolation by distance (IBD) with the help of the kinship coefficient by (Loiselle *et al.* 1995), implemented in the program SPAGeDi v.1.4c (Hardy & Vekemans 2002). We did this initially for the entire dataset with males and females combined, and then separately for each sex, using a distance class size of 40 kilometers. In order to explore the importance of scale in the statistical analysis of genetic data, we subsequently performed the analysis for each cluster and sex separately. For this second step we reduced the size of the distance class to 5 kilometers.

To assess the fine-scale genetic structure of male and female brown bears, we performed a spatial autocorrelation analysis with GenAlEx (Peakall & Smouse 2006, 2012), separately for males and females. Because we wanted to compare the extent and magnitude of sex-biased dispersal among the different clusters, we separated the data according to cluster, excluding individuals with a membership value $q<0.6$. We also chose this approach, because pooling data from genetically differentiated populations can lead to an inflation of the genetic correlation coefficient r (Banks & Peakall 2012). To tackle this issue within the clusters where we found substantial substructure, we reran the analysis of spatial autocorrelation using the multiple populations approach (Peakall *et al.* 2003), treating the detected subclusters and admixed individuals as population units. This approach sums the individual components of nominator and denominator of r across populations and then calculates r_c as a division of total nominator and denominator, rather than using a simple arithmetic mean, in order to give each pairwise comparison for each population equal contribution to r_c (Peakall *et al.* 2003).

Genetic diversity within and differentiation among clusters

We calculated number of alleles and observed and expected heterozygosity using GenAlEx (Peakall & Smouse 2006, 2012) and inbreeding coefficient using Genetix 4.05.2 (Belkhir *et al.* 1996-2004) for each cluster and sex separately. We also used GenAlEx (Peakall & Smouse 2006, 2012) to estimate F_{ST} to assess the genetic differentiation among clusters, for males and females separately and for sexes combined. We also estimated pairwise F_{ST} among subclusters, for the females of clusters 1 and 3, as well as for the males of cluster 3; unassigned individuals were included as an additional cluster in this calculation. We did not do this for the females and males of cluster 4, however, as the number of unassigned individuals was <5.

Though widely used in population genetic studies, the validity of using F_{ST} as a measure of population differentiation has been questioned, especially when applying it to highly polymorphic markers, such as microsatellites, due to its dependency on within-population diversity (Hedrick 1999; Balloux *et al.* 2000; Balloux & Lugon-Moulin 2002; Meirmans & Hedrick 2011). G_{ST} has been developed as an analogue to F_{ST} to be used specifically with highly polymorphic markers (Nei 1973, 1977; Nei & Chesser 1983) and has found wide application (e.g. Munthali *et al.* 2013; Haynes *et al.* 2014; Koch *et al.* 2014). G_{ST} has been criticized as suffering from the same limitations as F_{ST}, which has led to the development of the estimate D, which is supposed to measure differentiation independently of within-population diversity (Jost 2008). Since then, there have been discussions about which of the estimates measures population differentiation correctly (Ryman & Leimar 2008; Jost 2009; Ryman & Leimar 2009; Gerlach *et al.* 2010; Wang 2012). Unfortunately, the issue has not been resolved to date. Therefore, it has been proposed to estimate two to three different estimators and execute caution in their interpretation in order to avoid erroneous conclusions (Heller & Siegismund 2009; Meirmans & Hedrick 2011; Alcala *et al.* 2014; Verity & Nichols 2014). Following this

recommendation, we used GenAlEx (Peakall & Smouse 2006, 2012) to estimate G_{ST} as well as D. The program calculates G_{ST} using the corrections proposed by (Nei & Chesser 1983) and follows the formulae given in eq.2 in Meirmans & Hedrick (2011) to calculate D_{est} (Peakall & Smouse 2006, 2012).

Effective migration among clusters

In order to estimate migration and gene flow among clusters, we used Genepop v4.0 (Rousset 2008), which implements the private allele method to estimate the number of effective migrants N_m (Slatkin 1985) and corrects the estimate for number of samples by using a regression line according to Barton & Slatkin (1986). We made this estimate first for all samples combined as well as for males and females separately.

Results

Structure

Based on the estimated likelihood values and ΔK, as estimated with the Evanno approach, the initial STRUCTURE analysis suggested the existence of four genetic clusters for all three analyses, i.e., combined sexes and for each sex separately (Supplementary Material, Figure S1). Although the results were not as unambiguous for the males as for the females and the combined dataset, inspection of the bar plots for K=3 and K=4 showed a considerable number of individuals with high q-values for each of the four clusters (Supplementary Material, Figure S4c). We thus accepted K=4 as having biological relevance also for the males and used these cluster assignment result in the subsequent analyses. The number of unassigned individuals with membership value <0.7 was relatively low and similar for all three analyses; 12% (sexes combined), 10.9% (females), and 13.3% (males). More females than males had a membership

value of q≈0.9 (80.4% vs. 67.9%), but the percentages were very similar for q≈0.8 (females: 9.2%; males: 10.9%) and q≈0.7 (females: 6.1%; males: 7.8%).

The second round of STRUCTURE analyses was to test for substructure within each identified cluster, performed separately for each sex. The results for the females showed some support for the existence of two separate subclusters in the most southern cluster 1, no substructure in cluster 2, but strong evidence for the existence of four and three subclusters within clusters 3 and 4, respectively (Supplementary Material, Figure S2, S3 and S5). The percentage of unassigned individuals was higher within these two subclusters systems than in the first round of analyses; 17.6% of individuals in cluster 1 and 41.1% in cluster 3, but only 4.2% in cluster 4.

We plotted the location of the subclusters in each of the three locations (Figure 3), i.e., each individual was colored according to the subcluster it was assigned to, including unassigned individuals with a membership value q<0.7. Within cluster 1 the two subclusters showed relatively strong overlap (Figure 3a), whereas within cluster 2, the subclusters showed relatively strong geographic coherence and less overlap (Figure 3b). The subclusters in cluster 4 showed the least geographic overlap (Figure 3c). The STRUCTURE analysis for the males did not reveal any substructure in clusters 1 or 2, however, it did find substructure in cluster 3 (Figure 4a, Supplementary material Figure S3a) and cluster 4 (Figure 4b, Supplementary material Figure S3b). The proportion of individuals unassigned to a subcluster varied by cluster; 38.8% in cluster 3 and only 3.1% in cluster 4.

Population differentiation

Genetic diversity, i.e. expected and observed heterozygosity, was high and >0.7 in all clusters and for both sexes. Mean fixation index F_I was at the same time between -0.056 and -0.005 and

only one marker for the females in cluster 3 showed a slightly higher positive value of 0.094, which was significant (Table 1). We assessed the magnitude of genetic differentiation among clusters by calculating pairwise F_{ST}, G_{ST}, and Joost's measure of differentiation D_{est} among clusters, separately for males and females and for both sexes combined (Table 2). For the combined data, F_{ST} values ranged from 0.035 - 0.078, G_{ST} values were similar, ranging from 0.035-0.077, and D_{est} values ranged from 0.244-0.482. As expected under male-biased dispersal, the F_{ST} values calculated for males alone were slightly lower than for females alone (0.033-0.079 vs. 0.040-0.082). G_{ST} (0.032-0.077 vs. 0.049-0.079) and D_{est} (0.215-0.489 vs. 0.249-0.479) values showed the same relationship between male and female estimates. The smallest value was between clusters 2 and 3 for each of the three datasets, sexes combined, only females, and only males (0.035, 0.040, and 0.033, respectively) and the largest value was between clusters 1 and 4 (0.078, 0.082, and 0.079, respectively). For the G_{ST} estimation, the smallest value was also found between clusters 2 and 3 for all three estimates (0.035, 0.032, and 0.038, respectively) and the largest likewise between clusters 1 and 4 (0.077, 0.079, and 0.077, respectively). The estimates of D_{est} were smallest for the combined data between clusters 2 and 1 (0.244), but smallest for females and males analyzed separately were between clusters 2 and 3 (0.249 and 0.215, respectively). The largest values were found between clusters 1 and 4 (0.482, 0.479, and 0.489, respectively).

In the next step, we calculated pairwise F_{ST}, G_{ST}, and D_{est} for the detected subclusters (Tables 3-6). For males and females, the pairwise F_{ST} values between the group of unassigned individuals and the subclusters were low compared to the F_{ST} values among subclusters (0.012 - 0.040 vs. 0.034 - 0.154). We found the same relationship for G_{ST} and D_{est} (GST: 0.008-0.035 vs. 0.032-0.093; Dest: 0.042-0.194 vs. 0.161-0.495). For the females, we found the lowest, and thus weakest, genetic differentiation among subclusters in the southernmost cluster (Table 3:

cluster 1, F_{ST}=0.034; G_{ST}=0.031; D_{est}=0.161). The two more northern clusters, 3 and 4, showed comparably strong genetic differentiation among subclusters, ranging from F_{ST}=0.055-0.093 and F_{ST}=0.052-0.154, from G_{ST}=0.048-0.086 and G_{ST}=0.039-0.138, and from D_{est}=0.255-0.426 and D_{est}=0.190-0.495, respectively (Tables 4 and 5). For the males, the substructure we detected in cluster 3 was accompanied also by relatively high levels of genetic differentiation, with F_{ST} values ranging from 0.059-0.077, G_{ST} ranging from 0.053-0.070, and D_{est} ranging from 0.290-0.364 (Table 6). In cluster 4, the F_{ST} value between the two identified subclusters was 0.066 and G_{ST}- and D_{est}-values were 0.057 and 0.295, respectively.

Genetic and spatial distance

As expected in a population that is distributed across a relatively large area, we found a clear pattern of IBD (Figure 5), both for males and females. However, the females displayed a slightly steeper slope than the males and higher kinship coefficients in each distance class up to 280 km, after which their coefficients dropped below the male kinship coefficients. For the females, all kinship coefficients were significant; for the males, two kinship coefficients were not significant (at 560 km and 600 km).

In order to assess fine-scale genetic structure, we investigated the relationship between kinship and geographic distance also on a small scale, i.e. at distances ≤ 80 km within each cluster that had been determined by the STRUCTURE analysis, by performing an analysis of spatial autocorrelation (Figure 6). Females displayed higher genetic correlation coefficient (*r*) values than males in all four clusters, with an intercept of 42, 56, 89, and 82 km, in clusters 1 to 4, respectively. We observed significant positive spatial autocorrelation, i.e. values of *r* above the 95% confidence interval, for the females up to 35, 45 (with a nonsignificant value at 25 km), 60, and 60 km (with nonsignficant values at 25 and 55 km) in clusters 1 to 4, respectively.

Whereas the significant *r* values ranged from 0.287-0.042 and 0.329-0.074 in clusters 3 and 4, respectively, in clusters 1 and 2, the significant values were lower overall, ranging from 0.201-0.032 and 0.229-0.030, respectively.

For the males, the curve of genetic correlation intercepted the x axis at 9, 34, 69, and 39 km in clusters 1-4, respectively. The genetic correlation coefficient *r* in clusters 1 and 2 was significantly positive in only three distance classes each (in cluster 1 at 5, 20 and 25 km; in cluster 2 at 10, 25 and 30 km). However, *r* values were low compared to the females in the same clusters and in the same distance classes, thus giving clear evidence of sex-biased spatial autocorrelation. In cluster 4, a larger number of *r* values were significant, namely at the distance classes 5-25 km as well as at 35 km, thus showing weak positive spatial autocorrelation with slightly elevated $_c$ values as compared to the two more southern clusters 1 and 2 (0.122-0.44 vs. 0.088-0.035). In cluster 3, we observed positive spatial autocorrelation of an unexpectedly large magnitude, with significant *r* values up to a distance of 55 km and ranging from 0.220-0.030, comparable to the values displayed by females in clusters 1 and 2. To investigate whether genetic structuring within clusters had an effect on the magnitude of *r*, we reran the analysis for clusters 3 and 4, using the multiple populations approach provided with the software. In both clusters and for both sexes, r_c values were lower using this approach compared to pooling the data of the entire cluster (Figure 7). The males in cluster 3 still showed significantly positive spatial autocorrelation up to the 25 km distance class, albeit at a lower level than in the first analysis (Figure 6). It is notable that, especially for the females, the r_c estimates did not decrease as gradually as in the previous analysis, but showed more abrupt changes in magnitude, which was probably due to a low number of samples per distance class per subcluster. Especially for the females, this also resulted in fairly wide confidence intervals, so that in cluster 4 the r_c value became nonsignificant already at distance class 15 km.

To explore the importance of scale in the genetic analysis, we ran an estimation of the Loiselle kinship coefficient on the same scale as the analysis of spatial autocorrelation and plotted it against the previously estimated genetic correlation coefficients (Figure 8). In all clusters, the Loiselle kinship coefficient correlated with the *r* values, albeit most of the time at lower values. On some occasions, the Loiselle kinship coefficient displayed higher values than r_c, e.g., cluster 4 for the males at the distance classes 40-80 km (Figure 8b), however in the majority of distance classes for female and male clusters, this is not the case.

Effective migration among clusters

We estimated the number of effective migrants per generation among clusters using the private allele method (Slatkin 1985) implemented in Genepop v4.1.0 (Rousset 2008), first for the combined dataset and subsequently for males and females separately (Figure 9). The values ranged from 0.081-0.814 for the combined dataset, from 0.107-1.309 for the females, and from 0.168-1.114 for the males (Table 7), with the lowest value for all three analyses found between clusters 1 and 3. The highest value was found between clusters 2 and 3 for the combined sexes and the females alone, whereas the males showed the highest number of effective migrants between clusters 1 and 2. Overall the estimated number of effective migrants was surprisingly low and >1 in only two cases.

Discussion

Dispersal has a fundamental impact on the shape and structure of a population, demographically, spatially, and genetically. We have assessed the genetic population structure of the Scandinavian brown bear on a large and a small scale, separately for males and females, in order to gain new knowledge on the genetic effect of sex-biased dispersal in this continuously

distributed large mammal. Using the large database compiled during the course of the national brown bear monitoring programs of Sweden and Norway, we analyzed 1531 genotypes, determined by eight microsatellite markers and found substantial structuring of the population into four genetic clusters. Using a hierarchical analysis approach, we have found even further structuring into subclusters in the more northern regions, indicating limits to long-distance gene flow.

Population structure

Our result was not entirely unexpected, as previous studies showed similar genetic structure and regarded it as being caused by the historic bottleneck event and/or genetic drift (Waits *et al.* 2000; Manel *et al.* 2004). Such time-lag effects have been demonstrated in simulation studies, e.g. Landguth *et al.* (2010) have shown that the disappearance of barriers, which would allow gene flow among previously isolated populations, may take ~100 generations before it becomes detectable in population genetic analyses using global estimates like F_{ST}. The length of the time lag may be considerably shorter; <15 generations in species with long-distance dispersal (Landguth *et al.* 2010). However, most knowledge about the temporal scale of shifts in population genetic structure due to founder effects and/or bottleneck events in expanding populations is based on simulation studies. Only a limited number of empirical studies have been published and have shown both long-lasting effects on population structure (e.g. Castric & Bernatchez 2003; Schulte *et al.* 2013) and relatively quick changes in population differentiation (e.g. Nussey *et al.* 2005; Herborg *et al.* 2007) in founder effect settings. In contrast to this, admixture and a decrease in population differentiation progressed on a short temporal scale, i.e., 10 years, in the Finnish brown bear population (Hagen *et al.* 2015).

In Sweden, the average male natal dispersal distance is 118.9 km and the average distance between breeding males and females is 40 km (Støen *et al.* 2006; Zedrosser *et al.* 2007). Thus, male brown bears can be regarded as regularly dispersing across long distances, which should aid gene flow among regions. Our results of increasing spatial genetic structure towards the north, including the occurrence of positive spatial autocorrelation in male bears, suggest that the degree of population structure in the Scandinavian brown bear population may be enforced by ecological and/or behavioral mechanisms that restrict gene flow among clusters. This may include e.g. to territorial behavior of resident individuals, thus rendering it difficult for immigrant males to establish their own territory, and/or the nonacceptance of migrants by potential mates (Schulte *et al.* 2013).

Our data further suggested that there was an area effect regarding the degree of genetic population structure, with the southern regions being less structured than the northern ones, both in terms of the Bayesian population assignment test and estimated population differentiation among clusters and among subclusters within clusters. This kind of directional trends is often observed in expanding populations, with an increasing population differentiation towards the expansion front (Austerlitz *et al.* 1997; Excoffier *et al.* 2009; De Giorgio *et al.* 2011). However, this explanation can be ruled out for the Scandinavian population, as the population expanded from at least three core areas after the bottleneck event, one of which was situated in the area of cluster 3 in the north (Swenson *et al.* 1995). Furthermore, although we do not know whether the Anarjohka population (western part of cluster 4) had been eradicated during the extirpation period, in Pasvik (eastern part of cluster 4) bears apparently repopulated the area <20 years after they apparently had become functionally extinct around 1910 (Swenson *et al.* 1995). Thus, the expansion occurred much faster and from a different direction than for the rest of Sweden and Norway (Swenson *et al.* 1995). Therefore the northern clusters in this

study should not be regarded as expansion fronts and the structuring we documented cannot be explained by these mechanisms.

Spatial autocorrelation and sex-biased dispersal

In accordance with Støen at al. (2006), our analysis of spatial autocorrelation showed a strong positive relationship between genetic relatedness and spatial proximity in female brown bears. For the male brown bears, clusters 1 and 2 displayed low and mostly nonsignificant genetic correlation coefficients, suggesting a random distribution, which is in agreement with male-biased dispersal. However, in contrast to our expectations of random distribution of males, we found significant positive spatial autocorrelation for the males in cluster 3, even after applying the multipopulation approach to account for high levels of within-cluster structuring (Peakall *et al.* 2003; Banks & Peakall 2012). The fact that within-cluster differentiation had an inflating effect on the genetic correlation efficient in both northern clusters, as well as the comparably high levels of differentiation among subclusters, is all the more notable as the subclusters were not spatially distinct and overlapped to a large degree. Therefore, the amount of differentiation cannot be explained by geographic distance among them.

These results suggest the presence of mechanisms that prevent gene flow among clusters and, in the north, among subclusters. They might include kin recognition and territorial behavior, which may lead to a lower acceptance of nonkin immigrants. Such behavior is generally associated with the philopatric females (Støen *et al.* 2005; Støen *et al.* 2006) and we are not aware of studies suggestive of kin cooperation among male brown bears. Two population genetic studies conducted on Swedish brown bears also found indications of a stronger substructure in the northern parts of the population (Waits *et al.* 2000; Manel *et al.* 2004). Waits *et al.* (2000) used 19 STR markers to infer population structure and gene flow

among subpopulations by estimating pairwise F_{ST} and assignment tests. The two more southern subpopulations, defined in their study as S and M, correspond largely to clusters 1 and 2, respectively, in our study, and were confirmed by Manel *et al.* (2004), who reanalyzed the same data. They also found indications of substructure within the northern study area, with the two subpopulations NS and NN, which together correspond largely to cluster 3 in our study. Manel *et al.* (2004) later used the estimation of a neighbor-joining tree and STRUCTURE (Pritchard *et al.* 2000) to determine clusters without a prior assumption of substructure and concluded that the substructure within the northern subpopulation was probably due to matrilineal formations and should not to be considered as separate management units. However, Manel et al. (2004) found an additional cluster in the west of the northern cluster defined as NWN (their Figure 3), which may correspond to one of the three subclusters we found within cluster 3 (subcluster 3 (blue), Figure 4a). They speculated that this was formed by one dominant male having fathered 70% of the NWN group's members. In addition to this, Waits et al. (2000) stated that radio-tracking data showed that only two males migrated between the subpopulations NN and NS and another two between NS and M, with only one confirmed mating between members of NN and NS. These combined results point to stronger substructuring in the north, which may be due to kin-related structure not only on the maternal, but also on the paternal side.

Dispersal versus gene flow

The monitoring database contained 48 females and 151 males that had been detected in more than one county (data not shown), thus confirming more male than female dispersal movement. The vast majority of these movements were between neighboring areas, only two males moved across a longer distance. One male moved ~ 430 km from Hedmark County in southeastern Norway, were it was recorded in 2007, to Västerbotten County in northern Sweden, were it was

recorded 2009, essentially moving out of the core area of cluster 1 and into the core area of cluster 2 (compare Figure 1b). This individual showed a *q* value of 0.554 for cluster 2. The other male moved ~ 400 km from the Sør-Trøndelag County in central Norway in 2009 southwards to the Telemark County in southern Norway, which is effectively free of bears and where it was shot in 2012. These records match the results of the radio-tracking studies performed previously in the Swedish bear population, where the maximum dispersal distance observed in males was 467 km (Støen *et al.* 2006). However, physical movement apparently has not translated into gene flow, as the estimated number of effective migrants per generation among clusters was low (males: 0.168-1.114; females: 0.107-1.309; Table 7). In the Finnish brown bear population, where the genetic structure decreased relatively rapidly, the estimated number of effective migrants ranged from 1.60-3.63, displaying an increase over time (Hagen *et al.* 2015). An even higher number of effective migrants was estimated for the gene flow between Finland and Russian Karelia (7.64, Kopatz *et al.* 2014).

Large-scale genetic structure

In contrast to the patterns of genetic structure at the local scale, a study-area-wide analysis of kinship in relation to spatial distance revealed a clear pattern of IBD, which was similar between the sexes. When reducing the spatial extent of the analysis to the same scale as for the analysis of spatial autocorrelation, the comparison showed that, even though the Loiselle estimates were lower overall, they followed the same trend, showing that scale had a strong effect on the outcome of the spatial genetic analysis. Even though not numerous, other studies also have shown a similar effect of scale on sex-specific genetic dispersal estimates, especially in humans (Wilder *et al.* 2004; Wilkins & Marlowe 2006; Heyer *et al.* 2012), but also in red deer (*Cervus elaphus*) (Perez-Espona *et al.* 2010), the common vole (*Microtus arvalis*) (Gauffre *et al.* 2009)

and Cabanis's greenbuls (*Phyllastrephus cabanisi*), a cooperatively breeding bird (Vangestel *et al.* 2013). The reason for this may be the difference in costs/benefits of dispersal on the small versus the large scale: for males in polygamous species the benefits of dispersal, i.e. avoiding inbreeding and competition, may be larger on the small scale, whereas the costs of dispersal are likely to increase with distance (Perez-Espona *et al.* 2010). It is likely that different mechanisms apply to the evolution of long- versus short-distance dispersal (Lawson Handley & Perrin 2007; Ronce 2007).

In the study of sex-biased dispersal, the use of genetic approaches has been suggested to augment the findings of more ecological methods, e.g. radio-tracking (Lawson Handley & Perrin 2007; Ronce 2007; Driscoll *et al.* 2014). Our results show that this approach can uncover previously unavailable information. Several studies have found that dispersal decisions are condition dependent and consequently dispersal rate and frequency may vary among locations and individuals (Perez-Gonzalez & Carranza 2009; Solmsen *et al.* 2011; Debeffe *et al.* 2012; Gilroy & Lockwood 2012; Vercken *et al.* 2012; Hovestadt *et al.* 2014; Kentie *et al.* 2014). Given the pronounced differences in local population structure between the northern and southern areas of the Scandinavian brown bear population, further studies that take environmental and individual characteristics into account are needed to shed more light on what causes this gradient.

References

Aarnes SG, Tobiassen C, Broseth H*, et al.* (2014) Populasjonsovervåking av brunbjørn: DNA-analyse av prøver innsamlet i Norge i 2013. *Bioforsk rapport* **9 (48)**, 1-54.

Alcala N, Goudet J, Vuilleumier S (2014) On the transition of genetic differentiation from isolation to panmixia: What we can learn from G(ST) and D. *Theoretical Population Biology* **93**, 75-84.

Andreassen R, Schregel J, Kopatz A*, et al.* (2012) A forensic DNA profiling system for Northern European brown bears (Ursus arctos). *Forensic Science International-Genetics* **6**, 798-809.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: An empirical and simulation study. *Evolution* **54**, 1414-1422.

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

Banks SC, Peakall R (2012) Genetic spatial autocorrelation can readily detect sex-biased dispersal. *Mol Ecol* **21**, 2092-2105.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) *GENETIX 4.05, logiciel sous Windows pour la génétique des populations.* Laboratoire Génome, Populations, Interactions,CNRS UMR 5171, Université de Montpellier II, Montpellier.

Castric V, Bernatchez L (2003) The rise and fall of isolation by distance in the anadromous brook charr (Salvelinus fontinalis Mitchill). *Genetics* **163**, 983-996.

Coulon A, Fitzpatrick JW, Bowman R*, et al.* (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma coerulescens). *Mol Ecol* **17**, 1685-1701.

Cronin MA, Amstrup SC, Garner GW, Vyse ER (1991) Interspecific and intraspecific mitochondrial-DNA variation in North American Bears (Ursus). *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **69**, 2985-2992.

de Knegt HJ, van Langevelde F, Coughenour MB*, et al.* (2010) Spatial autocorrelation and the scaling of species-environment relationships. *Ecology* **91**, 2455-2465.

Debeffe L, Morellet N, Cargnelutti B*, et al.* (2012) Condition-dependent natal dispersal in a large herbivore: heavier animals show a greater propensity to disperse and travel further. *Journal of Animal Ecology* **81**.

Devillard S, Allaine D, Gaillard JM, Pontier D (2004) Does social complexity lead to sex-biased dispersal in polygynous mammals? A test on ground-dwelling sciurids. *Behavioral Ecology* **15**, 83-87.

Driscoll DA, Banks SC, Barton PS*, et al.* (2014) The Trajectory of Dispersal Research in Conservation Biology. Systematic Review. *Plos One* **9**, 18.

Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Eiken HG, Andreassen RJ, Kopatz A*, et al.* (2009) Population data for 12 STR loci in Northern European brown bear (Ursus arctos) and application of DNA profiles for forensic casework. *Forensic Science International: Genetics Supplement Series* **2**, 273-274.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**, 2611-2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Fontanillas P, Petit E, Perrin N (2004) Estimating sex-specific dispersal rates with autosomal markers in hierarchically structured populations. *Evolution* **58**, 886-894.

Gauffre B, Petit E, Brodier S, Bretagnolle V, Cosson JF (2009) Sex-biased dispersal patterns depend on the spatial scale in a social rodent. *Proceedings of the Royal Society B-Biological Sciences* **276**, 3487-3494.

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of population differentiation based on G(ST) and D: forget G(ST) but not all of statistics! *Mol Ecol* **19**, 3845-3852.

Gilroy JJ, Lockwood JL (2012) Mate-Finding as an Overlooked Critical Determinant of Dispersal Variation in Sexually-Reproducing Animals. *Plos One* **7**.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28**, 1140-1162.

Hagen SB, Kopatz A, Aspi J, Kojola I, Eiken HG (2015) Evidence of rapid change in genetic structure and diversity during range expansion in a large terrestrial carnivore. *Proceedings of the Royal Society B-Biological Sciences* **282**, 20150092.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.
Haynes BT, Marcus AD, Higgins DP*, et al.* (2014) Unexpected absence of genetic separation of a highly diverse population of hookworms from geographically isolated hosts. *Infection Genetics and Evolution* **28**, 192-200.
Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313-318.
Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation G(ST), D-EST and G'(ST): how wrong have we been? *Mol Ecol* **18**, 2080-2083.
Herborg L-M, Weetman D, Van Oosterhout C, Hanfling B (2007) Genetic population structure and contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, Eriocheir sinensis. *Mol Ecol* **16**, 231-242.
Heyer E, Chaix R, Pavard S, Austerlitz F (2012) Sex-specific demographic behaviours that shape human genomic variation. *Mol Ecol* **21**, 597-612.
Hovestadt T, Mitesser O, Poethke H-J (2014) Gender-Specific Emigration Decisions Sensitive to Local Male and Female Density. *American Naturalist* **184**, 38-51.
Jost L (2008) G(ST) and its relatives do not measure differentiation. *Mol Ecol* **17**, 4015-4026.
Jost L (2009) D vs. G(ST): Response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Mol Ecol* **18**, 2088-2091.
Keis M, Remm J, Ho SYW*, et al.* (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *Journal of Biogeography* **40**, 915-927.
Kentie R, Both C, Hooijmeijer JCEW, Piersma T (2014) Age-dependent dispersal and habitat choice in black-tailed godwits Limosa limosa limosa across a mosaic of traditional and modern grassland habitats. *Journal of Avian Biology* **45**, 396-405.
Kindberg J, Swenson JE, Ericsson G*, et al.* (2011) Estimating population size and trends of the Swedish brown bear Ursus arctos population. *Wildlife Biology* **17**, 114-123.
Koch K, Algar D, Schwenk K (2014) Population Structure and Management of Invasive Cats on an Australian Island. *Journal of Wildlife Management* **78**, 968-975.
Kojola I, Danilov PI, Laitala HM, Belkin V, Yakimov A (2003) Brown bear population structure in core and periphery: analysis of hunting statistics from Russian Karelia and Finland. *Ursus* **14**, 17-20.
Kojola I, Heikkinen S (2006) The structure of the expanded brown bear population at the edge of the Finnish range. *Annales Zoologici Fennici* **43**, 258-262.
Kojola I, Määtä E (2004) Suurpetojen lukumäärä ja lisääntuminen vuonna 2003. In: *Riistantutkimuksen tiedote*, Helsinki.
Kopatz A, Eiken HG, Aspi J*, et al.* (2014) Admixture and Gene Flow from Russia in the Recovering Northern European Brown Bear (Ursus arctos). *Plos One* **9**.
Kopatz A, Eiken HG, Hagen SB*, et al.* (2012) Connectivity and population subdivision at the fringe of a large brown bear (Ursus arctos) population in North Western Europe. *Conservation Genetics* **13**, 681-692.
Korsten M, Ho SYW, Davison J*, et al.* (2009) Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals? *Mol Ecol* **18**, 1963-1979.
Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol* **19**, 4179-4191.
Lawson Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Mol Ecol* **16**, 1559-1578.

Loiselle BA, Sork VL, Nason J, Graham C (1995) SPATIAL GENETIC-STRUCTURE OF A TROPICAL UNDERSTORY SHRUB, PSYCHOTRIA OFFICINALIS (RUBIACEAE). *American Journal of Botany* **82**, 1420-1425.

Manel S, Bellemain E, Swenson JE, Francois O (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Mol Ecol* **13**, 1327-1331.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F-ST and related measures. *Mol Ecol Resour* **11**, 5-18.

Munthali CRY, Chirwa PW, Changadeya WJ, Akinnifesi FK (2013) Genetic differentiation and diversity of Adansonia digitata L (baobab) in Malawi using microsatellite markers. *Agroforestry Systems* **87**, 117-130.

Murrell DJ, Travis JMJ, Dytham C (2002) The evolution of dispersal distance in spatially-structured populations. *Oikos* **97**, 229-236.

Naturvårdsverket (2014) *Björn, populationsutveckling i Sverige.* http://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-O/Bjorn-populationsutveckling-i-Sverige/. Access Date: 21.01.2015. In Swedish

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A* **70**, 3321-3323.

Nei M (1977) F-Statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics* **41**, 225-233.

Nei M, Chesser RK (1983) Estimation of fixation indexes and gene diversities. *Annals of Human Genetics* **47**, 253-259.

Nussey DH, Coltman DW, Coulson T*, et al.* (2005) Rapidly declining fine-scale spatial genetic structure in female red deer. *Mol Ecol* **14**, 3395-3405.

Peakall R, Ruibal M, Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. *Evolution* **57**, 1182-1195.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

Pelletier A, Obbard ME, Mills K*, et al.* (2012) Delineating genetic groupings in continuously distributed species across largely homogeneous landscapes: a study of American black bears (Ursus americanus) in Ontario, Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **90**, 999-1014.

Perez-Espona S, Perez-Barberia FJ, Jiggins CD, Gordon IJ, Pemberton JM (2010) Variable extent of sex-biased dispersal in a strongly polygynous mammal. *Mol Ecol* **19**, 3101-3113.

Perez-Gonzalez J, Carranza J (2009) Female-biased dispersal under conditions of low male mating competition in a polygynous mammal. *Mol Ecol* **18**, 4617-4630.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

Ronce O (2007) How does it feel to be like a rolling stone? Ten questions about dispersal evolution. In: *Annual Review of Ecology Evolution and Systematics*, pp. 231-253. Annual Reviews, Palo Alto.

Rousset F, Gandon S (2002) Evolution of the distribution of dispersal distance under distance-dependent cost of dispersal. *Journal of Evolutionary Biology* **15**, 515-523.

Ryman N, Leimar O (2008) Effect of mutation on genetic differentiation among nonequilibrium populations. *Evolution* **62**, 2250-2259.

Ryman N, Leimar O (2009) G(ST) is still a useful measure of genetic differentiation - a comment on Jost's D. *Mol Ecol* **18**, 2084-2087.

Saarma U, Kojola I (2007) Matrilineal genetic structure of the brown bear population in Finland. *Ursus* **18**, 30-37.

Schregel J, Kopatz A, Hagen SB*, et al.* (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Mol Ecol* **21**, 3474-3488.
Schulte U, Veith M, Mingo V, Modica C, Hochkirch A (2013) Strong genetic differentiation due to multiple founder events during a recent range expansion of an introduced wall lizard population. *Biological Invasions* **15**, 2639-2649.
Solmsen N, Johannesen J, Schradin C (2011) Highly asymmetric fine-scale genetic structure between sexes of African striped mice and indication for condition dependent alternative male dispersal tactics. *Mol Ecol* **20**, 1624-1634.
Støen O-G, Bellemain E, Sæbø S, Swenson JE (2005) Kin-related spatial structure in brown bears Ursus arctos. *Behavioral Ecology and Sociobiology* **59**, 191-197.
Støen OG, Zedrosser A, Saebo S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears Ursus arctos. *Oecologia* **148**, 356-364.
Swenson JE, Sandegren F, Söderberg A (1998) Geographic expansion of an increasing brown bear population: evidence for presaturation dispersal. *Journal of Animal Ecology* **67**, 819-826.
Swenson JE, Wabakken P, Sandegren F*, et al.* (1995) The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology* **1**, 11-25.
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, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* **7**, 453-464.
Underwood N, Hamback P, Inouye BD (2005) Large-scale questions and small-scale data: empirical and theoretical methods for scaling up in ecology. *Oecologia* **145**, 177-178.
Vangestel C, Callens T, Vandomme V, Lens L (2013) Sex-Biased Dispersal at Different Geographical Scales in a Cooperative Breeder from Fragmented Rainforest. *Plos One* **8**.
Vercken E, Sinervo B, Clobert J (2012) The importance of a good neighborhood: dispersal decisions in juvenile common lizards are based on social environment. *Behavioral Ecology* **23**, 1059-1067.
Verity R, Nichols RA (2014) What is genetic differentiation, and how should we measure it-G(ST), D, neither or both? *Mol Ecol* **23**, 4216-4225.
Vuilleumier S, Fontanillas P (2007) Landscape structure affects dispersal in the greater white-toothed shrew: Inference between genetic and simulated ecological distances. *Ecological Modelling* **201**, 369-376.
Waits L, Taberlet P, Swenson JE, Sandegren F, Franzen R (2000) Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (Ursus arctos). *Mol Ecol* **9**, 421-431.
Wang J (2012) On the measurements of genetic differentiation among populations. *Genetics Research* **94**, 275-289.
Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* **15**, 1419-1439.
Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nature Genetics* **36**, 1122-1125.
Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* **28**, 290-300.
Yannic G, Basset P, Buchi L, Hausser J, Broquet T (2012) Scale -specific sex-biased dispersal in the Valais shrew unveiled by genetic variation on the Y-chromosome, autosomes, and mitochondrial DNA. *Evolution* **66**, 1737-1750.

Zedrosser A, Støen O-G, Sæbø S, Swenson JE (2007) Should I stay or should I go? Natal dispersal in the brown bear. *Animal Behaviour* **74**, 369-376.

Table 1: Expected and observed heterozygosity and fixation index per genetic cluster estimated separately for males and female brown bears in Scandinavia. Each cluster contains only individuals with a membership value q ≥ 0.7 as estimated by STRUCTURE. First value = females; second, italic value = males; bold = the mean across all loci for each cluster; *) significant.

	Locus	No. allele	He	Ho	F_I
Cluster1	MU09	6 / *6*	0.732 / *0.698*	0.739 / *0.700*	-0.010 / *-0.003*
	MU10	6 / *6*	0.804 / *0.804*	0.791 / *0.834*	0.016 / *-0.037*
	MU23	7 / *7*	0.716 / *0.665*	0.719 / *0.687*	-0.003 / *-0.032*
	MU59	7 / *9*	0.684 / *0.686*	0.719 / *0.684*	-0.051 / *0.004*
	MU05	6 / *6*	0.617 / *0.604*	0.625 / *0.656*	-0.013 / *-0.087*
	G10L	6 / *6*	0.720 / *0.731*	0.763 / *0.787*	-0.061 / *-0.076*
	MU51	6 / *5*	0.794 / *0.787*	0.782 / *0.787*	0.015 / *0.000*
	MU50	6 / *7*	0.758 / *0.748*	0.795 / *0.775*	-0.050 / *-0.035*
	Mean	**6.5 / *6.3***	**0.715 / *0.728***	**0.738 / *0.742***	**-0.033 / *-0.020***
Cluster2	MU09	9 / *10*	0.803 / *0.825*	0.775 / *0.838*	0.035 / *-0.016*
	MU10	7 / *7*	0.685 / *0.660*	0.689 / *0.667*	-0.006 / *-0.010*
	MU23	7 / *8*	0.817 / *0.825*	0.870 / *0.820*	-0.065 / *0.006*
	MU59	9 / *10*	0.781 / *0.790*	0.764 / *0.837*	0.022 / *-0.059*
	MU05	6 / *6*	0.633 / *0.631*	0.646 / *0.665*	-0.021 / *-0.053*
	G10L	7 / *8*	0.629 / *0.716*	0.609 / *0.678*	0.032 / *0.053*
	MU51	7 / *7*	0.721 / *0.749*	0.720 / *0.764*	0.000 / *-0.020*
	MU50	8 / *7*	0.815 / *0.819*	0.845 / *0.828*	-0.036 / *-0.011*
	Mean	**7.9 / *7.5***	**0.752 / *0.735***	**0.762 / *0.740***	**-0.014 / *-0.005***
Cluster3	MU09	8 / *9*	0.828 / *0.848*	0.872 / *0.933*	-0.053 / *-0.100*
	MU10	6 / *7*	0.793 / *0.779*	0.773 / *0.736*	0.024 / *0.056*
	MU23	8 / *7*	0.691 / *0.713*	0.665 / *0.705*	0.038 / *0.011*
	MU59	10 / *10*	0.831 / *0.835*	0.887 / *0.808*	-0.067 / *0.031*
	MU05	7 / *7*	0.754 / *0.755*	0.783 / *0.684*	-0.039 / *0.094**
	G10L	8 / *8*	0.788 / *0.792*	0.729 / *0.860*	0.075 / *-0.086*
	MU51	7 / *7*	0.779 / *0.786*	0.823 / *0.833*	-0.056 / *-0.061*
	MU50	7 / *7*	0.746 / *0.728*	0.744 / *0.751*	0.002 / *-0.032*
	Mean	**7.8 / *7.6***	**0.779 / *0.776***	**0.789 / *0.784***	**-0.011 / *-0.009***
Cluster4	MU09	10 / *11*	0.829 / *0.867*	0.957 / *0.969*	-0.154 / *-0.118*
	MU10	6 / *8*	0.779 / *0.763*	0.681 / *0.810*	0.126 / *-0.061*
	MU23	7 / *9*	0.563 / *0.693*	0.574 / *0.688*	-0.021 / *0.008*
	MU59	11 / *11*	0.845 / *0.815*	0.830 / *0.767*	0.018 / *0.059*
	MU05	8 / *8*	0.829 / *0.802*	0.851 / *0.875*	-0.027 / *-0.091*
	G10L	5 / *6*	0.462 / *0.581*	0.511 / *0.656*	-0.106 / *-0.129*
	MU51	6 / *6*	0.786 / *0.787*	0.872 / *0.875*	-0.110 / *-0.111*
	MU50	7 / *8*	0.805 / *0.837*	0.787 / *0.844*	0.022 / *-0.008*
	Mean	**8.4 / *7.5***	**0.768 / *0.737***	**0.810 / *0.758***	**-0.056 / *-0.032***

Table 2: Genetic differentiation among genetic clusters of Scandinavian brown bears assessed by F_{ST}-, G_{ST}- and Joost's Dest-estimation. Values below the diagonal = F_{ST}, values above diagonal = first row: G_{ST}; second row/italic: Joost's Dest; the values were estimated for each sex separately and sexes combined; they are listed as follows: **combined** (females/males). All estimates are significant.

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**
Cluster 1		**0.042** (0.049/0.037) ***0.244*** *(0.285/0.215)*	**0.066** (0.066/0.0665) ***0.425*** *(0.435/0.417)*	**0.077** (0.079/0.077) ***0.482*** *(0.479/0.489)*
Cluster 2	**0.043** (0.050 / 0.038)		**0.035** (0.038/0.032) ***0.299*** *(0.249/0.215)*	**0.064** (0.067/0.065) ***0.423*** *(0.409/0.448)*
Cluster 3	**0.066** (0.067 / 0.066)	**0.035** (0.040 / 0.033)		**0.046** (0.050/0.045) ***0.327*** *(0.339/0.331)*
Cluster 4	**0.078** (0.082 / 0.079)	**0.065** (0.070 / 0.067)	**0.047** (0.054 / 0.048)	

Table 3: Genetic differentiation among subclusters and unassigned individuals of the female brown bears of cluster 1 in Scandinavia, assessed by F_{ST}-, G_{ST}- and Joost's Dest-estimation. Values below diagonal = F_{ST}, values above diagonal = G_{ST}/ *Joost's Dest*. All estimates are significant.

	Subcluster 1	**Subcluster 2**	**unassigned**
Subcluster 1		0.032/*0.161*	0.008/*0.042*
Subcluster 2	0.034		0.012/*0.054*
unassigned	0.012	0.016	

Table 4: Genetic differentiation among subclusters and unassigned individuals of the female brown bears of cluster 3 in Scandinavia, assessed by F_{ST}-, G_{ST}- and Joost's Dest-estimation. Values below diagonal = F_{ST}, values above diagonal = G_{ST}/ *Joost's Dest*. All estimates are significant.

	Subcluster 1	**Subcluster 2**	**Subcluster 3**	**Subcluster 4**	**unassigned**
Subcluster 1		0.071/*0.387*	0.048/*0.255*	0.073/*0.345*	0.021/*0.125*
Subcluster 2	0.078		0.059/*0.328*	0.086/*0.426*	0.024/*0.148*
Subcluster 3	0.055	0.066		0.064/*0.310*	0.021/*0.129*
Subcluster 4	0.080	0.093	0.072		0.035/*0.184*
unassigned	0.026	0.030	0.027	0.040	

Table 5: Genetic differentiation among subclusters of the female brown bears of cluster 4 in Scandinavia, assessed by F_{ST}-, G_{ST}- and Joost's Dest-estimation. Values below diagonal = F_{ST}, values above diagonal = G_{ST}/ *Joost's Dest*. All estimates are significant.

	Subcluster 1	**Subcluster 2**	**Subcluster 3**
Subcluster 1		0.138/*0.495*	0.093/*0.371*
Subcluster 2	0.154		0.039/*0.190*
Subcluster 3	0.109	0.052	

Table 6: Genetic differentiation among subclusters and unassigned individuals of the male brown bears of cluster 3 in Scandinavia, assessed by F_{ST}-, G_{ST}- and Joost's Dest-estimation. Values below diagonal = F_{ST}, values above diagonal = G_{ST}/*Joost's Dest*. All estimates are significant.

	Subcluster 1	**Subcluster 2**	**Subcluster 3**	**unassigned**
Subcluster 1		0.070/0.364	0.054/0.339	0.032/0.194
Subcluster 2	0.077		0.053/0.290	0.023/0.121
Subcluster 3	0.060	0.059		0.017/0.113
unassigned	0.036	0.028	0.021	

Table 7: Number of migrants per generation among the four genetic clusters of brown bears in Scandinavia, estimated with the private allele method using Genepop v4.1.0; estimates are given as: **sexes combined** (females/*males*)

	Cluster 1	Cluster 2	Cluster 3
Cluster 2	**0.741** (0.551 / *1.114*)		
Cluster 3	**0.081** (0.107 / *0.168*)	**0.814** (1.309 / *0.729*)	
Cluster 4	**0.107** (0.144 / *0.202*)	**0.187** (0.253 / *0.292*)	**0.153** (0.205 / *0.258*)

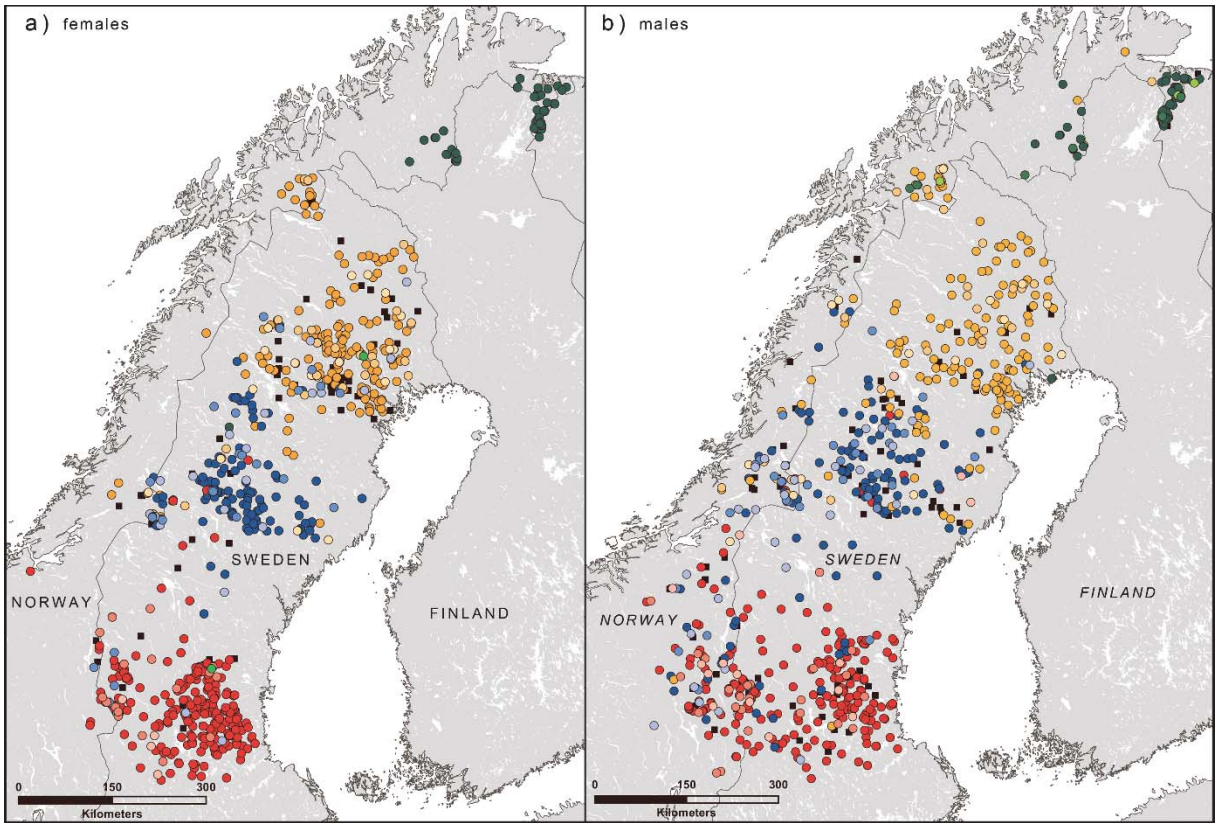


Figure 1: Location of genetic clusters for Norwegian and Swedish brown bears. a) only females, membership to a cluster is indicated by color, red = cluster 1, blue = cluster 2, yellow = cluster 3, green = cluster 4, darkest = membership value $q \geq 0.9$; medium = $q \geq 0.8$; lightest = $q \geq 0.7$; black squares = $q < 0.7$; b) only males, colors correspond to a).

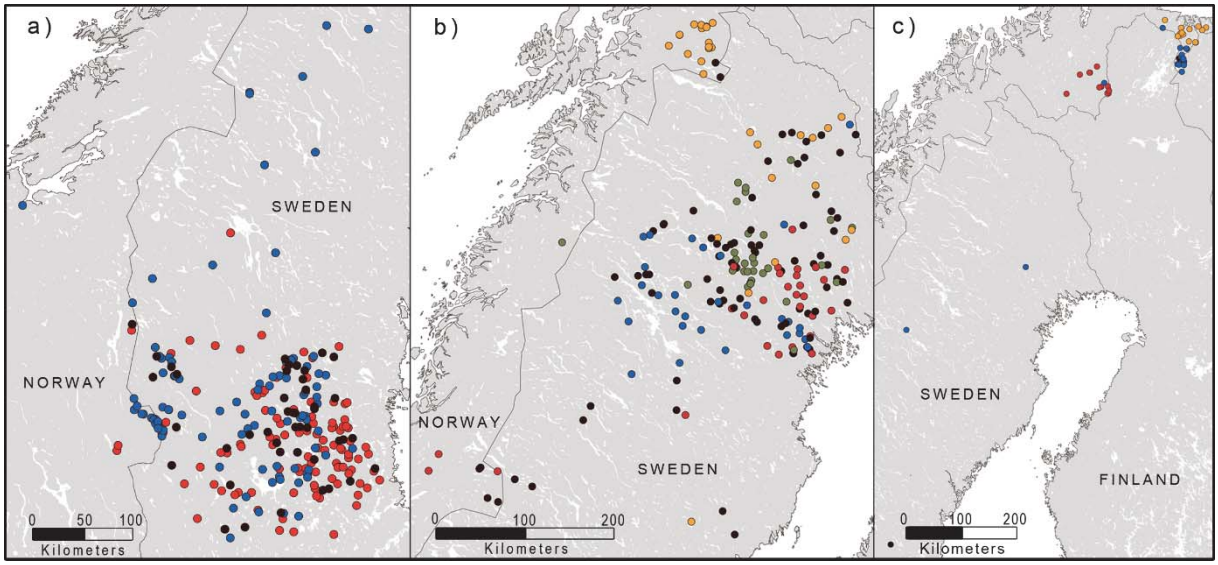


Figure 3: Location of female brown bears belonging to the different genetic subclusters in Scandinavia, determined by the reanalysis of each previously determined cluster. Individuals whose membership could not be determined ($q < 0.7$) are colored black. a) Subclusters within cluster 1: blue = subcluster 1, red = subcluster 2; b) subclusters within cluster 3: red= subcluster 1, yellow = subcluster 2, blue = subcluster 3, green = subcluster 4; c) subclusters within cluster 4: red = subcluster 1, yellow = subcluster 2, blue = subcluster 3. Note that the colors used to depict subcluster membership do not correspond to the color coding used in the other maps.

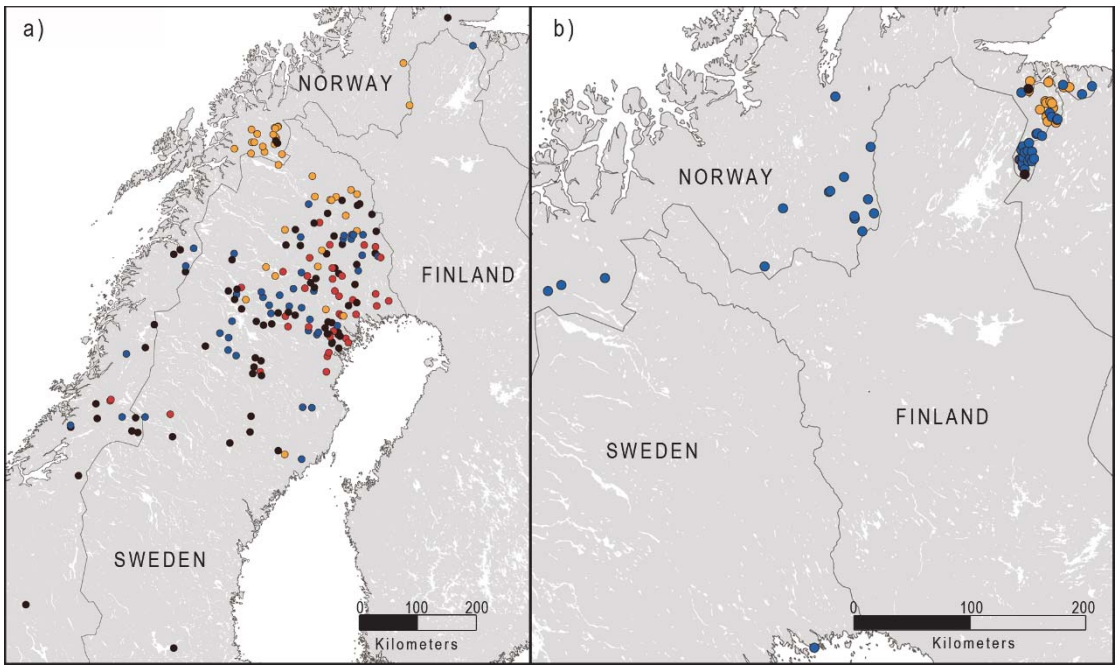


Figure 4: Location of male brown bears belonging to the different genetic subclusters in Scandinavia, determined by the reanalysis of each previously determined cluster. Individuals whose membership could not be determined ($q < 0.7$) are colored black. a) subclusters within cluster 3: yellow = subcluster 1, red = subcluster 2, blue = subcluster 3; b) subclusters within cluster 4: yellow = subcluster 1, blue = subcluster 2. Note that the colors used to depict subcluster membership do not correspond to the color coding used in the other maps.

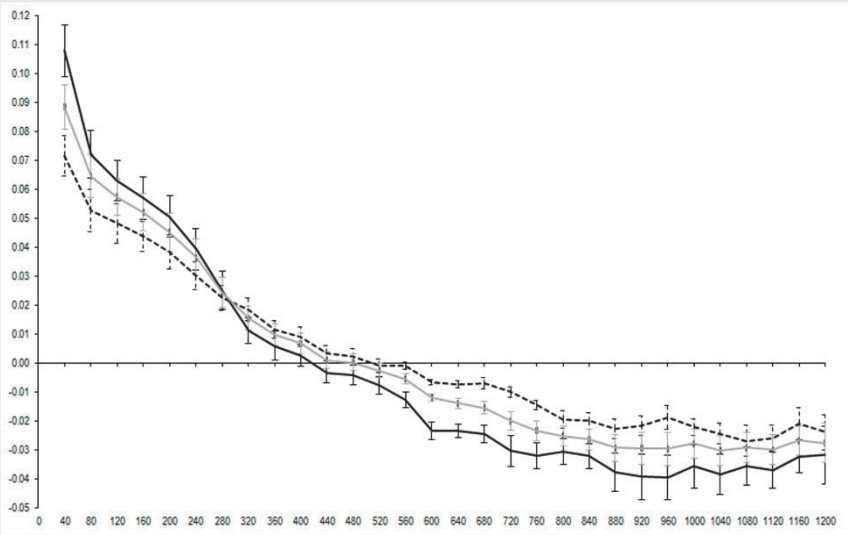


Figure 5: Loiselle kinship coefficient for brown bears in Scandinavia; distance class = 40 km, solid black line = females; broken black lines = males; solid gray lines = females and males combined.

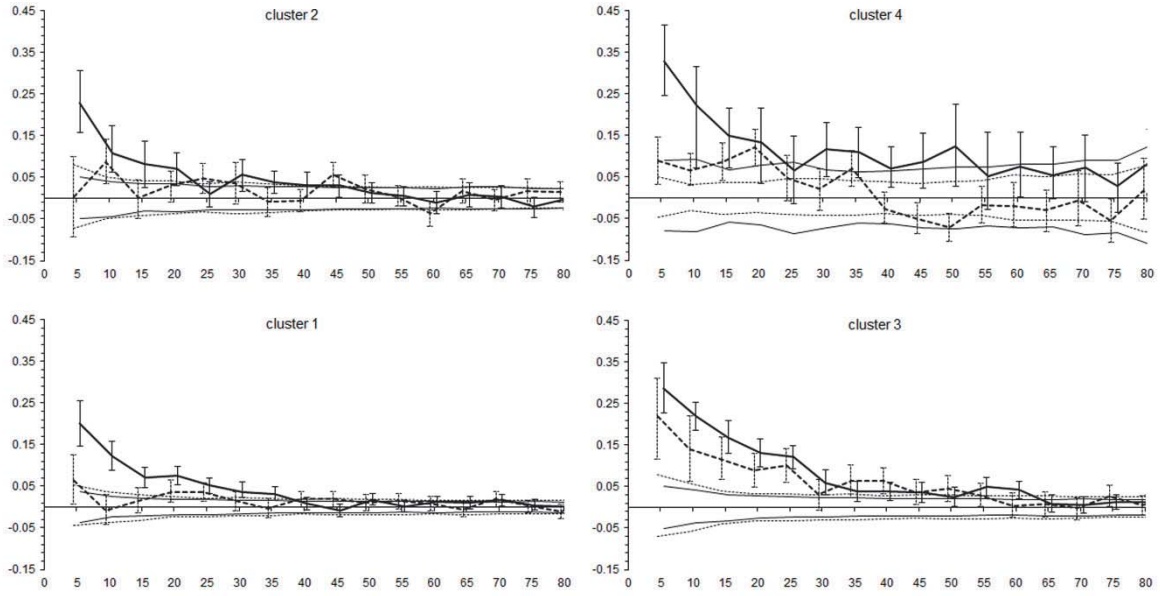


Figure 6: Spatial autocorrelation of female versus male brown bears compared within each cluster in Scandinavia, distance class of 5 km. The genetic correlation coefficient (r_c) is given as a solid line for females and a dashed line for males. The 95% confidence intervals for the null hypothesis of random distribution of genotypes, as well as bootstrap errors, are displayed in the same manner.

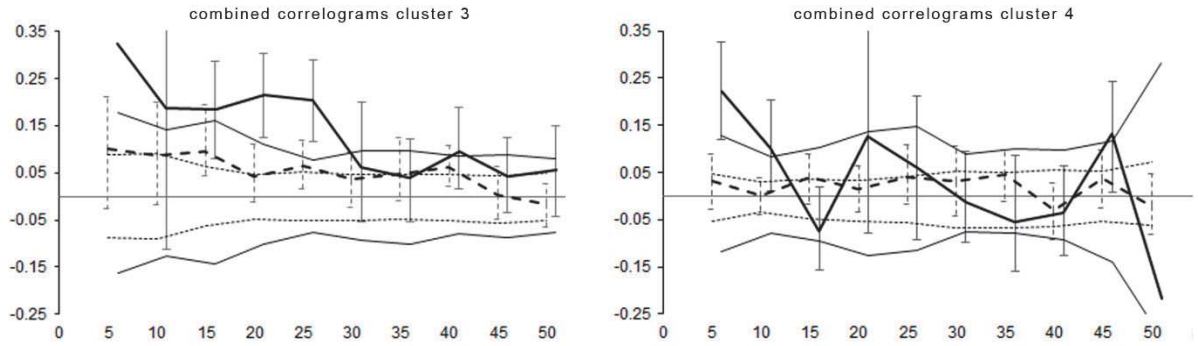


Figure 7: Combined spatial autocorrelation analysis for genetic clusters 3 and 4 for brown bears in Scandinavia, distance class of 5 km. The analysis uses the multiple population approach, which sums the individual components for calculating r_c as a division of the total numerator and denominator across populations, rather than the simple arithmetic mean. The genetic correlation coefficient (r_c) is given as a solid thick line for females and a dashed thick line for males. The 95% confidence intervals for the null hypothesis of random distribution of genotypes, as well as bootstrap errors, are displayed correspondingly in a thin line.

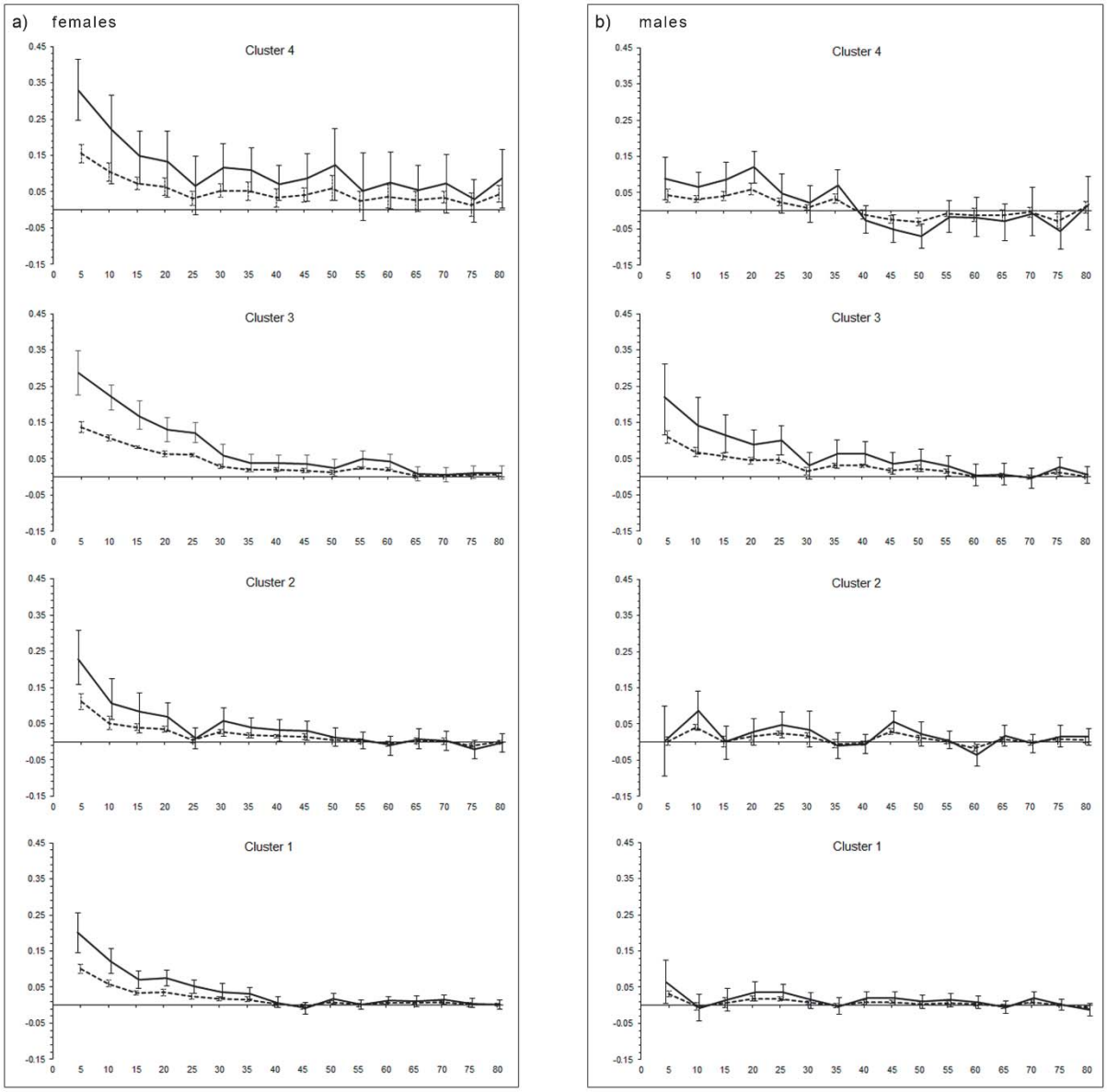


Figure 8: Loiselle kinship coefficient versus genetic correlation coefficient (*r*) for Scandinavian brown bears, estimated at distance intervals of 5 km up to a maximum distance of 80 km. Estimations were performed for each cluster and are displayed as follows: solid line = genetic correlation coefficient (*r*), dashed line = Loiselle kinship coefficient. For both coefficients the bootstrap error bars are displayed accordingly. a) females, genetic clusters 1 to 4; b) males, clusters 1 to 4.

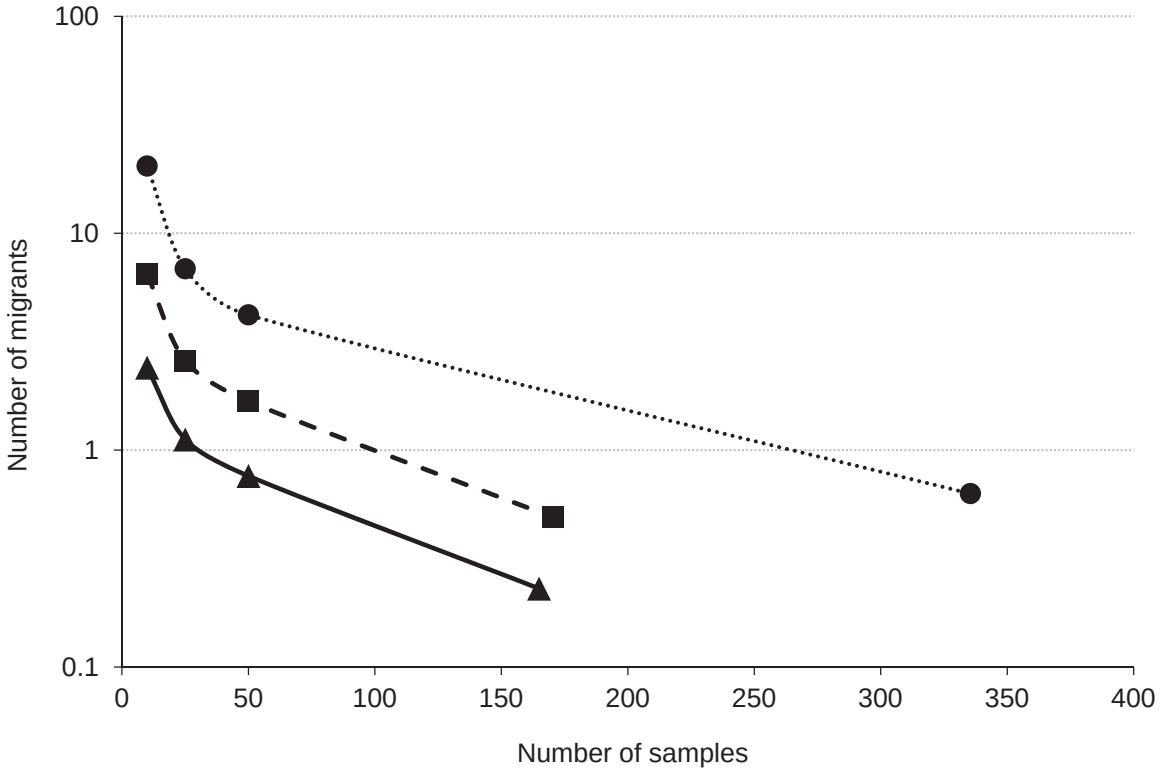


Figure 9: Number of migrant brown bears per generation among the four genetic clusters in Scandinavia, estimated by the private allele method. Given are the estimates for females (▲), males (■) and both sexes combined (●).

Appendix

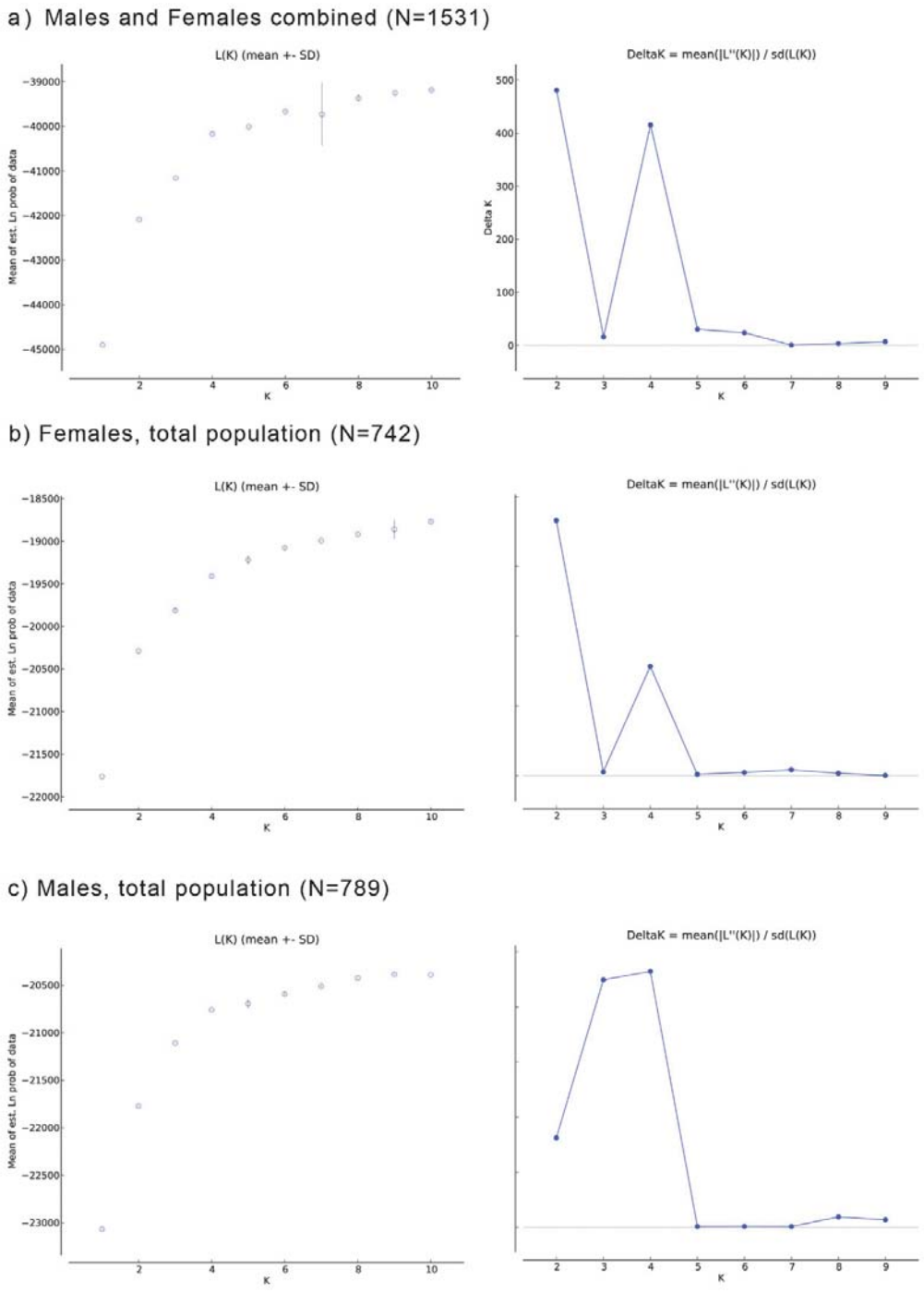


Figure S1: Results of the Bayesian clustering analysis of brown bears in Sweden and Norway with STRUCTURE (Pritchard et al. 2000), processed with the help of STRUCTURE HARVESTER (Earl & von Holdt 2012). a) results of the analysis performed on the total dataset; b) results of the analysis of only females; c) results of the analysis of only males.

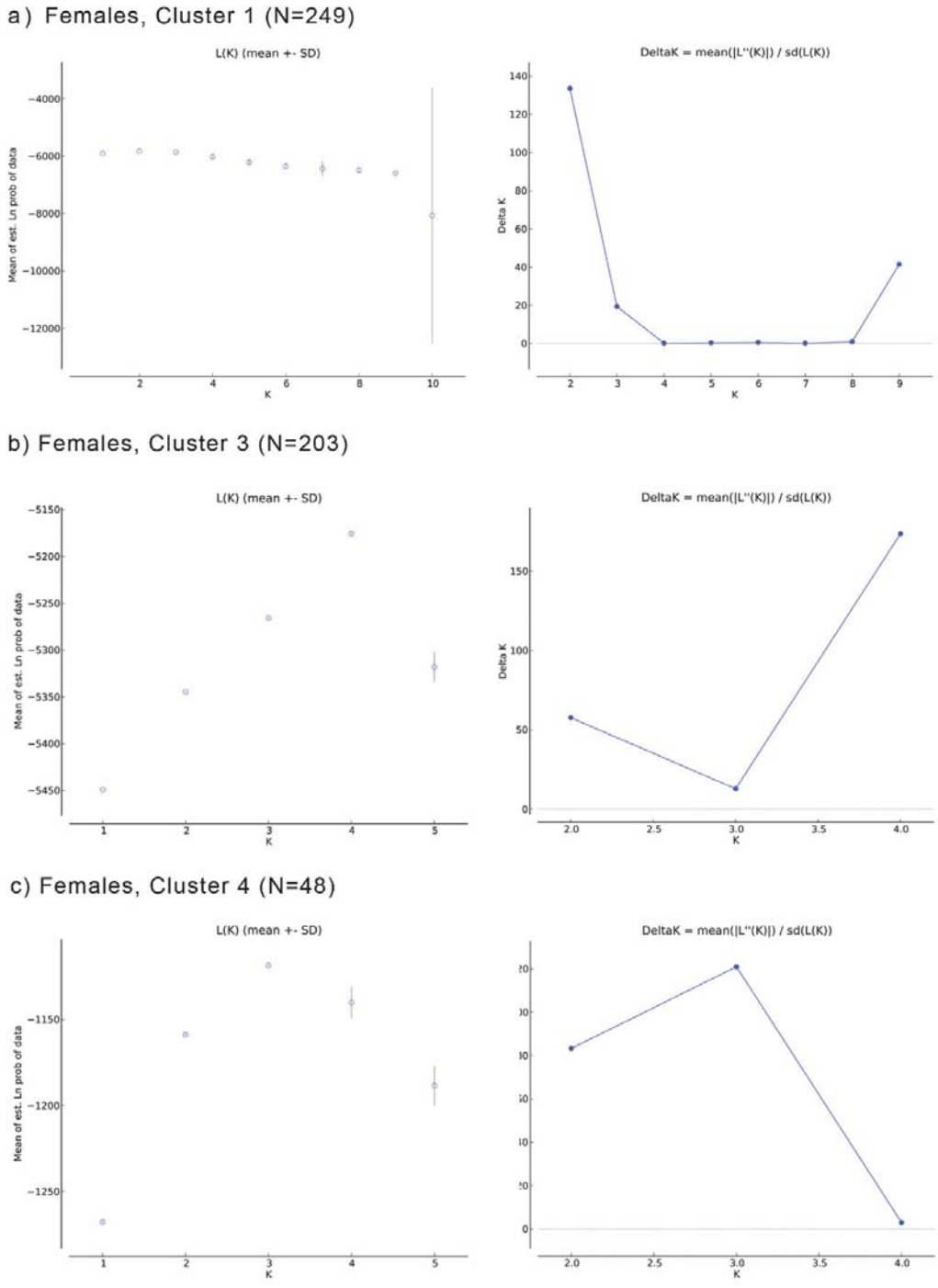


Figure S2: Results of the Bayesian clustering analysis of female bears in clusters a) one, b) three and c) four with STRUCTURE (Pritchard et al. 2000), processed with the help of STRUCTURE HARVESTER (Earl & von Holdt 2012).

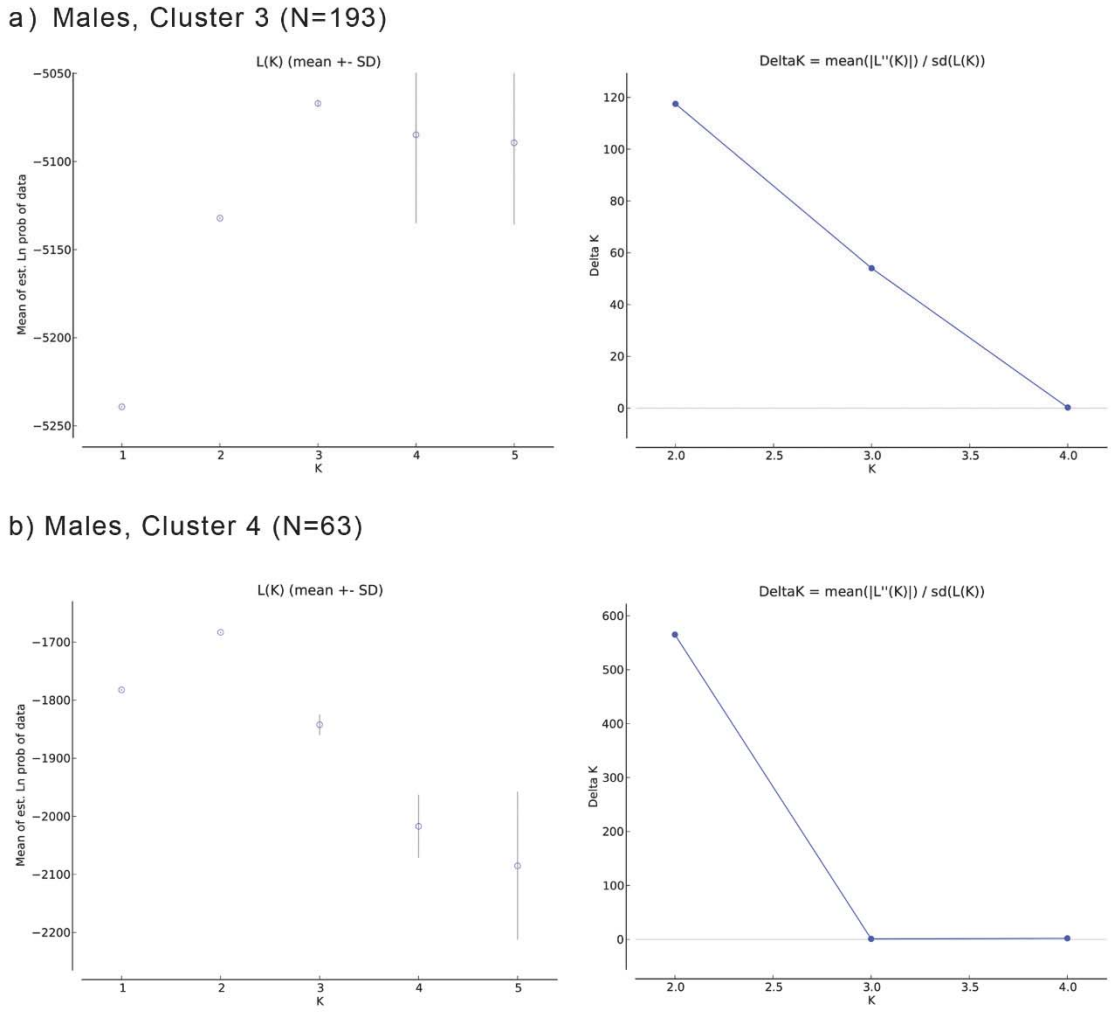


Figure S3: Results of the Bayesian clustering analysis of male bears in clusters a) three and b) four with STRUCTURE (Pritchard et al. 2000), processed with the help of STRUCTURE HARVESTER (Earl & von Holdt 2012).

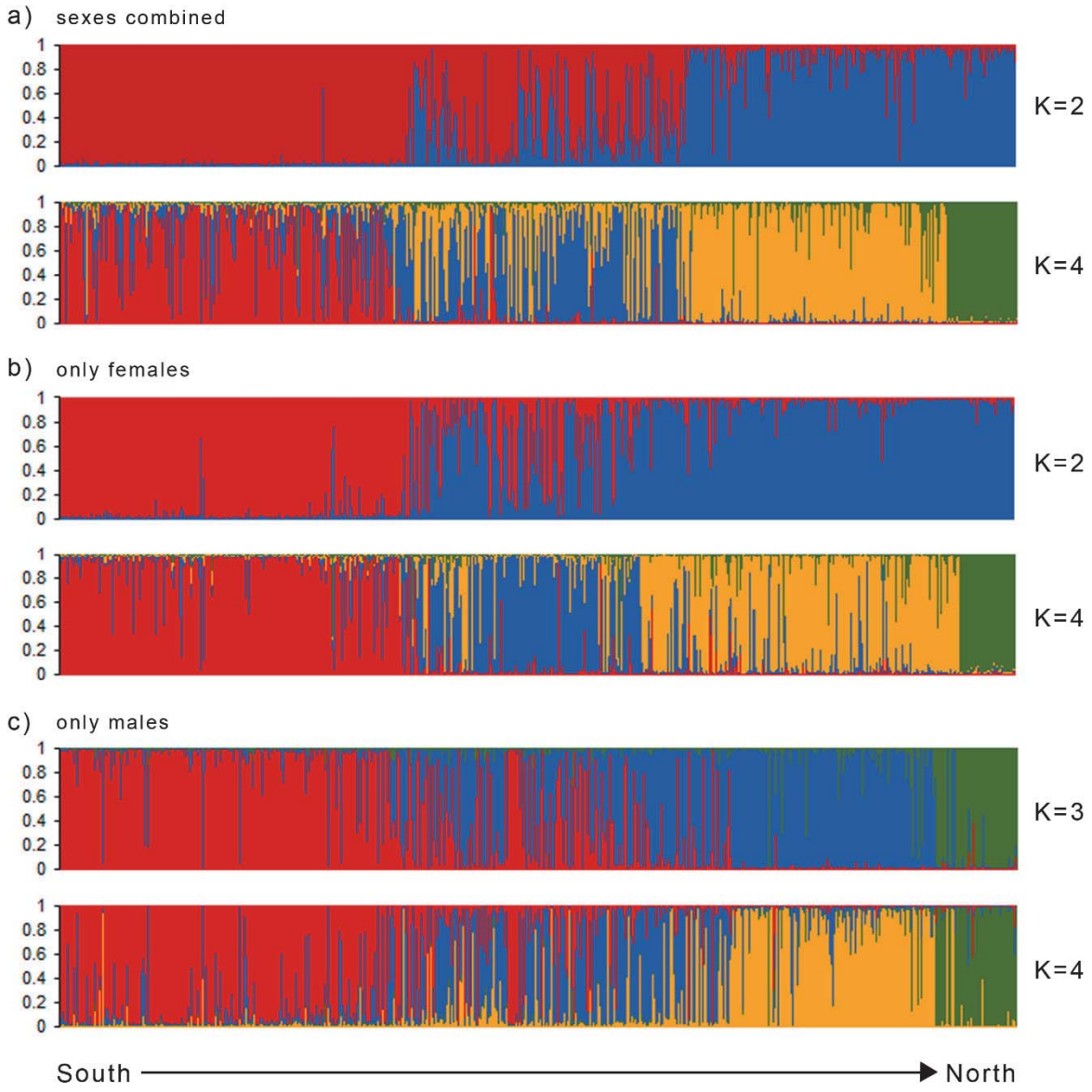


Figure S4: Results of the Bayesian clustering analysis of brown bears in Sweden and Norway with STRUCTURE (Pritchard et al. 2000). Each bear is represented by one bar, the segments of which are sized and colored according to the estimated assignment probability *q* for the given number of clusters K, the individuals are sorted from south to north. a) results for the analysis of males and females combined (n=1531) for K=2 and K=4; b) results for the analysis of only females (n=742) for K=2 and K=4; c) results for the analysis of only males (n=789) for K=3 and K=4.

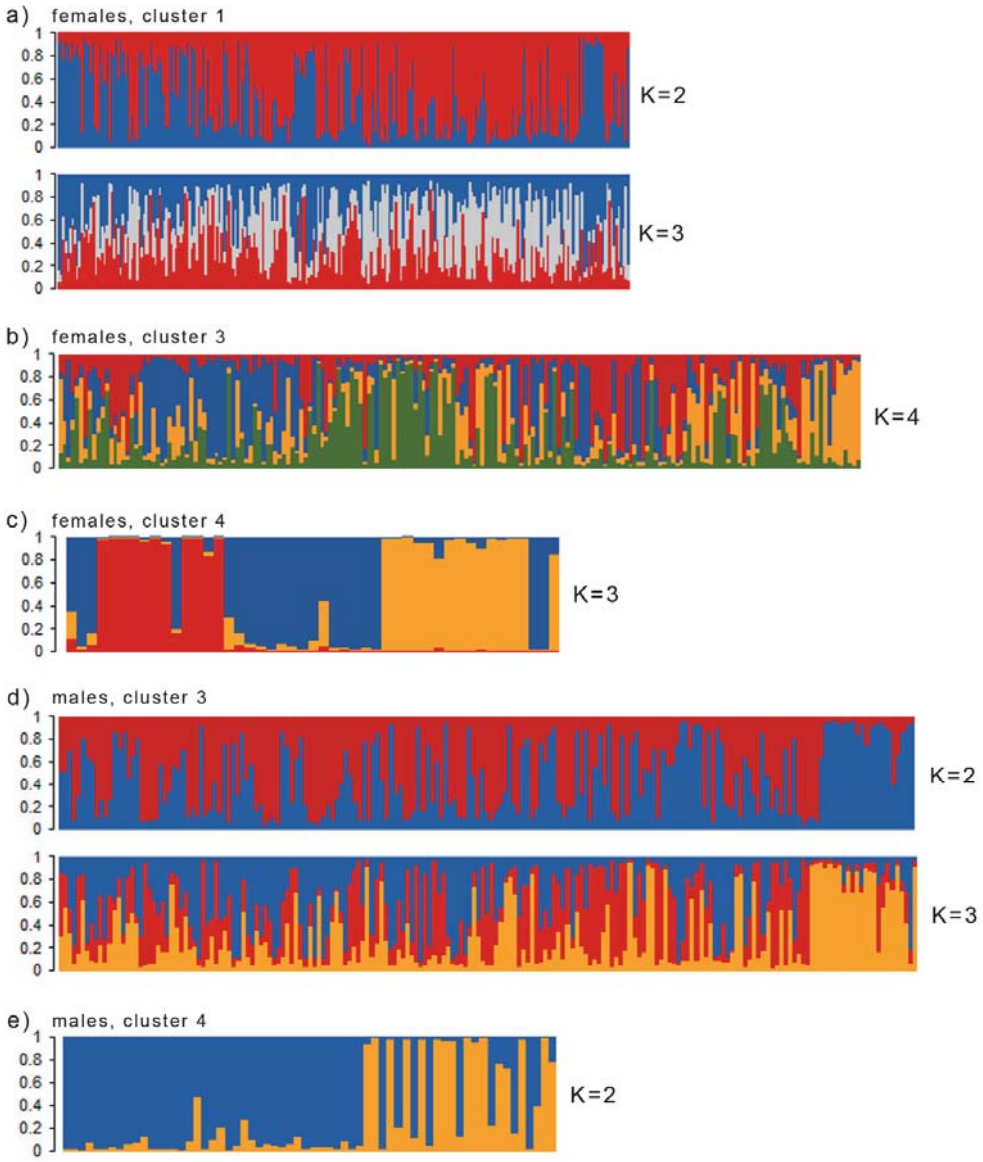


Figure S5: Results of the Bayesian clustering analysis of brown bears in Sweden and Norway within previously determined main clusters with STRUCTURE (Pritchard et al. 2000). Each bar equals one bar, the segments of which are sized and colored according to the estimated assignment probability *q* for the given number of subclusters K. Each barplot is ordered according to sampling region, either from south to north (a, b and d) or from west to east (c and e). a) results for the analysis of females in cluster 1 (n=249) for K=2 and K=3; b) results for the analysis of females in cluster 3 (n=203) for K=4; c) results for the analysis of females in cluster 4 (n=48) for K=3; d) results for the analysis of males in cluster 3 (n=193) for K=2 and K=3; e) results for the analysis of males in cluster 4 (n=63) for K=2.

PAPER V

The effect of scale and isolation by distance on the analysis of population structure in large, continuous populations: the Scandinavian brown bear (*Ursus arctos*)

Julia Schregel[1,2], Jaanus Remm[3], Hans Geir Eiken[1], Jon E. Swenson[2,4], Urmas Saarma[3], Snorre B. Hagen[1]

[1] Norwegian Institute of Bioeconomy, NIBIO - Svanhovd, 9925 Svanvik, Norway

[2] Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, 1432 Ås, Norway

[3] University of Tartu, Department of Zoology, Institute of Ecology and Earth Sciences, 51014 Tartu, Estonia

[4] Norwegian Institute for Nature Research, 7485 Trondheim, Norway

Keywords: spatial, population boundaries, gene flow, DResD, cluster placement, IBD, landscape genetics

Abstract

In conservation genetics it may be important to determine at which spatial and temporal scales population genetic processes operate and how these processes influence population genetic structure, connectivity, and diversity. The influence of isolation by distance (IBD) may bias the results of Bayesian clustering algorithms, thereby making the detection of population fragmentation and gene flow barriers challenging. We have studied the influence of IBD on the population structure of the recovered Scandinavian brown bear, which displays nearly continuous distribution across Norway and Sweden. We used data from 1531 individuals sampled between 2006 and 2013, genotyped with eight microsatellite markers. We performed a hierarchical population structure analysis and determined the spatial arrangement of identified clusters. Next, we used the DResD analysis, recently developed to correct pairwise genetic differences of individuals for the influence of IBD, to assess the population structure on two different scales. Our cluster placement results showed four non-overlapping core areas,

supporting the identification of genetically distinct clusters in the Scandinavian population. In contrast to this, the DResD analysis showed no spatial structuring for the central part of the Scandinavian brown bears, thus indicating that IBD has a stronger than assumed influence on the STRUCTURE analysis. Furthermore, the large-scale analysis revealed two distinct gene flow barriers in the south and the north for the females but not for the males, supporting the assumption of sex-specific spatial structure. On the small scale analysis, however, both sexes showed considerable amount of spatial genetic structure, suggestive of a scale-dependent influence of population structuring mechanisms.

Introduction

The successful management and conservation of wild animal populations require detailed information about population structure, connectivity, and gene flow. In many cases, traditional, invasive methods, such as GPS-tagging, are not feasible, due to logistic, financial, or animal welfare constraints, in particular for studies on larger scales, involving numerous individuals and populations. Therefore, genetic methods are increasingly used for studying many types of large-scale population patterns and processes in wildlife species (Schwartz *et al.* 2007; Luikart *et al.* 2010; Habel *et al.* 2015). For certain flagship species in conservation, such as large terrestrial carnivores, the application of genetic methods have contributed significantly to our understanding of country-wide population densities, as well as structure, connectivity, and diversity patterns among populations across national borders (Schwartz et al. 2007; Dalerum *et al.* 2009; Lowe & Allendorf 2010). Being at the top of the trophic chain, large terrestrial carnivores are often involved in livestock depredations and conflicts with humans (Breck 2004; Graham *et al.* 2005). Large home ranges, low population densities, and slow growth rates make them vulnerable to persecution, illegal hunting, and habitat fragmentation (Crooks 2002;

Cardillo *et al.* 2005; Dalerum *et al.* 2009). During the last century, large terrestrial carnivores experienced a global population bottleneck and declined dramatically both in abundance and distribution range, but due to conservation-minded management, the last decades have seen a comeback of large terrestrial carnivores into several areas of previous extirpation and some populations are again expanding (Chapron *et al.* 2014). Some countries, like Sweden and Norway, have used a genetic monitoring strategy for many years and have accumulated geographically extensive and high-resolution datasets across the distribution range of their recovering carnivore populations (Rovdata-Naturvårdsverket 2014; Rovdata 2015). This provides an opportunity for detailed population genetic analyses, which may be needed to determine the patterns, processes, threats, and linkages between local and regional scales that influence on these and other populations of large terrestrial carnivores in nature.

Two important and interrelated issues in genetic studies of large terrestrial carnivores and other widely distributed wildlife populations are to determine (1) at which spatial and temporal scales population genetic processes operate and (2) how these processes influence population genetic structure, connectivity, and diversity. Regarding the first issue, a growing number of studies suggests that scale may have a potentially strong influence on the importance of the mechanisms underlying observed population genetic patterns. For example, both gene flow patterns and genetic structure have been reported to show considerable variation depending on the spatial scale at which they are studied, indicating that the rate, and hence influence, of the respective underlying processes also vary in a scale-dependent manner, i.e. depend on the spatial and temporal extent at which they are considered (Wilder *et al.* 2004; Gauffre *et al.* 2009; Gabrielsen *et al.* 2013; Gorospe & Karl 2013; Vangestel *et al.* 2013). Identifying the scales on which genetic patterns arise and matching them to processes in terms of scale may therefore be important, both for the interpretation of results and for the

understanding of how the respective processes affect genetic structure, connectivity, and diversity. Ultimately, such knowledge also may be important for successful management and conservation e.g., for genetic risk assessments of differently scaled threats from various different anthropogenic and natural disturbances.

The second issue, i.e. how different processes give rise to genetic structure, is a central theme in theoretical, as well as applied, population genetics (Waples & Gaggiotti 2006). Understanding the scales and mechanisms of genetic structuring is important for predicting possible responses to change and hence for conservation and management. Isolation by distance (IBD), which describes a correlation between spatial and genetic distance among individuals due to limited dispersal distances (Wright 1943), is often assumed to be the "null-model", underlying observed patterns of genetic structuring among populations of large predators and other wildlife species (van Strien *et al.* 2015). This poses a challenge for their genetic assessment, as it may camouflage the influence of focal factors, such as landscape barriers, habitat fragmentation, isolation, and range expansion and contractions, on genetic structure. In addition, studies have shown that the presence of IBD can bias the results from Bayesian clustering algorithms, such as STRUCTURE (Pritchard *et al.* 2000) and Geneland (Guillot *et al.* 2005), by identifying more clusters than actually are present (Frantz *et al.* 2009), especially when coupled with noncontinuous sampling (Schwartz & McKelvey 2008; Tucker *et al.* 2014). Therefore, it is generally advised to test for IBD to help interpret the genetic clustering analysis and evaluate its biological significance (Schwartz & McKelvey 2008). However, this can be a challenge in continuously distributed populations, with no apparent spatial grouping of individuals that could aid in this evaluation process. First, even extensive knowledge about dispersal patterns may be of only limited help, partly because dispersal per see does not necessarily translate into gene flow (Johnson & Gaines 1990; Kitanishi &

Yamamoto 2015). Second, the results of STRUCTURE analysis are typically only used by assessing the individual posterior probabilities (membership value, *q*) of belonging to a particular cluster (typically the threshold value is q = 0.7, (Tammeleht *et al.* 2010; Pelletier *et al.* 2012; Kopatz *et al.* 2014; Valtonen *et al.* 2014). This approach does not allow further statistical evaluation of the posterior probabilities and thus the spatial expanse of clusters can only be assessed qualitatively, e.g. by plotting the individual results on a map of the study area.

Recent methodological advances allow further utilization and statistical evaluation of the posterior probability values from Bayesian clustering algorithms (Hindrikson *et al.* 2013) as well as correcting patterns of pairwise genetic distances for the influence of IBD, hereby identifying geographic areas in which genetic relatedness among individuals is either higher (constituting a corridor or enhanced dispersion) or lower (constituting a barrier or a transition zone) than expected under IBD (Keis *et al.* 2013). These approaches may be of considerable value for large-scale population genetic analyses of widely distributed species, where the influence of both scale dependency and IBD can be expected. Because many processes are scale dependent and because genetic structuring is often due to a concerted action of several mechanisms, studies that explicitly consider these issues in their sampling design and/or analytical approach are warranted. Hence, we applied these two new methods, combined with tests for scale dependency, to study the genetic structure of the Scandinavian brown bear population, which exhibits a more or less continuous distribution throughout most of the Scandinavian Peninsula, making distinct population boundaries nearly impossible to discern.

To this aim, we have utilized the genetic data bank compiled in the course of the Swedish and Norwegian national monitoring schemes for the brown bear (*Ursus arctos*), which ensures continuous and high resolution noninvasive genetic sampling across the distribution range. The data consist of 1531 individual genotypes of brown bears collected in Sweden and

Norway from 2006 to 2013. To analyze the data, we first applied the approach that utilizes the output generated by a STRUCTURE analysis in order to spatially delineate the identified genetic clusters (Hindrikson et al. 2013). Next, we tested for scale dependency by performing a variogram analysis. This was done to identify the spatial scales at which the spatial structuring exhibited maximal strength, thereby indicating the scale at which the appearance of a population genetic pattern is most probable and which might provide clues to the underlying processes. In a final step, we used the novel DResD method (Keis *et al.* 2013) to correct patterns of pairwise genetic distance for the effect of IBD on the various scales of genetic structuring.

Thus, in this study, we reassess the population structure of the continuously distributed brown bear population on the Scandinavian Peninsula, for the first time including also Norwegian individuals, in order to shed new light on the causes of the observed structure and the scale on which it arises. In doing so, we provide a comprehensive population genetic analysis of the Scandinavian brown bear population, which will be valuable in furthering the understanding of the mechanisms that shape its structure. In addition, our results can be used for evaluating the current and planning future management actions for the Scandinavian brown bear population.

Material and Methods

Study species

The brown bear was once abundant in large areas of the Scandinavian Peninsula. In both Sweden and Norway, it underwent a major population decline and range retraction during the 18th and 19th centuries, primarily due to human persecution and state-financed predator extermination programs (Swenson *et al.* 1994; Swenson *et al.* 1995). The Norwegian population was almost completely extirpated, whereas in Sweden ~130 individuals survived in three

refugee areas (Swenson *et al.* 1995; Servheen *et al.* 1999). Similar population declines were documented in Finland (Ermala 2003), Estonia (Valdmann *et al.* 2001), and Northwestern Russia (Danilov 2005). During the first half of the last century, attitudes towards large predators changed (Enserink & Vogel 2006) and protective measures were initiated during the 1960s and 1970s, leading to a gradual demographic recovery of the brown bear population (Swenson *et al.* 1995; Kindberg *et al.* 2011). Currently, the Swedish population is estimated to consist of ~2800 individuals (Naturvårdsverket 2014), whereas in Norway, the National Monitoring Program registered a minimum of 136 bears in 2014 (Aarnes *et al.* 2015). Previous studies have assessed the population structure of only the Swedish brown bear population with the help of microsatellite and single nucleotide polymorphism markers, identifying four (Waits *et al.* 2000) and later three subunits distributed along a north-south direction (Manel *et al.* 2004; Norman et al. 2013). Furthermore, studies using mitochondrial markers have shown that the southern part of the Swedish population represents a distinct lineage, probably originating from bears that recolonized Scandinavia from the south after the last glacial period, in contrast to the more northern areas, which received immigrants from the east (Taberlet *et al.* 1998; Hewitt 2000; Bray *et al.* 2013; Keis *et al.* 2013; Anijalg *et al.* in prep.). Dispersal in the brown bear is male-biased, with maximum dispersal distances of males being much larger than of females (467 vs 90 km in Scandinavia, (Støen *et al.* 2006)). Therefore, large-scale genetic connectivity is assumed to be driven by male brown bears (Manel *et al.* 2004; Bidon *et al.* 2014).

Sampling and genetic analysis

For this study, we utilized the extensive database compiled in the course of the genetic monitoring of the Swedish and Norwegian brown bear population. The database consists of individual genotypes obtained through the genetic analysis of georeferenced, noninvasively

collected samples (mostly fecal samples, but also some hair) and occasional tissue samples from so called "nuisance bears", shot legally. Since genetic monitoring was established in 2006, individual genotypes have been added yearly, resulting in a total of 1461 recorded brown bears (707 females, 754 males) from 2006 to 2013. These have been genotyped using eight microsatellite markers (STRs). In addition, we have added the genotypes of 35 males and 35 females from legally shot bears from the county of Jämtland in Sweden, the tissue samples of which were collected by the Swedish National Veterinary Institute. We have used this database also in another study to investigate the influence of sex-biased dispersal on the fine-scale genetic structure of the population (Schregel *et al.* in prep.) and a more detailed account of the database can be found there. The lab conducting the genetic analyses followed strict protocols to ensure unambiguous identification and is accredited according to the EN ISO/IEC 17025 standard (Norwegian accreditation: test 139). Details of the extraction, amplification and genotyping protocols are described in detail in Andreassen *et al.* (2012).

Assessing population structure and cluster placement

The first step of our population genetic analysis uses the individuals' posterior probabilities (membership values *q*) estimated in another study based on the same data (Schregel *et al.* in prep.), which included the application of the Bayesian assignment algorithm implemented in STRUCTURE v.2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003) in order to find the most likely number of genetic clusters and to assign individuals from Sweden and Norway to their likely origin. The parameters for this previous analysis were as follows. Maximum number of populations was set to K=10, assuming population admixture and correlated allele frequencies, as previous population genetic analyses using samples from Sweden resulted in K=3 as the most likely number of clusters (Manel *et al.* 2004). We performed ten independent runs for each K,

with a burn-in period of 100,000 Markov-Chain-Monte-Carlo (MCMC) iterations and a subsequent sampling of 1,000,000 MCMC iterations. The results were then processed with the help of Structure Harvester (Earl & vonHoldt 2012), which implements the calculation of the rate of change in the log probability of data between successive K values (ΔK) to determine the likely number of clusters (Evanno *et al.* 2005). The new approach by Hindrikson *et al.* (2013) enables the statistical evaluation of the obtained posterior probability values and the determination of cluster core areas as well as identification of areas with significantly low occupancy probability. This is done by first placing a grid over the study area, followed by estimating the probability of each grid point belonging to a particular cluster to be i) significantly higher, ii) significantly lower, or iii) no different to the expectations under random spatial structure of the entire population. The probability for each grid point is obtained by calculating the inverse distant weighted average (w=1/dist) of the membership value *q* from all samples for each cluster and subsequently applying bootstrap permutation test for statistical significance.

We performed the analysis separately for males and females, because the brown bear exhibits male-biased dispersal and females form matrilineal clusters (Støen *et al.* 2005; Zedrosser *et al.* 2007) and therefore males and females may show differences in the spatial extent of clusters. Based on the home range size, dispersal distance, and spatial extent of positive genetic spatial autocorrelation of brown bears in Sweden and Norway (Dahle & Swenson 2003; Zedrosser *et al.* 2007; Schregel *et al.* in prep.), we chose grid increments of 25 km. In order to obtain the statistical significance of q value estimations compared to spatial randomness (H0) at each grid point, we performed a bootstrapping procedure. Keeping the placement of sampling points, the q values were re-sampled randomly in 1000 permutations. Grid points displaying values outside the 2.5 and 97.5 percentile of the bootstrap distribution

were classified as statistically significant core area of the respective cluster and as significantly out of cluster range, respectively.

Identifying areas of high and low levels of individual admixture

The number of admixed individuals, i.e. individuals which cannot be assigned unambiguously to any one cluster, may be particularly high in areas where two or more clusters meet. In order to find those areas, we applied a modification of the method outlined above based on the Shannon Index ($H' = -\sum_{i=1}^{S} p_i ln p_i$), where p = membership value q, and S = cluster, so that an individual diversity index H' is calculated for each sample. We then applied the same algorithm to calculate the inverse distance weighted average for each grid. See Appendix S2 for the algorithm used for the cluster placement analysis and the estimation of individual diversity in R 3.0.2 language (R Core Team 2013).

Correcting population structure for the influence of IBD

IBD can have a strong effect on the performance of clustering algorithms and may render interpretation of analysis results challenging, especially when the sampling regime is not continuous (Schwartz & McKelvey 2008; Pritchard *et al.* 2010). We therefore performed a DResD analysis, which has been developed to detect subpopulation borders and areas of high dispersion based on IBD-corrected individual genetic distances (Keis *et al.* 2013). The corrected values are then interpolated throughout the study area using variogram-based distance weighting (kriging) and geographic midpoints between individuals. We estimated the Rousset's â (Rousset 2000), an analogue to $F_{ST}/(1-F_{ST})$, for genetic distance between pairs of individuals using SPAGeDi 1.4c (Hardy & Vekemans 2002). Rousset's â was developed specifically to study IBD on the individual level and in continuously distributed populations (Rousset 2000).

It has the advantage of not relying on a reference population (Vekemans & Hardy 2004) and has been used in various landscape and population genetic studies (Leblois *et al.* 2000; Coulon *et al.* 2004; Schwalm *et al.* 2014; Valtonen *et al.* 2014). We analyzed first males and females combined and then both sexes separately. For a detailed description of the DResD method and the algorithm, see Keis *et al.* (2012) and Hindrikson *et al.* (2013).

Identifying the scale at which population genetic structure arises

Different processes may act on different spatial scales when shaping the genetic structure of a population (Wilder *et al.* 2004; Gauffre *et al.* 2009; Gabrielsen *et al.* 2013; Gorospe & Karl 2013; Vangestel *et al.* 2013). We therefore performed a variogram analysis to identify the spatial scales at which the spatial structuring exhibited maximal strength, thereby indicating the scale at which the appearance of a population genetic pattern is most probable. Using deviations of pairwise genetic distance from the IBD-model, we conducted a series of variogram models that measure the appearance of a population pattern. The variogram models were calculated for sample pair geographic distance zones from 10 to 350 km with mean step of 20 km. The upper and lower limits of the distance zones were set as ± ⅓ of mid-zone distance. Hence, the short-distance zones were narrower, providing finer pattern indication. After variogram modeling for each distance zone, relative partial sill (partial sill / sill) was used to measure the degree of population spatial structuring at different scales. The locally weighted scatter plot smooth regression (LOESS) model was used to detect the peak values of pattern scales. The final DResD analysis for placement of subpopulations' borders and cores was performed at the scales indicated by the results of the variogram analysis. 200 bootstrap permutations were used to test statistical significance of the population border and core areas.

Results

Assessing population structure and cluster placement

The latest analysis of the population structure of Norwegian and Swedish brown bears identified four genetic clusters, arranged spatially following the Scandinavian Peninsula from south to northeast (Schregel *et al.* in prep.) (Figure 1). Using the estimated membership coefficient *q* for the analyzed individuals from each of the four clusters, we performed a cluster placement analysis to identify the core areas. The analysis identified four separate core areas for males and females, with males displaying slightly larger core areas than females. There was only minor overlap between clusters 3 and 4 (Figure 1c). For the clusters 1 to 3, the analysis identified areas significantly out of range for the respective clusters for both males and females (Figure S1 and S2), whereas for cluster 4 no such area could be determined. For the clusters 1 to 3 (except for the males from cluster 3), the out-of-range area included the easternmost part of cluster 4, which has been characterized by previous studies as belonging genetically to the brown bear population distributed across Finland and Western Russia (Tammeleht et al. 2010; Kopatz *et al.* 2012; Schregel *et al.* 2012; Kopatz *et al.* 2014). The largest area classified as out of range for both sexes was estimated for cluster 1, indicating limited bidirectional gene flow (Figure S1a and S2a). Likewise, gene flow from cluster 3 to cluster 1 seems to be limited, as this area was classified as out of range for both sexes (Figure S1c and S2c). In general, females showed a slightly larger amount of grid points classified as out of range, as would be expected under male-biased dispersal.

Areas of high and low levels of genetic admixture

To find areas with an elevated or reduced level of individual diversity of ancestors, we modified the cluster placement algorithm based on the Shannon index. The results were similar for both

sexes (Figures 2a and b), indicating a relatively large area with a high number of admixed individuals situated in Norway, supposedly between clusters 1 and 2. Similar, but smaller areas were indicated for the males farther north, between cluster 2 and 3, as well as at several locations in Sweden in clusters 2 and 3 for both sexes. These last areas were larger for the females than the males. We also found areas with a reduced level of admixture, i.e. a high proportion of individuals assigned to a particular cluster with a comparably high posterior probability ($q>0.7$). Those areas were the western and northeastern parts of cluster 4 for both sexes, as well as the southernmost region of cluster 1. Males displayed an additional small area of reduced admixture in the eastern coastal region of cluster 3.

Identifying the scale at which population genetic structure arises

The variogram analysis showed two peaks at which the strength of spatial patterning reached a maximum, thereby indicating the scale of analysis for which the appearance of population spatial structure was most probable (Figure 3). Based on this result, we chose to explore the population genetic structure at two spatial scales: on the small scale, we analyzed samples pairs at a distance of 23-47 km, and on the large scale, we considered sample pairs at a distance of 150-300 km.

Correcting population structure for the influence of IBD

IBD was similar for males and females (Figure S3). As expected, females showed a lower value of genetic dissimilarity than males at the lower end of the geographic scale, reflective of the formation of matrilineal assemblages (Støen *et al.* 2005). Males appeared to be slightly more similar to each other than females throughout the remainder of the spatial range, in agreement with the documented male-biased dispersal pattern (Støen *et al.* 2006; Zedrosser *et al.* 2007).

We performed the DResD analysis first with males and females combined and then separated by sex. On the small scale (23-47 km sample pair distance), there were some small but notable differences among the three different analyses (Figure 4a, c and e). Areas of relatively high genetic similarity were displayed for all three analyses for three regions; the northwestern part of cluster 3 and the western and northern-eastern parts of cluster 4. For the females, these identified regions encompassed a larger area, especially the ones in cluster 4, but also for the males these areas showed higher similarity than expected. At this scale, this probably points to large family groups in those regions, which is expected for females, but not for males. Some areas of high genetic similarity were identified in the coastal region of cluster 3 towards the Bothnian Bay, in the western part of cluster 2, and in several locations in cluster 1. Comparing males and females, it seems that matrilineal assemblages contributed the most to the genetic similarity found in the south, whereas males were most influential for the location identified at the Bothnian Bay in cluster 3. Furthermore, some small areas of higher-than-expected genetic dissimilarity were found scattered throughout the region in clusters 1, 2 and 3, indicating boundaries between genetically dissimilar groups.

The large-scale analysis revealed two major areas where positive deviation from IBD was significantly high: a stripe of approx. 27,800 km² in the North, and wide zone of 60,500 km² in the South. This barrier-like transition areas divides the brown bear population in three parts; cluster 1 in the south, a large cluster to the north of it (consisting of clusters 2 and 3), and cluster 4 to the northeast (Figure 4b, d and f). Unlike the small scale analysis, the difference between the results for males and females in the large-scale analysis was substantial (Figure 4d and f). Whereas females displayed an even more extensive partition between cluster 1 and 2, essentially limiting gene flow across the entire width of the peninsula, and an even stronger barrier towards the east, almost no significant barrier- or corridorlike signal was identified for

the males. Only a small number of grid points (approx. 3,000 km²) on the Bothnian Bay coastline in the area of cluster 2 were classified as barrier. The descriptive statistics for the entire DResD analysis are given in Table S1, showing that, on the 35-km scale, the proportion of areas identified as either barrier or corridor was similar between the total dataset (1.7%) and the females (1.8%) and only slightly lower for the males (1%). On the 225-km scale, however, males displayed a substantially lower proportion of areas showing departure from IBD expectations (0.3%) than females (13.5%) and the combined data (13.4%).

In summary, the results of the cluster placement and DResD analyses, as well as the analysis of individual-level admixture, substantially refined the knowledge about the population structure and gene flow of the Scandinavian brown bear population (Figure 5). Although the genetic clusters 1 and 4, identified with the STRUCTURE analysis, probably represented genuine subpopulations that are connected only weakly to the central part by mostly male gene flow, clusters 2 and 3 may represent the opposing poles of a large cluster under strong IBD influence. The small-scale analysis showed an unexpected amount of structure for male bears that seemed to increase towards the north. Gene flow from the western part of the population (clusters 1, 2, and 3) towards the eastern part (cluster 4) is less than vice versa and as such asymmetrical, whereas gene flow between cluster 1 and the population farther north does not occur uniformly along the contact zone, but mostly on the Norwegian side of it.

Discussion

A better understanding how the concerted action of small- and large-scale genetic processes influence population structure and gene flow of wildlife species has a multitude of applications, be it conservation-driven studies that attempt to understand responses to change or those that aim to refine and improve current population genetic theory. To obtain such integrated

perspectives at various spatial scales, it has been show that studies investigating the interrelationship between scale and population genetic processes should aim to increase the number and density of samples, rather than the number of markers (Banks & Peakall 2012). We have shown here, that data collected during the course of genetic monitoring projects may be of particular value, as they offer high sampling resolution across large spatial scale, even though the number of markers is often limited because of financial constraints. Provided that the genetic data is based on markers selected for their reliability to discern among individuals, our results show that a detailed population genetic study is possible with the lower number of markers typical for genetic monitoring projects.

Large-scale structure

The genetic structure of the recovered Swedish bear population has been assessed previously, first using an assignment test and the estimation of genetic distance among groups of bears determined *a priori* (Waits *et al.* 2000), followed by a re-assessment using STRUCTURE (Pritchard *et al.* 2000) and without prior definition of subgrouping (Manel *et al.* 2004). The latter study resulted in a reduction of the number of identified subpopulations within Sweden from four to three, demonstrating the importance of analyzing population structure without preimposing a particular spatial organization on the data. We have now corrected pairwise genetic dissimilarities for the influence of IBD, and our results indicate that the formation of the clusters 2 and 3 may be more influenced by IBD than previously assumed, as the DresD algorithm did not identify two separate clusters (Figure S5). One reason for the difference in results may lie in the strength of the markers used, as Waples & Gaggiotti (2006) have shown that a low number of loci can also lower the ability of structuring algorithms to correctly identify the number of genetic clusters. Whereas the two previous studies used 19 STRs (Waits *et al.*

2000; Manel *et al.* 2004), we used only eight. However, the number of genetic clusters, as well as their spatial arrangement that we found with our Structure analysis was similar to those previous studies. In addition, whereas the DResD analysis does not show a division between clusters 2 and 3, the barriers between clusters 1 and 2 and between cluster 3 and 4, which have been found in previous studies (Waits *et al.* 2000; Tallmon *et al.* 2004; Kopatz *et al.* 2012; Schregel *et al.* 2012; Bray *et al.* 2013; Kopatz *et al.* 2014) were detected by the algorithm. This indicates that the strength of the selected markers to detect population delineation should be sufficient.

The analysis with STRUCTURE showed a considerable number of individuals assigned to cluster 2 and 3 with a q value >0.7, supporting the interpretation of the existence of two separate clusters and rendering the reconciliation of the differing analysis results difficult. Several previous studies have shown how IBD can bias the results of a cluster analysis and generally recommend a critical evaluation of whether the obtained result is biological meaningful (Schwartz & McKelvey 2008; Frantz *et al.* 2009). As demonstrated by the contrasting results between Manel et al. (2004), our STRUCTURE analysis and our DresD results though, this may not be enough to uncover all relevant processes shaping the observed structure. Based on the spatial placement of female concentration areas and the refuge areas during and after the demographic bottleneck (Swenson *et al.* 1995), the interpretation by Manel *et al.* (2004) of two separate clusters in central Scandinavia makes biological sense. A recent study, which included prebottleneck brown bear samples from Scandinavia, proposed that the historic population may have been structured similarly as the modern one (Xenikoudakis *et al.* 2015), which may lend support to the interpretation of separate clusters 2 and 3. However, when Xenikoudakis *et al.* (2015) analyzed the historical samples separately with STRUCTURE, the software only identified a southern and a central-northern cluster. The number and spatial

density of historical samples was limited, especially in central Scandinavia, so that an unambiguous interpretation of results is not possible. However, in contrast to the interpretation by Xenikoudakis *et al.* (2015), we hypothesize that the STRUCTURE results for a single historical central Scandinavian population was correct and not caused mainly by limited sample size from central Scandinavia. If, furthermore, the prebottleneck population was under a similarly strong influence of IBD, as indicated by our DResD results for the contemporary brown bears, and then, during the bottleneck, mainly bears in the center between the southern end (corresponding to cluster 2) and the northern end (corresponding to cluster 3) were decimated (Swenson *et al.* 1995), the loss of these supposedly admixed individuals may have enhanced the signal of two separate clusters in the STRUCTURE analysis of contemporary brown bears.

Hagen *et al.* (2015) have demonstrated how rapidly population structure can change, in contrast to what is predicted from theory (Landguth *et al.* 2010), so that spatial structure of the Scandinavian population may have changed in the time between the previous study (Manel *et al.* 2004) and ours. However, because the results from the STRUCTURE analysis used in our analysis were similar to those in Manel et al. (2004), it is unlikely that the explanation for the merging of the two central genetic clusters by the DResD algorithm lies in a rapid admixture in the Scandinavian brown bear population, as Hagen *et al.* (2015) found in the Finnish population. Admixed individuals (q value <0.7) were not numerous and, most importantly, not located predominantly in the border area between the two clusters, as shown in our previous STRUCTURE analysis (Schregel *et al.* in prep.). Furthermore, we did not detect an extensive signal of individual diversity of ancestors in central Scandinavia, whereas an area of high admixed ancestry was identified between cluster 1 and 2. It seems, thus, that the most likely explanation for the detection of two separate clusters in central Scandinavia by STRUCTURE

is that IBD influences the spatial structure in that area to a larger degree than anticipated, an effect which may have been exacerbated by the bottleneck.

Small-scale structure

Small-scale DResD analysis has revealed an unexpected amount of local structure for male bears, both in terms of significantly higher and lower genetic dissimilarity than what would be expected under IBD. This may point to some degree of patrilocality, delayed natal dispersal, or territorial behavior, preventing immigrant males from successful mating. The results of a previous study using an analysis of spatial autocorrelation on the same samples as in this study (Schregel et al. *in prep.*), pointed to an unexpected level of structuring for male brown bears in northern Sweden/cluster 3. Also, a recent study using Y-chromosomal markers found a degree of male structuring in the Scandinavian brown bear population that was not anticipated, suggesting lower levels of male gene flow among regions than expected under male-biased dispersal (Schregel *et al.* submitted). Previous studies on the dispersal behavior of Scandinavian brown bears (Støen *et al.* 2006; Zedrosser *et al.* 2007) did not report deviant dispersal patterns of male brown bears that could explain the observed small-scale structure. However, Støen *et al.* (2006) concluded that inversely density-dependent natal dispersal, inversely density-dependent home range size (Dahle & Swenson 2003), and the ability to recognize kin, as demonstrated by the formation of female assemblages (Støen *et al.* 2005), suggest that brown bears are more territorial than previously assumed. Polygyny and short dispersal distances should promote the occurrence of fine-scale genetic structure (Quaglietta *et al.* 2013). Previous research has shown that, whereas natal dispersal distances are inversely density dependent (individual-based density), overall dispersal distances were longer in the southern than in the northern regions of the Swedish population (Støen *et al.* 2006). At the same time, Støen et al.

(2006) found that natal home ranges tended to be larger in the north than the south, so that dispersing males would have to cross longer distances to move out of the natal home range. The results of our genetic analysis show that these differences in behavior are reflected in the genetic structure. This in turn suggests that territorial behavior creating barriers to gene flow is more pronounced in the north than in the south, perhaps influenced by differences in carrying capacity of the habitat or other external factors, e.g. illegal killing of bears. It has been shown that the acceptance and occurrence of illegal killing of bears is higher in the north than in the south and has a strong effect on the large carnivore populations (Andrén et al. 2006; Persson et al. 2009; Liberg et al. 2011; Gangaas et al. 2013; Rauset 2013). Male bears leave the dens earlier than the females, often at a time when there is still snow on the ground, enabling poachers to use snowmobiles and thus kill more effectively (Rauset 2013). A targeted study should be conducted to shed more light on the underlying mechanisms causing the observed pattern of local structuring for male brown bears.

The influence of scale on genetic structure analysis

The results of our analyses show that scale plays an important role in the detection of genetic structure in continuous populations, even if the species is highly mobile, as is the brown bear. It is a well-known problem in ecology that the upscaling of results obtained at a small scale can lead to erroneous conclusions (Underwood *et al.* 2005). Here we show that this is equally true the other way around; even though male brown bears seem to be structured mostly by IBD at the large scale, at the small scale they display genetic structure that may be attributed to limits to gene flow and the formation of patrilineal groups. At the large scale, rare long-distance dispersal events and multigenerational gene flow may counteract the occurrence of genetic structure, whereas at the small scale, these have relatively little influence in comparison to natal

dispersal patterns and territoriality. Our results indicate two independent biological processes in the appearance of population genetic structure: home range related processes at scale of 35 km, and sub-population clustering at scale of 225 km. Spatial scales are likely to be related to the temporal variability of these processes and patterns. This is good evidence for the advantage of the DResD procedure over non-spatial clustering methods for individual based analyses to explore contemporary population processes and patterns with no limits for spatial scale.

The influence of IBD on genetic structure analysis

Our combined approach of STRUCTURE analysis with subsequent cluster placement and individual diversity of ancestors estimation and DResD analysis has resulted in a new and deeper understanding of how the brown bear population on the Scandinavian Peninsula is structured, including new information on small- and large-scale gene flow patterns. At the same time, our results highlight the difficulties of the more conventional approach of delineating population boundaries and identifying population structure. Earlier STRUCTURE analyses detected four clusters (Schregel *et al.* in prep.), with clusters 1 to 3 coinciding with the ones found by Manel *et al.* (2004). F_{ST} values among those four clusters ranged from 0.035 to 0.078 for the total data set, from 0.040 to 0.082 for the females, and from 0.033 to 0.079 for the males (Schregel *et al.* in prep.). These F_{ST} values did not allow for a differentiated interpretation of how reliable cluster identification was for the combined or for the data separated by sex. In particular, F_{ST} values between clusters 3 and 4 were very similar for males (0.048) and females (0.054), especially in regard to the values estimated for females between clusters 1 and 2 (0.050), and stand in strong contrast to the results of the large-scale DResD analysis, which identified a barrier to gene flow for the females for the clusters in question, but not for the males.

In the same study (Schregel *et al.* in prep.), each cluster was reanalyzed in a hierarchical manner to identify potential substructure. The result was the identification of an increasing amount of substructure for both males and females towards the northeast of the study area, with F_{ST} values increasing towards the northeast, from 0.034 between the two subclusters of the cluster 1 to 0.154 between two subclusters of cluster 4 (Schregel *et al.* in prep.). However, the large scale DResD analysis did not identify any barriers to gene flow between those previously detected subclusters in the northern parts of the study area. Instead of limited gene flow among those subclusters, it is more likely that the higher F_{ST} values among subclusters are caused by the strong and large family clusters identified in the small-scale DResD analysis. The location of these coincides with the location of the subclusters previously detected with the STRUCTURE analysis (Schregel *et al.* in prep.).

This study has demonstrated, that the application of the DResD algorithm to spatially explicit genetic data is a valuable contribution to population genetic studies, especially in cases where a specific landscape genetic approach is not feasible. Recently, several studies have shown that the spatial scale considered in a population or landscape genetic analysis can have a large influence on the results and on the ability to find the best fitting model to explain the observed pattern (Gabrielsen *et al.* 2013; Gorospe & Karl 2013; Keller *et al.* 2013). Here, the DResD approach may help in identifying areas that should be the focus of a targeted analysis.

Conservation aspects

Since the establishment of protection measures at the turn of the last century in Sweden and the 1930s in Norway (Swenson *et al.* 1995), the hunting of brown bears has been strictly regulated in both countries by the respective national agencies in order to ensure the establishment of a self-sustaining population (Servheen *et al.* 1999). Since 2001 in Sweden and 2006 in Norway,

monitoring of the population has been based primarily on the genetic identification of noninvasively collected samples (Rovdata-Naturvårdsverket 2014). The transboundary nature of the Swedish-Norwegian brown bears has prompted the development of a "methodology agreement" between the respective countries, so that monitoring will be conducted in a similar way to ensure comparability of results (Rovdata 2014). Our results indicate two areas that should receive special consideration in terms of management. One is the border area between the southernmost cluster (cluster 1) and the one north of it (cluster 2). Here it seems that especially male bears use Norway as a corridor, whereas on the Swedish side, communication between the two clusters is severely limited. Likewise, there seems to be a fairly good connection between Norway and Sweden in the northwestern most area of the central cluster (cluster 3). Previous results from a study using Y-chromosomal markers indicates that the Norwegian part of the area receives more gene flow from the east than the northernmost area in Sweden (Schregel *et al.* submitted). Together, these results indicate that that the brown bear population may benefit from coordinating management actions between the two countries in this area and our results may thus contribute to the promotion of joint management of this border-crossing species (Blanco 2012).

References

Andreassen R, Schregel J, Kopatz A*, et al.* (2012) A forensic DNA profiling system for Northern European brown bears (Ursus arctos). *Forensic Science International-Genetics* **6**, 798-809.

Andrén H, Linnell JDC, Liberg O*, et al.* (2006) Survival rates and causes of mortality in Eurasian lynx (Lynx lynx) in multi-use landscapes. *Biological Conservation* **131**, 23-32.

Bidon T, Janke A, Fain SR, *et al*. (2014) Brown and polar bear Y chromosomes reveal extensive male-biased gene flow within brother lineages. *Mol Biol Evol* **31**, 1353-1363.

Blanco JC (2012) Towards a population level approach for the management of large carnivores in Europe. Challenges and opportunities. European Comission.

Bray SCE, Austin JJ, Metcalf JL*, et al.* (2013) Ancient DNA identifies post-glacial recolonisation, not recent bottlenecks, as the primary driver of contemporary mtDNA phylogeography and diversity in Scandinavian brown bears. *Diversity and Distributions* **19**, 245-256.

Breck SW (2004) Minimizing carnivore-livestock conflict: the importance and process of research in the search for optimal solutions. In: Predators and People: from conflict to conservation (eds. Fascione A, Delach A, Smith ME). Island Press, Wahington D.C.

Cardillo M, Mace GM, Jones KE, et al. (2005) Multiple causes of high extinction risk in large mammal species. *Science* **309**, 1239-1241.

Chapron G, Kaczensky P, Linnell JDC, et al. (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* **346**, 1517-1519.

Coulon A, Cosson JF, Angibault JM*, et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Mol Ecol* **13**, 2841-2850.

Crooks KR (2002) Relative sensitivities of mammalian carnivores to habitat fragmentation. *Conservation Biology* **16**, 488-502.

Dahle B, Swenson JE (2003) Seasonal range size in relation to reproductive strategies in brown bears Ursus arctos. *Journal of Applied Ecology* **72**, 660-667.

Dalerum F, Cameron EZ, Kunkel K, Somers MJ (2009) Diversity and depletions in continental carnivore guilds: implications for prioritizing global carnivore conservation. *Biol Lett* **5**, 35-38.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Enserink M, Vogel G (2006) Wildlife conservation - The carnivore comeback. *Science* **314**, 746-749.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**, 2611-2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology* **46**, 493-505.

Gabrielsen CG, Kovach AI, Babbitt KJ, McDowell WH (2013) Limited effects of suburbanization on the genetic structure of an abundant vernal pool-breeding amphibian. *Conservation Genetics* **14**, 1083-1097.

Gangaas KE, Kaltenborn BP, Andreassen HP (2013) Geo-spatial aspects of acceptance of illegal hunting of large carnivores in Scandinavia. *PlosOne* **8**. doi: 10.1371/journal.pone.0068849

Gauffre B, Petit E, Brodier S, Bretagnolle V, Cosson JF (2009) Sex-biased dispersal patterns depend on the spatial scale in a social rodent. *Proceedings of the Royal Society B-Biological Sciences* **276**, 3487-3494.

Gorospe KD, Karl SA (2013) Genetic relatedness does not retain spatial pattern across multiple spatial scales: dispersal and colonization in the coral, Pocillopora damicornis. *Molecular Ecology* **22**, 3721-3736.

Graham K, Beckerman AP, Thirgood S (2005) Human-predator-prey conflicts: ecological correlates, prey losses and patterns of management. *Biological Conservation* **122**, 159-171.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes* **5**, 712-715.

Habel JC, Zachos FE, Dapporto L, et al. (2015) Population genetics revisited - towards a multidisciplinary research field. *Biological Journal of the Linnean Society* **115**, 1-12.

Hagen SB, Kopatz A, Aspi J, Kojola I, Eiken HG (2015) Evidence of rapid change in genetic structure and diversity during range expansion in a large terrestrial carnivore. *Proceedings of the Royal Society B-Biological Sciences* **282**, 20150092.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913.

Hindrikson M, Remm J, Maennil P*, et al.* (2013) Spatial Genetic Analyses Reveal Cryptic Population Structure and Migration Patterns in a Continuously Harvested Grey Wolf (Canis lupus) Population in North-Eastern Europe. *Plos One* **8**, e75765.

Johnson ML, Gaines MS (1990) Evolution of dispersal - theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21**, 449-480.

Keis M, Remm J, Ho SYW*, et al.* (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *Journal of Biogeography* **40**, 915-927.

Keller D, Holderegger R, van Strien MJ (2013) Spatial scale affects landscape genetic analysis of a wetland grasshopper. *Mol Ecol* **22**, 2467-2482.

Kitanishi S, Yamamoto T (2015) Comparison of genetic structure between juvenile and adult masu salmon indicates relatively low reproductive success of dispersers. *Environmental Biology of Fishes* **98**, 405-411.

Kopatz A, Eiken HG, Aspi J*, et al.* (2014) Admixture and Gene Flow from Russia in the Recovering Northern European Brown Bear (Ursus arctos). *Plos One* **9**, e97558.

Kopatz A, Eiken HG, Hagen SB*, et al.* (2012) Connectivity and population subdivision at the fringe of a large brown bear (Ursus arctos) population in North Western Europe. *Conservation Genetics* **13**, 681-692.

Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol* **19**, 4179-4191.

Leblois R, Rousset F, Tikel D, Moritz C, Estoup A (2000) Absence of evidence for isolation by distance in an expanding cane toad (Bufo marinus) population: an individual-based analysis of microsatellite genotypes. *Mol Ecol* **9**, 1905-1909.

Liberg O, Chapron G, Wabakken P*, et al.* (2011) Shoot, shovel and shut up: cryptic poaching slows restoration of a large carnivore in Europe. *Proceedings of the Royal Society of London Series B-Biological Sciences* **279**, 910-915.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Mol Ecol* **19**, 3038-3051.

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

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol Ecol* **7**, 963-974.

Manel S, Bellemain E, Swenson JE, Francois O (2004) Assumed and inferred spatial structure of populations: the Scandinavian brown bears revisited. *Mol Ecol* **13**, 1327-1331.

Pelletier A, Obbard ME, Mills K*, et al.* (2012) Delineating genetic groupings in continuously distributed species across largely homogeneous landscapes: a study of American black bears (Ursus americanus) in Ontario, Canada. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **90**, 999-1014.

Persson J, Ericsson G, Segerstrom P (2009) Human caused mortality in the endangered Scandinavian wolverine population. *Biological Conservation* **142**, 325-331.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

Pritchard JK, Wen X, Falush D (2010) *Documentation for structure software: Version 2.3* University of Oxford, University of Chicago, Oxford.

Quaglietta L, Fonseca VC, Hajkova P, Mira A, Boitani L (2013) Fine-scale population genetic structure and short-range sex-biased dispersal in a solitary carnivore, Lutra lutra. *Journal of Mammalogy* **94**, 561-571.

Rauset GR (2013) Life and Death in Wolverines (Doctoral Thesis). *Swedish University of Agricultural Sciences*, Uppsala.

Rcoreteam (2013) R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria.
Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.
Rovdata-Naturvårdsverket (2014) *Faktablad björn - Inventeringsmetodik.* http://www.rovdata.no/Portals/Rovdata/Dokumenter/Instrukser/Bj%C3%B6rn%20%C3%B6vervakningen%20i%20Skandinavien.pdf. 29.04.2015. in Swedish
Rovdata (2014) *Felles norsk-svenske metoder for bjørn og ulv er klare.* http://www.rovdata.no/Nyheter/Nyhetsartikkel/tabid/3764/ArticleId/3721/reftab/1999/Default.aspx. 20.04.2015. in Norwegian
Rovdata (2015) *Overvåking - Bestanden av brunbjørn i Norge blir primært overvåket ved å analysere DNA fra innsamlede ekskrementer og hår.* http://www.rovdata.no/Brunbj%C3%B8rn/Overv%C3%A5king.aspx. 13.05.2015
Schregel J, Eiken HG, Grøndahl FA*, et al.* (submitted) Present and past Y chromosomes reveal the demographic and genetic impact of male dispersal during the recovery of the Northern European brown bear (Ursus arctos).
Schregel J, Eiken HG, Swenson JE, Hagen SB (in prep.) A multiscale analysis of sex-dependent population structure and gene-flow: The case of the Northern European brown bear (Ursus arctos).
Schregel J, Kopatz A, Hagen SB*, et al.* (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Mol Ecol* **21**, 3474-3488.
Schwalm D, Waits LP, Ballard WB (2014) Little fox on the prairie: genetic structure and diversity throughout the distribution of a grassland carnivore in the United States. *Conservation Genetics* **15**, 1503-1514.
Schwartz MK, McKelvey KS (2008) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.
Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution* **22**, 25-33.
Servheen C, Herrero S, Peyton B (1999) Bear - Status Survey and Conservation Action Plan. IUCN, Cambridge.
Støen O-G, Bellemain E, Sæbø S, Swenson JE (2005) Kin-related spatial structure in brown bears Ursus arctos. *Behavioral Ecology and Sociobiology* **59**, 191-197.
Støen OG, Zedrosser A, Saebo S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears Ursus arctos. *Oecologia* **148**, 356-364.
Swenson JE, Sandegren F, Bjärvall A*, et al.* (1994) Size, trend, distribution and Conservation of the brown bear Ursus arctos population in Sweden. *Biological Conservation* **70**, 9-17.
Swenson JE, Wabakken P, Sandegren F*, et al.* (1995) The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology* **1**, 11-25.
Tallmon, DA, Bellemain E, Swenson JE, Taberlet P (2004) Genetic monitoring of Scandianvian brown bear: effective population size and immigration. *Journal of Wildlife Management* **68,** 960-965.
Tammeleht E, Remm J, Korsten M, *et al.* (2010) Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Mol Ecol* **19**, 5359–5370.
Tucker JM, Schwartz MK, Truex RL, Wisely SM, Allendorf FW (2014) Sampling affects the detection of genetic subdivision and conservation implications for fisher in the Sierra Nevada. *Conservation Genetics* **15**, 123-136.
Underwood N, Hamback P, Inouye BD (2005) Large-scale questions and small-scale data: empirical and theoretical methods for scaling up in ecology. *Oecologia* **145**, 177-178.

Valdmann H, Saarma U, Karis A (2001) The brown bear population in Estonia: current status and requirements for management. *Ursus* **12**, 31-36.
Valtonen M, Palo JU, Aspi J*, et al.* (2014) Causes and consequences of fine-scale population structure in a critically endangered freshwater seal. *Bmc Ecology* **14**.
van Strien MJ, Holderegger R, Van Heck HJ (2015) Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity* **114**, 27-37.
Vangestel C, Callens T, Vandomme V, Lens L (2013) Sex-Biased Dispersal at Different Geographical Scales in a Cooperative Breeder from Fragmented Rainforest. *Plos One* **8**.
Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**, 921-935.
Waits L, Taberlet P, Swenson JE, Sandegren F, Franzen R (2000) Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (Ursus arctos). *Mol Ecol* **9**, 421-431.
Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* **15**, 1419-1439.
Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nature Genetics* **36**, 1122-1125.
Xenikoudakis G, Ersmark E, Tison JL, Waits L, Kindberg J, Swenson JE, Dalén L (in press) Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear. Molecular Ecology. doi: 10.1111/mec.13239
Zedrosser A, Støen O-G, Sæbø S, Swenson JE (2007) Should I stay or should I go? Natal dispersal in the brown bear. *Animal Behaviour* **74**, 369-376.

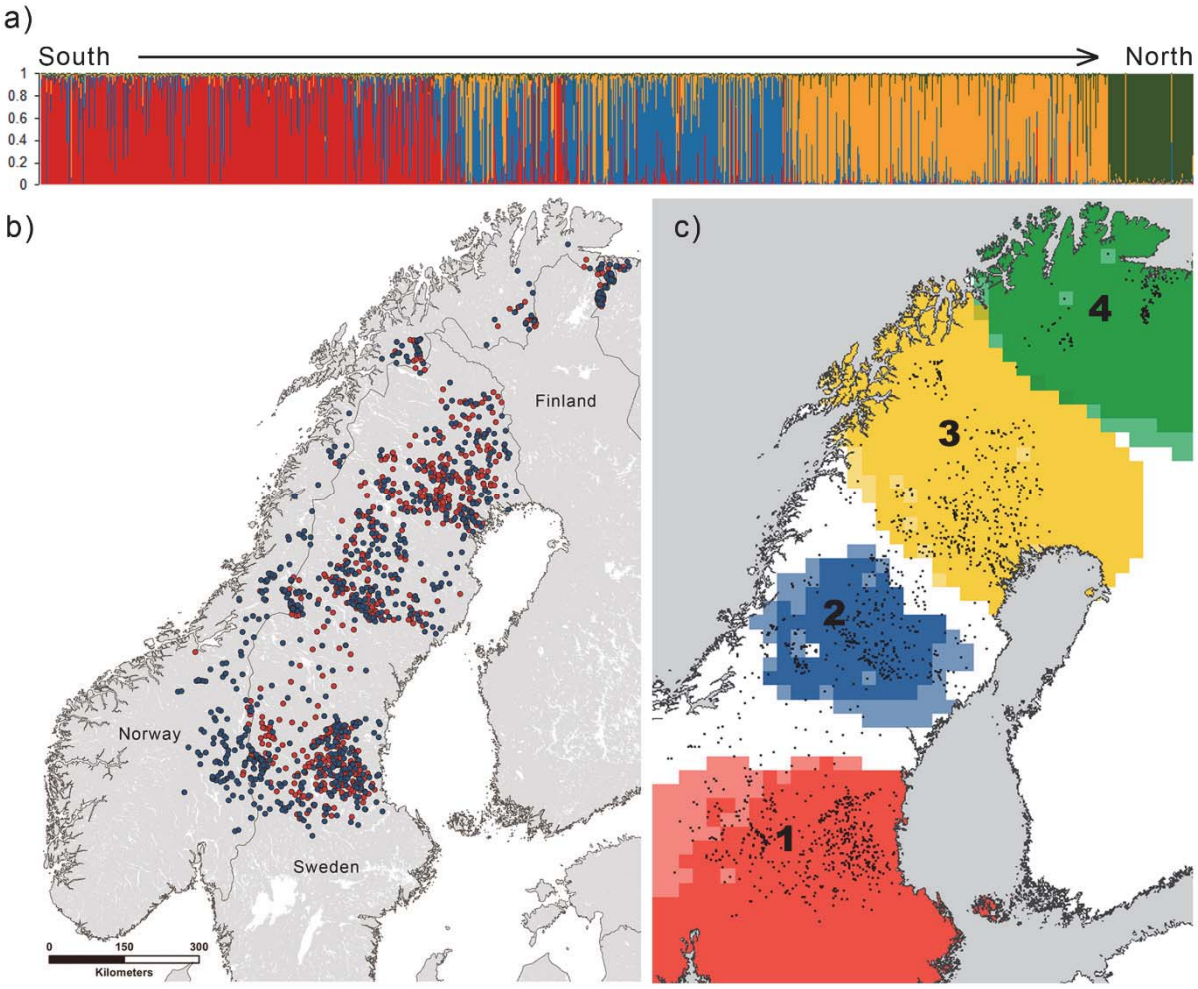


Figure 1: Population genetic structure of the Norwegian/Swedish brown bear population (n=1531). a) result of the Bayesian cluster assignment with the program Structure (Pritchard *et al.* 2000), each individual is represented by one bar, sorted by sampling location from South to North; each bar is divided into sections corresponding to the assignment probability (*q*, y-axis) for each of the four clusters. b) Overview over the sampled individuals, males are depicted blue, females red. c) Geographical placement of the statistically significant cluster core areas based on inverse distance weighted interpolation of assignment probability *q* determined by 100 bootstrap permutations. Colors correspond to the bar plot, the darker gridpoints (25x25km) were determined as core areas for males and females, the lighter ones for males only. The clusters are referred to in the text by the numbers given in the figure.

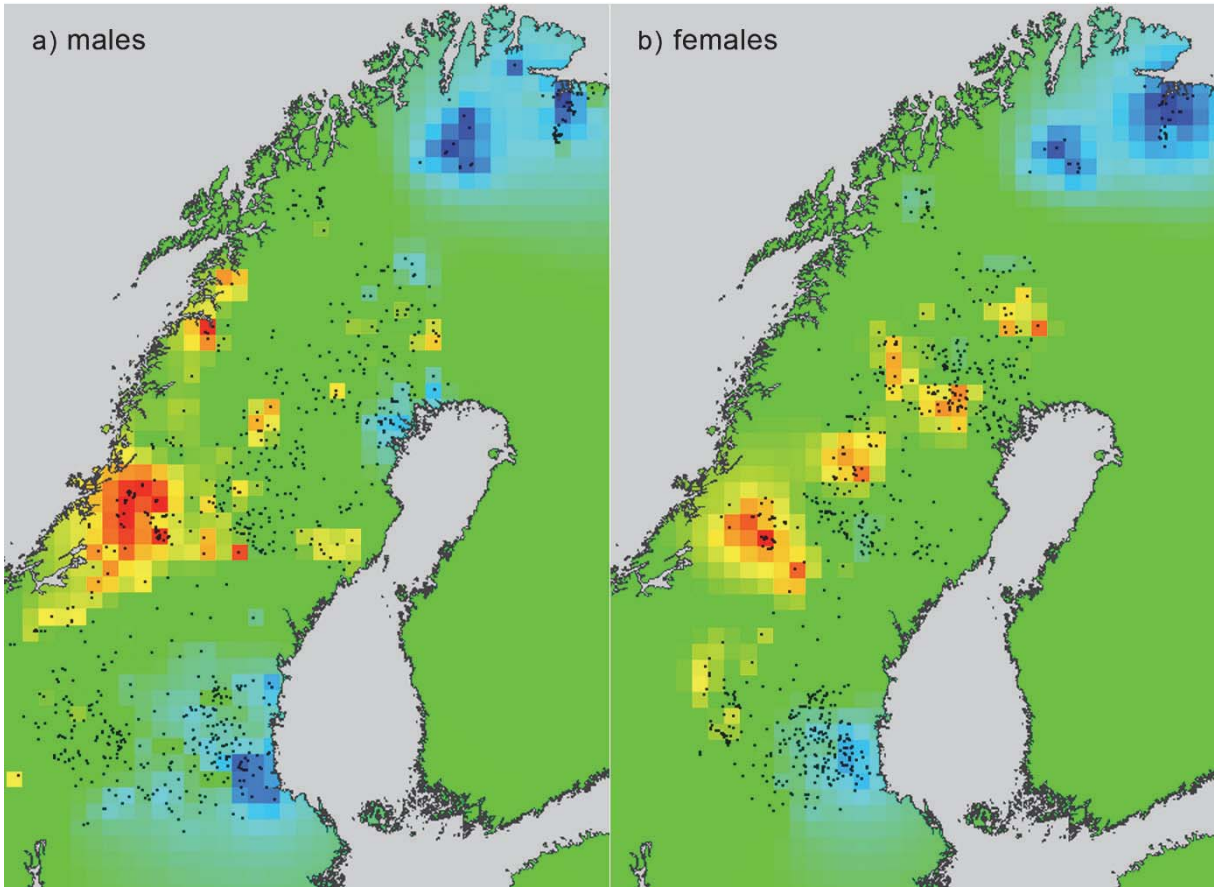


Figure 2: Individual genetic diversity of male and female brown bears of Norway and Sweden based on the membership value *q* for each of the four genetic clusters, estimated with the Structure analysis (Pritchard et al. 2000), and interpolated across the study area using inverse distance weighting. Grid points are color coded ranging from red = high individual genetic diversity, i.e. admixed individuals, to blue = low individual genetic diversity, i.e. most likely large family groups and/or little or no recent gene flow from outside the respective area. Samples are represented as black dots. a) male bears with n=789; b) female bears with n=742.

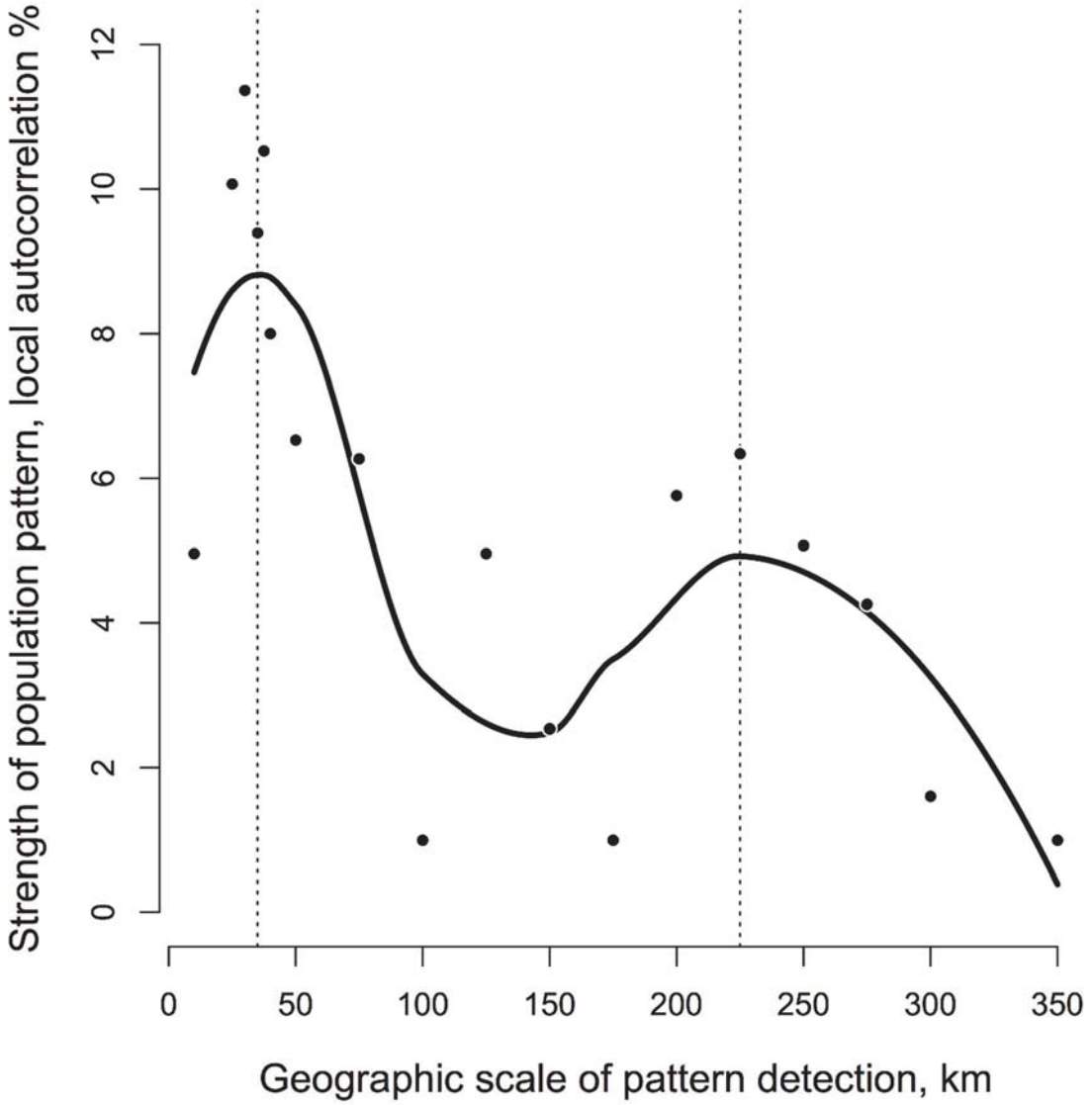


Figure 3: Strength of individual pairs based population spatial structure over geographic scales from 10 to 350 km. The analysis is based on series of variogram models, composed on values of pairwise deviations from IBD model and subsampled by geographic distance of pair (± ⅓ distance). The points represent strength of autocorrelation, measured as percentage of variogram partial sill. The curved line represents the LOESS model, used to define the scales of structuring peak values. The dashed vertical lines indicate the two scales of 35 and 225 km at X-axis, chosen for the final DResD analysis of barrier and corridor placement.

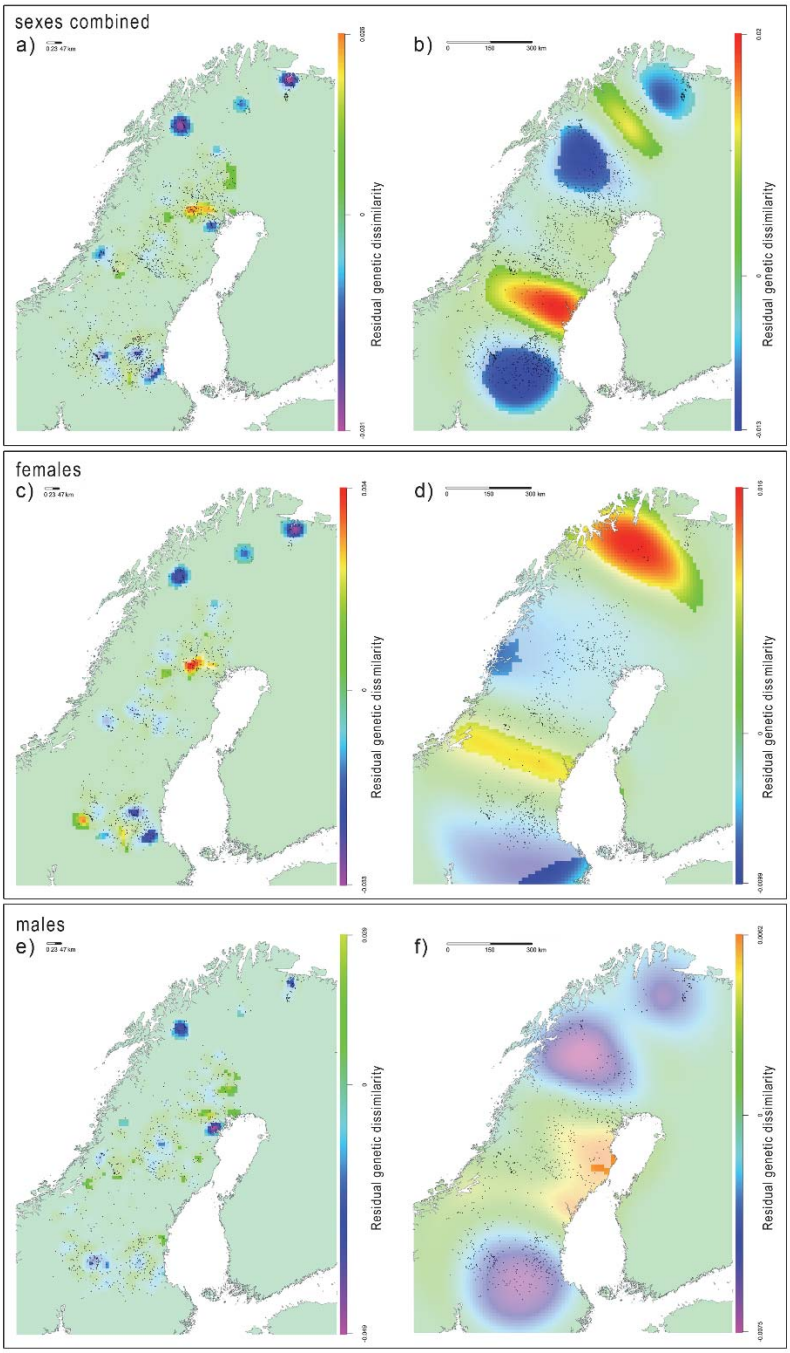


Figure 4: Areas of high and low genetic differentiation between sample pairs of brown bears in Norway and Sweden, estimated with the spatially explicit DResD method at two different scales, corrected for isolation-by-distance and interpolated across the study area using kriging procedure. The black and white scale bar indicates the distance ranges of sample pairs included in the analyses. Saturated colored grid points indicate significant departure from IBD expectations, magenta to blue indicating lower, and green to red higher than expected genetic dissimilarity; samples are represented as black dots. a) Local scale, males and females combined, including sample pairs between 23 to 47 km apart (n=26207); b) Population scale, males and females combined, including sample pairs between 150 to 300 km apart (n=243373); c) Local scale, only females (n=7913); d) Population scale, only females (n=53809); e) Local scale, only males (n=5907); b) Population scale, only males (n=65536).

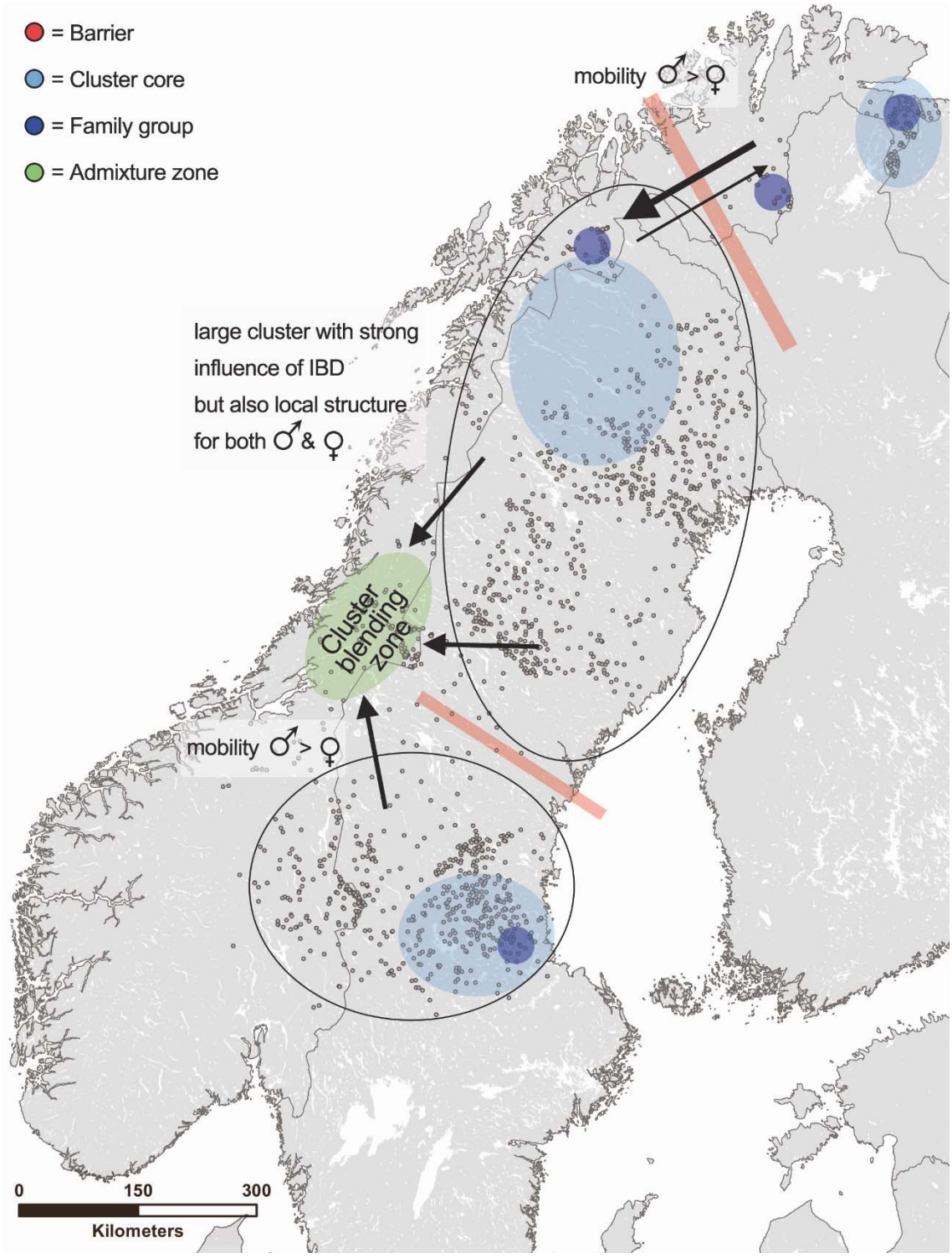


Figure 5: Summary of cluster core placement analysis, DResD analysis and analysis of individual diversity. Only the most pronounced results are depicted in this overview, for more detailed results refer to Figures 1, 3 and 4. Colors used to depict barriers, cluster cores, family groups and admixture zones are given in the legend.

Appendix S1

Table S1: Descriptive statistics of the DresD analysis.

		Global trend of variation between individuals			Local pattern of variance		n
		Direction low → high	R^2 trend	p	R^2 full model	Probability of p<0.05	
35 km	**Total**	SE → NW	0.013	0.005	0.039	0.017	26 207
	♀	SE → NW	0.016	0.010	0.059	0.018	7 913
	♂	SE → NW	0.030	0.030	0.067	0.010	5 907
225 km	**Total**	SW → NE	0.025	0.005	0.039	0.134	243 373
	♀	W → E	0.007	0.040	0.016	0.135	53 809
	♂	SW → NE	0.031	0.010	0.036	0.003	65 536

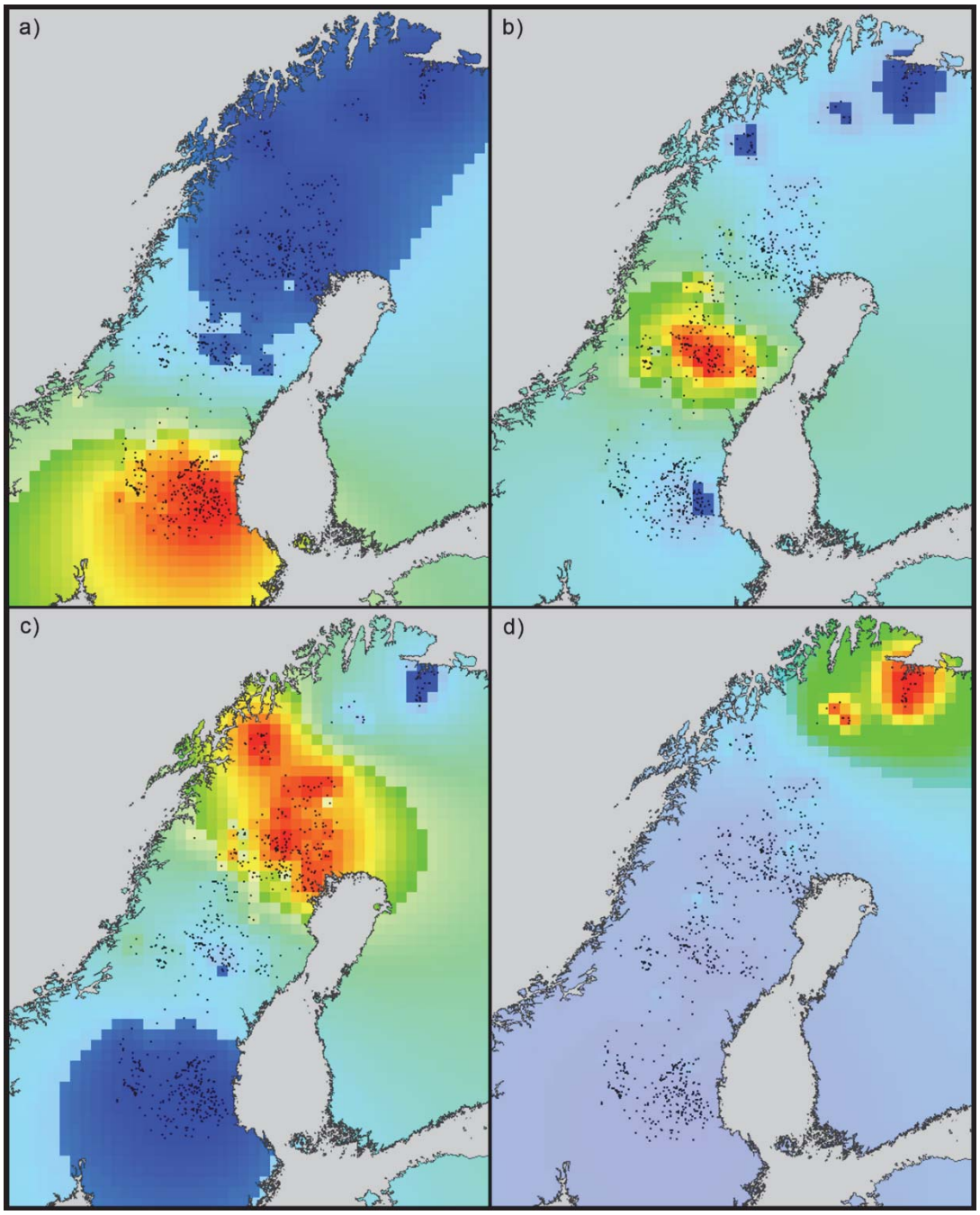


Figure S1: Geographic expansion of clusters of female bears as determined by the distance weighted interpolation of membership probabilities derived by the Bayesian clustering program Structure (Pritchard et al. 2000) across the study area. The saturated colored grids represent the cluster core (red to green) and areas significantly out of cluster range (blue), whereas the light colored grid points represent areas of random group probability. Samples are given as black dots. a) Expansion of cluster 1; b) Expansion of cluster 2; c) Expansion of cluster 3; d) Expansion of cluster 4.

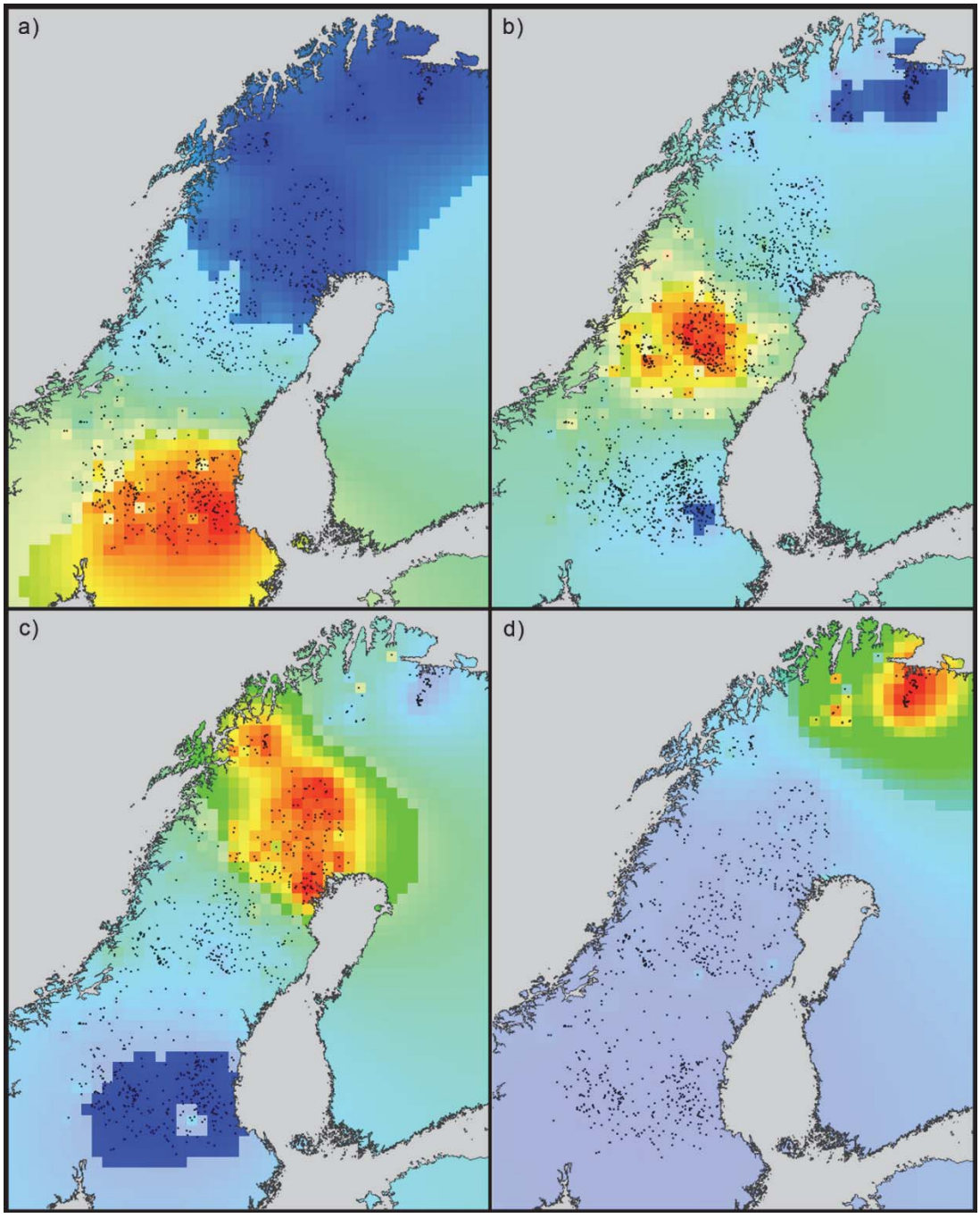


Figure S2: Geographic expansion of clusters of male bears as determined by the distance weighted interpolation of membership probabilities derived by the Bayesian clustering program Structure (Pritchard et al. 2000) across the study area. The saturated colored grids represent the cluster core (red to green) and areas significantly out of cluster range (blue), whereas the light colored grid points represent areas of random group probability. Samples are given as black dots. a) Expansion of cluster 1; b) Expansion of cluster 2; c) Expansion of cluster 3; d) Expansion of cluster 4.

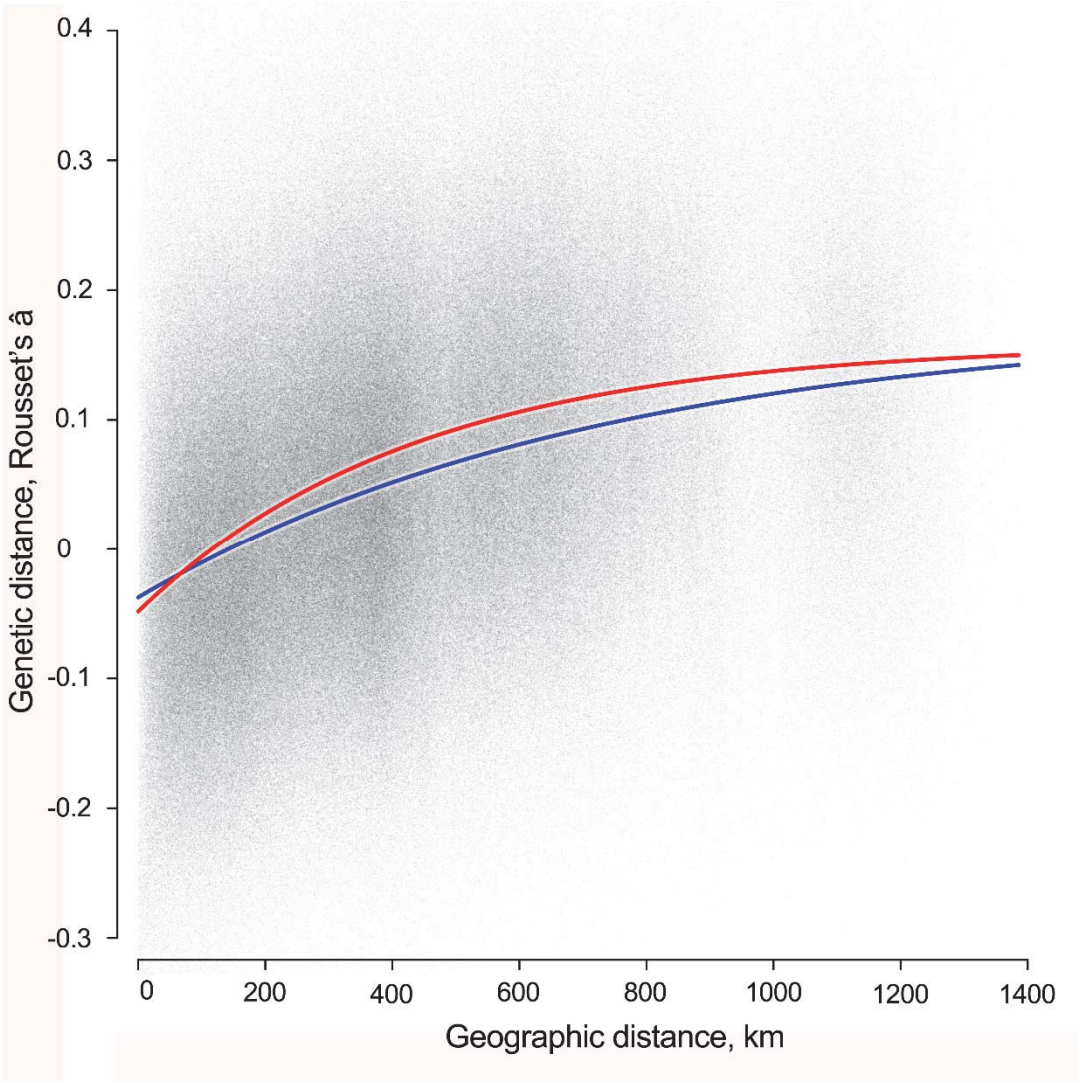


Figure S3: Pattern of isolation by distance (IBD), estimated separately for male (blue) and female (red) brown bears in Norway and Sweden. The shaded background shows density of sample pairs (n = 1,169,685), sexes were pooled.

Appendix S2

Script to calculate cluster placement and ancestral diversity based on Structure resulted Q-values with parameters values used in the analysis. The script was created using R 3.0.2. For further information and discussion, please contact Jaanus Remm (jaanus.remm@ut.ee)

```
# Section 0. Defintion of functions and constants

# Section 0.1. Constants

grid.step <- 25000                                  # Size of grid steps in sample coordinate units

weight.shift <- 10000                               # Correction for distance weighting in sample coordinate units. Used this distance value at 0-distance

n.resamp <- 100                                     # Number of resamplings of bootstrap to estimate statistical significance

# Section 0.2. Functions

Cluster.estimate <- function(grid.x, grid.y, sample.points, weight.shift){   # Function for inverse distance weighted average

    d <- sqrt((grid.x-sample.points[,1])**2+(grid.y-sample.points[,2])**2)   # Euclidean distance by Pythagorean theorem

    w <- 1/(d+weight.shift)**2                      # Inverse distance weight, corrected with weight.shift to avoid extremely large weight values at 0-distance

    return(weighted.mean(x=sample.points[,3], w=w, na.rm=T))   # Estimate value for a grid point

    }
```

 # Section 1. Input data

 # Section 1.1. Read and prepare data

```
full.data <- read.table("filepath", sep="\t", header=T)        # Read input data as tab-delimited table with header row and columns: 1-st ID, 2-nd sex (F, M), 3-rd X-coord, 4-th Y-coord, 5-th to last q-values of clusters

coordinates  <- full.data[,3:4]                                # Generate an object of just the coordinates; this ensures that same grid is generated for all subsamples, e.g. males and females

n.clusters <- ncol(full.data)-4                                # Number of clusters
```

 # Section 1.2. Shannon-Weaver diversity index H'

```
ln.p <- log(full.data[,5:(n.clusters+4)])                      # Estimation of first part of the Shannon index

ln.p.p <- ln.p*full.data[,5:(n.clusters+4)]                    # Estimation of the second part of the Shannon index

shannon.H <- -rowSums(ln.p.p, na.rm=T)                         # Estimation of the Shannon H' for each individual

full.data <- cbind(full.data, shannon.H)                       # Add the Shannon H' values as last column of the data table
```

 # Section 1.3. Select sub-sample

 # Section 1.3.1. If divide sample to males and females, command one of the following rows. If studying full sample, overpass this section, go to 1.3.2

```
full.data <- full.data[full.data[,2]=="M",]                    # Select a subsample to study only males. Not mandatory!
```

```
full.data <- full.data[full.data[,2]=="F",]                    # Select a subsample to study only females. Not mandatory!
```

 # Section 1.3.2. Command one of the following rows to study: a) cluster placement, b) ancestral diversity

```
"a"; data <- full.data[,3:(n.clusters+4)]                      # Select subsample to study placement of Q-values of clusters

"b"; data <- full.data[,c(3,4,ncol(full.data))]; n.clusters <- 1   # Select subsample to study placement of Shannon H' values
```

 # Section 2. Background grid

 # Section 2.1. Grid parameters

```
x.min <- min(coordinates[,1], na.rm=T)                         # Find westernmost sample coordinate

x.max <- max(coordinates[,1], na.rm=T)                         # Find easternmost sample coordinate

y.min <- min(coordinates[,2], na.rm=T)                         # Find southernmost sample coordinate

y.max <- max(coordinates[,2], na.rm=T)                         # Find northernmost sample coordinate

margin <- max(x.max-x.min, y.max-y.min)/10                     # Estimate margin for grid. 10% of sampling extent.
```

 # Section 2.2. Genrate grid

```
x.seq <- seq(from=x.min-margin,to=x.max+margin,by=grid.step)   # Make sequence of X-coordinate positions of grid steps

y.seq <- seq(from=y.min-margin,to=y.max+margin,by=grid.step)   # Make sequence of Y-coordinate positions of gridsteps

grid <- expand.grid(x=x.seq, y=y.seq)                          # Make grid as table of X and Y coordinates
```

```
# Section 3. Empirical estimates

for (j in 3:(n.clusters+2)){

    grid.estimate  <- mapply(FUN=Cluster.estimate, grid[,1], grid[,2],

    MoreArgs=list(sample.points=data[,c(1:2,j)], weight.shift=weight.shift))    # Calculate spatially interpolated estimates of q or H' for a cluster

    grid <- cbind(grid[,1:(j-1)], grid.estimate)                                # Combines the grid with the estimates of the q-values or H'

    }

names(grid)[3:(n.clusters+2)] <- names(data)[3:(n.clusters+2)]                  # Add column names in grid table, the same as data table


# Section 4. Statistical significance

# Section 4.1. Preparation for randomization test of significance

if(n.clusters>1) overall.means <- colMeans(data[,3:(n.clusters+2)], na.rm=T) else overall.means <-
                                                                                mean(data[,3],
                                                                                na.rm=T)   # Mean q-values or H' by clusters

grid.bts.values <- matrix(data=0, nrow=nrow(grid), ncol=n.clusters)             # Initiate empty grid table for bootstrap values

sample.var.slack <- rbind(sapply(X=data[3:(n.clusters+2)], FUN=sd), -sapply(X=data[3:(n.clusters+2)],
                                                                                FUN=sd))/2   # Slack that is used to eliminate meaningless statistical significance far from sample points, where biological significance is not reasonable


# Section 4.2. Bootstrapping

for(k in 1:n.resamp){
```

```
    random.row <- sample(x=1:nrow(data), size=nrow(data), replace=T)        # Vector of row numbers for bootstrap resampling. Random selection of all the rows of the data

    shuffle.data <- data[random.row,]                                       # Bootstrap sample

    grid.something.matrix  <- matrix(nrow=nrow(grid), ncol=n.clusters)      # Initiate temporary table for bootstrap estimates

    for (j in 3:(n.clusters+2)){                                            # Make it for evry cluster

        grid.estimate  <- mapply(FUN=Cluster.estimate, grid[,1], grid[,2],

        MoreArgs=list(sample.points=shuffle.data[,c(1:2,j)], weight.shift=weight.shift))   # Calculate estimated q or H' from randomized sample set

        grid.estimate <- grid.estimate+sample(sample.var.slack[,j-2], 1)    # Add random slack to the bootstrap estimate

        grid.something.vector <- (overall.means[j-2]<grid.estimate)/n.resamp # Count if the bootstrapped value is larger than overal mean

        grid.something.matrix[,j-2] <- grid.something.vector                # Combines the temporary grid table with the grid.estimates

        }

    grid.bts.values  <- grid.bts.values + grid.something.matrix             # Update count of bootstrap estimates larger than empirical estimate

    print(k/n.resamp)                                                       # Count bootstrap repetitions

    }

grid.p.values <- 1-abs(0.5-grid.bts.values)*2                               # Convert position of empirical estimate in bootstrap distribution to conventional p-value of significance (H0: randomness)
```

 # Section 5. Pooling of results

```
grid <- cbind(grid[,1:(n.clusters+2)], grid.p.values)                                    # Add column of p-value to grid table

names(grid) <- c(names(grid[,1:(2+n.clusters)]), paste("p.",1:n.clusters, sep=""))   # Add names for columns of p-values

write.table(grid, "filepath", sep="\t")                                                  # Save the grid as tab-delimited TXT file
```

ISBN: 978-82-575-1309-2
ISSN: 1894-6402

NIBIO
NORWEGIAN INSTITUTE OF BIOECONOMY RESEARCH

Postboks 115
NO-1431 Ås, Norway
+47 40 60 41 00
www.nibio.no

Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no