OULU 2006
A 473

ACTA UNIVERSITATIS OULUENSIS

Miika Tapio

ORIGIN AND MAINTENANCE OF GENETIC DIVERSITY IN NORTHERN EUROPEAN SHEEP

FACULTY OF SCIENCE,
DEPARTMENT OF BIOLOGY,
UNIVERSITY OF OULU

ACTA UNIVERSITATIS OULUENSIS
A Scientiae Rerum Naturalium 473

MIIKA TAPIO

ORIGIN AND MAINTENANCE OF GENETIC DIVERSITY IN NORTHERN EUROPEAN SHEEP

Academic dissertation to be presented, with the assent of the Faculty of Science of the University of Oulu, for public defence in Kuusamonsali (Auditorium YB210), Linnanmaa, on November 10th, 2006, at 12 noon

OULUN YLIOPISTO, OULU 2006

Acta Univ. Oul. A 473, 2006

Supervised by
Docent Juha Kantanen
Professor Pekka Pamilo

Reviewed by
Doctor Johannes A. Lenstra
Professor Craig Primmer

ISBN 951-42-8234-5 (Paperback)
ISBN 951-42-8235-3 (PDF) http://herkules.oulu.fi/isbn9514282353/
ISSN 0355-3191 (Printed)
ISSN 1796-220X (Online) http://herkules.oulu.fi/issn03553191/

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2006

Tapio, Miika, Origin and maintenance of genetic diversity in northern European sheep

Faculty of Science, Department of Biology, University of Oulu, P.O.Box 3000, FI-90014 University of Oulu, Finland

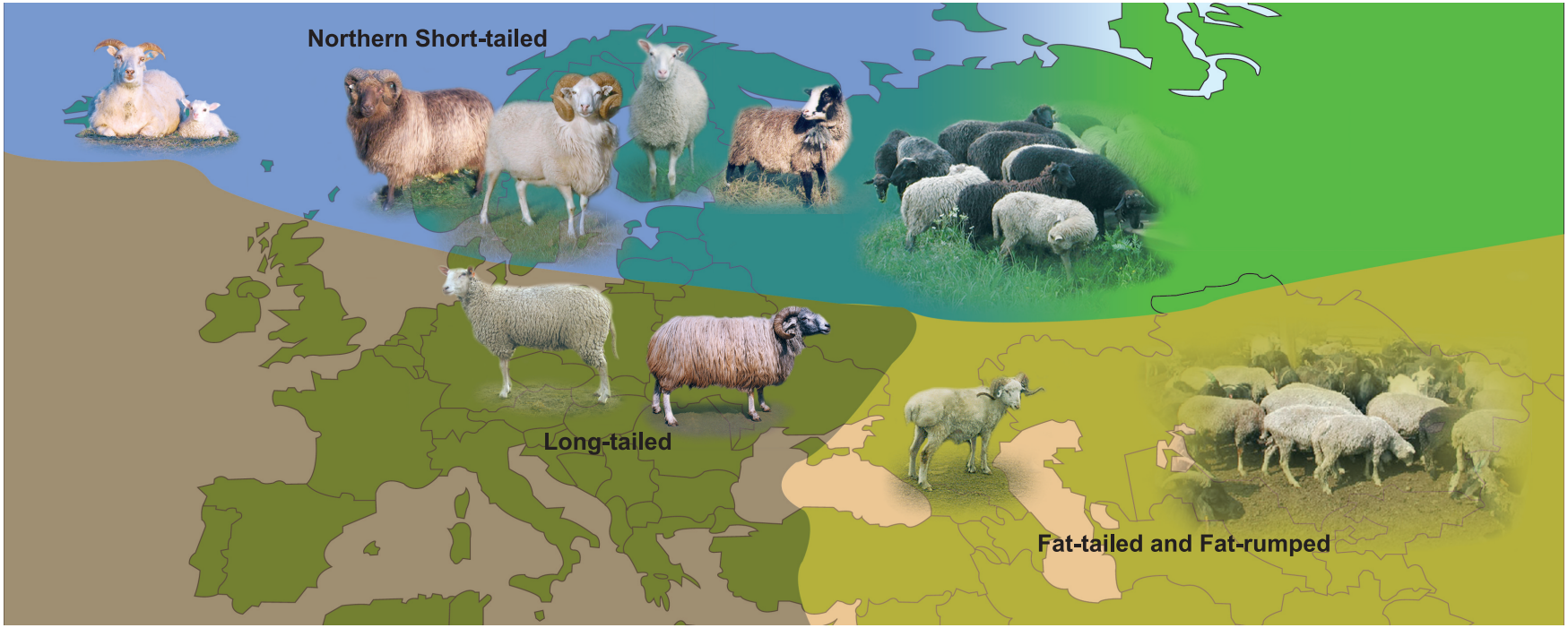
Acta Univ. Oul. A 473, 2006

Oulu, Finland

Abstract

The Nordic and Baltic countries and North-western Russia have >20 old native sheep breeds. These together with recently synthesized breeds and local populations of international breeds make up the northern European sheep diversity. Changes in agriculture threaten to erode genetic diversity in sheep. Molecular genetic variation was assessed to understand genetic diversity in northern European sheep. Distribution of maternal lineages were studied based on mitochondrial control region variation in 76 sheep breeds in northern Europe and in a wide neighbouring area extending to the Caucasus and Central-Asia. Autosomal microsatellite variation was studied in 37 northern European breeds, and autosomal blood protein variation was studied in six Finnish and Russian breeds. Four distinct maternal lineages were observed in Eurasian sheep. Their distribution agrees with sheep expansion starting from the Near East. Two most common distinct lineages were recorded in northern Europe. Majority of northern sheep have the lineage, which predominates in other parts of Europe. Results suggest that the main maternal origin of northern sheep is in the south. However, rare "Asian" lineage was observed in several old northern European breeds. The rare type in the Nordic sheep is descendant to the type observed in the Middle Volga region, which suggest that some sheep were brought to northern Europe from the east. Microsatellites showed clustering of geographically neighbouring sheep, when breed locations are corrected for the recent transportations. The analysis separated long and short-tailed sheep, although this macroscale structure explains a small proportion of breed differences. Differentiation among the northern European breeds is stronger than typically observed in sheep. Many native breeds are less inbred than the local populations of the international breeds, but some rare breeds and subpopulations of divided unofficial strains were inbred. Some breeds require more careful maintenance due to recent population size reduction. Maintaining prolificacy in breeds such as the Finnsheep and the Romanov may require efficient avoidance of inbreeding. The breeds were ranked for conservation using simultaneously within-breed variation and breed divergence. Set of important breeds included seven rare old native breeds or strains which merit efficient conservation measures urgently.

Keywords: conservation, domestic animals, genetic marker, Ovis aries



Photos:
Iceland (Magnus Hlynur Hreidarsson)
Norway (Bine Melby)
Sweden (Per Abrahamsson)
Finland (Ville Tuokko)
Russian Karelia (Juha Kantanen)
Central Russia (Juha Kantanen)
Poland (Maciej Murawski)
Ukraine (Tatyana Kiselyova)
Caucasus (Nurbiy Marzanov)
Altai (Juha Kantanen)

Acknowledgements

This research was carried out at the Biotechnology and Food Research, MTT Agrifood Research Finland, Jokioinen, Finland. The biggest thank-you goes to my principal supervisor, Adj. Prof. Juha Kantanen. This work would not have been possible without his enthusiasm to study genetic diversity and his willingness to travel around Eurasia sampling sheep for the study. I thank him for being able to study such an interesting and important research topic, and for the chance to participate in some sampling activities. I thank my auxiliary supervisor, Prof. Pekka Pamilo, for his supporting attitude to my every inquiry. I thank all the North-SheD project members. Project coordinators were Dr Emma Eythorsdottir (Iceland & general coordination between the countries), late Prof. Birgitta Danell (Sweden), Dr. Lars-Erik Holm (Denmark), Ass. Prof. Ilona Miceikiene (Lithuania), Ass. Prof. Ingrid Olsaker (Norway), Prof. Haldja Viinalass (Estonia) and Adj. Prof. Juha Kantanen (Finland & general coordination between the Nordic and the Baltic countries). I have not had the possibility to discuss directly with all the people that have contributed to the sample collection, but I'm grateful to all of them. Special thanks go to Prof. Nurbiy Marzanov at the Russian Academy of Agricultural Sciences, whom I was lucky to meet at the ISAG 2000 conference. He organized a large part of the Russian material and provided expertise on Russian sheep. Similarly, I want to thank Dr. Tatyana Kiseliova at the All-Russian Institute for Farm Animal Genetics and Breeding for arranging several sampling expeditions. The co-authors in the manuscripts have provided their insight and expertise on the sheep breeds and genetics. Some sheep experts and other researchers have offered their expertise without being included as co-authors. A large thanks for this goes to Prof. Asko Mäki-Tanila at MTT, who has commented on several versions of manuscripts. I'm also thankful to Zofia Kurowska at the Swedish Board of Agriculture, who offered me an introduction to the Swedish breeds, and to Bine Melby at the Norwegian Museum of Agriculture for the information about the Norwegian breeds. For technical assistance I thank Anneli Virta, Tiina Jaakkola, Sari Raiskio, Tuula-Marjatta Hamama (Finland), Bente Flügel (Denmark) and Ole-Albert Guttersrud (Norway), who helped with the large undertaking.

The Nordic-Baltic North-SheD project was financially supported by the Nordic Gene Bank for Farm Animals, whose director, Dr. Erling Fimland, was highly supportive from the beginning of the project. I thank the Academy of Finland and the Finnish Ministry of

Agriculture and Forestry, who financially supported the study presented in Paper I (SUNARE- and Russia In Flux-programmes). I thank the Department of Education in Finland for the graduate school grant (Finnish Graduate School in Population Genetics coordinated by Oulu University) and the Academy of Finland for the grant (#213712) for writing the summary part of the thesis. The Betty Väänänen Foundation grant provided me with the possibility to attend a workshop during the PhD work.

My close relatives were always ready to support and ask what was going on at work. I'm especially thankful to my wife, Ilma Tapio, who has been my closest companion and support at home and at work. Our lovely daughter, Marta, did her best to offer alternative activities, so that the father would forget the work stress. I thank colleagues at MTT for the discussions and their friendship at work and in extracurricular activities.

Jokioinen, September, 2006 Miika Tapio

"… sheep and other domestic animals are said to be extremely prolific in Lapland."
Darwin C (1868) Variation of Animals and Plants Under Domestication Volume 2

Abbreviations

DNA	deoxyribonucleic acid
mtDNA	mitochondrial DNA
N_e	Effective population size
PCR	polymerase chain reaction
r[g]	allelic richness, i.e. expected allele number in a sample of *g* chromosomes.
s	number of populations

List of original papers

This thesis is based on the following publications, which are referred to in the text by their Roman numerals.

I Tapio M, Marzanov N, Ozerov M, Ćinkulov M, Gonzarenko G, Kiselyova T, Murawski M, Viinalass H & Kantanen J (2006) Sheep mitochondrial DNA variation in European, Caucasian and Central Asian areas. Mol Biol Evol 23: 1776–1783.

II Tapio M, Tapio I, Grislis Z, Holm L-E, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2006) Geographical patterns in the variation of mitochondrial DNA in Northern European sheep. Manuscript.

III Tapio M, Tapio I, Grislis Z, Holm L-E, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2005) Native breeds demonstrate high contributions to the molecular variation in northern European sheep. Mol Ecol 14: 3951–3963.

IV Tapio I, Tapio M, Grislis Z, Holm LE, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2005) Unfolding of population structure in Baltic sheep breeds using microsatellite analysis. Heredity 94: 448–456.

V Tapio M, Miceikiene I, Vilkki J & Kantanen J (2003) Comparison of microsatellite and blood protein diversity in sheep: inconsistencies in fragmented breeds. Mol Ecol 12: 2045–2056.

In addition, unpublished data has been included.

Contents

1 Introduction: sheep genetic resources

Genetic resources refer to "genetic material of actual or potential value" to man (Rio Convention on Biological Diversity 1992). The term mainly refers to diversity in domesticated species and in wild species containing specific, potentially useful traits. It is also used as a synonym for genetic diversity, which is visible as inherited differences among species, populations and individuals. Genetic resources in sheep refer to genetic diversity among and within breeds or populations. The resources are mainly composed of living populations, but also include conserved material such as semen, oocytes and embryos. Maintenance of genetic resources in domestic animals is an urgent issue as one or two breeds become extinct each week (e.g. Simianer 2005a).

In domestic species conservation typically operates at the level of breeds, not at the level of species. The most important arguments for conservation of domestic breeds concern maintaining or increasing food production to keep pace with global environmental changes, opportunities to meet future market demands, possibilities to offer livelihoods for people, both locally and globally, together with cultural-historic and scientific reasons (Oldenbroek 1999). The practical conservation goals centre on maintaining the greatest possible genetic variation in the species, maintaining special populations and adaptations, and ensuring survival of the populations chosen for conservation without unnecessary loss of within-population variation and avoidance of inbreeding. These objectives are the same as those for conservation of wild animals. In the context of wild mammalian species, Garner *et al.* (2005) stressed that substantial losses of genetic diversity occur at the population or subpopulation level before the *species* becomes endangered. Studies in domestic species demonstrate how, for instance, management and population admixture can influence diversity and which kinds of populations are important to maintain variation in the species.

This thesis explores the origin, distribution, and conservation of genetic diversity in domestic sheep. The study concentrates on sheep genetic diversity in the northern areas of Europe, and also describes the variation closer to the sheep domestication centre to explain the origin of northern European sheep breeds. The variation is assessed using molecular genetic methods.

1.1 Wild sheep and early history of European domestic sheep

The common term “sheep” refers to species of the genus *Ovis*, although sometimes the term has been applied also to the genera *Ammotragus*, *Hemitragus* and *Pseudois*. There are several putative ancestral wild species of the domestic sheep (*Ovis aries*), because ovine populations can interbreed despite differences in chromosome numbers (Ryder 1983, Nadler *et al.* 1973). Typically, the wild sheep in the genus *Ovis* are divided into six species: *O. ammon* (argali or giant sheep) *O. orientalis* (Asiatic mouflon), *O. canadensis* (bighorn sheep), *O. dalli* (Dall's sheep), *O. nivicola* (snow sheep) and *O. vignei* (urial; International Comission of Zoological Nomenclature 2003, Myers *et al.* 2005). Archaeological evidence suggests that the sheep was domesticated ~9 000 years ago in the Near East, around the border region between Syria and Iraq (Fig. 1, Smith 1998).

The archaeological evidence suggests that the sheep arrived in south-eastern Europe 8 000 – 7 000 years ago and subsequently spread over Europe by two routes. One route was along the northern coast of the Mediterranean and further along the Atlantic shore. The second route was through the Danube Valley to the region of warm temperate climate in continental Europe. (Fig. 1, Smith 1998.) Agriculture had spread to the British Isles, Denmark and southern Scandinavia 6 000 years ago. The further spread to the region of cool temperate climate may have been slow and the expansion possibly paused for a millennium. (Milisauskas 2002, p. 156, 158, 194 and 199.) The evidence for sheep existing in Finland during the Stone Age (i.e. more than 3 500 years ago) is controversial, but evidence for sheep at the Bronze Age (2 500 – 3 500 years ago) is reliable (Huurre 2003). Sheep arrived in Iceland in Viking times 1 000 years ago (Adalsteinsson 1981).

New genetic variation was introduced to Europe after the initial arrival of sheep. Features such as a long tail, fine wool and larger size likely originated from sheep of southwestern Asia, although the characteristics have been developed further in Europe (Ryder 1983). Northern Europe has not been isolated either: Ryder (1983) considers the colour variants and wool types typical to the old northern native breeds to stem from the general European sheep of the Iron Age.

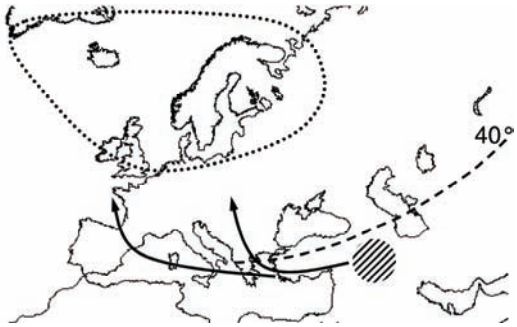


Fig. 1. The hypothesized Near East domestication centre (hatched area) and routes of spread of sheep to Europe (arrows). Northern Europe in the text refers to the area marked with the dotted line (detailed description in text). Sheep breeds from the area between northern Europe and the marked segment of the 40° northern latitude (broken line) was assessed for mtDNA variation only.

1.2 Northern European sheep breeds

1.2.1 General background

Sheep breeds in Europe are groups of animals of standard appearance (or management, in some cases). The old local breeds in Europe (so-termed landraces) were created from local sheep types adapted to diverse European agro-ecological zones (Maijala & Terrill 1991). The oldest defined sheep breeds are the fine-wooled Merino of Spain and several breeds in Britain, where systematic breeding began. Consequently, the majority of common breeds have either Merino or some British ancestry, and over one third of individual sheep are of the Merino type (e.g. Maijala 1997).

The breeds in sheep are more numerous than in other domestic animal species because sheep have modest needs, can be kept in diverse environments and the main product can be meat, milk, pelt or different types of wool (Maijala & Terrill 1991). In 1999, Europe had 629 breeds and nearly half of these breeds were endangered, i.e. there were less than 1 000 reproducing females and no more than 20 males within a breed. Furthermore, 142 European sheep breeds are known to have gone extinct. (Scherf 2000.)

The Nordic and Baltic countries have altogether 26 or 34 native breeds. The number depends on whether some sheep types are regarded as breeds or only as strains (i.e. special varieties within a breed or identifiable populations without a designed management programme). In addition, half a dozen imported breeds have large (>1 000 ewes) local populations in these countries (FAO 2005). Semionov and Selkin (1989) counted 58 national sheep breeds and seven imported breeds for the region of the former USSR, excluding the Baltic countries.

Breeds have been classified using the main production or phenotypic characteristics. The majority of European sheep breeds are thin-tailed fleece sheep, in which the

relatively long tail reaches the hocks (Mason 1991). In contrast, the old native landraces in northern Europe belong to the northern short-tailed group (Mason 1991). Northern Europe is here considered to include Denmark, Iceland, Norway, Sweden, Finland, north-western Russia, Estonia, Latvia and Lithuania (Fig. 1). A short overview of the northern European sheep is given below. Detailed breed descriptions can be found elsewhere (Semionov & Selkin 1989, Mason 1991, Maijala 1997, Eythorsdottir 2005, FAO 2005, Table 1).

1.2.2 The sheep breeds in northern Europe

The long-tailed sheep breeds, imported or synthesised locally from the modern imports, are the most important breeds in Denmark, the Baltic countries and Norway. These breeds (hereafter referred to as 'modern breeds') originate from the British Isles, the Netherlands or northern Germany. In Denmark, the breeds with more than 1 000 breeding individuals are *the Texel*, *the Shropshire*, *the Oxford Down* and *the Dorset* (FAO 2005). The common breeds in the Baltic countries (*the Estonian Whitehead* and *the Estonian Blackhead*, *the Latvian Darkheaded* and *the Lithuanian Blackface*) are of partly local origin, but they were graded to the Shropshire or the Cheviot, and later crossed with the Texel and the Ile-de-France along with other less extensively used breeds (FAO 2005). The most important breed in Norway is *the Norwegian White Sheep*, which was formed in 2001 by joining *the Dala Sheep*, *the Steigar Sheep* and *the Rygja Sheep*. These latter breeds still exist and are composites of local Norwegian sheep graded to the Leicester, the Cheviot or the North Country Cheviot originating from the U.K. and crossed later at least with the Texel, the Finnsheep and the East Friesian Milk Sheep. (Eythorsdottir 2005, FAO 2005.)

The assemblage of sheep in Russia is complex. There are no detailed statistics about the recent population sizes for the breeds. In 1980, 79% of the sheep stocks in state and collective farms of the USSR were considered purebred, and 63% of these purebred stocks belonged to fine-wooled breeds. These fine-wooled breeds are synthetics based on various Merino breeds and local sheep. The indigenous coarse-wooled breeds of the region include the short-tailed *Romanov*, originating from Yaroslavl region (north of Moscow) and several fat-tailed or fat-rumped breeds of the Caucasus and Central Asia. The proportion of these indigenous breeds of the purebred animals was 29%. The sheep in private holdings (18% of the total stock) and sheep of rare or imported breeds (7% of the total stock) were not included in these statistics. (Semionov & Selkin 1989.)

The native northern short-tailed breeds are the most numerous purebreds in Iceland (*the Icelandic Sheep*), Finland (*the Finnsheep*), Sweden (*the Gotland Sheep*) and the Faeroe Islands (*the Faeroe Island Sheep*) and Greenland (*the Greenland Sheep*). Although the native breeds are the most common purebreds in Finland and Sweden, various types of crossbred animals form the majority in these countries. Other unthreatened native short-tailed breeds are *the Lithuanian Native Coarsewooled*, *the Norwegian Spael* and *Feral Sheep*, *the Russian Romanov*, and *the Swedish Finewool*, *Gute* and *Rya Sheep*. The recognized rare breeds are *the Åland Sheep* (Finland), *the Danish Landrace Sheep*, and *the Norwegian Grey Troender* and *Old Spael Sheep*. In Sweden there is a collection of rare small groups of sheep termed *peasant breeds*, in

which *the Dala Fur Sheep*, *the Forest Sheep* (or the Wood Sheep) and *the Swedish Roslag Sheep* and some other rare sheep types are included. *The Finnish Grey Landrace* and *the Icelandic Leader Sheep* are rare strains within *the Finnsheep* and *the Icelandic Sheep*, respectively, containing unique traits. (Jeppsson 1998, Eythorsdottir 2005, FAO 2005.)

1.2.3 Need to organise the maintenance of northern European sheep genetic resources

The distinctive characteristic of northern Europe is the short summer season. This has been suggested to have caused the emergence of highly prolific northern breeds with rapid juvenile growth rates. Northern sheep are also considered to have adapted to humid or semi-humid environments and cool temperatures of the region. (Terril & Slee 1991.) Most of the northern breeds have been housed over winter with poor feed. Therefore, the local breeds are likely to be able to respond to fodder quality by changes in their rumen volume. This trait may be lost in the sheep selected under intensive management (Weyreter & Engelhardt 1984). Exotic wool or meat breeds were not preferred in northern Europe earlier because they were not adapted to local conditions and they were susceptible to diseases in the north. In addition, the wool of the local breeds was preferred in hand weaving. (Maijala 1988.)

Local breeds have become less popular during the last decades. European sheep production has concentrated on producing meat in increasingly intensified and standardized production systems. The modern breeds have been developed to respond to the high levels of feed, health care and management, which has made them more popular (Sherf 2000). However, the local breeds may have untested potential to increase production without loss of the local adaptations (Hall & Bradley 1995). These adaptations would be needed following a change in the available resources or in the desired form of production.

In addition to the recent change in preferred sheep types, the general decrease of the economic importance of European sheep threatens the native breeds. In 1999 Europe had 17.5% (equalling 185 million head) of the world sheep (Scherf 2000), but this proportion is decreasing by 0.4 percentage units *per annum*. This has resulted in a 23% decrease in the number of European sheep from 1995 to 2004. (FAO 2004.) In the near future, sheep and goat production in the EU is expected to continue decreasing, while the increasing European consumption is met by increasing imports from the global market, mainly New Zealand (Canali and Econogene Consortium in press).

The Nordic and the Baltic countries have 2 – 3% of the European sheep numbers. Including Russia, northern Europe has 10% of the European total. The number of sheep in the Baltic countries and Finland is 17 000 – 95 000. Denmark, Sweden and Iceland have 200 000 – 500 000 sheep each. Norway has 2.4 million sheep and the Russian Federation has 14.7 million sheep. (FAO 2004.) The number of ewes has fluctuated within the Nordic countries, but there has not been any drastic decrease in the number from 1950 to 1990. The only exception is Finland, where the number of sheep in this period collapsed from 0.6 million down to one tenth of the number (Kristensen 1992). In the last 10 years (1995 – 2004), the most drastic decrease in sheep number occurred in

the Russian Federation, where the number dropped by 54% or by 17 million, while in the rest of northern Europe the decrease was 5% or 0.2 million (FAO 2004).

The European sheep breeds are a very important genetic resource and globally they have been the most popular components of synthetic breeds. *The Finnsheep* and *the Romanov* have been the most popular breeds to increase fertility (Maijala & Terrill 1991) and their characteristics in crossbreeding use have been extensively studied (Fahmy *et al.* 1992, Littledike & Young 1993, Freking *et al.* 2000, Casas *et al.* 2004, Lupton *et al.* 2004, Freking & Leymaster 2004). Based on the valuable adaptations in the various breeds and major changes taking place in northern European agriculture, the conservation of sheep genetic resources is necessary. Many old breeds still exist and conservation actions can ensure maintenance of genetic diversity in the region. Large effective population sizes or low rates of inbreeding minimise the environmental sensitivity and maintain adaptation and breeding potential of the species (Lynch & Hill 1986, Frankham 2005, Garner *et al.* 2005, Kristensen *et al.* 2005). Understanding the structure and the origin of the genetic resources helps in planning the conservation actions more efficiently.

1.3 Molecular genetic variation

Pedigree records are the most readily exploitable source of information in recognizing crossbred individuals, planning matings, quantifying kinship and rate of inbreeding and estimating other population genetic parameters (Caballero & Toro 2000). However, surveys of cattle data show the known pedigrees to have information only about the recent history (<20 generations), to be incomplete (equal sized complete pedigree would have <10 generations), and to have 3 – 20% incorrectly recorded parentages (Baumung & Sölkner 2003, Gutiérrez *et al.* 2003). The pedigree data for a typical rare sheep breed cover an even smaller number of generations because sheep are kept on smallholdings without long-standing, coordinated breeding programmes (Goyache *et al.* 2003).

Molecular markers are needed when the pedigree records are missing or cover the considered time-scale only partially, or when different breeds are compared. Molecular markers are heritable traits of biomolecules (usually DNA, proteins or carbohydrates) that can be assayed for variation in organisms or populations. Most population genetic applications of molecular markers utilize so-called neutral variation, which is not affected by natural or artificial selection. Molecular markers can be used to estimate genetic variability, inbreeding and kinship, to detect structures in the distribution of genetic variation, and to infer details about the population history (e.g. Chikhi & Bruford 2005). This information can be used in planning population management. Molecular markers are also used in identifying genetically important populations for conservation (Weitzman 1992, 1993, Petit *et al.* 1998, Eding *et al.* 2002, Piyasatian & Kinghorn 2003, Bennewitz & Meuwissen 2005, Simianer 2005b).

The currently utilized marker types are described below. Y chromosomal variation was not studied because extensive sequencing has revealed very little usable variation in the sheep Y chromosome (Meadows *et al.* 2004).

1.3.1 Mitochondrial DNA

The mitochondrion is an organelle in the cell cytoplasm found outside the cell nucleus. It is the only animal organelle with its own DNA. In mammals, mitochondrial DNA (mtDNA) is transmitted to the progeny only from mother (Hutchison *et al.* 1974, Hayashi *et al.* 1978, Giles *et al.* 1980). The possibility for rare paternal inheritance and recombination among mitochondrial lineages has been suggested, but this remains controversial (Piganeau & Eyre-Walker 2004, Piganeau *et al.* 2004). In most mammals the sperm mitochondria are transferred to the oocyte during fertilization (Ankel-Simons & Cummins 1996), but the oocyte actively destroys these mitochondria (Sutovsky *et al.* 2000, 2003). In the hybrids of different mouse species, Shiatara *et al.* (1998) detected paternal mtDNA in the tissues of half of the hybrids, but the paternal mtDNA types were not transferred to the progeny. In mammalian sperm cells, the number of mtDNA molecules is low (50 – 75, Hecht *et al.* 1984) whereas in mammalian oocytes the copy number is extremely high ($\geq 10^5$, Michaels *et al.* 1982). Therefore, paternally inherited mitochondria are prone to disappear by drift in cell cycles after fertilization even if active destruction of the paternal mitochondria does not occur. The disappearance of within-individual mtDNA polymorphism (i.e. heteroplasmy) caused by spontaneous mutations has been shown to occur very quickly in cattle (Ashley *et al.* 1989). In sheep, one crossbred ewe was reported have repeatedly produced offspring with mitochondria derived from the father (Zhao *et al.* 2004). It is not know if the paternal mitochondria were present in the germ line cells. Slate and Phua (2003) did not detect signs of recombination in sheep mtDNA variation, which suggests that paternal inheritance is insignificant.

Structure and gene organization of mtDNA is conserved in mammals. The mitochondrial genome is a closed circular, double-stranded DNA molecule about 16 600 base pairs long that contains 13 protein-coding genes, 22 transfer RNA genes, and 2 ribosomal RNA genes. The molecule is very tightly organized, the genes have no introns and, except for the control region governing transcription and replication, intergenic sequences are absent or limited to a few bases. (Taanman 1999.) The average rate of synonymous substitutions in mtDNA is about 20 times higher than in nuclear DNA (Pesole *et al.* 1999). The control region can be divided into two peripheral highly variable segments (hypervariable regions I and II or ETAS and CSB domains) and a central conserved region (Fig. 2, e.g. Pesole *et al.* 1999). The sequences of the hypervariable region I have been widely used in molecular evolutionary studies. The average substitution rate for this domain is three quarters of the mitochondrial synonymous rate (Pesole *et al.* 1999). The substitution rate of some sites may be over four times the average rate for this domain (Excoffier & Yang 1999). Large numbers of variable sites in a short segment of DNA make the region very informative in intraspecific studies, while more conservative segments of the control region can be used to estimate more distant relationships.

Two distinct lineages were recorded in the first surveys of sheep mtDNA variation (Zardoya *et al.* 1995, Wood & Phua 1996, Hiendleder *et al.* 1998; 2002), and recently a new distinct haplotype group was reported (Guo *et al.* 2005, Pedrosa *et al.* 2005). Based on the similarity of the mtDNA sequences found in the European mouflon and the

sequences belonging to one of the haplotype groups found in domestic sheep, wild mouflons of western Asia have been suggested to represent the ancestral progenitor of domestic sheep, whereas mtDNA sequences from the two other previously hypothesized ancestors, Urial (*Ovis vignei* ssp.) and argali sheep (*Ovis ammon* ssp.), are too distinct from the domestic types to be the sources of their maternal lineages (Hiendleder *et al*. 1998).

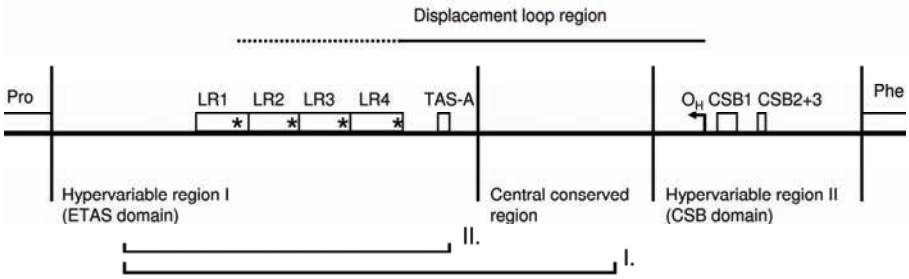


Fig. 2. Schematic representation of the sheep mitochondrial control region bounded by the tRNA genes for proline (Pro) and phenylalanine (Phe). Long tandem repeats (LR1, 2, 3 and 4) with putative termination signals (asterisk), the binding site for a protein regulating termination and transcription (TAS-A element), the origin of heavy strand replication (O_H), and conserved sequence blocks (CSB1 and CSB2+3) are presented. The line above indicates the three stranded displacement loop region between O_H and termination sites. Brackets below indicate the segments analysed in the papers I and II

1.3.2 Autosomal markers

Autosomal markers are located in autosomal chromosomes rather than in sex chromosomes or outside the nucleus. Diploid animals have two copies of these marker loci and one copy is inherited from both parents. The allelic status of both copies is known for co-dominant markers. For this study, variation in two types of co-dominant autosomal markers was explored: microsatellites and blood proteins.

Detection of the differences in the amino acid composition of proteins is based on the differences in their surface electric charges. The charge determines how proteins move in an electric field. In monomeric proteins, which contain a single polypeptide chain, each allele is characterized by a unique rate of movement. (e.g. Hartl & Clark 1989.)

Characterization of blood protein variation in sheep began in the 1950s (reviewed by Atroshi 1979) and the blood protein polymorphisms were the first genetic systems explored in northern European sheep (Evans *et al*. 1958, Efremov & Braend 1965, Bogdanov *et al*. 1972, Rasmusen & Tucker 1973, Atroshi 1979, Lazovskii 1983, Braend *et al*. 1987). Many early studies aimed at recognizing correlations between allele frequencies and phenotypic traits and only few aimed at estimating kinships (e.g. Lazovskii 1983).

A microsatellite is a region of the genome that contains tandem repeats of one to five base pairs. Variation occurs in the number of repeats within the region. The microsatellite

alleles are recognized by amplifying the repeat region using polymerase chain reaction (PCR) and determining the size of the amplified fragment with electrophoresis. Most studies use microsatellites with a repeat unit of two nucleotides. In these dinucleotide microsatellites the repeat unit generally occurs 10 - 30 times. (Jarne & Lagoda 1996, Ellegren 2004.) The mutation rate for microsatellites is two to three orders of magnitude higher than for electrophoretically genotyped proteins (Jarne & Lagoda 1996), but the rates vary between loci and organisms (Ellegren 2004). The mutation rate for sheep dinucleotide microsatellites has been estimated to be 1.4×10^{-4} mutations per gamete per locus (Crawford & Cuthbertson 1996).

A recent survey of diversity studies revealed microsatellites to be the main marker type in sheep (used in 89% of the research projects; Baumung *et al.* 2004). Use of microsatellites is straightforward because there are over 1 000 mapped microsatellite markers available in sheep (Maddox *et al.* 2001). Microsatellites have been shown to give results matching documented breed histories in sheep (Buchanan *et al.* 1994). To date the published sheep microsatellite studies have typically examined relationships among breeds within a country (e.g. Arranz 1998, Stahlberger-Saitbekova *et al.* 2001, Àlvarez *et al.* 2004) or relationships between populations of a breed (e.g. Diez-Tascón *et al.* 2000, Pariset *et al.* 2003). However, results from large-scale studies are becoming available, and some preliminary results have been published (e.g. Lenstra & the ECONOGENE Consortium 2005).

2 Outline of the study

The present studies were initiated by a practical need to estimate whether two Finnish sheep types recognized in the 1980s, *the Åland Sheep* and *the Grey Finnish Landrace*, were genetically distinct from the common *Finnsheep*. Soon it became clear that there were similar needs to characterize sheep breeds in the other northern European countries. This led to establishing a Nordic-Baltic collaborative project (North-SheD, Eythorsdottir 2005), which forms the main basis of the present thesis. Subsequently the investigations were extended to include the former Soviet Union and its neighbouring countries. From this only the mitochondrial study is included in this thesis.

The thesis aims at characterization of the genetic variation in the northern European sheep breeds, and at describing the domestication of the sheep to the extent that it is relevant for understanding this diversity. The study consists of five separate papers (numbered I-V) describing analyses of molecular genetic variation, and a summary section that describes the findings relevant to the following objectives (references to the original results in brackets):

1. Investigation of the basis of the genetic variation in northern European sheep. The place of domestication, and the route sheep spread after domestication. (I, II)
2. Characterizing structure in the distribution of genetic variation. How relevant is the macro-scale classification of breeds to breed-groups for genetic variation at molecular markers? How much of the variation is among the breeds or does maintenance of separate breeds matter? Are breeds homogeneous (i.e. panmictic) populations? (III, IV, V)
3. Assessment of past inbreeding and the possibility of recently increased rate of inbreeding in northern European sheep. Is there need to manage some breeds more carefully? (III, IV, V and original results in this summary section of the thesis)
4. Identification of the most important northern European breeds for conservation. (III and original results in this summary section of the thesis)

The results and discussion part of the thesis is organized according to these objectives.

The main contribution of this study to the literature on domesticated animal diversity consists of characterization of genetic resources of the well-known northern short-tailed breed-group, which has not been described collectively to a similar detailed degree before

publication of this thesis. The study provides novel insights into the history of sheep, and is relevant not only as a biological study, but also as a “culture-genetic” study of interest to archaeologists and historians. The management of genetic resources is a major motivation behind the thesis and a relatively large proportion of the thesis is devoted to identification of the most important breeds for genetic resource conservation.

3 Materials and methods

3.1 Sample collection and laboratory methods

In total, 76 different sheep breeds or strains were studied. The study area is presented in Fig. 1 and the breeds studied in the original papers are listed in Table 1. The assessed microsatellite loci are listed in Table 2, and the mtDNA segments analysed in papers I and II are presented in Fig. 2. Other details of the data sources, including the numbers of studied individuals and the laboratory methods can be found in the papers (I-V). Sampling of sheep aimed to cover the main stock of each breed or strain. This was achieved by sampling individuals avoiding horizontal and vertical relationships using information from two parental generations. In rare breeds and unofficial strains this was not possible due to small numbers of individuals or lack of pedigree records. When there were no pedigree records available, samples were collected from several farms.

Table 1. The list of sheep breeds and strains included in the study. The sheep are categorised to breed types [simplified from EAAP (2005), Mason (1991), and Semionov and Selkin (1989)]. For the breed type, tail morphology is given when it is known to deviate from the thin-tailed type. The breeds that have not been officially recognized are indicated as "unofficial strain". When the exact breed origin is uncertain, an approximate region is indicated in brackets.

Breed or strain	Breed type	Breed origin	Paper			
			I	II, III	IV	V
1.Azerbaijan Mazekh	Fat-tailed	Armenia[a]	x			
2.Azerbaijan Mountain Merino	Merino Finewooled type	Azerbaijan	x			
3.Gala	Fat-tailed	Azerbaijan	x			
4.Karabakh	Fat-tailed	Azerbaijan	x			
5.Azerbaijian Bozakh	Fat-tailed	Azerbaijan and Armenia	x			
6.Finnish Oxford Down	British shortwool type	Britain[a]	x			
7.Norwegian Cheviot Sheep	British shortwool type	Britain[a]		x		
8.Russian Romney Marsh	British longwool type	Britain[a]	x			
9.Danish Landrace Sheep	Nordic group	Denmark		x		
10.Faeroe Island Sheep	Nordic group/short-tailed	Denmark/Faeroe Islands		x		
11.Greenland Sheep	Nordic group/short-tailed	Denmark/Greenland		x		
12.Estonian Blackhead	British shortwool type	Estonia	x		x	
13.Estonian Ruhnu	Unofficial strain	Estonia	x	x	x	
14.Estonian Saaremaa	Unofficial strain	Estonia	x		x	
15.Estonian Whitehead	British shortwool type	Estonia	x		x	
16.Finnish Grey Landrace	Nordic group/short-tailed	Finland	x	x		x
17.Finnsheep	Nordic group/short-tailed	Finland	x	x		x[b]
18.Åland Sheep	Nordic group/short-tailed	Finland/Åland		x		x
19.Osetian Tushin	Fat-tailed	Georgia[a]	x			
20.Danish Whiteheaded Marsh Sheep	Continental longwool type	Germany[a]		x		
21.Icelandic Leader Sheep	Nordic group/short-tailed	Iceland		x		
22.Icelandic Sheep	Nordic group/short-tailed	Iceland		x		
23.Russian Edilbai	Fat-rumped	Kazakhstan[a]	x			
24.Latvian Darkheaded	British shortwool type	Latvia		x	x	
25.Lithuanian Blackface	British shortwool type	Lithuania		x	x	
26.Lithuanian Native Coarsewooled	Nordic group	Lithuania		x	x	
27Danish Texel	Continental longwool type	Netherlands[a]		x		
28.Dala Sheep	British type	Norway		x		
29.Feral Sheep	Nordic group/short-tailed	Norway	x	x		
30.Grey Troender Sheep	Nordic group	Norway		x		
31.Old Spael Sheep	Nordic group/short-tailed	Norway	x	x		x
32.Rygja Sheep	British type	Norway		x		
33.Spael Sheep	Nordic group/short-tailed	Norway	x	x		
34.Steigar Sheep	British type	Norway		x		
35.Norwegian Fuglestad Sheep	British type	Norway		x		
36.Olkuska	Continental longwool type	Poland	x			
37.Wrzosowka	Nordic group	Poland	x			
38.Polish Swiniarka		Poland and Central Europe	x			
39.Soviet Merino	Merino Finewooled type	Russia (Caspian Depression)	x			
40.Stavropol	Merino Finewooled type	Russia (N. Caucasus)	x			

Breed or strain	Breed type	Breed origin	Paper			
			I	II, III	IV	V
41.Caucasian	Merino Finewooled type	Russia (N. Caucasus)	x			
42.Karachai	Fat-tailed	Russia (N. Caucasus)	x			
43.Kulunda	Fat-tailed	Russia/Altay	x			
44.Aksaraisk type of Soviet Mutton-wool	British longwool type	Russia/Astrakhan	x			
45.Andi	Fat-tailed	Russia/Dagestan Republic	x			
46.Dagestan local	Unofficial strain	Russia/Dagestan Republic	x			
47.Dagestan Mountain Merino	Merino Finewooled type	Russia/Dagestan Republic	x			
48.Grozny	Merino Finewooled type	Russia/Dagestan Republic	x			
49.Dagestanian Lezgian	Fat-tailed	Russia/Dagestan Republic and Azerbaijan	x			
50.Gorno-Altay local	Unofficial strain/mixed tail types	Russia/Gorno-Altay	x			
51.Vepsia sheep	Unofficial strain, Nordic group/short-tailed	Russia/Karelia Republic	x			x
52.Viena sheep	Unofficial strain, Nordic group/short-tailed	Russia/Karelia Republic	x	x		x
53.Oparin	British shortwool type	Russia/Kirov	x			
54.Komi local	Unofficial strain	Russia/Komi Republic	x			
55.Mari local	Unofficial strain	Russia/MariEl Republic	x			
56.Mordovian local	Unofficial strain	Russia/Mordovia Republic	x			
57.Kuibyshev	British longwool type	Russia/Samara	x			
58.North Caucasus Mutton-Wool	British longwool type	Russia/Stavropol	x			
59.Udmurtian local	Unofficial strain	Russia/Udmurtia Republic	x			
60.Volgograd	Merino Finewooled type	Russia/Volgograd	x			
61.Kuchugur	Fat-tailed (semi)	Russia/Voronezh	x			
62.Russian Romanov	Nordic group/short-tailed	Russia/Yaroslavl	x			
63.Egyptian Romanov	Nordic group/short-tailed	Russia/Yaroslavl[a]		x[c]		x
64.Lithuanian Romanov	Nordic group/short-tailed	Russia/Yaroslavl[a]		x		x
65.Dala Fur Sheep	Nordic group/short-tailed	Sweden		x		
66.Finewool Sheep	Nordic group/short-tailed	Sweden		x		
67.Forest Sheep	Nordic group/short-tailed	Sweden		x		
68.Gotland Sheep	Nordic group/short-tailed	Sweden		x		
69.Gute Sheep	Nordic group/short-tailed	Sweden		x		
70.Roslag Sheep	Nordic group/short-tailed	Sweden		x		
71.Rya Sheep	Nordic group/short-tailed	Sweden		x		
72.Serbian Tsigai	Coarsewool	The Balkans	x			
73.Russian Tsigai	Coarsewool	The Balkans[a]	x			
74.Carpathian Mountain	Coarsewool	Ukraina	x			
75.Sokolsk	Coarsewool	Ukraina	x			
76.Russian Karakul	Fat-tailed	Uzbekistan[a]	x			
Sum			50	32	7	8

[a] Population sampled is located outside the original region of the breed. [b] The white, black and brown colour types of Finnsheep were studied as separate populations. [c] Not included in Paper III.

Table 2. Chromosomal locations (Chr.), primer sequences, annealing temperatures (T_a), GenBank accession numbers (Acc.#),reference and the observed size range (Size) for the 28 microsatellite markers in the original papers III – V. The BM, ILSTS and INRA markers were originally characterized for bovine, while the remaining markers were characterized for ovine.

Marker	Chr.	Primer sequences	T_a(°C)	Acc.#	Reference	Size (bp)	Papers
BM0757	9	5'-TGGAAACAAT GTAAACCTGG G-3' 5'-TTGAGCCACC AAGGAACC-3'	64-55	G18473	Bishop *et al.* (1994)	177-204	III, IV, V
BM0827	3	5'-GGGCTGGTCG TATGCTGAG-3' 5'-GTTGGACTTG CTGAAGTGAC C-3'	64-55	U06763	Bishop *et al.* (1994)	215-229	V
BM1314	22	5'-TTCCTCCTCT TCTCTCCAAA C-3' 5'-ATCTCAAACG CCAGTGTGG-3'	59-50	G18433	Bishop *et al.* (1994)	140-178	III, IV
BM1818	20	5'-AGCTGGGAAT ATAACCAAAG G-3' 5'-AGTGCTTTCA AGGTCCATGC -3'	64-55	G18391	Bishop *et al.* (1994)	250-286	III, IV
BM4621	6	5'-CAAATTGACT TATCCTTGGC TG-3' 5'-TGTAACATAT GGGCTGCATC -3'	64-55	G18529	Bishop *et al.* (1994)	127-175	III, IV
BM6506	1	5'-GCACGTGGTA AAGAGATGGC -3' 5'-AGCAACTTGA GCATGGCAC-3'	59-50	G18455	Bishop *et al.* (1994)	193-219	III, IV, V
BM6526	26	5'-CATGCCAAAC AATATCCAGC -3' 5'-TGAAGGTAGA GAGCAAGCAG C-3'	64-55	G18454	Bishop *et al.* (1994)	144-188	III, IV
BM8125	17	5'-CTCTATCTGT GGAAAAGGTG GG-3' 5'-GGGGGTTAGA CTTCAACATA CG-3'	64-55	G18475	Bishop *et al.* (1994)	106-124	III, IV, V
CSSM31	23	5'-CCAAGTTTAG TACTTGTAAG TAGA-3' 5'-GACTCTCTAG CACTTTATCT GTGT-3'	50-51	U03838	Moore & Byrne (1994)	128-174	III, IV, V
ILSTS002	14	5'-TCTATACACA TGTGCTGTGC -3' 5'-CTTAGGGGTG AAGTGACACG-3'	56-56	L23479	Kemp *et al.* 1992	120-142	III
INRA023	1	5'-GAGTAGAGCT ACAAGATAAA CTTC-3' 5'-TAACTACAGG GTGTTAGATG AACTC-3'	71-62	X67830	Vaiman *et al.* (1994)	197-231	III, IV
MAF214	16	5'-GGGTGATCTT AGGGAGGTTT TGGAGG-3' 5'-AATGCAGGAG ATCTGAGGCA GGGACG-3'	67-58	M88160	Buchanan & Crawford (1992)	134-262	III, IV
MAF36	22	5'-TTGCGAAAAG TTGGACACAA TTGAGC-3' 5'-CATATACCTG GGAGGAATGC ATTACG-3'	64-55	M80519	Swarbrick *et al.* (1991)	99-129	III, IV
MAF48	u.a.	5'-AGACGTGACT GAGCAACTAA GTACG-3' 5'-GGAAACCAAA GCCACTTTTC AGATGC-3'	67-58	M62645	Buchanan *et al.* (1991)	118-142	III, IV
MAF65	15	5'-AAAGGCCAGA GTATGCAATT AGGAG-3' 5'-CCACTCCTCC TGAGAATATA ACATG-3'	67-58	M67437	Buchanan *et al.* (1992)	119-141	III, IV

Marker	Chr.	Primer sequences	T_a(°C)	Acc.#	Reference	Size (bp)	Papers
McM527	5	5'-GTCCATTGCC TCAAATCAAT TC-3' 5'-AAACCACTTG ACTACTCCCC AA-3'	67-58	L34277	Hulme *et al.* (1994)	159-183	III, IV
OarAE129	5	5'-AATCCAGTGT GTGAAAGACT AATCCAG-3' 5'-GTAGATCAAG ATATAGAATA TTTTTCAACA CC-3'	59-50	L11051	Penty *et al.* 1993	137-163	V
OarCP20	21	5'-GATCCCCTGG AGGAGGAAAC GG-3' 5'-GGCATTTCAT GGCTTTAGCA GG-3'	65-56	U15695	Ede *et al.* (1995a)	70-96	III, IV, V
OarCP34	3	5'-GCTGAACAAT GTGATATGTT CAGG-3' 5'-GGGACAATAC TGTCTTAGAT GCTGC-3'	59-50	U15699	Ede *et al.* (1995b)	111-129	III, IV, V
OarCP38	10	5'-CAACTTTGGT GCATATTCAA GGTTGC-3' 5'-GCAGTCGCAG CAGGCTGAAG AGG-3'	65-56	U15700	Ede *et al.* (1995b)	112-134	III
OarFCB11	2	5'-GGCCTGAACT CACAAGTTGA TATATCTATC AC-3' 5'-GCAAGCAGGT TCTTTACCAC TAGCACC-3'	67-58	L01531	Buchanan & Crawford (1993)	121-143	III, V
OarFCB128	2	5'-CAGCTGAGCA ACTAAGACAT ACATGCG-3' 5'-ATTAAAGCAT CTTCTCTTTA TTTCCTCGC-3'	64-55	L01532	Buchanan & Crawford (1993)	95-131	III, IV, V
OarFCB304	19	5'-CCCTAGGAGC TTTCAATAAA GAATCGG-3' 5'-CGCTGCTGTC AACTGGGTCA GGG-3'	64-55	L01535	Buchanan & Crawford (1993)	146-198	III, IV, V
OarFCB48	17	5'-GAGTTAGTAC AAGGATGACA AGAGGCAC-3 5'-GACTCTAGAG GATCGCAAAG AACCAG-3'	67-58	M82875	Buchanan *et al.* (1994b)	136-172	III, IV, V
OarHH35	4	5'-AATTGCATTC AGTATCTTTA ACATCTGGC-3' 5'-ATGAAAATAT AAAGAGAATG AACCACACGG-3'	59-50	L12554	Henry *et al.* (1993)	116-136	V
OarHH47	18	5'-TTTATTGACA AACTCTCTTC CTAACTCCAC C-3' 5'-GTAACTTATT TAAAAAAATA TCATACCTCT TAAGG-3'	67-58	L12557	Henry *et al.* (1993)	126-152	III, IV, V
OarHH64	4	5'-CGTTCCCTCA CTATGGAAAG TTATATATGC-3' 5'-CACTCTATTG TAAGAATTTG AATGAGAGC-3'	67-58	L12558	Henry *et al.* (1993)	120-138	III
OarVH72	25	5'-CTCTAGAGGA TCTGGAATGC AAAGCTC-3' 5'-GGCCTCTCAA GGGGCAAGAG CAGG-3'	64-55	L12548	Pierson *et al.* (1993)	125-141	III, IV, V

3.2 Data analysis

3.2.1 Mitochondrial DNA

3.2.1.1 Phylogenetic relationships

Phylogenetic trees describing relationships among mtDNA haplotypes were constructed using the neighbour joining method (Saitou & Nei 1987) (I), Bayesian method (Ronquist & Huelsenbeck 2003) (I) and maximum likelihood method (Swofford 2003) (II). More detailed relationships among very similar haplotypes were estimated with median-joining (Bandelt *et al.* 1999) (I) and statistical parsimony methods (Templeton *et al.* 1992) (II). Differences between sequences were measured as fractions of different nucleotides between them (I, II). The population differentiation was estimated as $F_{st} = 1 - (d_w/d_b)$, where d_w is number of nucleotide differences between sequence pairs from the same populations and d_b is the number of differences between populations (Hudson *et al.* 1992).

3.2.1.2 Detecting population expansions

Presence of past population expansion was deduced from the distribution of pairwise differences between all pairs of sequences (i.e. mismatch distribution) (I, II). Population expansion creates a star-shaped genealogy, where the evolutionary paths coalesce immediately before the expansion. Therefore, if the expansion begins from a monomorphic population, the mean number of pairwise differences between sequences is equal to twice the number of mutations occurring after expansion in each lineage. When the original population was not monomorphic, the time since expansion can be estimated (together with the initial and final population sizes) using a least-squares approach (I). As a second method, Fu's Fs test was used to detect expansion (Fu 1997) (I). The test compares the observed haplotype number to the observed number of pairwise differences. This test is highly sensitive to population growth. The methods are described in detail in papers I and II.

3.2.1.3 Analysis of geographical distribution of mtDNA variation

In paper I, geographical distribution of mtDNA types grouped to nodes of a phylogenetic network (i.e. phylogeographic analysis of mtDNA variation) was studied using a permutational contingency test, where geographical area was used as a categorical variable. In paper II, a detailed Nested Clade Analysis was performed to distinguish significant spatial patterns in the distribution of mtDNA variation. More detailed descriptions are provided in papers I and II.

3.2.2 Autosomal markers

3.2.2.1 Analysis of genetic diversity within and among breeds

Three measures of molecular variation were used. Gene diversity or expected heterozygosity (Nei 1987) is the probability that two random alleles of a gene are different. Gene diversity reflects immediate potential response to selection. The second applied measure, the number of alleles or allelic richness, quantifies the extent to which traits can be changed in the long-term (e.g. Bataillon *et al.* 1996). The third measure, mean allele-sharing distance between individuals, explores the current distribution of variation based on multilocus genotypes. This measure can identify cases where individuals differ simultaneously at multiple unlinked loci.

In papers IV and V, gene diversity was estimated for the population k from: $H_k = (N/(N - 1))(1 - \Sigma_i p_i^2 - H_{ok}/2N_k)$, where N is the number of individuals in the sample, p_i is the sample frequency of the i-th allele, and H_{ok} is the observed homozygosity in the population k (Nei 1987, p. 164). The general within-population gene diversity (H_s or h_s) was the arithmetic average of population-wise estimates. The total gene diversity for the pool of studied populations was estimated from: $H_t = 1 - \Sigma i\ P_i^2 + H_s/(\tilde{N}s) - H_o/(2\tilde{N}s)$, where P_i is the mean allele frequency of i-th allele, $\tilde{N}$ is the harmonic mean of individuals in the samples, s is the number of subpopulations sampled, and H_o is the observed homozygosity (Nei 1987, p. 191).

In paper III, gene diversity estimation was modified in two ways. In this paper, the gene diversity was estimated assuming no within-breed subdivision. Secondly, the total diversity was estimated not only for the studied populations, but also for a theoretical wider set of populations (i.e. metapopulation). The studied populations were assumed to be a random sample from this metapopulation. In this case, gene diversity for population k was estimated from: $h_k = (n/(n - 1))(1 - \Sigma_i p_i^2)$, where n is the number of gene copies in the sample (Nei 1987, p. 191; Pons and Petit 1995, p. 463). The total gene diversity for the metapopulation was estimated from $h_t = 1 - (1 / s(s - 1))\ \Sigma_{k \neq l} \Sigma_i p_{ki} p_{li}$, where p_{ki} and p_{li} are the sample frequencies of the i-th allele in population k and l, respectively (Pons & Petit 1995, p. 463).

Sample size corrected allele number or allelic richness ($r[g]$) is the expected number of alleles in a sample of specified size (g), where g cannot exceed the actual population sample size (n). The expected allele numbers were calculated to correspond to the smallest actual breed sample in each study. Sample size correction uses binomial coefficients and in papers III, IV and V it was obtained from:

$$r(g) = \sum_i \left[1 - \binom{n - n_i}{g} \Big/ \binom{n}{g} \right],$$

where n_i is the number of occurrences of the i-th allele. For a population k this formula can be applied directly to obtain r_k. The general within-population allelic richness (r_s) was calculated as the average of population-wise values. The effect of sample size variation was removed in calculation of allelic richness for the total population (r_t): the

values of n and n_i for calculating r_t were weighted sums, weighting the populations inversely to their sample sizes. The n and n_i were rounded to the closest integer for the use of binomial coefficients. (El Mousadik & Petit 1996.)

The empirical allele-sharing distance between individuals is $d = 1 - P$, where the proportion (P) of common alleles (A) over L loci is $A/2L$ (Bowcock *et al*. 1994). Further details about analysis of mean allele-sharing distance can be found in paper III.

In paper III, the proportion of total variation due to genetic differentiation among populations was based on the above described diversity estimates of mean within-population diversity and total diversity. In papers IV and V, differentiation was quantified using the analysis of variance method of Weir and Cockerham (1984). Five genetic distance measures, Nei's standard distance (D_S; Nei 1972), D_A distance (Nei *et al.* 1983), Chord distance (Cavalli-Sforza & Edwards 1967) and D_{TL} distance (Tomiuk & Loeschcke 1995) and allele sharing distance (Bowcock *et al*. 1994), were calculated to estimate divergence of populations using different datasets (III, IV, V).

Genetic structures among the breeds were examined based on genetic distances between the breeds and on allele frequency data directly. Genetic distances were used to construct neighbour-joining trees (Saitou & Nei 1987) (V) and two-dimensional maps, which were constructed using multidimensional scaling (V) and principal coordinates methods (e.g. Everit & Dunn 1991) (IV). Allele frequencies were used to perform independent components analysis according to Hyvärinen (1999) (III).

The "genetic populations" or demes (i.e. a set of individuals that can be grouped without causing deviation from panmixia in the group) were constructed using Bayesian clustering methods (Prichard *et al.* 2000, Corander *et al.* 2003) (IV, V). In addition, clusters were searched using the neighbour-joining method (Saitou & Nei 1987) and independent components analysis (Hyvärinen 1999) for individuals (III, IV).

Estimation of single breed influence on the total variation of the metapopulation, within-population inbreeding (i.e. autozygosity or identity by descent) and within and between population co-ancestries reflect simultaneously features of within population variation (within a population gene diversity is the complement of co-ancestry) and population divergence. In paper III, the method of Ciofi *et al.* (1999) was used to estimate the level of within-population inbreeding from a molecular data. Paper III also includes assessment of single breed importance for the variation in the metapopulation. In this summary section of the thesis, the method of Caballero and Toro (2002) was used to calculate between breed molecular co-ancestry ($f_{kl} = \Sigma_i p_{ki} p_{li}$) and related statistics.

3.2.2.2 Detecting deviation from panmixia

Deviation from panmixia, where the genetic variation is distributed randomly to individuals, was studied using two methods: deviation from Hardy-Weinberg expectations for genotype frequencies within loci (III, IV, V), and deviation from random association of alleles in pairs of loci (i.e. genotypic linkage disequilibrium) (IV, V).

3.2.2.3 Detecting reduction in effective population size

Reduction in effective population size does not immediately have a large effect on gene diversity, but the allele number is strongly affected because the rare alleles are lost. The test of Cornuet and Luikart (1996) to compare gene diversity and allele number was used to detect recent reduction in effective population size (III, V). When the effective population size is reduced to size N_r, the allele number and gene diversity reach new equilibrium in 2–4 N_r generations (Luikart *et al.* 1998). During these generations, comparison of allele number and gene diversity may reveal the past population reduction.

Even long after the 2–4 N_r generations following the population reduction, the most recent common ancestor of the alleles in the reduced population is further in the past than in a population that has been of the same (N_r) size over a much longer period. General stepwise mutation model, where mutation changes the allele size by one repeat unit, is a simplified model describing the evolution of microsatellites. Under this model, the time to the most recent common ancestor of a random pair of microsatellite alleles, and the number of mutations in this genealogy, is proportional to the difference in the size of the alleles (Goldstein *et al.* 1995). The variance of the allele sizes is related to depth of the genealogy. Low gene diversity compared to allele size variance reveals population size reduction in the past. In paper III, the comparison of these two diversity measures was done according to Kimmel *et al.* (1998).

4 Results and discussion

4.1 Origin of sheep mtDNA lineages

4.1.1 Sheep domestication and arrival in Europe

Four distinct groups of mtDNA haplotypes (A, B, C and a novel group D) were observed in Eurasian sheep. The group structure is evidenced by well-supported clustering in the phylogenetic tree and clear structure in the nucleotide differences between haplotypes. Differences are over six times more extensive between than within the groups (I). The presence of highly diverged lineages has been inferred as evidence of multiple independent domestication events (e.g. Pedrosa *et al.* 2005). The geographical distribution of mtDNA variation in domestic sheep corresponds to expansion from the Near East, where they were derived from wild sheep. The derivation of both common groups (A and B) occurred at approximately the same time, ~9 000 years ago. Derivation of distinct groups approximately at the same time and place makes their fully independent derivation from wild sheep unlikely. The derivations may have been performed by the same people or the derivations may have been connected, for instance, through spread of the skills required in sheep management. Other haplotype groups are rare and they were likely introgressed into the domestic stock later. Mitochondrial DNA variation north of the hypothesised domestication centre can be explained by a single domestication area (the central part of the Fertile Crescent), but wider sampling in Asia is needed confirm this (I).

Domestication of small numbers of ewes or reduced population size immediately after domestication caused a strong bottleneck in the maternal lineages. This is seen as null estimates for the population size before the expansion (I) and suggests that the mitochondrial variation in domestic sheep originates from a small number of mtDNA lineages. This result differs from that based on autosomal microsatellites, where the inferred founder-population for northern sheep does not seem to have had a strong genetic bottleneck in its ancient past (III). The results fit to the idea that the original domesticated populations were repeatedly restocked and backcrossed with the wild

populations (Vilà *et al.* 2005). Moreover, the results suggest that no extreme autosomal bottleneck has separated the northern European sheep from the wild ancestor.

The two most common haplotype groups, A and B, are found in Europe. The hypothesis of an independent European domestication of sheep (e.g. Ryder 1983, p. 23-24) was not addressed by the earlier studies on sheep mtDNA variation (Hiendleder *et al.* 1998b, Hiendleder *et al.* 2002). These previous studies showed that Group B predominates in European sheep populations, and is rare in areas distant from Europe, and only Group B has been observed in the European mouflon (Hiendleder *et al.* 1998b, Hiendleder *et al.* 2002). However, the founding of a European population appears to have taken place approximately 3 000 years after the domestication, and after Group B had become more numerous in the Near East (I). The later expansion of this haplotype group in Europe is evident from smaller within-group differences in Europe than in the Caucasus. The predominance of Group B in Europe most likely resulted from a genetic bottleneck affecting the female lineages when the European sheep population was formed from the Near Eastern domesticated sheep. Furthermore, the European mouflon can be considered to be a feral sheep rather than a truly wild sheep (Poplin 1979). Thus the maternal lineages in European sheep originate from the domesticated sheep in the Near East (I)

4.1.2 The northern European mtDNA lineages

Group B represents the majority in northern Europe (II) as in the other parts of Europe (I, Hiendleder *et al.* 1998b, Hiendleder *et al.* 2002). Additionally, the reduced differences between sequences in Group B compared with those in Group A are evident in northern European sheep (II) resembling the finding in general eastern European sheep population (I). Similar findings regarding sheep from different areas of Europe suggest that the Group B lineages in Europe have a recent common ancestry in Europe, not only remotely in the Near East.

Group A was recorded as a rare type in several northern European sheep breeds (II). Large within-group differences suggest that the group originates from the Near Eastern expansion without an extreme genetic bottleneck. Group A has most likely arrived in the Nordic countries from the east. The Nordic haplotypes in Group A are closely related to the haplotypes detected in the Middle Volga region, which suggests that some maternal gene flow took place between the Near East and Nordic countries across Russia (I). Group A group has not been imported to Nordic countries very recently because the haplotype group was found in many old native breeds (II). Occurrence of this haplotype group may represent ancient influence from east or more recent influence of the Viking trade through the Volga River. Vikings are documented to have imported sheep from the Turkic speaking people that lived around the Middle Volga region (Logan 1991, p. 193). Wider investigation of sheep in central Europe is needed to determine if the haplotype group has arrived in all the Circum Baltic areas through the same route. Another route is possible because the Nordic haplotypes are different from those detected on the southern coast of the Baltic Sea (I).

Spatial analysis of mtDNA variation of sheep in northern Europe (II) did not identify a pattern caused by range expansion over the whole of northern Europe. The recognized patterns in the region commonly matched those caused through isolation by distance. There may be two explanations for this. Gene flow after the initial spread may have redistributed the variation, making detailed inferences about the more distant past very difficult. On the other hand, agriculture and animal husbandry may have spread almost instantaneously over the whole continental part of northern Europe, even if it did not become important in the entire area at that time (Zvelebil & Lillie 2000). Therefore, there may not have been time for emergence of new mtDNA variants, whose distribution could reveal the range expansion.

The spatial expansion of sheep to the North Atlantic islands is seen in the distribution of mtDNA types (II): the recent types (tips in the genealogical network) have a wide geographical distribution or the ancestor types (internal nodes in the network) have a reduced geographical distribution. The spatial analysis of mtDNA type distribution, as well as the migration rate analysis, supported the idea of sheep arriving in Iceland through the Faeroe Islands. More distant history is unclear. Sheep along all the coasts of the North Sea share mtDNA types that are absent or very rare elsewhere and have geographically reduced distribution compared with their immediate ancestor types. Thus the data support the idea of common recent ancestry for the sheep along the North Sea, but the analysis is not able to pinpoint a single location from where the expansion to the North Atlantic islands began. It is possible to demonstrate the historical geographic shift in the location of mitochondrial lineages by exploring the geographical centre points for related mtDNA types (a more detailed spatial analysis is presented in paper II). Starting from the most common haplotype in the Icelandic sheep (haplotype 51 in paper II) and stepwisely considering a group of more divergent haplotypes, the centre point that is initially in the Faeroe Islands moves to the Shetland Islands and finally to southern Scandinavia. However, the historical spread of sheep did not necessarily follow this route in reverse order: expansion may have included transportation of sheep in several directions even if the north-westward range expansion was the net result. Moreover, the final geographical centre point from this type of grouping is the centre of sampling sites.

4.2 Distribution of autosomal genetic variation

4.2.1 Breed relationships

Differentiation among the 32 northern European sheep breeds was explored with a multivariate analysis of microsatellite allele frequencies. The first component separates short-tailed and long-tailed breeds. The plot of breeds based on two components in variation match the geographical origins of breeds. Two clear discrepancies were the clustering of *the Swedish Finewool Sheep* with Finnish and Russian Karelian sheep and the clustering of *the Norwegian Spael Sheep* with *the Icelandic Sheep*. However, these correspond to well-known recent influences on *the Finewool* of *the Finnsheep*, and on *the*

Spael Sheep of *the Icelandic Sheep*. Clustering of the studied long-tailed breeds reflects their common origin in regions southwest of Scandinavia (III).

The two components explain only 16% of the differences among the breeds, and the macro-scale structure explains a substantial proportion of microsatellite variation only in the northernmost breeds. The explained proportion is two thirds of that observed in approximately corresponding European cattle studies where similar numbers of breeds were analysed (e.g. Cañón *et al.* 2001). The reason for a weak structure in the analysis of sheep breeds was obvious from the multivariate analysis done at the individual level. Influential origins causing divergence are frequently specific to a single breed or to a group of few breeds. In many cases, breeds within a single country or in neighbouring geographical areas demonstrate common origins (e.g. *the Icelandic Sheep* and related breeds, Finnish sheep and sheep in Russian Karelia, the Norwegian native breeds), while in other cases the shared origin results from recent crossbreeding (in particular, *the Rygja Sheep* shares an origin with several distinct breeds). The present-day breed differentiation results from relatively recent rather than from very ancient divergence (III).

In conclusion, although the macro-scale structure in breed relationships is weak, it corresponds to the historical locations of the breed origins. These locations of the founding populations represent the basis on which to understand the similarities among breeds.

4.2.2 Degree of breed differentiation

Molecular variation can be divided into within and among-breed variation. The standard measure of among-breed variation, or breed differentiation, is the proportion of total gene diversity that results from allele frequency differences among populations (G_{ST}, e.g. Nei 1987). The differentiation seen at microsatellites is extensive among the 32 northern European sheep breeds (III), and it explains 15% of total gene diversity (0.78). The proportion is larger than observed among the sheep breeds in Spain (6%, Àlvarez *et al.* 2004) or in East Europe and the Caucasus (5%, Ozerov *et al.* 2005) or in Africa (9%, Muigai 2002). Only one microsatellite study in sheep has reported stronger differentiation (17%), but this observation traces back to the mouflon population included in that study (Stahlberger-Saitbekova *et al.* 2001). Most microsatellite studies have suggested weaker differentiation in other domestic species (e.g. Kantanen *et al.* 2000, Laval *et al.* 2000, Cañón *et al.* 2001), but a recent Europe wide study on pig recorded even stronger differentiation (21% of total gene diversity ~0.70, SanCristobal *et al.* in press). We may conclude that differentiation accounts for an unusually large proportion of total variation in the northern European sheep breeds.

Evaluating the differentiation of the recently discovered sheep types was one of the practical questions motivating the current research. In Finland, *the Åland Sheep* and *the Grey Finnish Landrace* were discovered in the 1980s based on distinctively “primitive” appearances and geographically limited occurrences. Microsatellite and blood protein variation show a clear difference between the common *Finnsheep* and *the Åland Sheep* (III, V), and relatively clear differences also between *the Grey Finnish Landrace* and *the White*, *Black* or *Brown Finnsheep* types (V). Swedish rare sheep types are in a similar

situation. *The Dala Fur Sheep*, *the Forest Sheep* and *the Roslag Sheep* are "semi-recognized" breeds grouped under the term "peasant breeds". The peasant breeds diverge from each other and from the other studied breeds (III). Even the existence of a local sheep type in Estonia was unclear. The microsatellite study showed strong differentiation between flocks, and the two dimensional plot also separated part of the local flocks (including *the Ruhnu sheep*) from the more common modern production breeds (IV). This may indicate a local type. In general, the local sheep strains seem to be at least as "unique" (i.e. differentiated from the other breeds) as an average sheep breed in northern Europe (III).

Gene diversity is not the only measure of variation. For conservation purposes, mean number of alleles or allelic richness has been suggested to be a more important measure of variation than gene diversity (Bataillon *et al.* 1996, El Mousadik & Petit 1996). In northern European sheep, the breed differentiation is more than twice as important for allelic richness (*r[26]*) than for gene diversity (III): on average 37% of total allelic richness in microsatellites was due to breed differentiation. The median value of the proportion of total variability contained in a single breed was 68% for allelic richness (range over breeds 35 – 84%) and 85% for gene diversity (range 49 – 97%). The larger the sample size considered, the larger the expected difference between gene diversity and allelic richness based statistics as even rarer alleles will influence the allelic richness statistic (El Mousadik & Petit 1996). Breed differentiation is clearly more important if allelic variation rather than gene diversity is considered.

The average differentiation based on allele-sharing distance between individuals (17%) was very close to that based on gene diversity. Within-breed structure increases the within-breed estimate, and disappearance of substructure decreases it. Quantifying variation based on multi-locus genotypes may be the most descriptive method for the current generation because multi-locus data includes information about the distribution of alleles. However, considering the following generations, gene diversity and allelic richness are more suitable diversity measures because they are not as sensitive to the changing substructure.

4.2.3 Breed subdivision

Breeds are not always uniform panmictic populations completely separate from other such populations (i.e. demes). Several breeds were subdivided into groups based on a general excess of homozygotic genotypes (III, IV, V). The strongest subdivision was observed in unofficial local sheep strains. In the Estonian *Saaremaa sheep*, the strong within-farm inbreeding causes the differences among the flocks of the same strain to be larger than the differences among the officially recognized breeds in the area (IV). In the unofficial strains of Russian Karelia, the subdivision was less extreme because the sheep were subdivided into "village demes" rather than into farm flocks (V). The subdivision is not restricted to unofficial strains. In recognized breeds specific types may be distinguished from the main stock (e.g. *the Leadersheep* in the Icelandic Sheep, III). In *the Finnsheep*, maintenance of recessive wool colour in the flocks leads to assortative mating (V), which is too weak to cause a general excess of homozygotic genotypes (III,

V). In general, over one third of the northern European sheep breeds demonstrate strong subdivision that is detectable as deviation from Hardy-Weinberg proportions (III). Strong within-breed structure has been detected also in other studies of the European sheep breeds (e.g. Diez-Tascón *et al.* 2000, Pariset *et al.* 2003). Therefore, breed uniformity cannot be taken for granted in small ruminants. Subdivision, together with small population size and immigration, causes non-random associations both at and between loci (IV, V). The deviations from panmixia need to be taken into account in breed studies. They may cause, for example, false associations in linkage studies if population stratification is not properly addressed (Pritchard & Donelly 2001).

The between-breed differentiation is not, strictly speaking, between-population differentiation because breeds are not always panmictic populations (i.e. demes). Breed subdivision deflates differentiation estimates between them (e.g. Balloux & Lugon-Moulin 2002). In the breeds of the Baltic countries (IV), 9% of the total gene diversity is due to breed divergence. The degree of differentiation rose to 13% when the analysis was based on the panmictic demes, which were deduced from multilocus genotypes of individuals. The difference was due to both separating flocks of local strain into separate demes and joining undifferentiated breeds into a common deme (IV).

Deme specific divergence can be detected using multivariate analysis done at the individual level. In paper III, this type of multivariate analysis was used to dissect the variation among multi-locus genotypes. The explanatory components were presented as the proportion of explained variance in the allele frequency matrix for an individual (hereafter "variance"). The components reveal structuring. For example, in *the Swedish Forest Sheep* the main component represented less than 2.5% of the variance in ten sheep, while it represented on average 31.5% of the variance in the other 28 sheep (unpublished data). The ten less diverged sheep were sampled on a single farm. When the proportion of explained variance was averaged over all the studied *Forest Sheep* individuals, the component explained 23.5% of the variance. This result shows how averaging divergence over the breed sample may result in an imprecise characterization of gene pool structure. The multivariate analysis suggested also that there are diverged groups within *the Swedish Dala Fur Sheep* and the Norwegian short-tailed breeds, while the other breeds had approximately unimodal distributions of the component values (unpublished data).

Considering all the components of the multivariate analysis for individuals in paper III, the average proportion of explained variance for a breed correlates very closely ($r = 0.96$) with breed uniqueness measured as G_{ST}. Despite the different approach to the question, both measures describe how distinct the individuals in the breed are from a general pool of individuals. The ranges of the two measures differed: explained variance for a breed was between 1 and 56%, while G_{ST} ranged between 10 and 34%. The difference results from standardization of allele frequencies in multivariate analysis, which was done to increase the impact of rare alleles. Shared rare alleles are more informative about common origins than shared widely spread alleles. The proportion of within-individual variance explained by a single component is very large in highly inbred breeds, resembling the behaviour of G_{ST} (III).

The uniqueness and variability of the populations of imported international breeds are rarely discussed. Among the sheep in northern Europe, the Norwegian *Cheviot Sheep*, the Danish *Texel* and the Finnish *Oxford Down* are highly diverged from other breeds and

have less within-breed variation (III, V). At least in some cases, this is caused by a small number of imported sheep used in founding the population (V). Comparison of the Egyptian and Lithuanian *Romanov* and Finnish and Russian Karelian sheep showed that *Romanov* populations in different countries are diverged, but resemble each other more than any other breed (V). The same pattern has been demonstrated for a widely spread cattle breed (Blott *et al.* 1998).

4.2.4 Comparison between autosomal marker types

For very variable markers the proportion of gene diversity due to population divergence is expected to be smaller than for less variable loci. However, observations based on microsatellites and protein markers have suggested similar degrees of differentiation independent of variability. (Reviewed by Chikhi & Bruford 2005.) The sheep breeds in Finland and in Russian Karelia demonstrated stronger differentiation for blood proteins (15%) than for microsatellites (8%). This was caused by two bottlenecked breeds (*the Oxford Down* in Finland and *the Åland Sheep*) and was not a general feature. Most breeds showed similar degrees of differentiation for both marker types. *The Vepsia sheep* was clearly less diverged from other populations for proteins than for microsatellites. This is related to the subdivision of the strain causing underestimation of divergence. Immigration from distantly related breeds may also contribute because *the Vepsia sheep* had all the blood protein alleles that were observed in the other studied populations and two unique albumin alleles. The constructed breed relationships based on the two marker types agreed with each other, except for *the Vepsia sheep*. In general, the microsatellite data seem to match with the protein polymorphism with few exceptions caused by population history and structure.

4.3 Inbreeding in northern European sheep

4.3.1 Accumulated inbreeding and changes in the inbreeding rate

Avoiding inbreeding and the consequent loss of variation and undesired phenotypic effects from recessive alleles is one of the main issues in conservation genetics. Inbreeding may increase environmental sensitivity (Frankham 2005, Kristensen *et al.* 2005) and reduce the ability to respond to selection (Lynch & Hill 1986, Garner *et al.* 2005). Molecular data were used to infer allele frequencies for a founder-population of northern European sheep and to estimate the accumulated within-breed inbreeding since the split into different breeds. The median value for the autozygosity (i.e. proportion of allele pairs that are identical by descent within the breed) was 0.12 and the values ranged from 0.07 to 0.41 over breeds (III). This is similar to autozygosity estimated in Spanish sheep breeds, where the median was 0.10 and the range was 0.04 – 0.50 (Álvarez *et al.* 2004).

There is considerable variation in most of the northern European sheep breeds (III, IV, V), but some breeds are quite inbred. *The Roslag Sheep*, *the Ruhnu sheep*, and *the Dala Fur Sheep* exhibit the highest within-population inbreeding (III). Similarly, the flocks of an undefined Estonian sheep strain demonstrate low variability (IV). Additional variation from a related breed could be introduced if low variability or fixation of deleterious alleles (i.e. genetic load, Hedrick 2001) appears to decrease adaptation potential or vitality of a breed. In wild species conservation aims at maintaining networks of populations connected by "natural" gene flow, reducing inbreeding (Crandal *et al.* 2000). In principle, maintaining historical gene flow between geographically and genetically neighbouring breeds would be the ideal situation in domestic species as well. There is a practical evaluation problem and an adaptation maintenance problem in this. Inferring the extent of historical gene flow between breeds is difficult because the domestic populations are not generally in migration-drift equilibrium (III). Analysis based on an erroneous equilibrium assumption (resulting in estimates of 0.2 - 3.1 immigrants per generation across the 32 breeds, data not shown) can lead to overestimation of "natural" gene flow. Secondly, although breed conservation may aim at maintaining genetic adaptation to a specific environment, the selection pressure might not exist currently. This leaves the adaptation particularly vulnerable to erosion by admixture. Because introgression may further bring new deleterious alleles, it can be justified for rare breeds, at least in the short-term prior to restoration of a population, only if there is evidence of substantial impairment caused by inbreeding.

Not all highly inbred populations demonstrate reduced viability (Holm *et al*. 1999; Visscher *et al*. 2001). Slowly acquired inbreeding has a milder impact than the quickly acquired one (Kristensen *et al.* 2005), and rapid acceleration in the rate of inbreeding is the most threatening genetic change for a conserved population (e.g. Luikart *et al.* 1998). Rate of inbreeding is inversely related to the effective population size (Wright 1931). All northern European breeds have experienced a population size reduction, which is seen as imbalance between allele size variance and gene diversity (III). However, most of these reductions may be old, and do not pose a new threat to the breeds. Low allele number compared with gene diversity suggests that recent population size reduction has occurred in three breeds: the Estonian *Ruhnu sheep*, *the Grey Finnish Landrace* and the imported *Oxford Down* in Finland (III, V). The possible undesirable phenotypic changes caused by their recent population size reduction will be seen later, if new fixations occur. Reducing the risk of fixations of deleterious alleles through planned mating and enlarging the population are most urgently needed in these three populations.

4.3.2 Phenotypic effects of inbreeding

The phenotypic effects of ancient population reductions and the resulting differences in autozygosity should be visible, but relating phenotypic differences to inbreeding is not easy. The difference in the realized inbreeding between two populations is visible as a difference in the average gene diversity over loci. The difference varies across loci (V), which calls for the use of a large number of loci to establish a difference in the amount of gene diversity or inbreeding. Variation across the loci also implies that inbreeding

depression occurs repeatedly only in traits that can be distorted by fixation of deleterious alleles at various loci. These are particularly the complex or fitness traits.

Fixation of deleterious alleles can influence prolificacy. The exceptional prolificacy of the Finnsheep and the Romanov appears to have a polygenic origin (e.g. Maijala 1984), and inbreeding has a negative effect on fertility in domestic sheep (Ercanbrack & Knight 1991) as well as in wild species (DeRose & Roff 1999). Simple comparison between within-breed autozygosity and mean litter size across the studied breeds (Fig. 3) suggests that inbreeding explains 34% of the variation in prolificacy across the breeds. The relationship is not linear, most likely because litter size cannot be smaller than one, when considering successful gestations without requiring successful rearing. The decreasing trend implies that low inbreeding is needed for high prolificacy, and that very efficient maintenance of within-breed variation is required to preserve the exceptional fertility of the prolific northern European breeds. In principle, realized differences in the reproduction success affect the amount of genetic variation maintained (e.g. Wright 1938), but reproduction success is mainly a management question and does not need to be related to prolificacy. Since the effect of inbreeding has been demonstrated experimentally, it is more parsimonious to assume that inbreeding decreases prolificacy than that high prolificacy decreases inbreeding. At the same time it should be noted, that the degree of inbreeding is not the only determining factor. For instance, single gene variants with large positive effect on fertility have been found in many breeds (Davis 2004).

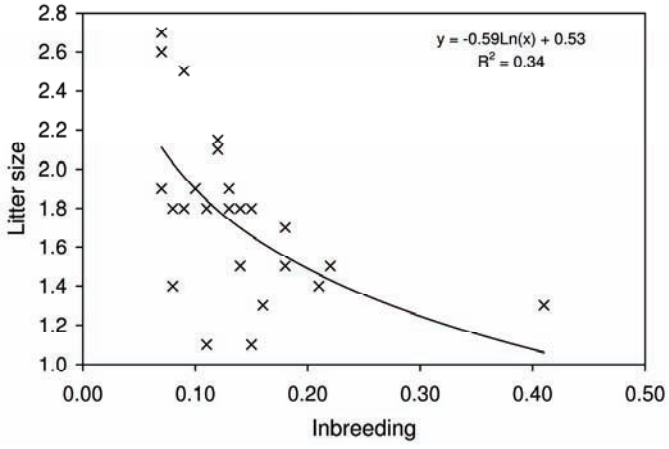


Fig. 3. Correlation between mean litter size and within-breed inbreeding in a subset of the breeds studied in paper III. No litter size estimates were available for nine breeds (Åland Sheep, Estonian Ruhnu sheep, Finnish Grey Landrace, Lithuanian Blackface, Norwegian Feral and Old Spael Sheep, Russian Viena sheep, Swedish Dala Fur Sheep).

4.4 Choices among the breeds for conservation

4.4.1 Unit of conservation

In wild species, statistically significant differentiation is used as a criterion to identify separate management units in conservation programmes (Moritz, 1994). Bayesian clustering (Corander *et al.* 2003) of breeds was performed to define similar types of unit in the 32 northern European sheep breeds. The analysis joins *the Finnsheep* and the Swedish *Finewool Sheep*, suggesting a weak differentiation between the breeds (unpublished data). The Swedish *Finewool Sheep* also contains ancestry that is foreign to *Finnsheep* (III). Based on these results, there is no need to restrict gene flow from *the Finnsheep* to the Swedish *Finewool Sheep*. Similarly, weak differentiation among the popular production breeds in Estonia and Lithuania would warrant their joint management (IV). For other recognized breeds, the breed itself is the appropriate unit of conservation.

The subdivision of unofficial strains of Estonia (IV) and Russian Karelia (V) illustrates why the above described idea of management unit based solely on molecular variation is not applicable in undefined strains. It is not desirable to define a flock as a separate endangered "breed" requiring intensive conservation measures because the differentiation may be very recent and there is no evidence that such a flock would contain unique adaptations. Molecular analysis of population structure and origin, combined with phenotypic evaluation, should be used to distinguish the sets of flocks that could form separate breeds carrying special alleles or traits.

4.4.2 Ranking of breeds by their contribution to total variability and choice of breeds for an enriched assemblage

There is a need to rank sheep breeds for conservation, especially when considering the most intensive or expensive conservation measures (e.g. cryo bank creation). In paper III, the most important breeds for conservation were identified by evaluating breed contribution to total variability, which is the influence of a single breed on the variability of the theoretical metapopulation composed of all the breeds (Petit *et al.* 1998, Caballero & Toro 2002). The assessment was based on several diversity measures (gene diversity, allelic richness and mean allele-sharing distance). An obvious feature was the opposite influence of breed divergence and within-breed variability. The effects from these two factors neutralized each other to a great degree, but some variation in the total breed contributions remained. Nine breeds (*the Greenland Sheep*, *the Icelandic Sheep*, *the Lithuanian Native Coarsewooled*, *the Norwegian Fuglestad*, *Rygja* and *Old Spael Sheep*, *the Russian Viena sheep*, and *the Swedish Gute* and *Roslag Sheep*) were consistently recognized as important breeds. A further six breeds (*the Ruhnu sheep*, *the Finnsheep*, *the Spael Sheep*, *the Romanov*, *the Dala Fur Sheep*, and *the Rya Sheep*) were recognized as important contributors to the total genetic variation when results from gene diversity,

allelic richness and mean allele-sharing distance were combined (III). These 15 breeds make up the "*diversity contribution set*".

The assessment of breed contributions to total variation resembles the use of low mean-kinship as a criterion to identify the most important individuals for maintaining variation within a population (Ballou & Lacy 1995). This is evident from comparison between the mean molecular co-ancestries of the breeds (Appendix I) and the contributions of the breeds to gene diversity (III). A highly ranked breed has a low mean co-ancestry with all the breeds, including itself. In other words, its genetic variation is under-represented in the metapopulation (Ballou & Lacy 1995). In general, such a breed can be used to increase the within-breed variability of the other breeds. On the other hand, the introgression from a breed making a negative contribution may lower the within-breed variation in other breeds. Reduction of within-population variation by gene flow may seem counterintuitive, but it is caused by replacement of the original gene pool with a less variable one. There is molecular evidence for such a development having occurred in cattle (Blott *et al.* 1998).

Assembling a concentrated collection, or an enriched assemblage, of breeds that would contain large amounts of variation is another possible objective in prioritizing breeds. It differs from the previous goal because a population is not chosen according to its potential to increase within-breed variability of the other (possibly non-conserved) breeds. The simplest method to construct the set is to exclude stepwisely a breed that has the highest mean co-ancestry and repeat this over the breeds. The stepwise method was applied to the 32 northern European sheep breeds, giving each breed the same weight. The mean of $s(s+1)/2$ co-ancestries for s breeds reached the minimum with $s = 15$ (Fig. 4). The 15 breeds in the set in preference order were: *the Finnsheep*, *the Lithuanian Native Coarsewooled*, *the Norwegian Old Spael Sheep*, *the Greenland Sheep*, *the Norwegian Fuglestad Sheep*, *the Russian Viena sheep*, *the Norwegian Rygja Sheep*, *the Icelandic Sheep*, *the Swedish Gute Sheep*, *the Norwegian Grey Troender Sheep*, *the Finnish Grey Landrace*, *the Norwegian Spael Sheep*, *the Lithuanian Romanov*, *the Norwegian Dala Sheep* and *the Faeroe Island Sheep*. This "*enriched set*" contained 93% of the observed microsatellite alleles and the total allelic richness ($r_t[50]$) would not have increased if additional breeds were included (Fig. 4). The set largely overlapped with the "diversity contribution set". Compared with the "diversity contribution set", four new breeds were included (*the Faeroe Island Sheep*, *the Finnish Grey Landrace* and *the Norwegian Dala* and *Grey Troender Sheep*), and four breeds were excluded (*the Estonian Ruhnu sheep* and *the Swedish Dala Fur, Roslag* and *Rya Sheep*).

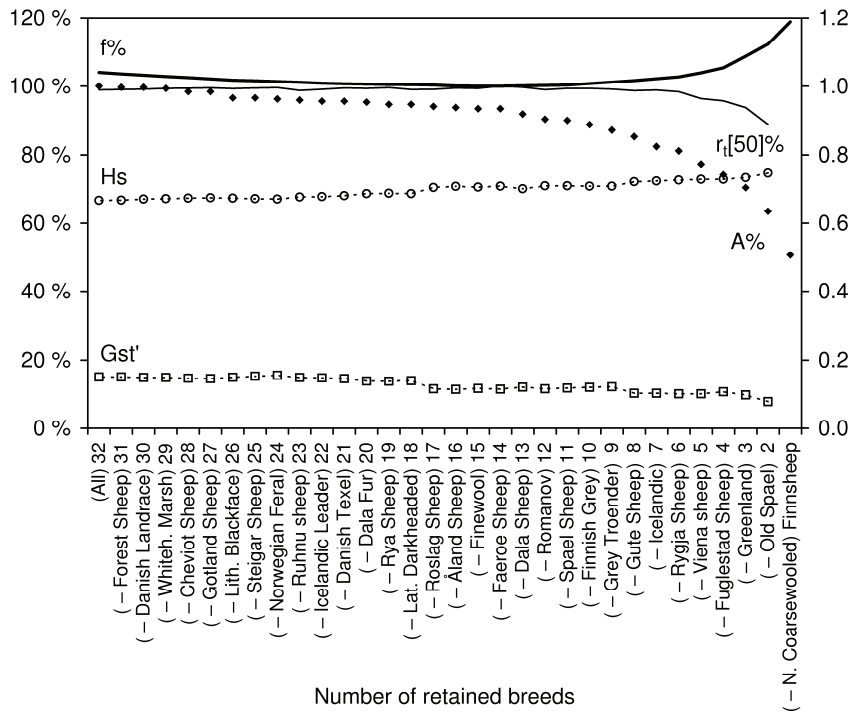


Fig. 4. Stepwise exclusion of breeds demonstrating highest mean kinship (on left all breeds are included, on right only Finnsheep is included). Three measures are presented as percentages in relation to the best observed value (left axis) excluding breeds stepwisely from left to right: Mean kinship in all simple combinations of breeds (f%), expected number of alleles in a random sample of 50 chromosomes over breeds (r_t[50]%), and proportion of observed alleles included (A%). Below, Hs and Gst' (Nei 1987), respectively, indicate mean within-breed gene diversity and breed divergence (right axis).

4.4.3 Conclusions on breed ranking

The "diversity contribution set" contains breeds that could be used to increase within-breed variability of other breeds, while the "enriched set" contains diverse breeds, which could be used to create variable synthetic breeds. Summarising the results from the two different approaches, there are eleven breeds common to the two methods (*the Finnsheep*, *the Greenland Sheep*, *the Icelandic Sheep*, *the Lithuanian Romanov* and *Native Coarsewooled*, *the Norwegian Fuglestad*, *Spael*, *Old Spael* and *Rygja Sheep*, *the Russian Viena sheep*, and *the Swedish Gute Sheep*). One of the two methods suggests five additional endangered breeds (*the Estonian Ruhnu sheep*, *the Finnish Grey Landrace*, *the Norwegian Grey Troender*, *the Swedish Dala Fur* and *Roslag Sheep*) and three numerous breeds (*the Faeroe Island Sheep*, *the Norwegian Dala Sheep*, and *the Swedish Rya Sheep*) should be also included. Thus, the "*combined set*" includes 19 out of 32 breeds and

contains over 95% of the observed microsatellite alleles. This set, and in particular the seven rare native breeds in the set, merit intensified conservation measures first.

4.4.3.1 Comparison with other ranking methods

Estimating conservation importance of breeds is a complicated task because there are short and long-term interests, which would require different methods to measure diversity for breed ranking. Additionally, there are several ways to utilize variability of a breed, spanning from pure-breeding to complete amalgamation of breeds. These affect the breed rankings.

Weitzman (1992 and 1993) methods or their derivatives have been the most widely used methods to prioritize breeds, but they are also widely criticised (reviewed by Simianer 2005a) The Weitzman diversity is the sum of branches in a maximum likelihood tree based on genetic distances between populations. Weitzman methods were originally designed for ranking *species*. Applying this criterion among *breeds* might favour inbred populations, which contribute very little to the gene pool of the metapopulation (Caballero & Toro 2002, Grigaliunaite *et al.* 2002). Even though developments of the method consider breed extinction probabilities (Reist-Marti *et al.* 2003), integrate other conservation arguments (Simianer *et al.* 2003), and can encompass within-breed variation through the idea of "genetic extinction" referring to loss of all within-breed variation (García *et al.* 2005), the method assumes pure-breeding in conserving the *breeds*. For these reasons Weitzman methods were not preferred in this study.

The core-set methods are designed for optimizing collection of germplasm and consider simultaneously among-breed and within-breed variation. Unfortunately, they have a narrow conservation goal, aiming at maximising diversity (Caballero & Toro 2002, Eding *et al.* 2002) or the selection response (Bennewitz & Meuwissen 2005) in the *undivided* population formed from the collected germplasm in precisely refined proportions. It is difficult to apply the defined proportions (i.e. contributions to the core-set) in conservation beyond the specific sampling question forming the theoretical basis of the core-set methods. Breed fusion is rarely the aim in conservation programmes, because maintaining several breeds is a more secure way to maintain genetic variation (e.g. Hall & Bradley 1995). Therefore, the core-set methods were not used.

The "diversity contribution method" (Petit *et al.* 1998, Caballero & Toro 2002) is distinct from the Weitzman methods and the core-set methods, while the construction of "enriched assemblage" is closely related to core-set methods, minimizing co-ancestry (Caballero & Toro 2002, Eding *et al.* 2002). However, here the breed assessment is simplified to a decision to keep or to exclude a breed. Furthermore, the minimized mean co-ancestry is calculated for a *triangle* of $s(s+1)/2$ of co-ancestries for s breeds, which refers to the intended use of genetic variation in separately managed purebred or crossbred populations. The core-set methods assume fusion of breeds and use the *square* s^2 of co-ancestries.

For comparison, an optimization similar to the method of Eding *et al.* (2002) was done using the molecular co-ancestry estimates ($f_{kl,}$ Appendix I). This was obtained by adjusting the proportions of germplasm, c, of breeds so that they minimize the co-

ancestry of the collection ($\Sigma_{k,l} f_{kl} c_k c_l$, Caballero & Toro 2002). First, 25% of the collection was filled with an equal contribution (0.8%) from each breed, and then the remaining 75% of the collection was optimized. All the additional deposits came from the "combined set" and the total deposits from the most important breeds 1–7 and 9 identified in Fig. 4 were at least five times the initial deposit, whereas seven breeds from the "combined set" had no further deposits (unpublished data). The "combined set" is satisfactory also for the conservation goal defined by Eding *et al.* (2002).

4.4.3.2 Integrating other conservation arguments

Single important alleles may exist in any breed and strong selection causes loci to deviate from the neutral diversity pattern shaped by mere demography. Furthermore, the socio-economic and cultural reasons for conservation are not taken into account in the analysis. These may influence conservation preferences. The stepwise method can integrate other arguments. For the analysis this means that if other conservation reasons exist, the breed is not removed from the co-ancestry matrix even if it has the highest mean co-ancestry value (or makes the most negative contribution to diversity if stepwise exclusion is performed using contributions). This may have an effect on the breed ranking. We may consider two examples of breeds with special adaptations, *the Icelandic Leader Sheep* and *the Norwegian Feral Sheep* (e.g. Eythorsdottir 2005), which were not included in the "combined set". If *the Leader Sheep* is kept in the matrix through the analysis, the exclusion of *the Icelandic Sheep* is not suggested at step 26 (Fig. 4), but as early as at step 14 because it is closely related to *the Leader Sheep*. On the other hand, keeping *the Norwegian Feral Sheep* has a minor effect on subsequent breed rankings.

5 Concluding remarks

In the investigations forming the basis of this thesis it was established that there are four, not three, distinct mtDNA lineages in domestic sheep, and that the mtDNA variation among sheep in Europe, the Caucasus and Central Asia can be explained by expansion starting from single domestication area, the Near East. Studied European sheep populations seem to have their maternal origin mainly in a common European founder population. These results support previous archaeological hypotheses on sheep domestication and spread, and they fit with the limited range for the Asiatic mouflon, which is considered to be the truly wild ancestral form of the domestic sheep. A novel finding is that the arrival of few distinct rare types of mtDNA in the Nordic countries appears to have occurred across Russia. This may be related to Viking trade through the River Volga. More certainly, the Vikings caused the minor range expansion that began from the Faeroe Islands. Most of the observed *Icelandic Sheep* mtDNA lineages originate from this expansion, but the descendants of the expansion are scattered widely across northern Europe. Considering autosomal variation, the present results imply that the basis for it has been wider than for the mitochondrial variation. This suggests that the domestic and wild varieties were not strictly isolated after domestication.

The study showed that an unusually large proportion of total variation in northern European sheep is among the breeds. Microsatellite data can distinguish the northern short-tailed breeds from the long-tailed breeds originating from areas southwest of Scandinavia, although this macro-scale structure explains only a small proportion of among-breed variation. The clustering of breeds corresponds to geographical locations of breed origins. The effect of small-scale geographical closeness is very clearly visible in clustering of individuals. For instance, the local sheep strain in the Vepsian Karelia in Russia was shown to be an intermediate or a composite between its two neighbours, *the Finnish Grey Landrace* and *the Romanov*. The relatively weak macro-scale structure is not surprising when considering that the landraces were formed from the local strains a hundred years ago. The subdivision in present-day Estonian and Russian Karelian strains was clear, and we can assume that similar subdivision was common in the old strains preceding landrace formation. Such subdivision can outweigh the large-scale structure.

The native sheep breeds are sometimes assumed to be highly inbred and in need of “refreshing” with crossbreeding. The present results show that many landraces are more

variable than the “modern” breeds. Substantial inbreeding is limited to a few of the rarest breeds. Similarly, recent population reductions, which may require special attention in breed management, were shown to have occurred in rare breeds and in imported populations of a “modern” breed. Thus the potential problems caused by inbreeding are not specific to the native breeds, but need to be considered in all types of population management. It was further noted that the prolificacy and within-breed variability are correlated among the breeds. This together with earlier results suggests that within-breed inbreeding has a negative impact on breed prolificacy. Therefore, special attention should be paid to the maintenance of variation in the prolific breeds.

In general, the northern sheep genetic resources should be managed at breed level. A major practical exception is that joint management of the “modern” Estonian breeds could be appropriate based on low differentiation. A set of the most important breeds for conservation was identified in the thesis. This set comprised 19 breeds, including seven endangered old native breeds or strains (*the Estonian Ruhnu sheep*, *the Finnish Grey Landrace*, *the Norwegian Old Spael* and *Grey Troender Sheep*, *the Russian Viena sheep*, and *the Swedish Dala Fur* and *Roslag Sheep*). Securing these breeds represents a high priority among the conservation actions needed in the near future. In the long term, conservation measures are needed also for more abundant breeds in the set.

The analysis of genetic variation was based on molecular markers and loci affected by selection were excluded from the analyses (III). Even if conservation aims to maintain potentially advantageous variation, neutral variation is a preferable and cost-efficient tool to evaluate breeds. Neutral variation reveals the overall patterns in the genome modified by the demographic processes. In contrast, the effect of selection is specific to a chromosomal region. In the future, genome-wide studies, using single-nucleotide polymorphisms (SNP), will become possible and enable extensive diversity characterization, including functional genes affecting adaptations to specific environment, management or production types. This will give even more precise information about the variation within and among the breeds, and if the genome can be extensively assessed, such studies are preferable to the use of neutral variation in breed evaluation.

There are also “non-genetic” reasons for conservation. Rare breeds are important in providing livelihoods in rural areas, and they can be used in creation of special niche products or, for instance, they can be used in history education and tourism. Most cultural, historical or socio-economic arguments for conservation do not rely on genetic information, but molecular markers can reveal the degree of past crossbreeding, which may influence the value of a breed as a historical object. This type of widened analysis goes beyond genetics and requires a broad multidisciplinary approach.

The data produced for the thesis can be used as a reference in follow-up studies evaluating success of established conservation programmes of sheep. The data further enable identification of unintentional crossbreeding. Appendix I contains co-ancestry parameters, which can be used in detailed conservation value estimation within countries or can be transformed to genetic distances for constructing representations of kinships. They can also be used in optimising contributions, if synthetic breeds are created from the studied breeds. The mitochondrial sequences can be included in later analyses of domestication history and maternal gene flow in sheep. Thus the raw data and the presented parameters can be reused in several ways if specific needs arise.

Some general lessons learned about the northern European sheep can be applied to wild species. For example, a genetic bottleneck in the maternal lineages is not necessarily paralleled by a bottleneck in autosomal variation. The study also points out the danger of losing genetic variation when populations are relocated. A large proportion of total allelic richness derives from among population differences, and thus allelic variation is supported by the existence of separate populations. Especially in this sense, small populations may be disproportionately important for the variation of the larger metapopulation. The most striking observation that has application to wild species is that extremely low or high within-population variability or population divergence cannot be separately used as evidence about the genetic importance of the population for the metapopulation.

References

Adalsteinsson S (1981) Origin and conservation of farm animal populations in Iceland. J Anim Breed Genet 98: 258–264.

Atroshi F (1979) Phenotypic and genetic association between production/reproduction traits and blood biochemical polymorphic characters in Finnsheep. PhD. thesis, University of Helsinki, Helsinki. Ann Agric Fenn 18.

Ankel-Simons F & Cummins JM (1996) Misconceptions about mitochondria and mammalian fertilization: Implications for theories on human evolution. Proc Natl Acad Sci USA 93: 13859–13863.

Àlvarez I, Royo LJ, Fernández I, Gutiérrez JP, Gómez E & Goyache F (2004) Genetic relationships and admixture among sheep breeds from Northern Spain assessed using microsatellites. J Anim Sci 82: 2246–2252.

Arranz JJ, Bayón Y & San Primitivo F (1998). Genetic relationships among Spanish sheep using microsatellites. Anim Genet 29: 435–440.

Ashley MV, Laipis PJ & Hauswirth WW (1989) Rapid segregation of heteroplasmic bovine mitochondria. Nucleic Acids Res 17: 7325–7331.

Ballou JD & Lacy RC (1995) Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou JD, Gilpin M & Foose TJ (eds) Population management for survival and recovery. Columbia University Press, New York, p 76–111.

Balloux F & Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. Mol Ecol 11: 155–165.

Bandelt, H-J., Forster P & Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37–48.

Bataillon TM, David JL & Schoen DJ (1996) Neutral genetic markers and conservation genetics: simulated germplasm collections. Genetics 144: 409–417.

Baumung R. & Sölkner J (2003) Pedigree and marker information requirements to monitor genetic variability. Genet Sel Evol 35: 369–383.

Baumung R., Simianer H. & Hoffmann I. (2004) Genetic diversity studies in farm animals – a survey. J Anim Breed Genet 121: 361–373.

Bennewitz, J & Meuwissen THE (2005) A novel method for the estimation of the relative importance of breeds in order to conserve the total genetic variance. Genet Sel Evol 37: 315–337.

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SL, Hawkins GA Toldo SS, Fries R., Grosz MD, Yoo J & Beattie CW (1994). A genetic linkage map for cattle. Genetics 136: 619–639.

Blott SC, Williams JL & Haley CS (1998) Genetic variation within the Hereford breed of cattle. Anim Genet 29: 202–211.

Bogdanov LV, Polyakovskii VI, Lazovskii AA, Martsinkevich IS, Pogonyailo DI, Nikonchik LI, Zotova VS, Troshchenko LV & Dzhumkov VA (1972) Genetic polymorphism of some blood and milk proteins in livestock bred in White Russia (in Russian). Byulleten' Nauchnykh Rabot Vsesoyuznyi Nauchno Issledovatel' ski Institut Zhivotnovodstva. 26: 42–48.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR & Cavalli-Sforza LL (1994). High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368: 455–457.

Braend M, Aursjo J & Austbo O (1987) Haemoglobin types of old Norwegian short tail land race sheep. Acta Vet Scand 2: 121–123.

Buchanan FC & Crawford AM (1992). Ovine dinucleotide repeat polymorphism at the MAF214 locus. Anim Genet 23: 394.

Buchanan FC & Crawford AM (1993). Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. Anim Genet 24: 145.

Buchanan FC, Swarbrick PA & Crawford AM (1991) Ovine dinucleotide repeat polymorphism at the MAF48 locus. Anim Genet 22: 379–380.

Buchanan FC, Swarbrick PA & Crawford AM (1992). Ovine dinucleotide repeat polymorphism at the MAF65 locus. Anim Genet 23: 85.

Buchanan FC, Adams LJ, Littlejohn RP, Maddox JF & Crawfrod AM (1994a) Determination of evolutionary relationships among sheep breeds using microsatellites. Genomics 22: 397–403.

Buchanan FC, Galloway SM & Crawford AM (1994b). Ovine microsatellites at the OarFCB5, OarFCB19, OarFCB20, OarFCB48, OarFCB129 and OarFCB226 loci. Anim Genet 25: 60.

Caballero A & Toro MA (2002) Analysis of genetic diversity for the management of conserved subdivided populations. Conserv Genet 3: 289–299.

Canali G & Econogene Consortium (in press) Common agricultural policy reform and its effects on sheep and goat market and rare breeds conservation, Small Ruminant Res, Corrected proof, Available online 19 October 2005.

Cañón J, Alexandrino P, Bessa I, Carleos C, Carretero Y, Dunner S, Ferran N, Garcia D, Jordana J, Laloë D, Pereira A, Sanchez A & Moazami-Goudarzi K. (2001) Genetic diversity measures of local European beef cattle breeds for conservation purposes. Genet Sel Evol, 33: 311–332.

Casas E, Freking BA & Leymaster KA (2004) Evaluation of Dorset, Finnsheep, Romanov, Texel, and Montadale breeds of sheep: II. Reproduction of F1 ewes in fall mating seasons. J Anim Sci 82: 1280–1289.

Cavalli-Sforza LL & Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evolution 21: 550–570.

Ciampolini R, Moazami-Goudarzi K, Vaiman D, Dillmann C, Mazzanti E, Foulley JL, Leveziel H & Cianci D (1995) Individual multilocus genotypes using microsatellite polymorphisms to permit the analysis of the genetic variability within and between Italian beef cattle breeds. J Anim Sci 73: 3259–3268.

Ciofi C, Beaumont MA, Swingland IR & Bruford MW (1999) Genetic divergence and units for conservation in the Komodo dragon Varanus komodoensis. Proc R Soc Lond B Biol Sci 266: 2269–2274.

Chikhi L & Bruford MW (2005) Mammalian population genetics and genomics. In: Ruvinsky A & Marshall Graves J (eds) Mammalian Genomics, CABI Publishers, UK, p 539–584.

Corander J, Waldmann P & Sillanpää J (2003). Bayesian analysis of genetic differentiation between populations. Genetics 163: 367–374.

Cornuet JM & Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014.

Crandall KA, Bininda-Emonds OR, Mace GM & Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15: 290-295.

Crawford AM & Cuthbertson RP (1996) Mutations in sheep microsatellites. Genome Research 6: 876–879.

Davis GH (2004) Fecundity genes in sheep. Anim Reprod Sci 82-83: 247–253.

DeRose MA & Roff DA (1999) A comparison of inbreeding depression in life-history and morphological traits in animals. Evolution 53: 1288–1292.

Diez-Tascón C, Littlejohn RP, Almaeida PAR & Crawford AM (2000) Genetic variation within the Merino sheep breed: analysis of closely related populations using microsatellites. Anim Genet 31: 243–251.

EAAP (2005) EAAP-Animal Genetic Data Bank [online, cited June 10 2005]. Available from: http://www.tiho-hannover.de/einricht/zucht/eaap/index.html, last updated April 27 2005.

Ede AJ, Pierson CA & Crawford AM (1995a). Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26 loci. Anim Genet 26: 129–130.

Ede AJ, Pierson CA & Crawford AM (1995b). Ovine microsatellites at the OARCP34, OARCP38, OARCP43, OARCP49, OARCP73, OARCP79 and OARCP99 loci. Anim Genet 26: 130–131.

Eding H, Crooijmans RPMA, Groenen MAM &.Meuwissen THE (2002) Assessing the contribution of breeds to genetic diversity in conservation schemes. Genet Sel Evol 34: 613–633.

Efremov G & Braend M (1965) Hemoglobins, Transferrin and Albumins of Sheep and Goats. Proc. 9th European Animal Blood Group Conference Prague, August 18-22, p 313–320.

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

El Mousadik A & Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the arvan tree [Arvania spinosa (L.) Skeels] endemic to Morocco. Theor Appl Genet 92: 832–839.

Ercanbrack SK & Knight AD (1991) Effects of inbreeding on reproduction and wool production of Rambouillet, Targhee, and Columbia sheep. J Anim Sci 69: 4734–4744.

Evans JV, Harris H & Warren FL 1958. Haemoglobin and potassium blood types in some non-British breeds of sheep and in certain rare British breeds. Nature 182: 320–321.

Everitt BS & Dunn G (1991) Applied Multivariate Data Analysis. Edward Arnold, London.

Excoffier L & Yang Z (1999) Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. Mol Biol Evol 16: 1357–1368.

Eythorsdottir E (ed.) (2005) North-SheD –homepage [online, cited August 18 2005]. Available from: http://www.lbhi.is/northshed.

FAO (2004) FAOSTAT data [online, cited May 13 2005]. Available from: http://faostat.fao.org/faostat/, last updated December 20 2004.

FAO (2005) DAD-IS 2.0 Domestic animal diversity information system [online, cited August 4 2005]. Available from http://www.fao.org/dad-is/.

Fahmy MH, Boucher JM, Poste LM, Gregoire R, Butlers G & Comeau JE (1992) Feed efficiency, carcass characteristics, and sensory quality of lambs, with or without prolific ancestry, - Fed diets with different protein supplements. J Anim Sci 70: 1365–1374.

Frankham R (2005) Stress and adaptation in conservation genetics. J Evol Biol 18: 750–755.

Freking BA, Leymaster KA & Young LD (2000) Evaluation of Dorset, Finnsheep, Romanov, Texel, and Montadale breeds of sheep: I. Effects of ram breed on productivity of ewes of two crossbred populations. J Anim. Sci 78: 1422–1429.

Freking BA & Leymaster KA (2004) Evaluation of Dorset, Finnsheep, Romanov, Texel, and Montadale breeds of sheep: IV. Survival, growth, and carcass traits of F1 lambs. J. Anim. Sci. 82: 3144–3153.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915–925.

Garner A, Rachlow JL & Hicks J (2005) Patterns of genetic diversity and its loss in mammalian populations. Conserv Biol 19: 1215–1221.

García D, Corral N & Cañón J (2005) Combining inter- and intrapopulation information with the Weitzman approach to diversity conservation. J Hered 96: 704–712.

Giles RE, Blanc H, Cann HM & Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci USA 77: 6715–6719.

Goldstein DB, Linares AR, Cavalli-Sforza LL & Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics 139: 463–471.

Goyache F, Gutiérrez JP, Fernández I, Gomez E, Alvarez I, Díez J & Royo LJ (2003) Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. J Anim Breed Genet 120: 95–105.

Grigaliunaite I, Tapio M, Holm L-E, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2002) Weitzman approach and components of diversity in Northern European sheep breeds. Proceedings 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, CDRom Communication No. 26-37.

Guo J, Du L-X, Ma Y-H, Guan W-J, Li H-B, Zhao Q-J, Li X & Rao S-Q (2005) A novel maternal lineage revealed in sheep (Ovis aries). Anim Genet 36: 331–336.

Gutiérrez JP, Altarriba J, Díaz C, Quintanilla R, Cañón J & Piedrafita J (2003) Pedigree analysis of eight Spanish beef cattle breeds. Genet Sel Evol 35: 43–63.

Hall SJG & Bradley DG (1995) Conserving livestock breed biodiversity. Trends Ecol Evol 10: 267–270.

Hartl DL & Clark AG (1989) Principles of Population Genetics, 2nd edition. Sinauer, Sunderland, Mass.

Hayashi JI, Yonekawa H, Gotoh O, Watanabe J & Tagashira Y (1978) Strictly maternal inheritance of rat mitochondrial DNA. Biochem Biophys Res Commun 83: 1032–1038.

Hecht NB, Liem H, Kleene KC, Distel RJ & Ho SM. (1984) Maternal inheritance of the mouse mitochondrial genome is not mediated by a loss or gross alteration of the paternal mitochondrial DNA or by methylation of the oocyte mitochondrial DNA. Dev Biol 102: 452–461.

Hedrick PW (2001) Conservation genetics: where are we now? Trends Ecol Evol 16: 629–636.

Henry HM, Penty JM, Pierson CA & Crawford AM (1993). Ovine microsatellites at the OARHH35, OARHH41, OARHH44, OARHH47 and OARHH64 loci. Anim Genet 24: 223.

Hiendleder S, Mainz K, Plante Y & Lewalski H (1998) Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. J Hered 89: 113–120.

Hiendleder S, Kaupe B, Wassmuth R & Janke A (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. Proc R Soc Lond B Biol Sci 269: 893–904.

Holm L-E, Forchhammer MC & Boomsma JJ (1999) Low genetic variation in muskoxen (Ovibos moschatus) from western Greenland using microsatellites. Mol Ecol, 8: 675–679.

Hudson RR, Slatkin M & Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. Genetics 132: 583–589.

Hulme DJ, Silk JP, Redwin JM, Barendse W, Beh KJ (1994). Ten polymorphic ovine microsatellites. Anim Genet 25: 434.

Hutchison CA, Newbold JE, Potter SS & Edgell MH (1974) Maternal inheritance of mammalian mitochondrial DNA. Nature 251: 536–538.

Huurre M (2003) Maatalouden alku Suomessa. In: Rasila V, Jutikkala E & Mäkelä-Alitalo A (eds) Suomen maatalouden historia, I : Perinteisen maatalouden aika esihistoriasta 1870-luvulle. Gummerus, Jyväskylä, p 38–66.

Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE T Neural Netw 10: 626–634.

International Commission of Zoological Nomenclature (2003) Opinion 2027 (Case 3010). Usage of 17 specific names based on wild species which are pre-dated by or contemporary with those based on domestic animals: conserved. Bull Zool Nomencl 53: 28–37.

Jarne P & Lagoda P (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11: 424–429.

Jeppsson S (1998) Swedish Breeds of Domestic Animals: Animals in Agricultural Production. Bratts Tryckeri AB Jönköping, Sweden.

Kantanen J, Olsaker I, Holm LE, Lien S, Vilkki J, Brusgaard K, Eythorsdottir E, Danell B & Adalsteinsson S (2000). Genetic diversity and population structure of 20 North European cattle breeds. J Hered 91: 446–457.

Kemp SJ, Brezinsky L & Teale AJ (1992) ILSTS002: a polymorphic bovine microsatellite. Anim Genet 23: 184.

Kimmel M, Chakraborty R, King JP, Bamshad M, Watkins WS & Jorde LB (1998) Signatures of population expansion in microsatellite repeat data. Genetics, 148: 1921–1930.

Kristensen K (1992) Produktionsudviklingen for de forskellige dyrearter. In Adalsteinsson S (ed) Husdyr i Norden, Vår arv - vårt ansvar. Årbok for Norsk landbruksmuseum, Jord og Gjerning 1992/93, Landbruksforlaget, Oslo, p 24–32.

Kristensen TN, Sorensen AC, Sorensen D, Pedersen KS, Sorensen JG & Loeschcke V (2005) A test of quantitative genetic theory using Drosophila- effects of inbreeding and rate of inbreeding on heritabilities and variance components. J Evol Biol 18: 763–770.

Laval G, Iannuccelli N, Legault C, Milan D, Groenen MA, Giuffra E, Andersson L, Nissen PH, Jorgensen CB, Beeckmann P, Geldermann H, Foulley JL, Chevalet C & Ollivier L. (2000). Genetic diversity of eleven European pig breeds. Genet Sel Evol 32: 187–203.

Lazovskii AA (1983) The origin of Romanov sheep. Zhivotnovodstvo 8: 29.

Lenstra JA & The ECONOGENE Consortium (2005) Evolutionary and demographic history of sheep and goats suggested by nuclear, mtDNA and Y-chromosomal markers. Proc. Int. Workshop 'The role of biotechnology for the characterization of crop, forestry, animal and fishery genetic resources'. Turin, March 5–7, 2005. [online, cited August 31 2006]. Available from http://www.fao.org/biotech/docs/lenstra.pdf.

Littledike ET & Young LD (1993) Effect of sire and dam breed on copper status of fat lambs. J Anim Sci 71: 774–778.

Logan FD (1991) The Vikings in History, 2nd edition. Routledge, London.

Luikart GL, Allendrof FW, Cornuet J-M & Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89: 238–247.

Lupton CJ, Freking BA & Leymaster KA (2004) Evaluation of Dorset, Finnsheep, Romanov, Texel, and Montadale breeds of sheep: III. Wool characteristics of F1 ewes. J Anim Sci 82: 2293–2300.

Lynch M & Hill WG (1986) Phenotypic evolution by neutral mutation. Evolution 40: 915–935.

Maddox JF, Davies KP, Crawford AM, Hulme DJ, Vaiman D, Cribiu EP, Freking BA, Beh KJ, Cockett NE, Kang N, Riffkin CD, Drinkwater R, Moore SS, Dodds KG, Lumsden JM, van Stijn TC, Phua SH, Adelson DL, Burkin HR, Broom JE, Buitkamp J, Cambridge L, Cushwa WT, Gerard E, Galloway SM, Harrison B, Hawken RJ, Hiendleder S, Henry HM, Medrano JF, Paterson KA, Schibler L, Stone RT & van Hest B. (2001) An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Res 11: 1275–1289.

Maijala K (1984) Review of experiences about the use of Finnsheep in improving fertility. Proc 2nd Wld Cong Sheep and Beef Cattle Breeding, Pretoria, South Africa 1: 519–531.

Maijala K (1988) History, recent development and uses of Finnsheep. J Agr Sci Finland 60: 449–454.

Maijala K (1997) Genetic aspects of domestication, common breeds and their origin. In: Piper L & Ruvinsky A (eds) The Genetics of Sheep. Cab International, Wallingford, p13–49.

Maijala K & Terrill CE (1991). Breed structure, dynamics, and new breed development in sheep. In: Maijala K (ed) Genetic Resources of Pig, Sheep and Goat. Elsevier, Amsterdam, p 305–326.

Mason IL. (1991) Classification and distribution of sheep breeds. In: Maijala K (ed) Genetic resources of pig, sheep and goat. Elsevier, Amsterdam, p 179–194.

Meadows JRS, Hawken RJ & Kijas JW (2004) Nucleotide diversity on the ovine Y chromosome. Anim Genet 35: 379–385.

Michaels G.S., Hauswirth WW & Laipis PJ (1982) Mitochondrial DNA copy number in bovine oocytes and somatic cells. Dev Biol 94: 246–251.

Milisauskas S (2002) European Prehistory, A Survey. Kluwer Academic/Plenum Publishers, New York.

Moore SS & Byrne K (1994). Characterisation of 65 bovine microsatellites. Mamm Genome 5: 84–90.

Moritz C (1994). Defining 'Evolutionary Significant Units' for conservation. Trends Ecol Evol 10: 373–375.

Muigai AWT (2002) Characterisation and conservation of indigenous animal genetic resources: Genetic diversity and relationships of fat-tailed and thin-tailed sheep of Africa. PhD Thesis, Department of Biochemistry, Jomo.

Kenyatta University of Agriculture and Technology, Juja, Kenya.

Myers P, Espinosa R, Parr CS, Jones T, Hammond GS & Dewey TA (2005) The Animal Diversity Web [online, cited June 8 2005]. Available from: http://animaldiversity.org.

Nadler CF, Korobitsina KV, Hoffmann RS & Vorontsov NN (1973) Cytogenetic differentiation, geographic distribution and domestication in palaearctic sheep (*Ovis*). Zeit f. Saugetierkunde, 38: 109–125.

Nei M (1972) Genetic distance between populations. Am Nat 106: 283–292.

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

Nei M, Tajima F & Tateno Y (1983) Accuracy of genetic distances and phylogenetic trees from molecular data. J Mol Evol 19: 153–170.

Oldenbroek JK (1999) Genebanks and the Conservation of Farm Animal Genetic Resources. DLO Institute for Animal Sciences and Health, Lelystad, The Netherlands.

Ozerov M, Tapio M, Marzanov N, Kiseleva T & Kantanen J (2005) Genetic differentiation of East European and Caucasian Sheep breeds by microsatellites. Proceedings of the 11th Baltic Animal Breeding and Genetics Conference : May 2005, Palanga, Lithuania. p 132–136.

Pariset L, Savarese MC, Cappuccio I & Valentini A (2003) Use of microsatellites for genetic variation and inbreeding analysis in Sarda sheep flocks of central Italy. J Anim Breed Genet 120: 425–432.

Pedrosa S, Uzun M, Arranz JJ, Gutiérrez-Gil B, San Primitivo F & Bayón Y. (2005) Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. Proc Biol Sci 272: 2211–2217.

Penty JM, Henry HM, Ede AJ & Crawford AM (1993) Ovine microsatellites at the OarAE16, OarAE54, OarAE57, OarAE119 and OarAE129 loci. Anim Genet 24: 219.

Pesole G, Gissi C, De Chirico A & Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol 48: 427–434

Petit RJ, El Mousadik A & Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12: 844–855.

Pierson CA, Hanrahan V, Ede AJ & Crawford AM (1993). Ovine microsatellites at the OARVH34, OARVH41, OARVH58, OARVH61 and OARVH72 loci. Anim Genet 24: 224.

Piganeau GV & Eyre-Walker A (2004) A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. Heredity 92: 282–288.

Piganeau GV, Gardner M & Eyre-Walker A (2004) A broad survey of recombination in animal mitochondrial DNA. Mol Biol Evol 21: 2319–2325.

Piyasatian N & Kinghorn BP (2003) Balancing genetic diversity, genetic merit and population viability in conservation programmes J Anim Breed Genet 120: 137–149

Pons O & Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity. I. Haploid locus. Theor Appl Genet 90: 462–470.

Poplin F (1979) Origine du mouflon de Corse dans une nouvelle perspective paleontologique, par marronnage. Ann Génét Sél Anim 11:133–143.

Pritchard JK & Donnelly P. 2001 Case-control studies of association in structured or admixed populations. Theor Popul Biol 60: 227–237.

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

Rasmusen BA & Tucker EM (1973) Transferrin type and reproduction in sheep. Anim Blood Grps Biochem Genet 4: 207–220.

Reist-Marti SB, Simianer H, Gibson J, Hanotte O & Rege JEO (2003) Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. Conserv Biol 17: 1299–1311.

Ronquist F & Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Ryder ML (1983) Sheep and man. Duckworth, London.

Saitou N & Nei M (1987) The neighbour joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.

SanCristobal M, Chevalet C, Haley CS, Joosten R, Rattink AP, Harlizius B, Groenen MAM, Amigues Y, Boscher M-Y, Russell G, Law A, Davoli R, Russo V, Désautés C, Alderson L, Fimland E, Bagga M, Delgado JV, Vega-Pla JL, Martinez AM, Ramos M, Glodek P, Meyer JN, Gandini GC, Matassino D, Plastow GS, Siggens KW, Laval G, Archibald AL, Milan D, Hammond K & Cardellino R (in press) Genetic diversity within and between European pig breeds using microsatellite markers. Anim Genet, Available online 3 May 2006.

Scherf BD (ed.) (2000) World watch list for domestic animal diversity, 3rd edition. FAO. Rome.

Semyonov SI & Selkin II (1989) Sheep. In: Dmitriev NG & Ernst LK (eds) Animal genetic resources of the USSR . Food Animal Production and Health paper No. 65. Food and Agriculture Organization of the United Nations, Rome, Italy, p 143–254.

Shiatara H, Hayashi J-I, Takahama S, Kaneda H & Yonekawa H (1998) Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. Genetics 148: 851–857.

Simianer H (2005a) Decision making in livestock conservation. Ecol Econ 53: 559– 572.

Simianer H (2005b) Using expected allele number as objective function to design between and within breed conservation of farm animal biodiversity. J Anim Breed Genet, 122: 177–187.

Slate J & Phua H (2003) Patterns of linkage disequilibrium in mitochondrial DNA of 16 ruminant populations. Mol Ecol 12: 597–608.

Smith BD (1998) The Emergence of Agriculture,Scientific American Library, New York

Stahlberger-Saitbekova N, Schläpfer J, Dolf G & Gaillard C (2001) Genetic relationships in Swiss sheep breeds based on microsatellite analysis. J Anim Breed Genet 118: 379–387.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C & Schatten G (2000) Ubiquitinated Sperm Mitochondria, Selective Proteolysis, and the Regulation of Mitochondrial Inheritance in Mammalian Embryos. Biol Reprod 63: 582–590.

Sutovsky P, McCauley TC, Sutovsky M & Day BN (2003) Early Degradation of Paternal Mitochondria in Domestic Pig (Sus scrofa) Is Prevented by Selective Proteasomal Inhibitors Lactacystin and MG132. Biol Reprod 68: 1793–1800.

Swarbrick PA, Buchanan FC, Crawford AM (1991). Ovine dinucleotide repeat polymorphism at the MAF36 locus. Anim Genet 22: 377.

Swofford DL (2003) PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b 10. Sinauer, Sunderland, Mass.

Taanman J-W (1999) The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta 1410: 103–123.

Templeton AR, Crandall KA & Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619–633.

Terrill CE & Slee J (1991) Breed differences in adaptation of sheep. In: Maijala K (ed) Genetic Resources of Pig, Sheep and Goat. Elsevier, Amsterdam, p 225–233.

Tomiuk J & Loeschcke V (1995) Genetic identity combining mutation and drift. Heredity 74: 607–615.

Vaiman D, Mercier D, Moazami-Goudarzi K, Eggen A, Ciampolini R, Lepingle A, Velmala R, Kaukinen J, Varvio SL, Martin P, Leverziel H and Guerin G (1994). A set of 99 cattle microsatellites: characterisation, synteny mapping, and polymorphism. Mamm Genome 5: 288–297.

Vilà C, Seddon J & Ellegren H (2005) Genes of domestic mammals augmented by backcrossing with wild ancestors. Trend Genet 21: 214–218.

Visscher PM, Smith D, Hall SJG & Williams JA (2001) A viable herd of genetically uniform cattle. Nature, 409: 303.

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

Weitzman ML (1992) On diversity. Q J Econ 107: 363–405.

Weitzman ML (1993) What to preserve? An application of diversity theory to crane conservation. Q J Econ 108: 157–183.

Weyreter H & Engelhardt WV (1984). Adaptation of Heidschnucken, Merino and Blackhead sheep to a fibrous roughage diet of poor quality. Can J Anim Sci 64: 152–153.

Wood NJ & Phua SH (1996) Variation in the control region sequence of the sheep mitochondrial genome. Anim Genet 27: 25–33.

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

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

Zardoya R, Villalta M, Lopez-Perez MJ, Garrido-Pertierra A, Montoya J & Bautista JM (1995) Nucleotide sequence of the sheep mitochondrial D-loop and its flanking tRNA genes. Curr Genet 28: 94–96.

Zhao X, Li N, Guo W, Hu X, Liu Z, Gong G, Wang A, Feng J & Wu C (2004) Further evidence for paternal inheritance of mitochondrial DNA in the sheep (Ovis aries). Heredity 93: 299–403.

Zvelebil M & Lillie M (2000) Transition to agriculture in eastern Europe. In: Price TD (ed.) Europe's first farmers. Cambridge University Press, Cambridge, UK, p 57–92.

Appendix 1

Molecular mean self-coancestries within each population *k* (*s[k]*), mean kinship distance (*Mean Dk*) between individuals from population *k* to the individuals in all studied populations (including the source population *k*), mean molecular co-ancestry (*Mean f*) of population to all populations (including population *k* itself). The remaining rows of the table represent the symmetric square matrix of molecular co-ancestries between populations (f_{kl}).

	Åland Sheep	Danish Landrace Sheep	Danish Texel	Danish Whiteheaded Marsh Sheep	Estonian Ruhnu sheep	Faeroe Island Sheep	Finnish Grey Landrace
S[k]	0.679	0.704	0.693	0.688	0.730	0.644	0.648
Mean Dk	0.462	0.459	0.463	0.452	0.483	0.451	0.45
Mean f	0.219	0.235	0.225	0.233	0.224	0.212	0.215
Åland Sheep	0.349	0.230	0.205	0.221	0.241	0.207	0.219
D.Landrace Sheep	0.230	0.412	0.265	0.252	0.219	0.199	0.211
D.Texel	0.205	0.265	0.389	0.278	0.215	0.196	0.236
D.Wh.Marsh Sheep	0.221	0.252	0.278	0.378	0.208	0.236	0.212
E.Ruhnu sheep	0.241	0.219	0.215	0.208	0.479	0.183	0.240
Faeroe Island Sheep	0.207	0.199	0.196	0.236	0.183	0.341	0.203
F.Grey Landrace	0.219	0.211	0.236	0.212	0.240	0.203	0.294
Finnsheep	0.228	0.204	0.203	0.197	0.228	0.214	0.231
Greenland Sheep	0.192	0.213	0.212	0.218	0.187	0.234	0.185
Icelandic Sheep	0.200	0.213	0.213	0.214	0.187	0.247	0.199
Icelandic Leader Sheep	0.197	0.210	0.227	0.240	0.183	0.258	0.202
Latvian Darkheaded	0.203	0.236	0.251	0.241	0.224	0.216	0.228
Lithuanian Blackface	0.226	0.257	0.244	0.258	0.234	0.212	0.217
Lith.Nat.Coarsewooled	0.208	0.196	0.185	0.217	0.192	0.197	0.186
Lith. Romanov	0.195	0.233	0.229	0.219	0.210	0.206	0.208
N.Cheviot Sheep	0.236	0.252	0.237	0.261	0.233	0.226	0.222
N.Dala Sheep	0.206	0.249	0.255	0.254	0.212	0.195	0.217
N.Feral Sheep	0.219	0.239	0.203	0.225	0.233	0.222	0.217
N.Fuglestad Sheep	0.226	0.235	0.205	0.235	0.203	0.190	0.192
N.Grey Troender Sheep	0.214	0.215	0.214	0.232	0.232	0.215	0.197
N.Rygja Sheep	0.214	0.232	0.209	0.240	0.216	0.181	0.195
N.Spael Sheep	0.200	0.218	0.193	0.209	0.204	0.230	0.188
N.Old Spael Sheep	0.204	0.229	0.196	0.209	0.205	0.196	0.184
N.Steigar Sheep	0.213	0.238	0.261	0.266	0.224	0.212	0.223
R.Viena sheep	0.214	0.222	0.229	0.215	0.236	0.191	0.242
S.Dala Fur Sheep	0.226	0.216	0.217	0.243	0.228	0.212	0.222
S.Finewool Sheep	0.222	0.222	0.217	0.201	0.240	0.205	0.225
S.Gotland Sheep	0.259	0.268	0.217	0.241	0.249	0.188	0.224
S.Gute Sheep	0.220	0.220	0.165	0.187	0.187	0.160	0.205
S.Forest Sheep	0.238	0.244	0.239	0.234	0.240	0.212	0.247
S.Roslag Sheep	0.170	0.224	0.201	0.208	0.157	0.191	0.202
S.Rya Sheep	0.200	0.243	0.191	0.212	0.233	0.215	0.210

	Finn-sheep	Green-land Sheep	Icelandic Sheep	Icelandic Leader Sheep	Lat. Dark-headed	Lith. Blackface	Lith. Native Coarse-wooled
s[k]	0.614	0.650	0.661	0.688	0.661	0.648	0.640
Mean Dk	0.439	0.460	0.459	0.465	0.45	0.434	0.462
Mean f	0.210	0.207	0.213	0.221	0.223	0.231	0.200
Åland Sheep	0.228	0.192	0.200	0.197	0.203	0.226	0.208
D.Landrace Sheep	0.204	0.213	0.213	0.210	0.236	0.257	0.196
D.Texel	0.203	0.212	0.213	0.227	0.251	0.244	0.185
D.Wh.Marsh Sheep	0.197	0.218	0.214	0.240	0.241	0.258	0.217
E.Ruhnu sheep	0.228	0.187	0.187	0.183	0.224	0.234	0.192
Faeroe Island Sheep	0.214	0.234	0.247	0.258	0.216	0.212	0.197
F.Grey Landrace	0.231	0.185	0.199	0.202	0.228	0.217	0.186
Finnsheep	0.254	0.190	0.197	0.207	0.203	0.211	0.190
Greenland Sheep	0.190	0.310	0.255	0.263	0.199	0.205	0.178
Icelandic Sheep	0.197	0.255	0.303	0.290	0.212	0.213	0.193
Icelandic Leader Sheep	0.207	0.263	0.290	0.362	0.219	0.222	0.208
Latvian Darkheaded	0.203	0.199	0.212	0.219	0.315	0.243	0.214
Lithuanian Blackface	0.211	0.205	0.213	0.222	0.243	0.297	0.229
Lith.Nat.Coarsewooled	0.190	0.178	0.193	0.208	0.214	0.229	0.276
Lith. Romanov	0.201	0.204	0.218	0.234	0.209	0.239	0.200
N.Cheviot Sheep	0.227	0.208	0.208	0.212	0.242	0.255	0.205
N.Dala Sheep	0.202	0.201	0.196	0.212	0.245	0.247	0.196
N.Feral Sheep	0.222	0.201	0.211	0.219	0.219	0.242	0.208
N.Fuglestad Sheep	0.196	0.179	0.186	0.197	0.213	0.225	0.205
N.Grey Troender Sheep	0.197	0.208	0.204	0.218	0.207	0.222	0.193
N.Rygja Sheep	0.195	0.178	0.177	0.183	0.194	0.237	0.205
N.Spael Sheep	0.199	0.230	0.240	0.232	0.209	0.216	0.182
N.Old Spael Sheep	0.189	0.179	0.201	0.203	0.199	0.213	0.179
N.Steigar Sheep	0.215	0.205	0.206	0.227	0.242	0.249	0.206
R.Viena sheep	0.221	0.185	0.193	0.199	0.210	0.221	0.179
S.Dala Fur Sheep	0.217	0.204	0.216	0.223	0.207	0.216	0.188
S.Finewool Sheep	0.227	0.208	0.214	0.216	0.215	0.230	0.199
S.Gotland Sheep	0.224	0.191	0.199	0.211	0.238	0.246	0.223
S.Gute Sheep	0.204	0.168	0.187	0.188	0.201	0.202	0.181
S.Forest Sheep	0.236	0.213	0.210	0.217	0.249	0.243	0.207
S.Roslag Sheep	0.167	0.241	0.219	0.203	0.215	0.199	0.193
S.Rya Sheep	0.209	0.182	0.194	0.187	0.212	0.231	0.187

	Lith. Romanov	N.Chev. Sheep	N.Dala Sheep	Norwegian Feral Sheep	N.Fugle-stad Sheep	N.Grey Troender Sheep	N.Rygja Sheep
s[k]	0.662	0.689	0.665	0.668	0.661	0.672	0.678
Mean Dk	0.458	0.454	0.453	0.451	0.462	0.462	0.471
Mean f	0.214	0.232	0.221	0.224	0.210	0.215	0.210
Åland Sheep	0.195	0.236	0.206	0.219	0.226	0.214	0.214
D.Landrace Sheep	0.233	0.252	0.249	0.239	0.235	0.215	0.232
D.Texel	0.229	0.237	0.255	0.203	0.205	0.214	0.209
D.Wh.Marsh Sheep	0.219	0.261	0.254	0.225	0.235	0.232	0.240
E.Ruhnu sheep	0.210	0.233	0.212	0.233	0.203	0.232	0.216
Faeroe Island Sheep	0.206	0.226	0.195	0.222	0.190	0.215	0.181
F.Grey Landrace	0.208	0.222	0.217	0.217	0.192	0.197	0.195
Finnsheep	0.201	0.227	0.202	0.222	0.196	0.197	0.195
Greenland Sheep	0.204	0.208	0.201	0.201	0.179	0.208	0.178
Icelandic Sheep	0.218	0.208	0.196	0.211	0.186	0.204	0.177
Icelandic Leader Sheep	0.234	0.212	0.212	0.219	0.197	0.218	0.183
Latvian Darkheaded	0.209	0.242	0.245	0.219	0.213	0.207	0.194
Lithuanian Blackface	0.239	0.255	0.247	0.242	0.225	0.222	0.237
Lith.Nat.Coarsewooled	0.200	0.205	0.196	0.208	0.205	0.193	0.205
Lith. Romanov	0.328	0.219	0.206	0.221	0.194	0.202	0.197
N.Cheviot Sheep	0.219	0.396	0.254	0.241	0.224	0.233	0.249
N.Dala Sheep	0.206	0.254	0.282	0.220	0.220	0.228	0.233
N.Feral Sheep	0.221	0.241	0.220	0.311	0.213	0.225	0.200
N.Fuglestad Sheep	0.194	0.224	0.220	0.213	0.297	0.214	0.226
N.Grey Troender Sheep	0.202	0.233	0.228	0.225	0.214	0.306	0.206
N.Rygja Sheep	0.197	0.249	0.233	0.200	0.226	0.206	0.309
N.Spael Sheep	0.208	0.201	0.192	0.242	0.183	0.217	0.184
N.Old Spael Sheep	0.198	0.201	0.195	0.240	0.202	0.211	0.188
N.Steigar Sheep	0.212	0.260	0.259	0.221	0.226	0.227	0.229
R.Viena sheep	0.235	0.218	0.212	0.227	0.197	0.199	0.202
S.Dala Fur Sheep	0.202	0.203	0.203	0.245	0.202	0.209	0.196
S.Finewool Sheep	0.220	0.223	0.217	0.231	0.199	0.219	0.193
S.Gotland Sheep	0.214	0.216	0.230	0.222	0.235	0.207	0.235
S.Gute Sheep	0.166	0.204	0.195	0.206	0.183	0.168	0.210
S.Forest Sheep	0.230	0.251	0.226	0.240	0.222	0.221	0.208
S.Roslag Sheep	0.185	0.194	0.204	0.183	0.183	0.215	0.164
S.Rya Sheep	0.221	0.219	0.219	0.215	0.215	0.217	0.225

	N.Spael Sheep	N.Old Spael Sheep	N.Steigar Sheep	Russian Viena sheep	Swedish Dala Fur Sheep	S.Fine-wool Sheep	S.Gotland Sheep
s[k]	0.666	0.718	0.655	0.722	0.770	0.656	0.693
Mean Dk	0.461	0.496	0.441	0.489	0.505	0.451	0.456
Mean f	0.213	0.204	0.228	0.214	0.222	0.218	0.232
Åland Sheep	0.200	0.204	0.213	0.214	0.226	0.222	0.259
D.Landrace Sheep	0.218	0.229	0.238	0.222	0.216	0.222	0.268
D.Texel	0.193	0.196	0.261	0.229	0.217	0.217	0.217
D.Wh.Marsh Sheep	0.209	0.209	0.266	0.215	0.243	0.201	0.241
E.Ruhnu sheep	0.204	0.205	0.224	0.236	0.228	0.240	0.249
Faeroe Island Sheep	0.230	0.196	0.212	0.191	0.212	0.205	0.188
F.Grey Landrace	0.188	0.184	0.223	0.242	0.222	0.225	0.224
Finnsheep	0.199	0.189	0.215	0.221	0.217	0.227	0.224
Greenland Sheep	0.230	0.179	0.205	0.185	0.204	0.208	0.191
Icelandic Sheep	0.240	0.201	0.206	0.193	0.216	0.214	0.199
Icelandic Leader Sheep	0.232	0.203	0.227	0.199	0.223	0.216	0.211
Latvian Darkheaded	0.209	0.199	0.242	0.210	0.207	0.215	0.238
Lithuanian Blackface	0.216	0.213	0.249	0.221	0.216	0.230	0.246
Lith.Nat.Coarsewooled	0.182	0.179	0.206	0.179	0.188	0.199	0.223
Lith. Romanov	0.208	0.198	0.212	0.235	0.202	0.220	0.214
N.Cheviot Sheep	0.201	0.201	0.260	0.218	0.203	0.223	0.216
N.Dala Sheep	0.192	0.195	0.259	0.212	0.203	0.217	0.230
N.Feral Sheep	0.242	0.240	0.221	0.227	0.245	0.231	0.222
N.Fuglestad Sheep	0.183	0.202	0.226	0.197	0.202	0.199	0.235
N.Grey Troender Sheep	0.217	0.211	0.227	0.199	0.209	0.219	0.207
N.Rygja Sheep	0.184	0.188	0.229	0.202	0.196	0.193	0.235
N.Spael Sheep	0.304	0.232	0.209	0.199	0.227	0.220	0.216
N.Old Spael Sheep	0.232	0.306	0.197	0.197	0.219	0.202	0.210
N.Steigar Sheep	0.209	0.197	0.289	0.224	0.220	0.225	0.234
R.Viena sheep	0.199	0.197	0.224	0.298	0.221	0.226	0.223
S.Dala Fur Sheep	0.227	0.219	0.220	0.221	0.447	0.213	0.231
S.Finewool Sheep	0.220	0.202	0.225	0.226	0.213	0.285	0.230
S.Gotland Sheep	0.216	0.210	0.234	0.223	0.231	0.230	0.380
S.Gute Sheep	0.192	0.188	0.189	0.178	0.210	0.178	0.261
S.Forest Sheep	0.223	0.209	0.254	0.248	0.287	0.239	0.274
S.Roslag Sheep	0.207	0.150	0.202	0.164	0.167	0.211	0.189
S.Rya Sheep	0.223	0.202	0.223	0.228	0.172	0.215	0.252

	S.Gute Sheep	S.Forest Sheep	S.Roslag Sheep	S.Rya Sheep
s[k]	0.711	0.738	0.802	0.690
Mean Dk	0.497	0.474	0.535	0.47
Mean f	0.201	0.236	0.207	0.217
Åland Sheep	0.220	0.238	0.170	0.200
D.Landrace Sheep	0.220	0.244	0.224	0.243
D.Texel	0.165	0.239	0.201	0.191
D.Wh.Marsh Sheep	0.187	0.234	0.208	0.212
E.Ruhnu sheep	0.187	0.240	0.157	0.233
Faeroe Island Sheep	0.160	0.212	0.191	0.215
F.Grey Landrace	0.205	0.247	0.202	0.210
Finnsheep	0.204	0.236	0.167	0.209
Greenland Sheep	0.168	0.213	0.241	0.182
Icelandic Sheep	0.187	0.210	0.219	0.194
Icelandic Leader Sheep	0.188	0.217	0.203	0.187
Latvian Darkheaded	0.201	0.249	0.215	0.212
Lithuanian Blackface	0.202	0.243	0.199	0.231
Lith.Nat.Coarsewooled	0.181	0.207	0.193	0.187
Lith. Romanov	0.166	0.230	0.185	0.221
N.Cheviot Sheep	0.204	0.251	0.194	0.219
N.Dala Sheep	0.195	0.226	0.204	0.219
N.Feral Sheep	0.206	0.240	0.183	0.215
N.Fuglestad Sheep	0.183	0.222	0.183	0.215
N.Grey Troender Sheep	0.168	0.221	0.215	0.217
N.Rygja Sheep	0.210	0.208	0.164	0.225
N.Spael Sheep	0.192	0.223	0.207	0.223
N.Old Spael Sheep	0.188	0.209	0.150	0.202
N.Steigar Sheep	0.189	0.254	0.202	0.223
R.Viena sheep	0.178	0.248	0.164	0.228
S.Dala Fur Sheep	0.210	0.287	0.167	0.172
S.Finewool Sheep	0.178	0.239	0.211	0.215
S.Gotland Sheep	0.261	0.274	0.189	0.252
S.Gute Sheep	0.415	0.204	0.202	0.192
S.Forest Sheep	0.204	0.380	0.181	0.237
S.Roslag Sheep	0.202	0.181	0.622	0.213
S.Rya Sheep	0.192	0.237	0.213	0.349

Original papers

I Tapio M, Marzanov N, Ozerov M, Ćinkulov M, Gonzarenko G, Kiselyova T, Murawski M, Viinalass H & Kantanen J (2006) Sheep mitochondrial DNA variation in European, Caucasian and Central Asian areas. Mol Biol Evol 23: 1776–1783.

II Tapio M, Tapio I, Grislis Z, Holm L-E, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2006) Geographical patterns in the variation of mitochondrial DNA in Northern European sheep. Manuscript.

III Tapio M, Tapio I, Grislis Z, Holm L-E, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2005) Native breeds demonstrate high contributions to the molecular variation in northern European sheep. Mol Ecol 14: 3951–3963.

IV Tapio I, Tapio M, Grislis Z, Holm LE, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H & Eythorsdottir E (2005) Unfolding of population structure in Baltic sheep breeds using microsatellite analysis. Heredity 94: 448–456.

V Tapio M, Miceikiene I, Vilkki J & Kantanen J (2003) Comparison of microsatellite and blood protein diversity in sheep: inconsistencies in fragmented breeds. Mol Ecol 12: 2045–2056.

Original publications are not included in the electronic version of the dissertation.

457. Iivari, Netta (2006) Discourses on ‘culture’ and ‘usability work’ in software product development

458. Vähäoja, Pekka (2006) Oil analysis in machine diagnostics

459. Mutanen, Marko (2006) Genital variation in moths—evolutionary and systematic perspectives

460. Bhaumik, Prasenjit (2006) Protein crystallographic studies to understand the reaction mechanism of enzymes: α-methylacyl-CoA racemase and argininosuccinate lyase

461. Korkalo, Tuomo (2006) Gold and copper deposits in Central Lapland, Northern Finland, with special reference to their exploration and exploitation

462. Pahnila, Seppo (2006) Assessing the usage of personalized web information systems

463. Puhakainen, Petri (2006) A design theory for information security awareness

464. Rytkönen, Anna (2006) The role of human replicative DNA polymerases in DNA repair and replication

465. Rönkä, Antti (2006) Dynamics, genetic structure and viability of a small and declining Temminck's stint (*Calidris temminckii*) population

466. Wäli, Piippa (2006) Environment and genetic background affecting endophyte-grass symbiosis

467. Broggi, Juli (2006) Patterns of variation in energy management in wintering tits (*Paridae*)

468. Mykrä, Heikki (2006) Spatial and temporal variability of macroinvertebrate assemblages in boreal streams: implications for conservation and bioassessment

469. Kekonen, Teija (2006) Environmental information from the Svalbard ice core for the past 800 years

470. Ojala, Pasi (2006) Implementing a value-based approach to software assessment and improvement

471. Saarenketo, Timo (2006) Electrical properties of road materials and subgrade soils and the use of Ground Penetrating Radar in traffic infrastructure surveys

472. Jokikokko, Erkki (2006) Atlantic salmon (*Salmo salar* L.) stocking in the Simojoki river as a management practice

Book orders:
OULU UNIVERSITY PRESS
P.O. Box 8200, FI-90014
University of Oulu, Finland

Distributed by
OULU UNIVERSITY LIBRARY
P.O. Box 7500, FI-90014
University of Oulu, Finland

UNIVERSITY OF OULU P.O. Box 7500 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

ISBN 951-42-8234-5 (Paperback)
ISBN 951-42-8235-3 (PDF)
ISSN 0355-3191 (Print)
ISSN 1796-220X (Online)