Cryptic population structure in a large, mobile mammalian predator: the Scandinavian lynx

E. K. RUENESS,* P. E. JORDE,*† L. HELLBORG,‡ N. C. STENSETH,* H. ELLEGREN‡ and K. S. JAKOBSEN*

*Centre for Ecology and Evolutionary Synthesis, Department of Biology, University of Oslo, PO Box 1031 Blindern 0315 Oslo, Norway; †Institute of Marine Research, Flødevigen, Norway; ‡Department of Evolutionary Biology, Uppsala University, Uppsala, Sweden

Abstract

The Eurasian lynx (*Lynx lynx*) is an example of a species that has gone through a severe bottleneck, leading to near extinction in Scandinavia around 1930 — a pattern shared with several other large carnivorous mammals. Here we extend previous genetic analyses of northern European lynx, confirming that lynx from the Scandinavian Peninsula represent a distinct clade differing clearly from European conspecifics. Furthermore, and despite a recent bottleneck and subsequent range expansion, we detect marked genetic differentiation within Scandinavia. This differentiation is largely manifested as a north–south gradient, with a linear increase in the quantity $F_{ST}/(1 − F_{ST})$. Aided by computer simulations we find that this pattern is unlikely to have arisen by random genetic drift in the short time since lynx started to expand in the 1950s, suggesting that the spatial structure may predate the bottleneck. Individual-based analyses indicate that, instead of a continuous gradient, Scandinavian lynx may be structured into three more or less distinct groups, possibly corresponding to northern, central and southern subpopulations. The presence of such structuring was unknown previously and was unexpected from general considerations on the mobility of the species, historical data and the absence of geographical barriers. Our study demonstrates how molecular markers may be used to detect cryptic population structure, invisible using traditional methods.

Keywords: genetic structure, isolation-by-distance, *Lynx lynx*, microsatellites, migration, population history

Received 3 April 2003; revision received 10 July 2003; accepted 10 July 2003

Introduction

Local extinctions and subsequent recolonizations are recurrent events in many species, and during the last century populations of large predators have been particularly exposed to such processes. The Eurasian lynx (*Lynx lynx*) went extinct in most parts of western Europe during the 20th century (Kratochvíl 1968), a process in which hunting probably played an important role. Despite a strong decline, in particular, during the 1930s (cf. Fig. 1), the lynx in the Scandinavian Peninsula (Norway and Sweden) survived and is currently also found in regions where it was not

observed earlier (Olstad 1945; Curry-Lindahl 1968; Kratochvíl 1968; Myrberget 1970; Kvam 1997b; Liberg 1997). No reproduction of lynx was reported outside the central region of Scandinavia (cf. Fig. 2) for several decades, and it has been assumed that the present-day distribution results from range expansion from this area (Kvam 1997b). It is presumed that northward dispersal first took place, starting around 1950, exceeding the former northern boundary of the distribution area (Myrberget 1970; Heggberget & Myrberget 1980; Kvam 1997a). Re-establishment in southern Norway was not reported until the 1980s (Kvam 1997a). One important reason for the recent increase in lynx abundance is thought to be enhanced abundance of prey, primarily ungulates (Kvam 1997a, b; Kvam *et al.* 1997; Liberg 1997). Lynx is a solitary living mammal and each adult animal exploits a

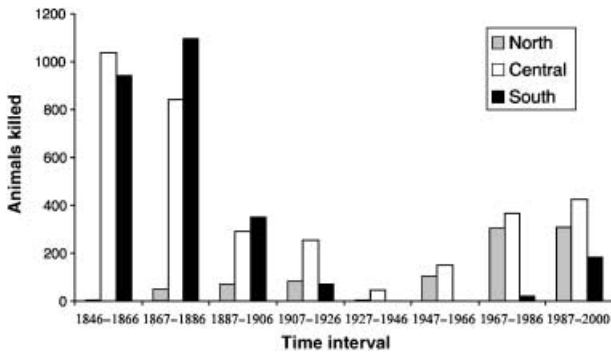Correspondence: K.S. Jakobsen. Fax: + 47 22 85 46 05; E-mail: k.s.jakobsen@bio.uio.no

Fig. 1 Hunting statistics for Norway 1846–2000, based on information available from Statistics Norway (http://www.ssb.no). Although lynx were protected in Sweden between 1928 and 1943 a bounty was kept in Norway until 1980. Thus, the hunting statistics are presumed to roughly reflect the distribution of lynx in Norway. Here, the animals are divided into three groups: North, Central, South — in accordance with the terminology throughout the paper (North = N1 + N2, Central = N3 + N4 and South = N5, see Fig. 2). Information about animals killed outside our sampling range is not shown.
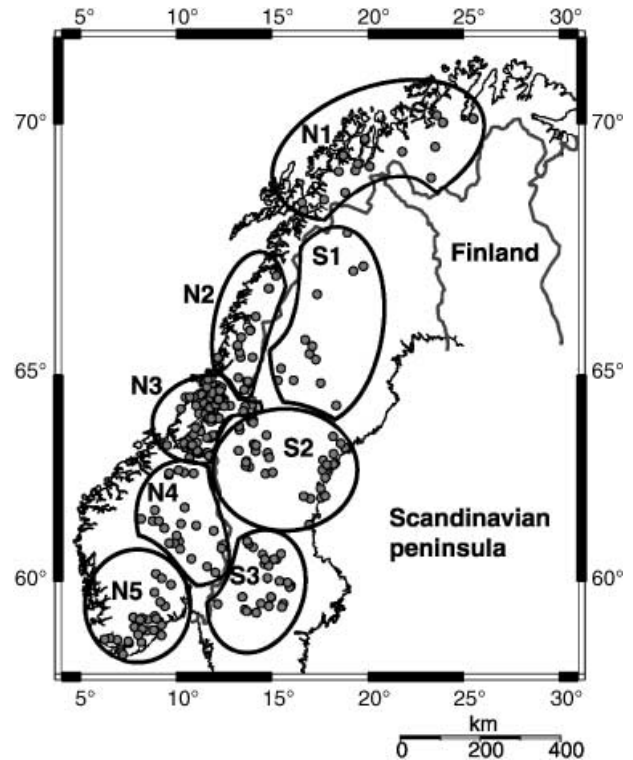

Fig. 2 Distribution of analysed Scandinavian lynx (dots) with sample groupings outlined (N, Norwegian samples; S, Swedish samples). Around 1930, when the abundance of lynx was lowest, reproducing populations were anticipated to be restricted to the areas covered by the sampling regions N3 and S2. The Scandinavian lynx was reported to occur only very sporadically north of the Arctic Circle (66.3° N) up to the 1950–60s (Olstad 1945; Curry-Lindahl 1968; Pulliainen 1968; Myrberget 1970; Kvam 1997b). The map was drawn using the online map creator at http://www.aquarius.geomar.de/omc/.

huge territory of several hundreds of square kilometres (Moa *et al*. 2001). The dispersal capacity is very high and radio-collared animals have moved > 450 km during one year (Andersen *et al*. 2002). In particular, juveniles often cover large areas until they find a home range of their own (Andersen *et al*. 2002). For such highly mobile species living in continuous habitats, the definition of subpopulations is intricate. Boundaries between 'subpopulations' for management purposes are traditionally set on the basis of spatial clustering and political borders are often used where there is a lack of knowledge of population substructure.

A recent study of northern European lynx using molecular markers [mitochondrial DNA (mtDNA) and microsatellites] revealed that Scandinavian lynx is distinct from populations in Finland and the Baltic Republics (Hellborg *et al*. 2002). Moreover, genetic differentiation within Scandinavia was suggested through comparisons among regions. These findings are interesting in view of the high dispersal capacity of lynx (i.e. making isolation less likely), and the presumed recent origin of lynx populations within Scandinavia.

The population genetic processes and subsequent demographic models describing population subdivision in lynx may have relevance for other mobile predators showing recent range expansions (see, e.g. Flagstad *et al*. 2003 for wolf and Waits *et al*. 2000 for brown bear). Because a large number of lynx samples in Scandinavia, and in particular from Norway, were available through years of legal hunting, lynx may be one of the best models for studying population differentiation in a rapidly expanding species. The aim of this study was to characterize the spatial structure of Scandinavian lynx in further detail, and thereby provide

more insight into the causes and mechanisms of spatial structuring in this and other large predators that have experienced a recent bottleneck. In particular, we aimed to evaluate the hypothesis of local extinction followed by recent recolonizations from a single, surviving Scandinavian core population. In order to investigate the spatial genetic structure of lynx we applied both frequency- and individual-based approaches. Furthermore, we compared the observed spatial genetic structuring with computer simulations depicting the genetic consequences of isolation-by-distance (which is the expected cause of differentiation assuming that the current lynx population originates from one source) both under equilibrium and nonequilibrium conditions (see Slatkin 1993 and Hutchison & Templeton 1999). Finally, we discuss the validity of the computer simulations and evaluate the possible influences of alternative demographic scenarios in creating and maintaining genetic divergence in a highly mobile species like lynx.

Materials and methods

Samples and DNA analyses

A total of 303 lynx from localities covering the majority of the lynx distribution range on the Scandinavian Peninsula (with the exception of the southernmost part of Sweden) were included in the analyses (Table 2, Fig. 2). This was partly material genotyped previously by Hellborg *et al.* (2002). To provide a more detailed analysis of population structure we included an additional 137 individual samples from Norway. These were muscle samples collected from animals that had been shot legally during the 1994–2000 hunting seasons. We also included 98 lynx from other European populations (see Table 2) to compare amounts of genetic variation and phylogenetic relationships. Samples from Finland (Fin) and the Baltic Republics (Estland and Latvia: Bal), were genotyped by Hellborg *et al.* (2002). Samples from the Czech Republic (Cze) consisting of muscle tissue (five specimens, collected in 2000) or claws (two specimens, from 1969 and 1996, respectively), were provided by local managers. Teeth from Russian (Rus) specimens (11), collected during the 1980s, were obtained from a museum collection. Tissue samples were stored in 96% ethanol at room temperature until analysis; claws and teeth were kept dry. DNA was extracted from tissue and claws following a standard phenol–chloroform protocol (Sambrook *et al.* 1989) and from teeth using the Isoquick DNA extraction kit (Orca Research Inc.) according to the procedure described in Flagstad *et al.* (2003).

Ten polymorphic microsatellite loci were scored using primers developed for the domestic cat (Menotti-Raymond & O'Brien 1995; Menotti-Raymond *et al.* 1999). All 10 loci are unlinked in the domestic cat (Menotti-Raymond *et al.* 1999). Polymerase chain reaction (PCR) conditions were as described in Hellborg *et al.* (2002). Genotyping was carried out in different laboratories and three different labelling techniques were applied to identify alleles: (i) fluorescent-labelled primers with PCR products analysed on an automated sequencer (ABI377); (ii) radioactive-labelled primers (γ P32) and auto-radiography (Sambrook *et al.* 1989); or (iii) unlabelled primers with products visualized using silver staining (Bassam *et al.* 1991). The techniques were used randomly with regard to the geographical origin of individuals and a few individuals were screened using all three techniques at each locus, which confirmed consistent scoring across methods.

Scandinavian lynx population structure was inferred from the spatial distribution of microsatellite genotypes, combining information from all 10 loci. In order to cope with the fairly continuous lynx distributions we applied two different statistical approaches. First, we characterized the amount and distribution of genetic variation within and among groups or 'samples' and performed the analyses on sample allele and genotype frequencies. These 'samples' were predefined according to the geographical proximity of individuals and borders between countries and counties (cf. Fig. 2). Next, we ignored any a priori notation of groupings within Scandinavia and rather based the analyses on genotypes of individuals.

Frequency-based analyses

Measures of genetic variability (gene diversity, H_S, and average number of alleles per locus) were made according to Nei & Chesser (1983) and Nei (1987: Eqn 7.39), and expected heterozygosity according to Nei (1987: Eqns 8.4 and 8.6). Deviations from Hardy–Weinberg equilibrium (HWE), expressed as F_{IS}, were calculated and tested for significance by an exact test (Guo & Thompson 1992) using GENEPOP 3.3 software (Raymond & Rousset 1995). Average genetic differentiation, F_{ST}, among samples was estimated according to Weir & Cockerham (1984), and an exact test of genotypic differentiation was performed in GENEPOP 3.3 (Raymond & Rousset 1995). Pair-wise F_{ST} values were estimated after Reynolds *et al.* (1983) and Slatkin (1995), and tested for significant deviations from the null hypothesis of no differentiation between samples by permutations (10 000) of individuals among samples using ARLEQUIN 2.000 (Schneider *et al.* 1999). F_{ST} was chosen over R_{ST} because drift can be assumed to be the main cause of allelic divergence considering the small population sizes and the short time perspective (Balloux & Lugon-Moulin 2002).

Phylogenetic inference among European populations was based on Nei's genetic distance measure Dm (Nei *et al.* 1983), which was calculated between each pair of populations using the population-to-population calculator found at http://www.biology.ualberta.ca/jbrzusto/GeneDist.php. A neighbour-joining tree (NJ) was constructed from the Dm distance matrix and bootstrapped over loci using PHYLIP 3.5c (Felsenstein 1993), and visualized using TREEVIEW (Page 1996) available at http://taxonomy.zoology.gla.ac.uk/rod/treeview.html.

The linear regression of $F_{ST}/(1 - F_{ST})$ against the geographical distance between pairs of Scandinavian samples was calculated, as suggested by Rousset (1997), and tested for significance by a Mantel test (Mantel 1967), using 10 000 permutations in GENEPOP 3.3 (Raymond & Rousset 1995). The 95% confidence interval (CI) for the regression coefficient was calculated on the basis of the *t*-distribution (Sokal & Rohlf 1981; box 14.3.3). To investigate whether the observed linear relationship could have arisen by random genetic drift and restricted gene flow over the time since the population bottleneck in the 1930s, we performed a number of computer simulations. These simulations were directed towards testing the hypothesis that present-day lynx in Scandinavia represent a recent expansion from a single source (from the 1950s and onwards). Noting that the Scandinavian Peninsula is an elongated habitat and that genetic differentiation is observed mainly in the north–south

direction, we chose a one-dimensional stepping-stone model (cf. Rousset 1997, p. 1223) as a practical, but simplified model of the lynx population in Scandinavia (see Discussion). In this model, geographical, structure is build up so that, at equilibrium between drift and migration, the slope of $F_{ST}/(1 - F_{ST})$ against distance equals $b = 1/(4Nm)$, where Nm (the product between effective population size and migration rate) is the effective number of migrants exchanged among neighbouring populations. Rousset (1997) gives the expression for the slope in the corresponding spatially continuous situation as $1/(4D\sigma^2)$, were D is the population density and σ^2 is the variance of the migration distance. Translated to the spatially discrete situation, $D\sigma^2 = Nm\varepsilon$, where ε is the distance between neighbouring subpopulations, or 1 in the stepping stone case (Rousset 1997, p. 1224)

Simulations were carried out by constructing a linear lattice of 100 subpopulations, each consisting of N individuals ($2N$ genes) that are of either of two types (alleles). Initially, the frequency of both alleles was set to 0.5 for all subpopulations, so that $F_{ST} = 0$ at all distances. Each generation a number, Nm, of individuals ($2Nm$ genes) was collected from each subpopulation, chosen such that the quantity $1/(4Nm)$ was equal to the observed regression of $F_{ST}/(1 - F_{ST})$ on the average distance among samples (283 km), or 0.043 (see below). In this case $Nm = 1/(4 * 0.043) = 5.81$. As this is not a whole number we rounded off to nearest even number (6) and subtracted the difference (0.19) from the next generation's migrants, and so on. Nm thus refers to the average number of migrants per generation. Emigrants are transferred to nearest neighbours only, in accordance with the stepping-stone model (Kimura & Weiss 1964), so that $Nm/2$ individuals (usually 3) are transferred to the neighbour in the north and in the south. After reproduction all individuals die and are immediately replaced by their offspring (i.e. the generation interval is discrete). Offspring are generated by random drawing, with replacement, of $2N$ genes from the $2N$ genes present in the subpopulation in the preceding (parental) generation. Census is taken each generation immediately after death/birth, and used to calculate F_{ST} (Weir & Cockerham 1984). F_{ST} was calculated and averaged over all populations belonging to the same distance classes, located one (i.e. nearest neighbour), two, three and four steps from each other (four steps corresponds to a real distance of $4 * 283 = 1132$ km: close to the maximum distance among Scandinavian lynx samples). The cycle is repeated for several generations and we ran different simulations with different subpopulation sizes, $N = 12, 25, 50, 100$ and 200, while keeping Nm constant (5.81). For the number of generations we used a value of $t = 25$, corresponding to the maximum number of generations (assuming a generation interval of 2 years) since the 1950s when the lynx population started to expand. We also ran a number of simulations for considerably longer ($t = 250$) in order to compare observed and simulated genetic differ-

entiation patterns to those expected at or near equilibrium. All simulations were repeated 1000 times and we present the results as averages over replicate runs, i.e. over a total of $(100 - d) * 1000$ replicates at each distance step ($d = 1, 2, 3$ and 4). Numerical evaluations verify that this level of replication is sufficient to give highly accurate simulation results (the standard errors of the mean simulated F_{ST} values are of the order of 10^{-5} or less.)

Individual-based analyses

Genetic mixture analysis (GMA), as implemented in the software STRUCTURE (Pritchard *et al*. 2000), is a parametric cluster analysis to infer population structure from multilocus genotypes. We applied GMA to investigate the degree of substructuring within Scandinavia. The program calculates the log likelihood, $Pr(X \mid K)$, i.e. the probability of the data given a certain number of subpopulations (K). Simultaneously, it calculates population allele frequencies, attempting to maximize HWE within groups (assuming all loci are unlinked). Various values of K were tried, and several runs were performed ['burn-in period' 100 000, number of Markov Chain Monte Carlo (MCMC) replicates 100 000] for each value of K to check if the results were consistent. We also estimated immigration into Scandinavia from the other European samples by treating Scandinavia as a single predefined population before the program was run. The migration rate was set to 0.05 (values from 0.001 to 0.1 recommended), and the number of generations backwards to 2 (i.e. to grandparents).

In order to further confirm the subpopulations suggested by GMA, we performed a more traditional, maximum likelihood-based assignment test (Paetkau *et al*. 1995) on Scandinavian individuals, using WHICHRUN 3.2 software (Banks & Eichert 2000).

Results

The amount of genetic variability varied substantially among loci, both with respect to gene diversity (H_S) and number of alleles (Table 1). The number of alleles ranged from 2 (*Fca045*) to 11 (*F115*). For most loci the number of alleles was lower in Scandinavia than in the total European sample. The average gene diversity over Scandinavian samples varied between 0.057 (*Fca031*) and 0.768 (*F115*), with most loci being close to the mean of 0.5. All 10 loci displayed statistically significant differentiation within Scandinavia, with F_{ST} values falling within a fairly narrow range of 0.015 (*Fca008*) to 0.076 (*Fca391*).

### Genetic diversity within and between samples

The average amount of genetic variability in the samples was estimated as the number of alleles per locus and as

Table 1 Summary statistics for microsatellite variability in Scandinavian lynx. Number of alleles (total European sample in parentheses), H_S = gene diversity and F_{ST} = genetic differentiation

Locus	No. of alleles	Gene diversity H_S	F_{ST}
Fca043	3 (6)	0.4563	0.0182
Fca045	2 (5)	0.4557	0.0614
Fca149	3 (3)	0.2908	0.0200
Fca001	5 (11)	0.5318	0.0613
Fca008	3 (7)	0.4220	0.0155
Fca391	4 (4)	0.6134	0.0764
Fca506	6 (7)	0.6300	0.0722
Fca559	6 (6)	0.7342	0.0228
F115	11 (15)	0.7685	0.0555
Fca031	2 (8)	0.0574	0.0238
Average	4.1 (7.3)	0.4960	0.0480

expected heterozygosity, H (Table 2). Both measures were lower in each of the Scandinavian samples than in any of the others, except for the very small sample from the Czech Republic ($n = 7$; Table 2). The sample from Russia (believed to come from a long-standing large population, e.g. Heptner & Sludskii 1972) had the highest gene diversity (0.699) of all, and also showed six unique alleles, despite its small size ($n = 11$).

The differentiation among European samples is illustrated in Fig. 3 as a NJ tree based on Nei's genetic distance (Dm). Clearly, Scandinavian samples are separated from the other European populations and the distances within Scandinavia are relatively shorter. However, N1 and N5 are separated by strikingly longer branches compared to the other samples
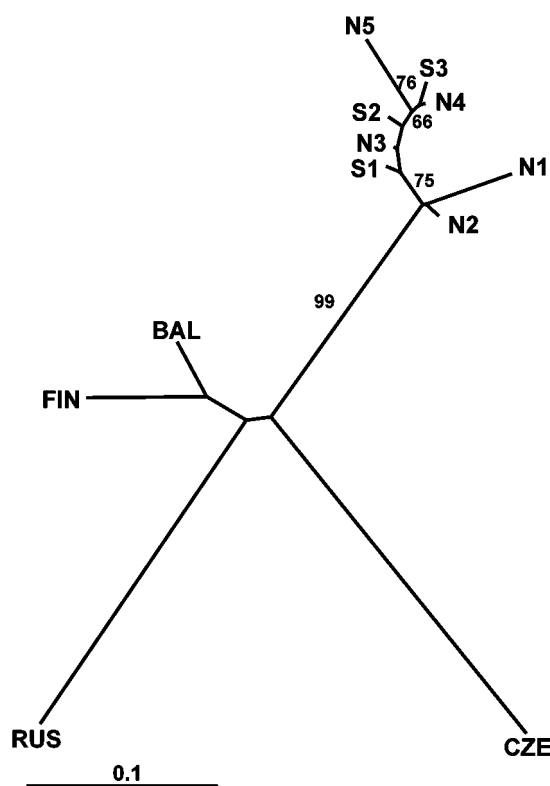
Fig. 3 Unrooted neighbour joining (NJ) tree, based on Nei's genetic distance measure Dm, between samples. Bootstrap values > 50 are given next to the branches (100 replicates).

showing that lynx from the northernmost and southernmost part of the Scandinavian distribution range are the most divergent, not only compared with each other, but also relative to their neighbouring samples.

Table 2 Diversity indices of the samples analysed

Country	Sample name	n	Alleles per locus	SD	H	SD	F_{IS}
Norway	N1	29	3.60	(1.58)	0.521	(0.067)	0.103
	N2	15	3.80	(1.75)	0.538	(0.078)	−0.027
	N3	104	4.10	(2.23)	0.511	(0.073)	0.072**
	N4	25	2.90	(1.45)	0.474	(0.076)	−0.023
	N5	37	3.10	(1.37)	0.427	(0.056)	−0.047
	Total	210	4.40	(2.46)	0.517	(0.073)	0.081***
Sweden	S1	24	3.70	(1.64)	0.527	(0.059)	−0.031
	S2	37	3.60	(1.71)	0.505	(0.069)	0.066
	S3	32	3.40	(1.43)	0.465	(0.075)	0.026
	Total	93	4.10	(1.85)	0.506	(0.068)	0.045
Finland	Fin	48	5.40	(2.32)	0.633	(0.057)	0.010
Balticum	Bal	32	4.70	(2.34)	0.612	(0.075)	−0.007
Czech Republic	Cze	7	2.80	(1.14)	0.510	(0.064)	−0.162
Russia	Rus	11	5.00	(1.70)	0.699	(0.037)	0.059
Grand total		401	7.30	(3.53)	0.606	(0.056)	0.153***

n, Number of individuals per sample. H, Expected heterozygosity. SD, Standard deviation. F_{IS}, Deviations from Hardy–Weinberg proportions. **Significant at 0.01, ***significant at 0.001.

Table 3 Pair-wise F_{ST} values among eight samples of Scandinavian lynx (below the diagonal)

| | | Norway | | | | | Sweden | | |
	n	N1	N2	N3	N4	N5	S1	S2	S3
N1	29		0.010	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
N2	15	0.0266		0.026	< 0.001	< 0.001	0.049	< 0.001	< 0.001
N3	104	0.0615	0.0167		0.019	< 0.001	0.071	0.007	< 0.001
N4	25	0.0922	0.0607	0.0121		< 0.001	0.069	0.203	0.959
N5	37	0.1027	0.1087	0.0814	0.0438		< 0.001	< 0.001	< 0.001
S1	24	0.0501	0.0145	0.0077	0.0107	0.0640		0.015	< 0.001
S2	37	0.0979	0.0481	0.0119	0.0055	0.0820	0.0185		0.012
S3	32	0.1129	0.0812	0.0283	−0.0103	0.0566	0.0309	0.0173	

n, Number of individuals. F_{ST} values were tested against the null hypothesis of no genetic structuring between populations (10 000 permutations). P-values given above the diagonal.

Within Scandinavia both the average number of alleles and the gene diversity tended to be slightly higher in the northern than the southern samples (despite lower sample sizes in the north), a trend paralleled in Norway and Sweden (cf. Table 2). This indication of population differentiation within Scandinavia is strengthened by the positive F_{IS} values (i.e. heterozygote deficiencies), most likely representing a Wahlund effect. Significant deficiencies were found in the combined sample from Norway (six loci) and within Norway also in the sample N3 (two loci) indicating substructuring within a restricted area. In Sweden significant deficiencies were observed at 2 loci, but the result of the global test over all 10 loci was not significant ($P = 0.24$).

All Norwegian samples are significantly differentiated from each other (at the 5% level): average pair-wise F_{ST} values (Table 3) range from 0.012 (between N3 and N4) to 0.109 (between N2 and N5). Similarly, all three Swedish samples are significantly differentiated and pair-wise F_{ST} values range from 0.018 (between S2 and S3) to 0.031 (between S1 and S3). In other words, there is quite strong genetic differentiation along the north–south axis within Scandinavia. There is apparently less east–west differentiation as indicated by the lower F_{ST} values between the Norwegian and Swedish samples (Table 3). Nonsignificant values were observed between N3 and S1 (0.008). Moreover, N4 was not significantly differentiated from any of the Swedish samples. However, the southern Norwegian sample N5 is highly differentiated from the southern Swedish one S3 ($F_{ST} = 0.057$; $P < 0.001$) suggesting a more complex spatial genetic structure within Scandinavian lynx than a simple north–south gradient.

Isolation-by-distance

Figure 4 shows the regression of $F_{ST}/(1 − F_{ST})$ against geographical distance, clearly conforming to a linear, or almost linear, pattern. A Mantel test showed that the positive association between genetic differentiation and geographical distances separating samples ($b = 0.00015$/km) is highly signi-
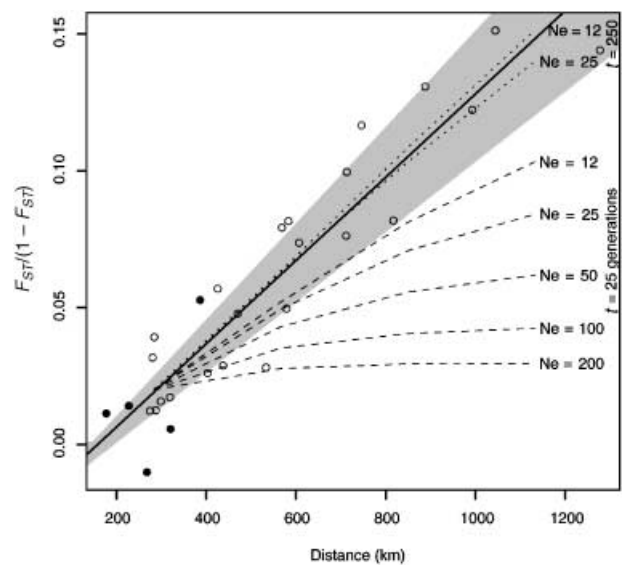
Fig. 4 Plot of $F_{ST}/(1 − F_{ST})$ among pairs of Scandinavian lynx samples against their geographical distance. (●) Pairs ordered in an east–west direction (i.e. samples N2–S1, N3–S2, N4–S2, N4–S3 and N5–S3). Solid line: regression of $F_{ST}/(1 − F_{ST})$ on distance (per km: $b = 0.00015$, or $b' = 0.043$ per average intersample distance of 283 km). Shaded area: 95% CI for b. Dotted lines: result of computer simulations of a one-dimensional stepping-stone model with various effective subpopulation sizes (N_e). The results are shown for generation $t = 25$ (dashed lines), corresponding to the maximum number of generations since the 1950s (assuming a generation interval of 2 years), and $t = 250$ (dotted lines), and represent averages over nearly 100 000 replicate F_{ST} values at each distance step. When preparing the plots, we multiplied the simulated distance classes with the average distance (283 km) among samples, which was used when estimating pair-wise F_{ST}. We also adjusted the absolute magnitude of the simulated $F_{ST}/(1 − F_{ST})$ to yield the observed value (0.02) for the shortest distance class (at 283 km). This was done because the simulated values were higher, and dependent on subpopulations size. As discussed by Rousset (1997), it is mainly the slope of $F_{ST}/(1 − F_{ST})$ that are of interest and we restrict the comparisons between simulations and observations accordingly.

Table 4 Likelihood of the Scandinavian lynx data (303 individuals), assuming different numbers of subpopulations. ln $PrX \mid K$ = Ln likelihood of the data for different values of K; mean after five runs (burn-in 100 000, MCMC 100 000) SD = standard deviation. The probability of $K = 3$, given these estimates, was calculated to be 1.00, following Bayes' Rule

K	ln $PrX \mid K$ (SD)	PrK
1	−5697 (0.07)	0.00
2	−5579 (3.29)	0.00
3	−5506 (1.45)	1.00
4	−5533 (4.92)	0.00
5	−5554 (18.65)	0.00

Fig. 5 Frequency of individuals from the different Scandinavian samples that was assigned to the three subpopulations N, C and S with a probability > 0.80 by the GMA (burn-in 100 000, MCMC 100 000). The samples are distributed on a north–south gradient (cf. Fig. 2) along the *x*-axis.

ficant ($P < 0.0001$). Results from the computer simulations, using different effective subpopulation sizes (N_e) and time (*t*) in generations are also plotted (Fig. 4). Using realistic separation times (i.e. assuming $t = 25$ lynx generations since the 1950s) did not yield a sufficiently steep slope of $F_{ST}/(1 − F_{ST})$ against distance for any population size (simulated $N_e = 12–200$). The simulated slopes range from 0.00001 / km for $N_e = 200 − 0.00010$ / km for $N_e = 12$ and none of them overlap the 95% CI for the observed slope (from 0.000129 to 0.000176: shaded region in Fig. 4). However, there is a good fit to the observed slope for small populations when the separation time is much longer (here $t = 250$). This latter concordance between theoretical expectations and simulation results verifies that the computer simulations performed as expected and that the appropriate number of migrants (*Nm*) was used. Hence, although the linear relationships between genetic differentiation and distance within Scandinavian lynx superficially look like the results of an 'isolation-by-distance' mechanism (cf. Wright 1943; Kimura & Weiss 1964; Rousset 1997) the computer simulations do not lend support to the notion that the observed pattern could have developed by this mechanism alone in the time available since the lynx started expanding in the 1950s. Instead, the simulations demonstrate that it would take considerable time to build up a slope as steep as that observed, even with small population sizes, and that additional factors are likely to have played a significant role in shaping the present genetic structure of Scandinavian lynx.

Distribution of individuals based on multilocus genotypes

In order to characterize substructuring in Scandinavian lynx without using information about geographical origin, all 303 individuals were included in GMA calculating posterior probabilities [$n Pr (X \mid K)$] for different assumed number of subpopulations (*K*s). The likelihood increased up to $K = 3$ and then decreased for higher *K*s (being significantly higher for $K = 3$; Table 4), showing that our data are best explained by a division on three groups, possibly corresponding
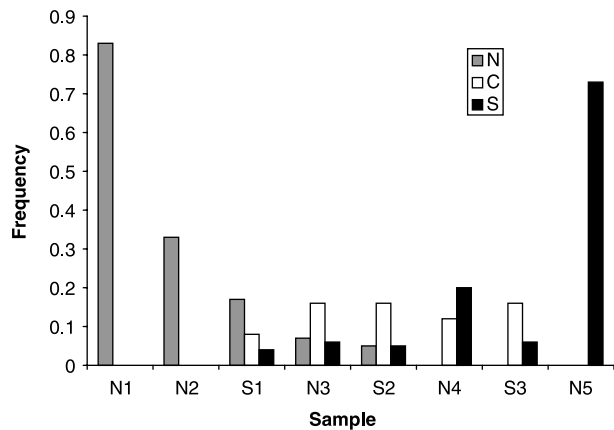
to three subpopulations of lynx within Scandinavia. We call these three inferred subpopulations north (N), central (C) and south (S). Individuals being assigned to N were abundant in the north and diminished gradually southwards, whereas individuals that assigned to S followed an opposite gradient (cf. Figure 5). Interestingly, no individuals from N1, N2 or N5 were assigned to the central subpopulation. This reflects considerable genetic differentiation of the northernmost and southernmost samples. The samples from Sweden (S1, S2, S3) and Central Norway (N3, N4), which included the majority of the individuals, were more heterogeneous and many of the individuals did not assign to any of the clusters with a high probability.

Using the hypothetical subpopulations N, C and S, as inferred above, we performed a standard assignment test. The geographical sampling localities of individuals that were assigned to any of the subpopulations are displayed on the map in Fig. 6. This illustrates how lynx, based on genotypic similarity, divides into three groups with core areas in northern, central and southern Scandinavia, respectively. However, the three groups are not clearly demarcated, but rather blend gradually into each other, as indicated by the low assignment for most individuals in the intervening areas (cf. the open circles in Fig. 6).

Testing for migrants between Scandinavia and the eastern samples yielded no cross-assigned individuals, but three Scandinavian individuals were detected as possible first-generation migrants, of these one was from S3, one from N1 and one from N2.

Discussion

According to historical information, lynx existed only in the central region of Scandinavia for a period during the 20th
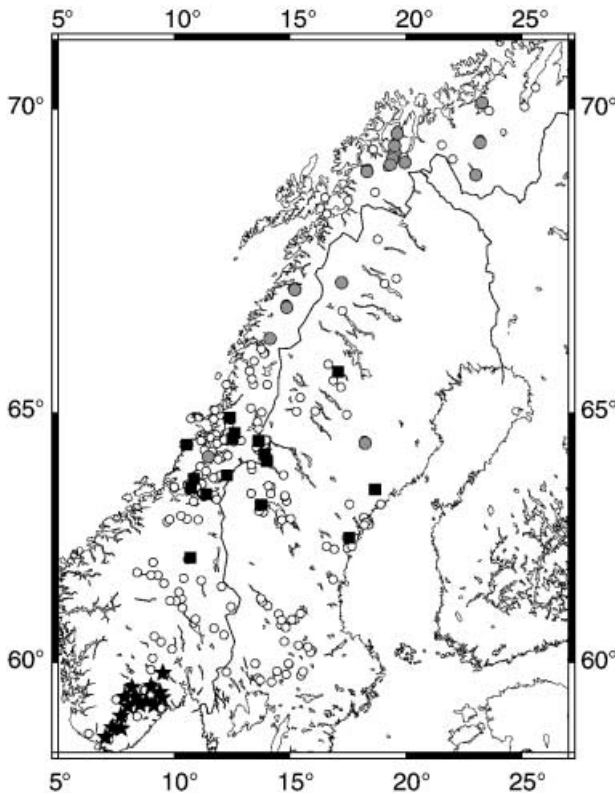
Fig. 6 The geographical distribution of individuals that assigned to the three inferred subpopulations N (circles), C (squares) or S (stars); individuals that did not assign to any subpopulation are shown as open circles. The criteria used for assignment of individuals was: (i) > 0.80 probability for being assigned to N, C or S by GMA; and (ii) a likelihood of at least 100/1 for being assigned to the same population by the assignment test. The fractions of individuals from the different samples that met these criteria were as follows, N: N1 (13/29), N2 (3/15) S1 (1/24), NT (1/104), S2 (1/37) C: N31 (12/104), N4 (2/25), S1 (1/24), S2 (3/37), S: N5 (15/37).

century (cf. Figs 1 and 2). Following this scenario, the current, widely distributed lynx population has originated from a recent radiation from this central region, and the spatial genetic differentiation observed has accumulated within the last few decades. The finding of genetic substructuring by Hellborg *et al.* (2002) suggests that the Scandinavian lynx might have had a more complex history and demography than hitherto thought. Our results revealed a marked genetic differentiation in the north–south direction; lynx from the northernmost and southernmost part of the distribution range are the most strongly differentiated and suggest the existence of (at least) three Scandinavian subpopulations. An important question is whether this observed differentiation has arisen after recolonization, thereby being consistent with the single-source hypothesis, or at earlier times. Apparently, the differentiation follows a pattern of isolation-by-distance (Fig. 4), but the computer simulations seem to reject this explanation as the sole cause of differentiation

because there is not sufficient time for building up the observed pattern. In the following we evaluate the validity of the computer simulations and discuss alternative and/ or additional demographic scenarios that might explain the cryptic population structure in Scandinavian lynx.

Isolation-by-distance, computer simulations and lynx biology

The simulations are based on the assumptions implicit in the null hypothesis as well as a number of auxiliary assumptions. Under H_O, which we wish to test, the population structure follows a stepping-stone model in one dimension with a finite number of equally sized, ideal populations with an equal number of migrants among nearest neighbours. The choice of a one-dimensional model is dictated by the narrowness of the Scandinavian Peninsula in the north–south direction, and the observation that differentiation in the orthogonal, east–west, direction (i.e. between Swedish and Norwegian samples at similar latitudes) is very modest (filled circles in Fig. 4).

The simulated stepping-stone model limits migration to nearest neighbours only (parameter m_1 in Kimura & Weiss 1964). By using very large patches (averaging 283 km across) restricting migration to nearest neighbours seems reasonable. Furthermore, it is known from mathematical analysis that additional long-distance movements (parameter m_∞ in Kimura & Weiss 1964) have similar effects as migration in the island model, namely, to reduce differentiation among populations, regardless of distance. By ignoring this effect in the computer simulations we allow a more rapid build-up of differentiation with distance, and thus create a situation favourable for H_O. We also, implicitly, assume that the observed slope, *b*, is representative for the actual rate of migration, and use the corresponding estimate of *Nm* in the simulations. If equilibrium between drift and migration has not yet been reached, we are underestimating *b* and, thus, probably using too large a value for *Nm* in the simulations. However, we observe a linear trend for $F_{ST}/(1 - F_{ST})$ with distance, and this is expected only after equilibrium has been reached. For nonequilibrium situations, as also depicted in the simulation plots (cf. Fig. 4 for short separation times, *t* = 25), linearity does not hold and $F_{ST}/(1 - F_{ST})$ instead flattens with distance. This is because equilibrium is reached first for short distances and gradually propagates outward (Slatkin 1993; Hardy & Vekemans 1999). Although minor deviations from linearity may pass undetected because of scatter in the data points, it is unlikely that our estimate of the slope is far off because there is no indication that the slope may be steeper at the shorter distances (Fig. 4). Hence, to the extent that a stepping-stone model is an appropriate description of H_O (and not necessarily the real situation), the simulations appear quite robust to a number of important assumptions.

Given that the geographical structure of the Scandinavian Peninsula largely limits us to consider one-dimensional

models (above), there is an immense restriction on alternative migration patterns or models. Clearly, when migration is limited to one-dimension the only free variables are the amount of migration and its distance distribution. The amount of migration is taken care of by using the estimated slope, which is directly related to migration both in continuous (isolation-by-distance) and discrete (stepping-stone) models (Rousset 1997). With regard to the distance distribution of migration Ibrahim *et al*. (1996) demonstrated through simulations that, in a model with leptokurtic dispersal, rare long-distance dispersers (colonizers) could found patches of similar genotypes during a range expansion. For the northern subpopulation this option can be ruled out because of the high level of genetic variability (Table 2) and the occurrence of private alleles (see below). The relatively lower diversity of the southern subpopulation could, however, fit such a model, but the fact that long-distance dispersal is relatively frequent in lynx makes it less likely that patches of deviating genotypes arise by colonization. Radio telemetry has showed that distances of several hundred kilometres (i.e. large parts of Scandinavia) might be covered within a year by a dispersing lynx (Andersen *et al*. 2002). Hence, the dispersal capacity represents no limitation to the recent colonization depicted above, and Scandinavia may be considered as a continuous (although patchy) habitat for the species. In view of this, the observed strong genetic differentiation of Scandinavian lynx is remarkable. Clearly, although the linear distribution of $F_{ST}/(1 - F_{ST})$ may suggest isolation-by-distance, the high dispersal ability of lynx casts doubt upon distance itself causing differentiation within this restricted geographical area. In this respect, common biological reasoning is in line with the results of the computer simulations in rejecting this hypothesis. Nevertheless, the strong differentiation remains, and there are obvious restrictions to gene flow within this species.

Historical discontinuity

For some Scandinavian mammals, for example, the brown bear (*Ursus arctos*), differentiation has been shown to be caused by different postglacial colonization routes (Taberlet *et al*. 1995), and in this species two divergent (7%) mtDNA clades have been found. Among 196 Scandinavian lynx sequenced by Hellborg *et al*. (2002) only one single mtDNA (D-loop) haplotype was found suggesting that present-day Scandinavian lynx most likely origin from one glacial refugium. The Scandinavian brown bear reached, as did the lynx, a low point in its abundance around 1930, but survived in four restricted regions in Sweden. Analysis of the microsatellite variation (19 loci) of the bear revealed four subpopulations presumably shaped by independent drift effects (Waits *et al*. 2000). Differentiation between subpopulations, expressed as F_{ST} values, ranged from 0.0151 to 0.1393, comparable with that reported here for lynx (Table 1).

In the southernmost part of Norway no lynx were killed in the years 1925–78 (Fig. 1), yet there were reports on the occurrence of stray animals in southern Norway in the 1940s (Olstad 1945), and reports of increased abundance in the 1960s and 1970s (Heggberget & Myrberget 1980). The sample from southern Norway, N5, showed a relatively high degree of differentiation from its neighbouring populations N4 and S3 (Table 3, Fig. 3). This suggests that territories, being favourable in this region, were most likely occupied before the observed increase in the lynx population (cf. Figure 1), hindering immigration (see below for discussion of the impact of territorial behaviour). This further implies that reproducing lynx probably inhabited this region even during periods when lynx were not registered.

We observed relatively high diversity in the northern Scandinavian samples N1, N2 and S1 (Table 1) even though considerably fewer individuals were analysed from these regions than from N3, and three alleles were unique to the northern samples (all of these alleles were found in at least one other European population). These results cannot be explained only by genetic drift following a recent founding event from central Scandinavia. Rather, they suggest the existence of a historical lynx population north of the Arctic Circle, and/or immigration from the East (see below). In the northernmost part of Scandinavia only stray animals were observed until the presumed establishment of a reproducing population during the 1950–60s (Olstad 1945; Myrberget 1970; Kvam 1997a). In the northernmost county, Finnmark, only a single individual was registered shot during the 17th century (Myrberget 1970). Competition from wolves for semidomestic reindeer has been suggested to be the reason for the historical lack of lynx in northern Scandinavia (Myrberget 1970). The wolf population became extinct in the 1960s (Wabakken *et al*. 2001). It should, however, be noted that human settlement in this huge region (50 000 km^2) was concentrated mainly along the coast. Inland, where lynx would be expected to be found, was (and still is) very sparsely populated and it is reasonable to assume that the occurrence of lynx may have been underreported in this area. Mazák (1970) discussed the possibility of a more northerly distribution of lynx than commonly considered, and Heggberget & Myrberget (1980) pointed out how the expansion of the northern Norwegian population in the 1970s was underestimated in reports from game boards. These authors also claimed that hunting in the 1970s was responsible for only a minor part of the annual mortality and, hence, do not reliably reflect the population size.

Our results suggest that the current lynx population in Scandinavia may originate from more than one core area, which has not been detected by historical data and census estimates. Recognition of genetically divergent groups is obviously significant for the maintenance of diversity and our study emphasizes the usefulness and importance of the application of molecular techniques for management

purposes when combined with knowledge of a species' ecology.

Immigration into northern Scandinavia

Hellborg *et al.* (2002) found one individual from northern Norway assigned to the Baltic sample. Using GMA we detected two first-generation immigrants in the northern samples (N1 and N2) when comparing with the other European samples. Northern Scandinavia share borders with Finland and Russia, and these countries are therefore the most likely sources for immigrants. Our northern Scandinavian samples (N1, N2 and S1) did not show particular similarity with either the Finnish or the Russian samples. However, the Russian individuals analysed here are from Ural, very far from the Norwegian border. The status of lynx in parts of Russia closer to Norway (i.e. the Kola Peninsula) is unclear. However, according to Heptner & Sludskii (1972) the Russian lynx range may extend all the way to the Norwegian border (at the Varanger Fjord), and thus Russian animals should be able to migrate into Norway. It is also reported that lynx may undertake long-distance migrations in years when prey is scarce following Norwegian lemmings (*Lemmus lemmus*): a species only present within Russia at the Kola Peninsula. Samples from Kola (if available) should be analysed to investigate this further.

Underlying mechanisms for cryptic population structure

This study unravels population structure in lynx, invisible using traditional methods, and dispersal of this highly mobile predator within the Scandinavian Peninsula. What then are the mechanisms involved in establishing and maintenance of such population differentiation? The answer to this is likely to involve intrinsic factors preventing gene flow. Territorial behaviour and natal philopatry have previously been suggested to promote genetic structuring (see Sugg *et al.* 1996; Piertney *et al.* 1998), but for solitary carnivores little is known about such behaviour and its effects on genetic differentiation (Ratnayeke *et al.* 2002). Scandinavian lynx is known to occupy very large territories (Moa *et al.* 2001) and to display strong home range fidelity (Andersen *et al.* 2002), and we propose that this may be an explanation for the maintenance of the cryptic population structure. This hypothesis may be tested by studying the kinship between individuals occupying neighbouring territories.

Acknowledgements

References

Andersen R, Kvam T, Linnell JDC *et al.* (2002) Møt gaupa; uproblematisk rovdyr-eller notorisk lystmorder? *NINA, Temahefte*, **18**, 78–85 (in Norwegian).

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

Banks MA, Eichert W (2000) WHICHRUN Version 3.2: a computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity*, **91**, 87–89.

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry*, **196**, 80–83.

Curry-Lindahl K (1968) The lynx population in Sweden. *Acta Scientiarum Naturalia Academia Scientiarium Bohemoslovacae-Brno*, **2**, 21–26.

Felsenstein J (1993) PHYLIP (*Phylogeny inference package*), *Version 3.5c.* Distributed by the author. Department of Genetics, University of Washington, Seattle.

Flagstad Ø, Walker CW, Vilà C *et al.* (2003) Two centuries of the Scandinavian wolf population: patterns of genetic variability and migration during an era of dramatic decline. *Molecular Ecology*, **12**, 869–880.

Guo S, Thompson E (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83**, 145–154.

Heggberget TM, Myrberget S (1980) The Norwegian *Lynx lynx* population in the 1970s. *Fauna Norvegica, Series A*, 29–33.

Hellborg L, Walker CW, Rueness EK *et al.* (2002) Differentiation and levels of genetic variation in northern European lynx (*Lynx lynx*) populations revealed by microsatellites and mitochondrial DNA analysis. *Conservation Genetics*, **3**, 97–111.

Heptner VG, Sludskii AA (1972) Lynx. In: *Mammals of the Soviet Union*, Vol. 2 (ed. Hoffmann RS), pp. 524–636. Vysshaya Shkola, Moscow.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of population genetic correlation with distance. *Genetics*, **49**, 561–576.

Kratochvíl J (1968) Survey of the distribution of the genus lynx in Europe. *Acta Scientiarum Naturalia Academia Scientiarium Bohemoslovacae-Brno*, **2**, 5–12.

Kvam T (1997a) *Bestandsestimat for Gaupe 1995–96 og 1996–97.* Rapport til DN (in Norwegian).

Kvam T (1997b) Hva er årsaken til økningen i gaupestammen? *NINA, Temahefte*, **8**, 173–178 (in Norwegian).

Kvam T, Sunde P, Overskaug K (1997) Matvaner hos gaupe i Nord-Trøndelag. *NINA, Temahefte*, **8**, 94–104 (in Norwegian).

Liberg O (1997) Historik och utbredning. In: *Lodjuret* (ed. Svenska Jägareförbundet), pp. 9–19. Almqvist & Wiksell tryckeri, Uppsala (in Swedish).

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Mazák V (1970) Comments on the distribution of the lynx, *Lynx lynx lynx* (Linneaus, 1758) in Scandinavia. *Vestnik Ceskoslovenske Spolecnosti Zoologicke*, **33**, 227–233.

Menotti-Raymond M, David VA, Lyons LA *et al.* (1999) A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). *Genomics*, **57**, 9–23.

Menotti-Raymond MA, O'Brien SJ (1995) Evolutionary conservation of ten microsatellite loci in four species of Felidae. *Journal of Heredity*, **86**, 319–322.

Moa PF, Linnell JDC, Andersen R *et al.* (2001) Home range size and choice of management strategy for lynx in Scandinavia. *Environmental Management*, **27**, 869–879.

Myrberget S (1970) The Norwegian population of wolverine, *Gulo gulo* (L.), and lynx *Lynx lynx* (L.). *Papers of the Norwegian State Game Research Institute, Series 2*, **33**, 16–28.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. *Annals of Human Genetics*, **47**, 253–259.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution*, **19**, 153–170.

Olstad, *O* (1945) *Jaktzoologi*. J.W. Cappelen, Oslo (in Norwegian).

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354.

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, **12**, 357–358.

Piertney SB, MacColl ADC, Bacon PJ, Dallas JF (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Molecular Ecology*, **7**, 1645–1654.

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

Pulliainen E (1968) The lynx population in Finland. *Acta Scientiarium Naturalia Academia Scientiarium Bohemoslovacae-Brno*, **2**, 27–34.

Ratnayeke S, Tuskan GA, Pelton MR (2002) Genetic relatedness and female spatial organization in a solitary carnivore, the racoon, *Procyon lotor*. *Molecular Ecology*, **11**, 1115–1124.

Raymond M, Rousset F (1995) GENEPOP (v. 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, **105**, 767–779.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbour Laboratory Press, New York.

Schneider S, Kueffer JM, Roessli D, Excoffier L (1999) ARLEQUIN, *Version 2.000: A Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Sokal RR, Rohlf FJ (1981) *Biometry*, 2nd edn. Freeman, New York.

Sugg DW, Chesser RK, Dobson FS *et al.* (1996) Population genetics meets behavioral ecology. *Trends in Ecology and Evolution*, **11**, 338–342.

Taberlet P, Swenson J, Sandegren F, Bjärvall A (1995) Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear (*Ursus arctos*) in Scandinavia. *Conservation Biology*, **9**, 1255–1261.

Wabakken P, Sand H, Liberg O, Bjärvall A (2001) The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978–98. *Canadian Journal of Zoology*, **79**, 710–725.

Waits LP, 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.

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

Wright S (1943) Isolation by distance. *Genetics*, **28**, 139–156.