Università Degli Studi Della Tuscia Di Viterbo

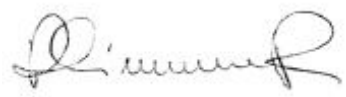
Dipartimento per l' Innovazione nei sistemi Biologici, Agroalimentari e Forestali (DIBAF)

DOTTORATO di RICERCA IN
ECOLOGIA E GESTIONE DELLE RISORSE BIOLOGICHE
XXIV CICLO

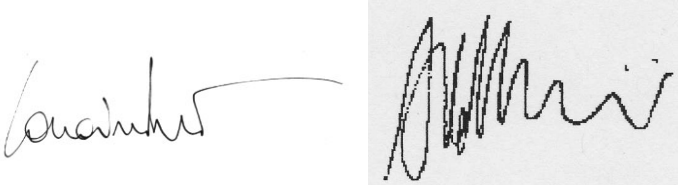
Assessing livestock genetic variability by applying molecular markers: mtDNA and SNPs can help in evaluating genetic structure and breeds relationship

AGR/17

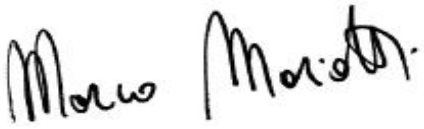
Coordinatore:
Dott. Roberta Cimmaruta



Tutor:
Dott. Lorraine Pariset – Prof. **Alessio Valentini**



Dottorando: **Marco Mariotti**



1. Introduction

1.1. Conservation of biodiversity in livestock

The domestication of plants and animals was a crucial step in human demographic and cultural development, the Neolithic Revolution has radically changed the behaviour of the man who has turned from hunter-gatherer to farmer-breeder. Recently, the use of molecular genetics had a greater impact on our understanding of how these events took place, compared to traditional research of anthropologists and archaeobiologists. In particular, recent studies made it possible to identify the wild ancestor of modern livestock and the nature of livestock expansion in past millennia (Salamini et al. 2002; Taberlet et al. 2008). Currently, these data have more direct importance for the management and conservation of animal and plant genetic resources. The loss of agricultural diversity in the face of increasing pressures from modern farming is a cause for global concern (Bruford et al. 2003).

After the domestication, the main evolutionary forces of mutation, adaptation, genetic drift, isolation and selective breeding have created an enormous diversity of local livestock populations. During the last centuries, highly productive breeds have replaced local ones across the world, due to phenomena such as the rise of the concept of breed, the progress of selection programmes, the artificial insemination and the embryo transfer; this development has inevitably led to a significant loss of genetic variability (Taberlet et al., 2011).

While the conservation problem in domestic animals may appear as a paradox for the presence of extremely high population size of farm animals (Table 1), the issue was already raised in the seventies by the Food and Agriculture Organization of the United Nations (FAO) (Taberlet et al. 2011).

	Cattle	**Sheep**
Population size	1,367,335	1,060,606
Number of breeds	1311	1409
Number of extinct breeds	209	180

Table 1: Population sizes, current number of breeds, and number of extinct breeds for cattle and sheep at the worldwide level (statistics concerning 169 countries, Taberlet et al. 2011).

In Europe, it is necessary to protect the existing great heritage of genetic diversity; in fact in our continent there is a gradual disappearance of indigenous breeds that are substituted with more productive cosmopolitan breeds. This trend is accelerated by the abandonment of farming in marginal areas.

The biodiversity conservation of farm animals is universally recognized as a priority. The ideal approach would preserve the genetic diversity function of breeds, that is, the variability of the genes that control characters expressed by the animals that are or may be useful in the future. Indeed, the knowledge of these genes and often also of phenotypic traits of many breeds, considered at risk of extinction, is still very fragmented and insufficient. Even less if considering the characters that may be useful in the future. In recent years molecular markers were used to compensate this lack of information.

Neutral Molecular markers are not affected by selection and are then used to reconstruct the species evolutionary history as indicators of genetic diversity existing within and between breeds and to locate and identify useful genes. The understanding of the evolutionary history of livestock species of interest and the estimates of genetic variability within and between breeds are fundamental for decision-making on the overall theme of conservation. For example, data on genetic variation within breeds obtained by molecular markers can be used to estimate and control, through proper breeding management, the level of consanguinity whose excessive increase would undermine the survival. The molecular information on genetic variability among breeds allows instead the identification of breeds carrying specific genotypes that is worth preserving.

1.2. Molecular markers

Molecular markers are representative of heritable and polymorphic characteristics at level of species, population and individual. A molecular marker is any chromosomal locus whose allelic variants can be easily identified by directly analyzing the DNA. A good molecular marker must be polymorphic, codominant, easily identifiable, repeatable and must have a simple mechanism of transmission (Mendelian or Uniparental).

In recent decades, advances in molecular biology since the development of PCR technology (Mullis & Faloona, 1987) have contributed to the development of new molecular markers. The main markers, that are based on PCR technology, are microsatellites (Weber & May, 1989), RAPD (Random Amplification Polymorphic DNA, Williams et al., 1990), AFLP (Amplified Fragment Length Polymorphism, Vos et al., 1995) and SNP (Single Nucleotide Polymorphism, Landegren et al., 1998), while the RFLP are based on cutting enzymes restriction.

In recent domestication studies the mitochondrial DNA (Meadows et al., 2011; Bonfiglio et al., 2010) and the variable sequences on the Y chromosome (Meadows et al., 2008; Bollongino et al., 2008) have been used as markers.

1.2.1. SNP (Single Nucleotide Polymorphisms)

The SNP (Single Nucleotide Polymorphism) markers identify, as point mutations, the single nucleotide polymorphisms, but also short insertions and deletions in DNA sequence (Figure 1). The SNPs are very abundant (in humans representing more than 90% of the total polymorphisms), and randomly distributed in the genome. There are two types of nucleotide base substitutions:

- Transition, occurring between purines (A, G) or between pyrimidines (C, T). This type of substitution constitutes two thirds of all SNPs.
- Transversion, occurring between a purine and a pyrimidine.

The majority of SNPs reside in non-coding regions in which the probability of finding point mutations is 4 times greater than in coding regions. A SNP in a coding region may have two different effects on the resulting protein:

- Synonymous, the substitution causes no amino acid change to the protein (silent mutation).
- Non-synonymous, the substitution results in an alteration of the encoded amino acid. A missense mutation changes the protein by causing a change of codon. A nonsense mutation results in a misplaced termination codon.

Also, SNPs may reside in regulatory region of genes, and these are capable of changing the protein's production.

The SNPs are considered biallelic markers (two alleles that may differ in a given nucleotide position in a diploid cell) and codominant (allow discrimination of homozygote from heterozygote individuals).

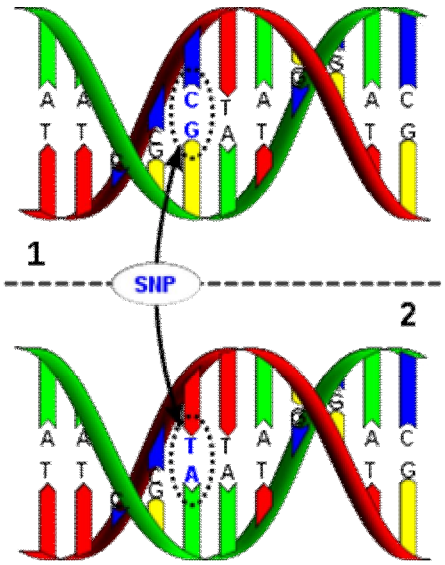


Figure 1: C/T polymorphism

There are different techniques to identify the polymorphisms, including the sequencing of the fragment carrying the mutation. recently high throughput technologies are used for the identification of hundreds of SNPs simultaneously, based on hybridization as the BeadArray (Shen et al., 2005), or tens of thousands to millions, with the Bead-chip (Illumina).

These markers are used for evolutionary studies but also for mapping and association with QTLs (Quantitative Trait loci). In livestock, markers are also used to study the traceability of products and they are used in diseases association studies.

In cattle, for example, SNPs are used for the identification of genetic variation that modulates the corresponding change in economically important production traits, differential susceptibility to disease and favourable host response to vaccines (VanRaden et al., 2009).

1.2.2. Mitochondrial DNA (mtDNA)

To investigate the origins of humans or animals an ideal marker should have several characteristics (Bruford et al., 2003):

- It must be sufficiently evolutionary conserved to allow the identification of the wild taxon from which the species descends
- It must be variable and sufficiently structured across the geographical range of the species so that the approximate location of domestication can be identified
- It should evolve at a rapid but constant rate.

The mitochondrial DNA has these features, and it is used in many evolutionary studies. In mammals, the mtDNA is the nucleic acid located in mitochondrial organelle; it is about 20 kb long,

exclusively maternally inherited, haploid and does not undergo recombination (Figure 2). For example, in mtDNA of sheep there are 13 protein-coding genes, 2 sr-RNAs, control region and 22 t-RNAs (Table 2, Hiendleder et al., 1998).

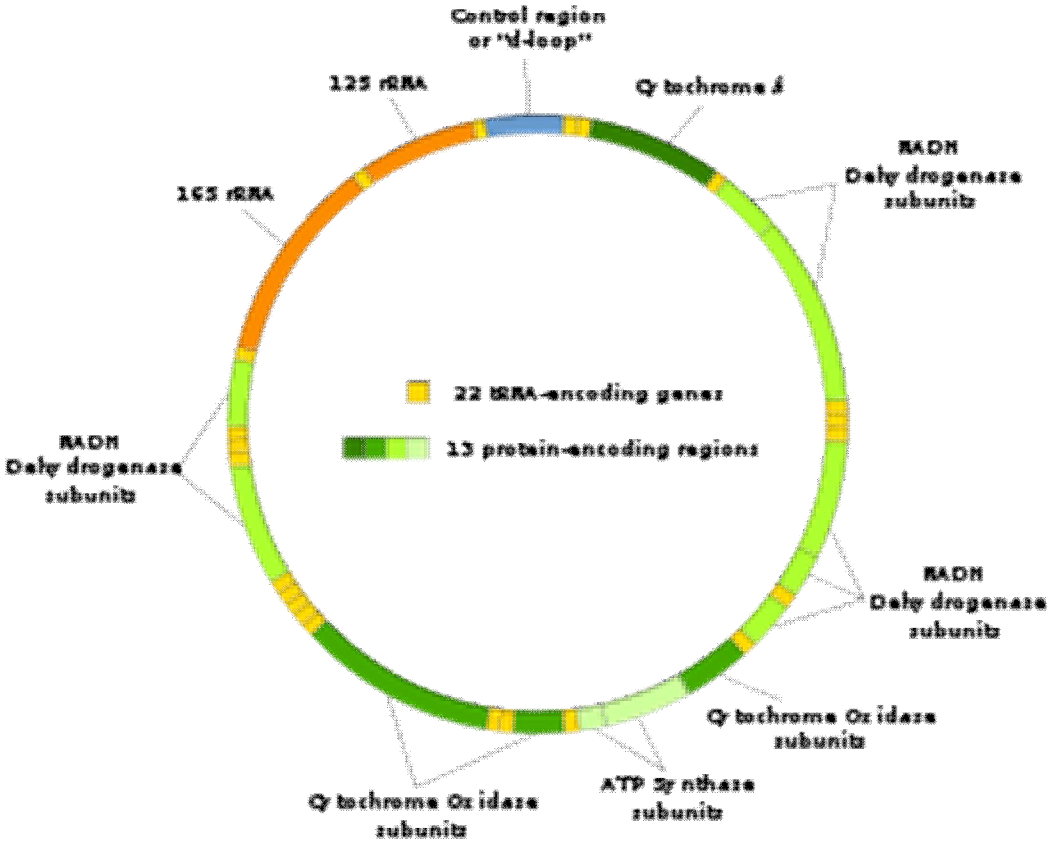


Figure 2: Mitochondrial DNA

Feature	From	To	Size	Start codon	Stop codon[b]	3′ spacer
tRNA-Phe	1	68	68			
12S rRNA	69	1,026	958			
tRNA-Val	1,027	1,093	67			
16S rRNA	1,094	2,667	1,574			
tRNA-Leu (UUR)	2,668	2,742	75			AA
NADH1	2,745	3,700	956	ATG	TAa	
tRNA-Ile	3,701	3,769	69			
tRNA-Gln (L)	3,767	3,838	72			AT
tRNA-Met	3,841	3,909	69			
NADH2	3,910	4,951	1,042	ATA	Taa	
tRNA-Trp	4,952	5,018	67			A
tRNA-Ala (L)	5,020	5,088	69			A
tRNA-Asn (L)	5,090	5,162	73			
Origin of L-strand repl.	5,163	5,194	32			
tRNA-Cys (L)	5,195	5,262	68			
tRNA-Tyr (L)	5,263	5,330	68			C
COI	5,332	6,876	1,545	ATG	TAA	
tRNA-Ser (UCN) (L)	6,874	6,944	71			TAAAC
tRNA-Asp	6,950	7,017	68			T
COII	7,019	7,702	684	ATG	TAA	AAT
tRNA-Lys	7,706	7,773	68			T
ATPase8	7,775	7,975	201	ATG	TAA	
ATPase6	7,936	8,615	680	ATG	TAa	
COIII	8,616	9,399	784	ATG	Taa	
tRNA-Gly	9,400	9,468	69			
NADH3	9,469	9,815	347	ATA	TAa	
tRNA-Arg	9,816	9,884	69			
NADH4L	9,885	10,181	297	ATG	TAA	
NADH4	10,175	11,552	1,378	ATG	Taa	
tRNA-His	11,553	11,621	69			
tRNA-Ser (AGY)	11,622	11,681	60			A
tRNA-Leu (CUN)	11,683	11,753	71			
NADH5	11,754	13,574	1,821	ATA	TAA	
NADH6 (L)	13,558	14,085	528	ATG	TAA	
tRNA-Glu (L)	14,086	14,154	69			ACTA
Cyt *b*	14,159	15,298	1,140	ATG	AGA	CAA
tRNA-Thr	15,302	15,371	70			
tRNA-Pro (L)	15,371	15,436	66			
Control region	15,437	16,616	1,180			

[a] Nucleotide number 1 is the 5′ end of the tRNA-Phe-specifying gene. Anticodons for the two tRNA-Leu and the two tRNA-Ser are given in parentheses. (L) denotes light-strand sense. Positions include the 5′ and 3′ nt of each feature. ATPase6 and ATPase8, genes encoding subunits 6 and 8 of ATPase; COI-III, genes encoding subunits I–III of cytochrome *c* oxidase; Cyt *b*, gene encoding cytochrome *b;* NADH1–6, genes encoding subunits 1–6 of nicotinamide adenine dinucleotide dehydrogenase.

[b] Incomplete stop signals are denoted by lowercase letters.

Table 2: Features of the *Ovis aries* mitochondrial genome.

Mitochondrial polymorphisms have been used to explain the origins of many modern domestic species. The mtDNA control region shows the highest evolutionary rate, and by analyzing its variability it is possible to define maternal lineages. Meadows et al. (2011) showed that the control region with removal of the hypervariable repeat is the most suitable for phylogenetic analysis. Maternal lineages can be grouped into phylogenetically resolved haplogroups and in domestic species haplogroups have been used to infer domestication events (Bruford et al 2003). Recent studies in goats (Naderi et al 2007), cattle (Achilli et al 2009), and pigs (Larson et al 2010) have revealed a high number of maternal lineages. In sheep, two (Wood and Phua et al 1996; Hiendleder

et al 1998a; Hiendleder et al 2002), three (Guo et al 2005; Pedrosa et al 2005), and more recently five (Meadows et al 2007; Meadows et al 2011) lineages have been identified.

1.3. The goals of the thesis

The research has focused on different analyses of genetic diversity in livestock. Specifically, genetic variability in different European breeds belonging to two livestock species, *Bos taurus* and *Ovis aries,* were examined by applying different molecular markers.

The three main projects followed during the thesis are:

- The identification of new SNPs in TLR2, TLR4 and TLR6 genes in cattle. The Toll-like receptors (TLRs) play an important role in the recognition of components of pathogens and subsequent activation of the innate immune response We screened nucleotide sequences of bovine TLR2, TLR4 and TLR6 genes to identify SNPs can be used in diseases resistance studies in cattle. The frequencies of the SNPs were assessed in 16 different bovine European cattle breed (Jersey, South Devon, Aberdeen Angus and Highlands, from Great Britain; Holstein, Danish Red and Simmental, from Denmark; Asturiana de los Valles, Casina, Avilena and Pirenaica, from Spain; Piemontese, Marchigiana and Maremmana, from Italy; Limousin and Charolais, from France) and used to carry out a phylogenetic analysis to describe the relationships between the breeds.
- The analysis of the relationship between Maremmana, Turkish Grey and Hungarian Grey breeds belonging to the same Podolic group of cattle and to verify whether their genetic state reflects their history. We genotyped about 100 SNPs on individuals belonging to these breeds and compared them to genotypes of individuals of two Italian beef breeds, Marchigiana and Piemontese, which underwent different selection and migration histories.
- The investigation the genetic diversity of sheep breeds of three countries of the Mediterranean basin: Albania, Greece, and Italy. We aimed at investigating the geographic distribution of the genetic diversity of sheep breeds in Albania (Bardhoka, Ruda, and Shkordane), Greece (Kalarritiko, Orino, Pilioritiko, Kefalleneas, Lesvos, Kymi, Karagouniko, Skopelos, Anogeiano, and Sfakia), and Italy (Bergamasca, Delle Langhe, Laticauda, Altamurana, and Gentile di Puglia). To accomplish that, we employed sequence data from the mitochondrial D-loop and 37 previously described SNPs (nuclear loci).

2. Articles

2.1. Polymorphisms within the Toll-Like Receptor (TLR)-2, -4 and -6 Genes in Cattle

Diversity **2009**, *1*, 7-18; doi:10.3390/d1010007

OPEN ACCESS

diversity

ISSN 1424-2818

www.mdpi.com/journal/diversity

Article

Polymorphisms within the Toll-Like Receptor (TLR)-2, -4, and -6 Genes in Cattle

Marco Mariotti [1], John L. Williams [2], Susana Dunner [3], Alessio Valentini [1] and Lorraine Pariset [1,*]

[1] Dipartimento di Produzioni Animali, Università della Tuscia, Viterbo, Italy; E-Mails: mariotti@unitus.it (M.M.); alessio@unitus.it (A.V.)

[2] Parco Tecnologico Padano, Polo Universitario, Lodi, Italy; E-Mail: john.williams@tecnoparco.org

[3] Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense, Madrid, Spain; E-Mail: dunner@vet.ucm.es

* Author to whom correspondence should be addressed: E-Mail: pariset@unitus.it; Tel.: +39 0761 357454; Fax: +39 0761 357434

Received: 4 May 2009 / Accepted: 17 July 2009 / Published: 30 July 2009

Abstract: In mammals, members of the TLR gene family play a primary role in the recognition of pathogen-associated molecular patterns from bacteria, viruses, protozoa and fungi. Recently, cattle TLR genes have been mapped to chromosomes using a radiation hybrid panel. Nucleotide sequences of bovine TLR2, TLR4 and TLR6 genes were screened to identify novel SNPs that can be used in studies of cattle resistance to diseases. In total, 8 SNPs were identified and were submitted to the NCBI dbSNP database. The frequencies of the SNPs were assessed in 16 different bovine European cattle breeds and a phylogenetic analysis carried out to describe the relationships between the breeds. Even if from our analysis the SNPs do not appear located in loci under selection, a deviation of three SNPs from Hardy Weinberg equilibrium was observed, and we hypothesize that some of the polymorphisms may be fixated since many generations. The described variations in immune function related genes will contribute to research on disease response in cattle. In fact, the SNPs can be used in association studies between polymorphisms and cattle resistance to diseases.

Keywords: toll-like receptors; cattle; SNPs; diversity; genetic distances

1. Introduction

The immune system in mammals consists of innate and adaptive immune responses. Adaptive immunity is mediated by antigen specific T and B cells responses, and is observed only in vertebrates. Innate immunity, however, is conserved between invertebrates and vertebrates [1]. Toll-like Receptors (TLRs) play an important role in the recognition of components of pathogens and subsequent activation of the innate immune response, which then leads to development of adaptive immune responses [2,3]. The TLRs are an ancient gene group which is found both in invertebrates and vertebrates; related genes are found also in plants [4]. In mammals, members of the TLR gene family play a primary role in the recognition of pathogen-associated molecular patterns (PAMPs) in proteins from bacteria, viruses, protozoa and fungi [5,6]. Mammalian TLRs derive their name from the Drosophila Toll protein, with which they share sequence similarity. The drosophila Toll protein was shown to be involved in dorsal-ventral pattern formation in fly embryos and was also implicated as a key component of host immunity against fungal infection [7-9].

The TLRs consist of a large extracellular domain responsible for PAMP binding, a transmembrane domain and an intracellular Toll/interleukin-1 receptor (TIR) domain which binds molecules and initiates cellular immune responses [10]. The extracellular domains are composed of about 20 leucine-rich repeats (LRRs) motifs of 20–30 amino acids (AA) and form a solenoid shape with the potential to bind the TLR specific PAMP [11].

Ten TLRs, which recognize molecular patterns from all major classes of pathogens, have been identified in mammals, eleven in mice [12,13]. TLRs operate with diverse variety of ligands ranging from hydrophilic nucleic acid to LPS, furthermore the heterodimerization expands the ligand spectrum [14]. TLR2 and TLR4 recognize bacterial cell components, and are critical in the immune response against Gram positive and negative bacteria [15]. TLR6 in association with TLR2 recognizes a wide variety of bacterial cell wall components including lipopolysaccharides, teichoic acid and lipoproteins [16,17] and induce NFkB signalling pathway [18].

Recently, all 10 TLR genes have been mapped in cattle using a radiation hybrid panel: TLR2 and TLR4 have been previously mapped to the proximal end of *Bos taurus* chromosome (BTA) 17 and the distal end of BTA 8, respectively [19]. TLR6, TLR1 and TLR10 cluster on BTA 6 [20], as observed on human chromosome 4; this organization is most likely the result of gene duplication [21].

Several studies have shown that mutations in the TLR may reduce the ability of the protein to recognise PAMP and hence interfere with innate immune activation. Describing genetic variation in these loci in relation to resistance against specific diseases in livestock may be useful in guiding genetic selection for disease resistance. Single nucleotide polymorphisms (SNPs) within TLR genes in humans seem to be associated with susceptibility to infection by specific diseases [22]. Among cattle genes TLR1, TLR5 and TLR10, 98 polymorphisms have been identified, 14 of which are non synonymous SNPs positioned in domains considered to be functionally significant [23]. Eighty three polymorphisms have been also identified for bovine TLR2 and TLR6 [24]. The initiation of the innate response to bovine respiratory syncytial virus (BRSV) requires the interaction of the viral F protein with TLR4, which leads to activation of NFkB via the Myd88-dependent pathway [25]. A recent study showed an association between TLR mutations and increased susceptibility to MAP (*Mycobacterium avium*

paratubercolosis) infection in cattle, exactly two missense mutations in TLR4 (LRR domain) were associated with MAP infection [26].

In this study, we screened nucleotide sequences of bovine TLR2, TLR4 and TLR6 genes to identify SNPs that can be used in disease resistance studies in cattle. Eight new SNPs were identified and their frequency assessed in 16 different European cattle breeds.

Materials and Methods

Samples:

A total of 951 individuals belonging to the following European breeds were analysed: Jersey (50), South Devon (43), Aberdeen Angus (45) and Highlands (48), from Great Britain; Holstein (60), Danish Red (59) and Simmental (30), from Denmark; Asturiana de los Valles (66), Casina (66), Avilena (65) and Pirenaica (73), from Spain; Piemontese (67), Marchigiana (36) and Maremmana (91), from Italy; Limousin (72) and Charolais (80), from France. Genomic DNA was isolated from blood using conventional methods and concentration and quality were evaluated by agarose gel electrophoresis.

Polymerase Chain Reaction (PCR) Conditions:

PCR primers for TLR2, 4 and 6 were designed using *PolyPrimers* [27] from the sequences available in Genebank (TLR2: AY634629, TLR4: DQ839567, TLR6: AJ618974) to amplify genomic fragments of approximately 1 kb (Table 1, Figure S1) covering most of the gene sequence. Each polymerase chain reaction (PCR) was performed in a total volume of 30 μL containing 30 ng of genomic DNA, 1.6 pMol of each primer (Sigma-Aldrich), 200 μM dNTPs, 1X PCR buffer and 0.2 units of *Taq* DNA polymerase (Promega) on a PCR Express cycler (Hybaid), using the annealing temperatures reported in Table 1. A 5 minutes denaturation step was followed by 14 cycles of denaturation at 94 °C (30 sec), annealing starting from T.A. + 7 °C and decreasing 0.5 °C per cycle (45 sec) and extension at 72 °C (40 sec), then by 20 cycles of denaturation at 94 °C (30 sec), annealing at T.A. (45 sec) and extension at 72 °C (40 sec); the final extension step was carried out at 72 °C for 5 minutes.

Sequence analysis:

PCR products were purified through ExoSap-IT (USB Corporation) to remove residual primers and dNTPs and used as templates for forward and reverse sequencing reactions. Sequencing was performed by means of a ceq 8,800 sequencer using DTCS QuickStart Kit and purifying with Agencourt CleanSEQ 96 (Beckman Coulter), according to manufacturer instructions. To identify SNPs, sequences of at least one individual each of six different breeds (Maremmana, Charolais, Jersey, Holstein, Pirenaica and Piemontese) were analysed and aligned with *Bioedit* software [28]. The putative SNPs identified by sequencing were confirmed and allele frequencies estimated by genotyping 951 individuals. SNP genotyping was performed by Kbiosciences using the patented technology KASPar (www.Kbioscience.com).

Data analysis:

Allelic frequencies, Gene Diversity, Heterozygosity and PIC were calculated using *Powermarker* software [29]. Genotypes were analysed using *Fdist2* software to verify whether any of the loci were under selection [30]. Hardy-Weinberg equilibrium and *Nei* genetic distances [31] between populations pairs were calculated using *Powermarker*. The *Neighborjoining algorithm* was used to calculate the phylogeny relationship which was visualised using *Treeview* [32].

Table 1. Sequence of Forward (Fw) and Reverse (Rw) primers, annealing temperature (T.A.), amplicon size and amplicon position relatively to Genbank sequences.

Locus	Sequence (5'→3')	T.A. (°C)	Amplicon size (bp)	Amplicon position	Genbank Accession #
TLR2	Fw: CTGTCCAACAATGAGATCACCT Rw: AATTCTGTCCAAACTCAGTGCT	49	735	311-1045	AY634629
TLR2	Fw: GTTCAGGTCCCTTTATGTCTTG Rw: ATGGGTACAGTCATCAAACTCT	47	509	493-1003	AY634629
TLR2	Fw: ACTACCGCTGTGACTCTCCCTC Rw: GACCACCACCAGACCAAGACT	55	711	1818-2530	AY634629
TLR2	Fw: CTCCCTTTCTGAATGCCACA Rw: AAAGTATTGGAGCTTCAGCA	47	754	1876-2631	AY634629
TLR4	Fw: GTGTGGAGACCTAGATGACTGG Rw: GTACGCTATCCGGAATTGTTCA	50	705	7938-8644	DQ839567
TLR4	Fw: CTACCAAGCCTTCAGTATCTAG Rw: GGCATGTCCTCCATATCTAAAG	47	741	8880-9623	DQ839567
TLR4	Fw: TCAGGAACGCCACTTGTCAGCT Rw: TGAACACGCCCTGCATCCATCT	55	710	9635-10346	DQ839567
TLR6	Fw: AAAGAATCTCCCATCAGAAGCT Rw: GAAGGATACAACTTAGGTGCAA	46	515	228-745	AJ618974
TLR6	Fw: CTGCCCATCTGTAAGGAATTTG Rw: GATAAGTGTCTCCAATCTAGCT	47	739	624-1382	AJ618974
TLR6	Fw: TTGGAAACACTGGATGTTAGCT Rw: ACTGGAGAGTTCTTTGGAGTTC	49	710	1428-2138	AJ618974
TLR6	Fw: CTGCCTGGGTGAAGAATGAATT Rw: TGTAGTTGCACTTCCGGGCT	50	715	2173-2888	AJ618974

2. Results and Discussion

To discover SNPs in the three TLR genes, 12 PCR fragments were amplified and sequenced, five for TLR2, three for TLR4 and four for TLR6. One of the TLR2 primer pairs was soon discharged because of BLASTing problems. We then choose nine of 12 fragments, giving better results in terms of amplification and sequencing. In total eight SNPs were identified, three in TLR2, three in TLR4 and two in TLR6 [33] and were deposited in NCBI dbSNP (the ss# identities are listed in Table 2). These three genes are very important because they could be involved in immune response against various bovine diseases. In fact, TLR2 and TLR6 are critical in the immune response against Gram positive bacteria, TLR4 against Gram negative bacteria and virus. The polymorphisms in TLRs may reduce the ability of the protein to recognise ligands.

Table 2. Characterization of the detected SNPs.*: also described by Seabury and Womack [24].

SNPs	Position in GeneBank sequence[1]	aa change	Position in the protein[2]	SNP ID number
TLR2_591G>A	591	non - coding		ss107911951
TLR2_738A>G	738	non - coding		ss107911952
*TLR2_767G>A**	767	non - coding		ss107911953
TLR4_254G>A	254	non - coding		ss107911954
TLR4_1678C>T	1678	Synonymous (Ser)	552: LRR domain	ss107911955
TLR4_2043T>C	2043	non - coding		ss107911956
*TLR6_855G>A**	855	Asp/Asn	214	ss107911957
*TLR6_2315T>C**	2315	Synonymous (Phe)	315: TIR domain	ss107911958

[1] Nucleotide positions are numbered relatively to the first base of the sequence in GeneBank.
[2] Aminoacid position are numbered according to protein sequence in GeneBank (TLR4: NP776623, TLR6: NP001001159).

The allele frequencies are reported in Table 3 and major allele frequencies ranged from 0.557 (locus TLR2_767) to 0.969 (locus TLR2_738). Except for the latter, in all SNPs the frequency of the minor allele is greater than 5%. Observed heterozygosity (Ho) and Expected heterozygosity (He) of the loci determined from SNP frequencies ranged from 0.051 to 0.466 and from 0.060 to 0.493, respectively. Polymorphism Information Content (PIC) ranged from 0.058 to 0.372 (Table 3).

Eight of the breeds analysed were polymorphic at all the SNPs (Holstein, Asturiana de los Valles, Casina, Avilena, Pirenaica, Piemontese, Charolais) and five SNPs were polymorphic in all the breeds (*TLR2_767, TLR4_254, TLR4_1678, TLR6_855, TLR6_2315*). Both SNPs identified in TLR6 gene were polymorphic in all the breeds.

Table 3. Frequencies of the major allele (M.A.F.), expected heterozygosity (He), observed heterozygosity (Ho), Polymorphism Information Content (PIC) of the 8 characterized SNPs.

SNP	M.A.F.	He	Ho	PIC
TLR2_591	0.866	0..33	0.205	0.206
TLR2_738	0.969	0.060	0.051	0.058
TLR2_767	0.557	0.493	0.461	0.372
TLR4_254	0.595	0.482	0.466	0.366
TLR4_1678	0.655	0.452	0.421	0.350
TLR4_2043	0.843	0.264	0.234	0.230
TLR6_855	0.608	0.477	0.444	0.363
TLR6_2315	0.688	0.429	0.425	0.337

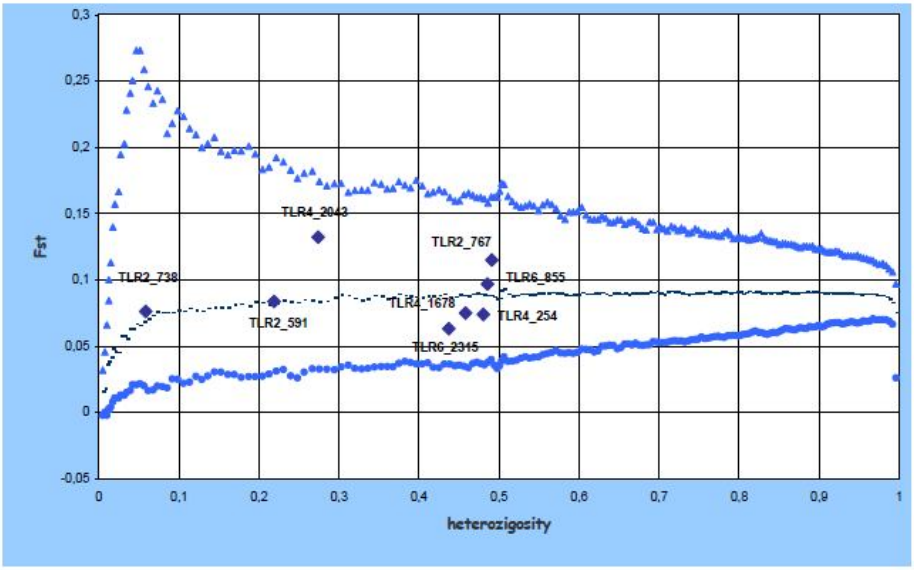
Some breeds were fixed at a number of SNPs, particularly *TLR2_738* is fixed in seven breeds (Highlands, Jersey, Limousine, Marchigiana, Maremmana, Simmenthal and South Devon), as shown in Table 4. *TLR2_591* and *TLR4_2043* are fixed in one breed (South Devon and Highlands, respectively). Interestingly, the SNP *TLR2_738* is fixed in three breeds from Great Britain: Highlands, Jersey and South Devon. Indeed, the only two breeds with two fixed alleles are South Devon and Highlands. Moreover, its rare allele frequency is lower than 0.05 in seven other breeds (Aberdeen Angus, Asturiana de los Valles, Avilena, Charolais, Danish Red, Piemontese and Pirenaica), being higher than 0.05 in Holstein and Casina breeds only (Table 4). This suggests an involvement of the gene in some important roles which prevents its polymorphism.

Table 4. Allelic frequencies in the 16 European cattle breeds.

SNP	*TLR2_591*		*TLR2_738*		*TLR2_767*		*TLR4_254*		*TLR4_1678*		*TLR4_2043*		*TLR6_855*		*TLR6_2315*	
ALLELE	**A**	**G**	**A**	**G**	**A**	**G**	**A**	**G**	**C**	**T**	**C**	**T**	**A**	**G**	**C**	**T**
A. Angus	0.011	0.989	0.012	0.988	0.693	0.307	0.837	0.163	0.683	0.317	0.989	0.011	0.581	0.419	0.349	0.651
A. Valles	0.086	0.914	0.016	0.984	0.548	0.452	0.581	0.419	0.635	0.365	0.898	0.102	0.373	0.627	0.234	0.766
Avilena	0.250	0.750	0.031	0.969	0.317	0.683	0.712	0.288	0.533	0.467	0.913	0.087	0.437	0.563	0.353	0.647
Casina	0.327	0.673	0.152	0.848	0.432	0.568	0.538	0.462	0.649	0.351	0.848	0.152	0.302	0.698	0.196	0.804
Charolais	0.114	0.886	0.044	0.956	0.487	0.513	0.545	0.455	0.671	0.329	0.840	0.160	0.353	0.647	0.336	0.664
Danish Red	0.035	0.965	0.017	0.983	0.595	0.405	0.632	0.368	0.582	0.418	0.796	0.204	0.500	0.500	0.324	0.676
Highlands	0.011	0.989	**0**	**1**	0.932	0.068	0.415	0.585	0.564	0.436	**1**	**0**	0.260	0.740	0.239	0.761
Holstein	0.147	0.853	0.154	0.846	0.759	0.241	0.417	0.583	0.740	0.260	0.855	0.145	0.366	0.634	0.364	0.636
Jersey	0.170	0.830	**0**	**1**	0.372	0.628	0.707	0.293	0.888	0.112	0.413	0.587	0.776	0.224	0.582	0.418
Limousine	0.271	0.729	**0**	**1**	0.479	0.521	0.583	0.417	0.616	0.384	0.819	0.181	0.340	0.660	0.326	0.674
Marchigiana	0.030	0.970	**0**	**1**	0.529	0.471	0.894	0.106	0.338	0.662	0.833	0.167	0.557	0.443	0.559	0.441
Maremmana	0.093	0.907	**0**	**1**	0.517	0.483	0.572	0.428	0.742	0.258	0.884	0.116	0.330	0.670	0.331	0.669
Piemontese	0.215	0.785	0.038	0.962	0.405	0.595	0.675	0.325	0.538	0.462	0.817	0.183	0.500	0.500	0.281	0.719
Pirenaica	0.132	0.868	0.014	0.986	0.650	0.350	0.418	0.582	0.690	0.310	0.963	0.037	0.215	0.785	0.174	0.826
Simmenthal	0.096	0.904	**0**	**1**	0.717	0.283	0.522	0.478	0.591	0.409	0.875	0.125	0.519	0.481	0.395	0.605
S. Devon	**0**	**1**	**0**	**1**	0.838	0.163	0.638	0.363	0.907	0.093	0.643	0.357	0.116	0.884	0.085	0.915

Selection can leave, in the genes under its influence, a set of signatures that can be analyzed to identify genes or chromosomal regions which are likely targets of positive selection. We used *FST* statistic to assess if the variation of SNP allele frequencies among populations leads to signatures of selection. For each *locus*, the allele frequencies are used to compute *FST* values conditional on heterozygosity and to calculate P-values for each *locus*. This method provides evidence for divergent selection by looking for outliers with *FST* values higher than expected, controlling for heterozygosity. The analysis performed using *FDist2* software to identify outlier loci revealed that none of the SNPs lied outside the 95% confidence limits assumed for conditional joint distribution of *FST* vs. mean heterozygosity. Analysis was performed by bootstrapping 200,000 replications on real data using a coalescent model (Figure 1).

Figure 1. Upper (▲) and lower (●) confidence limits of 95% quantiles; median (-) of 200,000 replications of expected Fst and heterozygosity using the coalescent model.



None of the identified SNPs is located in loci under selection according to the model of Beaumont and Nichols [30]. Anyway, significant deviations from Hardy-Weinberg equilibrium over all populations (p-value < 0.01) were observed in three SNPs at two loci: *TLR2_591*, *TLR2_738* and *TLR4_2043* (Table 5). We hypothesize that polymorphisms are fixed in the analysed breeds since many generations, and that the coalescent model employed is not powerful enough to identify selection events happened too far in the past.

Distance based phylogenetic analysis was used to describe the relationships between breeds regarding the investigated TLRs. Table 6 presents the Nei genetic distances relating the 16 breeds studied. The lowest distance values are observed between Charolais and Asturiana de los Valles (0.002), while the highest distance is observed between Highlands and Jersey (0.117). Furthermore, the Jersey breed results very distant from all the other breeds of Great Britain, confirming the results obtained by AFLP and suggesting isolation within the Jersey island as the major cause of distinctiveness [34]. Indeed, Nei distances show that the highest genetic diversity is

observed in the geographically isolated breeds: it is suggestive (Figure 2) that the breeds of Great Britain (Aberdeen Angus, Highlands, South Devon and Jersey), using the analysed polymorphisms, are distributed accordingly to their geographic provenience.

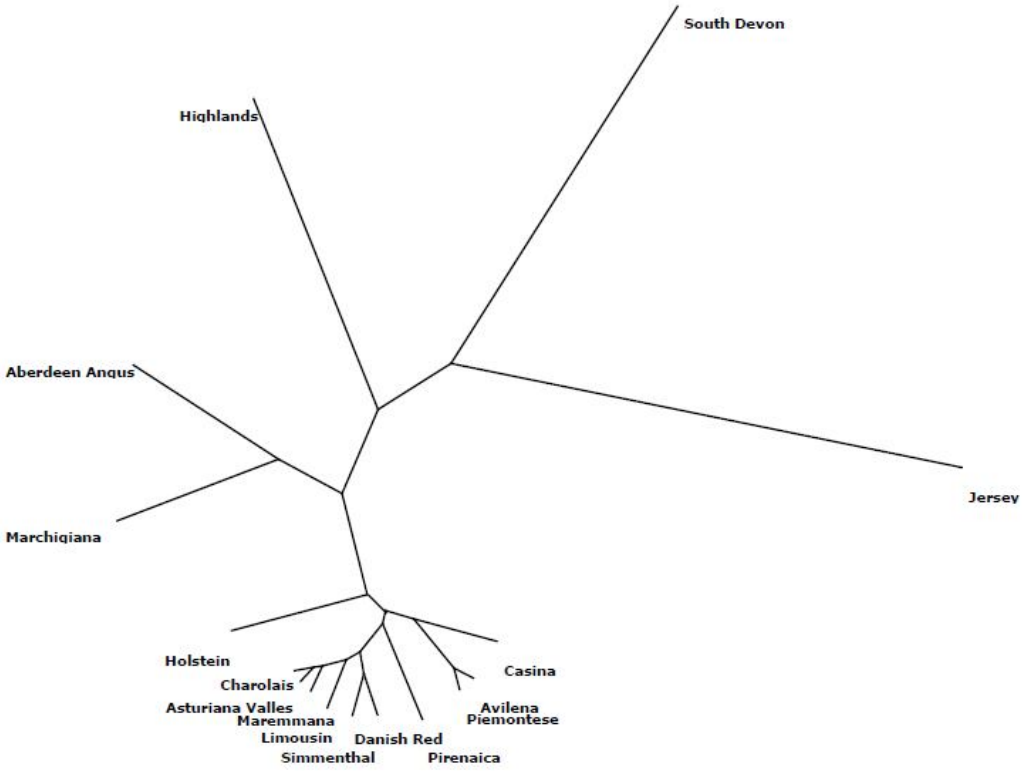
Table 5. Hardy-Weinberg equilibrium Test.

Locus	*Hw Test*
TLR2_591	0.0007
TLR2_738	0.0011
TLR2_767	0.0542
TLR4_254	0.2693
TLR4_1678	0.0345
TLR4_2043	0.0012
TLR6_855	0.0303
TLR6_2315	0.7460

Table 6. Nei genetic distances. GB: Great Britain; ES: Spain; FR: France; DK: Denmark; IT: Italy.

	Geographic location of the breed	Aberdeen Angus	Asturiana de los Valles	Avilena	Casina	Charolais	Danish Red	Highlands	Holstein	Jersey	Limousin	Marchigiana	Maremmana	Piemontese	Pirenaica	Simmenthal	South Devon
Aberdeen Angus	GB	0															
Asturiana de los Valles	ES	0.016	0														
Avilena	ES	0.027	0.010	0													
Casina	ES	0.044	0.012	0.011	0												
Charolais	FR	0.023	0.002	0.008	0.009	0											
Danish Red	DK	0.014	0.004	0.015	0.022	0.005	0										
Highlands	GB	0.029	0.027	0.057	0.058	0.037	0.034	0									
Holstein	DK	0.032	0.013	0.027	0.014	0.009	0.015	0.033	0								
Jersey	GB	0.061	0.048	0.044	0.058	0.039	0.036	0.117	0.056	0							
Limousin	FR	0.032	0.007	0.007	0.012	0.006	0.012	0.045	0.020	0.038	0						
Marchigiana	IT	0.020	0.026	0.023	0.052	0.028	0.015	0.060	0.050	0.048	0.029	0					
Maremmana	IT	0.019	0.003	0.013	0.018	0.004	0.007	0.030	0.016	0.040	0.005	0.028	0				
Piemontese	IT	0.026	0.007	0.003	0.009	0.006	0.008	0.055	0.020	0.035	0.006	0.022	0.011	0			
Pirenaica	ES	0.029	0.006	0.023	0.016	0.010	0.018	0.018	0.013	0.076	0.014	0.052	0.009	0.021	0		
Simmenthal	DK	0.016	0.007	0.019	0.027	0.009	0.005	0.023	0.014	0.041	0.010	0.021	0.007	0.014	0.015	0	
South Devon	GB	0.053	0.035	0.073	0.060	0.040	0.037	0.046	0.044	0.082	0.049	0.076	0.033	0.060	0.035	0.044	0

Figure 2. Phylogenetic relationship among the 16 breeds studied. The genetic distances were calculated from allelic frequencies by using Nei distances. The reconstruction was done with UPGMA (Sneath and Sokal, 1973).



0.01 (meaning 0.01 nucleotide substitutions per site)

3. Conclusions

We could identify eight SNPs in genes of great interest in cattle management by screening the nucleotide sequences of bovine TLR2, TLR4, and TLR6 genes. These variations in immune function related genes will contribute to research on disease response in cattle. In fact, the newly identified SNPs can be used in association studies between polymorphisms and cattle resistance to diseases.

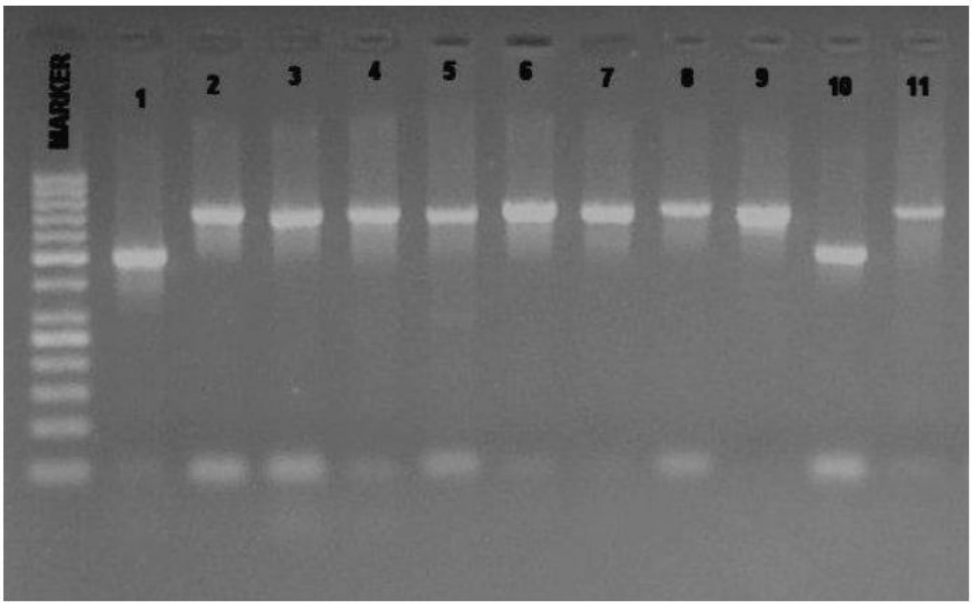
The SNPs characterization was performed by analysing a conspicuous number of individuals from 16 European breeds, and the main statistics were calculated. Even if from our analysis the SNPs do not appear located in loci under selection, a deviation of three SNPs from Hardy Weinberg equilibrium was observed. We hypothesize that some of the polymorphisms were fixated many generations ago within breed and the coalescent model could not be powerful enough to reveal selection events so far in the past. It would be interesting to apply a more powerful model to confirm the absence of selection in the SNPs and their suitableness as neutral markers.

Acknowledgements

The authors wish to thank Paolo Ciorba and Gabriella Porcai for technical assistance, the Eu-GeMQual project QLRT-1999-30147 for providing some of the data and Flora Jeane Dause for language revision. The research was partially carried out by funds of Regione Lazio PRAL 2003, "Valorizzazione e sviluppo dell'allevamento bovino di razze locali nel Lazio attraverso la caratterizzazione genetica per la resistenza alle malattie, per la longevità e per la qualità della carne". JLW thanks the Cariplo Foundation for financial support.

Electronic Supplementary Information

Figure S1. PCR amplified fragments: 1. TLR6_1; 2. TLR6_2; 3. TLR6_3; 4. TLR6_4; 5. TLR4_1; 6. TLR4_2; 7. TLR4_3; 8. TLR2_5; 9. TLR2_4; 10. TLR2_3; 11. TLR2_2.



References

1. Takeda, K.; Akira, S. Roles of Toll-like receptors in innate immune responses. *Genes Cells* **2001**, *6*, 733-742.
2. Medzhitov, R.; Janeway, C.A., Jr. Innate immunity. *N. Engl. J. Med.* **2000**, *343*, 338-344.
3. Aderem, A.; Ulevitch, R.J. Toll-like receptors in the induction of the innate immune responses. *Nature* **2000**, *406*, 782-787.
4. Takeda, K.; Kaisho, T.; Akira, S. Toll-like receptors. *Annu. Rev. Immunol.* **2003**, *21*, 335-376.
5. Vasselon, T.; Detmers, P.A. Toll receptors: a central element in innate immune responses. *Infec. Immunity* **2002**, *70*, 1033-1041.
6. Kaisho, T.; Akira, S. Toll-like receptor function and signaling. *J. Allergy Clin. Immunol.* **2006**, *117*, 979-987.

7. Stein, D.; Roth, S.; Vogelsang, E.; Nusslein-Volhard, C. The polarity of the dorsoventral axis in the Drosophila embryo is defined by an extracellular signal. *Cell* **1991**, *65*, 725-735.
8. Lemaitre, B.; Nicolas, E.; Michaut, L.; Reichhart, J.M.; Hoffman, J.A. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell* **1996**, *86*, 973-983.
9. Hashimoto, C.; Hudson, K.I.; Anderson, K.V. The toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a trans-membrane protein. *Cell* **1998**, *52*, 269-279.
10. Muzio, M.; Polentarutti, N.; Bosisio, D.; Prahladan, M.K.; Mantovani, A. Toll-like receptors: a growing family of immune receptors that are differentially expressed and regulated by different leukocytes. *J. Leukocyte Biol.* **2000**, *67*, 450-456.
11. Jin, M.S.; Kim, S.E.; Heo, J.Y.; Lee, M.E.; Kim, H.M.; Paik, S.G.; Lee, H.; Lee, J.O. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **2007**, *130*, 1071-1082.
12. Menzies, M.; Ingham, A. Identification and expression of Toll-like receptors 1-10 in selected bovine and ovine tissues. *Vet. Immunol. Immunopathol.* **2006**, *109*, 23-30.
13. De Schepper, S.; De Ketelaere, A.; Bannerman, D.D.; Paape, M.J.; Peelman, L.; Burvenich, C. The toll-like receptor-4 (TLR-4) pathway and its possible role in the pathogenesis of Escherichia coli mastitis in dairy cattle. *Vet. Res.* **2008**, *39*, 1-5.
14. Werling, D.; Jann, O.C.; Offord, V.; Glass, E.J.; Coffey, T.J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends Immunol.* **2009**, *30*, 124-130.
15. Underhill, D.M.; Ozinsky, A.; Smith, K.D.; Aderem, A. Toll-like receptor-2 mediates mycobacteria-induced proinflammatory signaling in macrophages. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14459-14463.
16. Buwitt-Beckmann, U.; Heine, H.; Wiesmuller, K.H.; Jung, G.; Brock, R.; Akira, S.; Ulmer, A.J. TLR1- and TLR6-independent recognition of bacterial lipopeptides. *J. Biol. Chem.* **2006**, *281*, 9049-9057.
17. Kirschning, C.J.; Schumann, R.R. TLR2: cellular sensor for microbial and endogenous molecular patterns. *Curr. Top. Microbiol. Immunol.* **2002**, *270*, 121-144.
18. Akira, S.; Uematsu, S.; Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **2006**, *124*, 783-801.
19. White, A.N.; Kata, S.R.; Womack, J.E. Comparative fine maps of bovine toll-like receptor 4 and toll-like receptor 2 regions. *Mamm. Genome* **2003**, *14*, 149-155.
20. McGuire, K.; Jones, M.; Werling, D.; Williams, J.L.; Glass, E.J.; Jann, O. Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors. *Anim. Genet.* **2005**, *37*, 47-50.
21. Beutler, B.; Rehli, M. Evolution of the TIR, tolls and TLRs: functional inferences from computational biology. *Curr. Top. Microbiol. Immunol.* **2002**, *270*, 1-21.
22. Pandey, S.; Agrawal, D.K. Immunobiology of Toll-like receptors: Emerging trends. *Immun. Cell Biol.* **2006**, *84*, 333-341.
23. Seabury, C.M.; Cargill, E.J.; Womack, J.E. Sequence variability and protein domain architectures for bovine Toll-like receptors 1, 5 and 10. *Genomics* **2007**, *90*, 502-515.

24. Seabury, C.M.; Womack, J.E. Analysis of sequence variability and protein domain architectures for bovine peptidoglycan recognition protein 1 and Toll-like receptors 2 and 6. *Genomics* **2008**, *92*, 235-245.
25. Valarcher, J.F.; Taylor, G. Bovine respiratory syncytial virus infection. *Vet. Res.* **2007**, *38*, 153-180.
26. Mucha, R.; Bhide, M.R.; Chakurkar, E.B.; Novak, M.; Mikula, I., Sr. Toll-like receptors TLR1, TLR2 and TLR4 gene mutations and natural resistance to Mycobacterium avium subsp. paratuberculosis infection in cattle. *Vet. Immunol. Immunopathol.* **2009**, *128*, 381-388.
27. Valentini, A. *PolyPrimers, version 1.00*; 2006. Available online at: http://www.unitus.it/SAG/primers.zip (accessed on 23 July, 2009).
28. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95-98.
29. Liu, K.; Muse, S.V. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **2005**, *21*, 2128-2129.
30. Beaumont, M.A.; Nichols, R.A. Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B.* **1996**, *263*, 1619-1626.
31. Nei, M.; Tajima, F. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics* **1983**, *105*, 207-217.
32. Page, R.D. TreeView: an application to display phylogenetic trees on personal computer. *Comput. Appl. Biosci.* **1996**, *12*, 357-358.
33. Mariotti, M. Polimorfismi nei geni Toll-like receptor 2, 4 e 6 in Bos Taurus. Università degli studi della Tuscia: Viterbo, Italy, 2008. Available online: http://dspace.unitus.it/handle/2067/679 (accessed on 23 July, 2009).
34. Negrini, R.; Nijman, I.J.; Milanesi, E.; Moazami-Goudarzi, K.; Williams, J.L.; Erhardt, G.; Dunner, S.; Rodellar, C.; Valentini, A.; Bradley, D.G.; Olsaker, I.; Kantanen, J.; Ajmone-Marsan, P.; Lenstra J.A. European Cattle Genetic Diversity Consortium. Differentiation of European cattle by AFLP fingerprinting. *Anim. Genet.* **2007**, *38*, 60-66.

2.2. Relationships between Podolic cattle breeds assessed by single nucleotide polymorphisms (SNPs) genotyping

Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

Relationships between Podolic cattle breeds assessed by single nucleotide polymorphisms (SNPs) genotyping

L. Pariset[1], M. Mariotti[1], A. Nardone[1], M.I. Soysal[2], E.Ozkan[2], J.L. Williams[3], S. Dunner[4], H. Leveziel[5], A. Maróti-Agóts[6], I. Bodò[7] & A. Valentini[1]

1 Dipartimento di Produzioni Animali, Università degli Studi della Tuscia, Viterbo, Italy
2 Namık Kemal University, Agriculture Faculty of Tekirdag, Department of Animal Science, Tekirdag, Turkiye
3 Parco Tecnologico Padano, Via Einstein, Polo Universitario, Lodi, Italy
4 Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense, Madrid, Spain
5 Unité de Génétique Moléculaire Animale, INRA/Université de Limoges, Facultédes Sciences et Techniques, Limoges Cedex, France
6 Szent István University, Department of Animal Breeding, Nutrition and Laboratory Animal Science, Faculty of Veterinary Science, Budapest, Hungary
7 University of Debrecen, Faculty of Agriculture, Department of Animal Breeding and Nutrition, Debrecen, Hungary

Keywords
Bayesian inference; diversity; migration; Podolic cattle; SNPs.

Correspondence
Lorraine Pariset, Dipartimento di Produzioni Animali, Università degli Studi della Tuscia, Via S. Camillo de Lellis, 01100 Viterbo, Italy.
Tel: +39 0761 357447; Fax: +39 0761 357434;
E-mail: pariset@unitus.it

Received: 12 May 2009;
accepted: 25 March 2010

Summary

Italian Maremmana, Turkish Grey and Hungarian Grey breeds belong to the same Podolic group of cattle, have a similar conformation and recently experienced a similar demographic reduction. The aim of this study was to assess the relationship among the analysed Podolic breeds and to verify whether their genetic state reflects their history. To do so, approximately 100 single nucleotide polymorphisms (SNPs) were genotyped on individuals belonging to these breeds and compared to genotypes of individuals of two Italian beef breeds, Marchigiana and Piemontese, which underwent different selection and migration histories. Population genetic parameters such as allelic frequencies and heterozygosity values were assessed, genetic distances calculated and assignment test performed to evaluate the possibility of recent admixture between the populations. The data show that the physical similarity among the Podolic breeds examined, and particularly between Hungarian Grey and Maremmana cattle that experienced admixture in the recent past, is mainly morphological. The assignment of individuals from genotype data was achieved using Bayesian inference, confirming that the set of chosen SNPs is able to distinguish among the breeds and that the breeds are genetically distinct. Individuals of Turkish Grey breed were clearly assigned to their breed of origin for all clustering alternatives, showing that this breed can be differentiated from the others on the basis of the allelic frequencies. Remarkably, in the Turkish Grey there were differences observed between the population of Enez district, where *in situ* conservation studies are practised, and that of Bandirma district of Balikesir, where *ex situ* conservation studies are practised out of the original raising area. In conclusion, this study demonstrates that molecular data could be used to reveal an unbiased view of past events and provide the basis for a rational exploitation of livestock, suggesting appropriate cross-breeding plans based on genetic distance or breeding strategies that include the population structure.

 • J. Anim. Breed. Genet. **127** (2010) 481–488 doi:10.1111/j.1439-0388.2010.00868.x

Introduction

Over the past 15 years, 300 out of 6000 breeds of all farm animal species identified by FAO have gone extinct (Scherf 2000). It is argued in the World Watch List for Domestic Animal Diversity (DAD) that 1350 farm animal breeds currently face extinction. The danger of extinction of farm animal breeds is mainly attributable to three factors: the first is genetic erosion because of artificial insemination. The second factor is strong economic pressure on the farmer to focus on single traits, such as milk production. The third factor is unrestricted and indiscriminate cross-breeding, especially in developing countries (Soysal *et al.* 2004).

Podolic cattle include a very ancient group of breeds, considered to be straight descendant from the Auroch (*Bos primigenius*). Podolic breeds are present in various European areas, and many of them are seriously endangered of extinction.

Hungarian Grey, Maremmana and Turkish Grey belong to the same Podolic group of cattle and show similar external conformation. These breeds recently underwent a similar demographic reduction. All three of these breeds face risks for their future survival because of inbreeding, indiscriminate cross-breeding and substitution with cosmopolitan more productive breeds.

The breeds

Maremmana

The breed can be traced back to Grey Steppe cattle which entered Italy in large numbers during the 14th to 18th centuries. Herdbook registration started in 1935 (total breed population 274 000 head). Since 1945, head numbers have declined dramatically because of the changes in land use and mechanization, and by the mid-1960s extinction was predicted. The breed recovered between 1965 and 1975 because of its ability to adapt to the environmental constraints of the hilly areas of the Maremma, reaching 60 000 head in 1975. In 1992, 10 000 head, of which 4000 females and 120 sires, were registered in the herdbook. Since then the number remained constant.

Maróti-Agóts *et al.* (2005) report that the Maremmana cows are significantly bigger in every body measurement than Hungarian Grey cattle, and in particular there are significant differences of rump length/body length index.

Hungarian Grey

The breed was imported by the Hungarian conquerors who came from the Podolic area in the south, in the 9th century. There is evidence of the presence of similar cattle in Egypt and Italy. Explicit reference to long-horned, 'magnus cornutes boves Hungaricos', first appears in a 16th century document. The breed became common by the 17–18th century.

A radical decline followed World War II and the breed nearly became extinct in 1947–1967. During the late 1950s, 1800 of the 2000–3000 cows were mated with sires of the Kostroma dairy breed. By 1962 only 200 purebred Hungarian Grey cows and six bulls were saved (Bodó *et al.* 1996; Bartosiewicz 1997). Inbreeding was avoided by using a rotational mating scheme based on six local Hugarian Grey sires, two imported sires of the same breed and three Maremmana sires introduced during the early 1970s. After 7–9 generations, the initial lines became completely randomized.

Turkish Grey

Found in north-western Turkey, this breed is a tri-purpose breed: kept for milk and meat as well as being used as a work animal. The breed came from the Grey Steppe type and originated from the Iskar breed of Bulgaria (Mason 1996). It is believed that Turkish Grey cattle are the ancestor or relatives of European Grey cattle found in Italy, Bulgaria and Hungary. The breed has a 'postendangered' status; however, efforts are still required to maintain the grey cattle population. Native grey cattle represent the most interesting cattle breed in Turkey, especially because of its resistance to parasites and its ability of using wetland pasture, and as a consequence its sustainable use of marginal lands (Soysal & Kok 2006).

The aim of this study was to assess the relationship between Hungarian Grey, Maremmana and Turkish Grey. To do so, single nucleotide polymorphisms (SNPs) were genotyped on individuals belonging to the three Podolic breeds as well as on individuals of two Italian beef breeds, Marchigiana and Piemontese, which have different selection and migration histories, and were chosen as example of unrelated breeds.

SNPs are abundant in the genome; genotyping results are easy to reproduce in different laboratories and are simple to score. The usefulness of SNPs in population genetics has been demonstrated in several studies on the last decades, and they have been recently used to discovering signatures of selection (Akey *et al.* 2002; Kelley *et al.* 2006; Luikart *et al.* 2003; Pariset *et al.* 2006a; Pariset *et al.* 2009) and in evaluating population structure (Pariset *et al.* 2006b; Negrini *et al.* 2008).

Materials and methods

Samples

Animals were sampled from their native regions: the Hungarian Grey (HU) population in Hortobagy (Hungary), the Maremmana (MM) from the population of Castelporziano (Rome), and the Turkish Grey (GS) in the Enez district of the Edirne Province of Trakya and in the Bandirma Province of Balikesir (Turkey). Sixty-three (HU), 93 (MM) and 93 (GS) individuals were sampled in each population; DNA from blood samples was extracted using Wizard Genomic DNA Purification kit (Promega, Madison, WI, USA) following the manufacturer instructions. Samples and data of Piemontese (PD) and Marchigiana (MR) cattle were obtained from a previous project (EU-GemQual QLRT CT2000-0147).

SNPs genotyping

Ninety-nine SNPs, selected from a panel of 701 SNPs in candidate genes for meat quality previously characterized (Williams *et al.* 2009), were genotyped on individuals belonging to the five breeds. SNPs were discovered by sequencing a panel of eight individuals each from a different breed.

Genotyping was performed by outsourcing to Kbiosciences (http://www.Kbioscience.co.uk), using quality control criteria as negative controls, interplate and intraplate duplicate testing of a known DNA. Generally, a genotyping repeatability >99% was achieved.

Statistical analysis

Allele frequencies were calculated using *Fstat* 2.93 (Goudet 1995, 2001). *Fis*, Ho and He were estimated for each locus using *Powermarker* (Liu & Muse 2001). The same software was used to test for deviation from Hardy–Weinberg equilibrium (HWE) for each locus and population and for loci over all populations using a Markov chain of 100 000 steps and 1000 dememorization steps. GENEPOP version 4.0 (Raymond & Rousset 1995) was used to calculate *F* statistics per breed, following Weir & Cockerham (1984) and to characterize genetic differentiation among breeds by estimating overall and pairwise *Fst* values. Nei (1972) genetic distances between populations were calculated using *Powermarker*. The same software was used to infer haplotypes.

To evaluate whether some of the loci supposed neutral in this study could be identified as outliers, and therefore genotyped SNPs were unsuitable to be used as markers, the approach proposed by Beaumont & Nichols (1996), further developed by Beaumont & Balding (2004), and implemented in the *FDIST2* software (http://www.rubic.rdg.ac.uk/~mab/software.html) was used. For each locus, the allele frequencies were used to compute *Fst* values conditional on heterozygosity and to calculate P-values for each locus. Each simulation included five populations, 91 loci and an expected *Fst* of 0.104. Population data sets were built using 50 000 coalescent simulations on real data using the infinite alleles model.

Individuals were clustered by applying a parametric genetic admixture analysis implemented in the *Structure* 2.0 software (Pritchard *et al.* 2000). This software uses a model-based clustering method that employs a Markov Chain to estimate the posterior distribution (q) of the admixture coefficient of each individual, to characterize parental populations, to assign individuals to these populations, to detect admixed individuals and to estimate individual admixture starting from allele frequency. Results were obtained using a burn-in period of 100 000 followed by 200 000 Markov chain Monte Carlo (MCMC) repeats and considering SNPs frequencies independent among populations. A number of genetic clusters (K) ranging from 2 to 6 was tested using the admixture model; four runs for each K were performed. Graphical reconstruction of *Structure* results were produced by using *Distruct* 1.1 (Rosenberg 2004).

Results

SNPs genotyping

Ninety-nine polymorphic SNPs were used to genotype 311 individuals belonging to the five breeds. Of these, 91 SNPs resulted polymorphic and suitable for the analyses. A total of 26 006 genotypes were produced, and the frequencies of the major alleles are reported in Table S1. The SNP data were analysed to assess their neutrality using *Fdist* (Beaumont & Nichols 1996), and all SNPs resulted within the 99% upper and lower limits of distribution. When more than one SNP within the same gene was genotyped, haplotypes have been inferred and used for the subsequent analyses (Table S2), for a final data set of 67 markers. Major allele frequency, Expected and Observed heterozygosity, Polimorfic information content (PIC), *Fis*, p-value relative to HW test and chromosomal location relative to Btau 4.0 for each polymorphic marker are reported in Tables S1 and S2, respectively.

Genetic diversity and differentiation of cattle breeds

Observed heterozygosity of the 67 markers determined from SNP frequencies ranged from 0.557 to 0.017, with a mean of 0.346. Expected heterozygosity ranged from 0.649 to 0.024, with a mean of 0.386. The frequencies of the major alleles ranged from 0.988 to 0.428. Frequency of the minor allele was >5% in all but three SNPs (POMC_b1_63T, VCL_a1_160T, PRKAA2_a1_88C). *Fis* value of the markers ranged from 0.483 to −0.138, with a mean of 0.138. Significant deviations from HWE over all populations (p-value <0.05) were observed in 38 markers.

Within population variance estimate (*Fis*) per population ranged from −0.020 (PD) to 0.186 (GS) (Table 1). Positive *Fis* was observed in Turkish Grey breed. A moderately positive *Fis* value was observed also in Marchigiana breed (Table 1).

The pairwise *Fst* between populations showed a maximum (0.124) between Hungarian Grey and Maremmana, and a minimum (0.081) between Marchigiana and Piemontese (Table 2).

The Nei standard genetic distance (Nei 1972) indicates a maximum distance between Marchigiana and Turkish Grey (0.070) and a minimum distance between Marchigiana and Maremmana (0.051). Considering only the three Podolic breeds, the greater distance is observed between Hungarian Grey and Maremmana (0.066) (Table 3).

Individual assignment

To estimate the number of genetic clusters among the 311 individuals, a parametric genetic mixture analysis implemented in the *Structure* 2.0 software (Pritchard *et al.* 2000) was performed. Between two and six clusters (K values) were tested using the admixture model, assuming that each individual does not necessarily have a genetic background originating from one of the K populations. Consistent results across runs were obtained.

To identify the optimal K value and hence identify the most reliable result we applied the methodology described in Evanno *et al.* (2005), concluding that 5 was the optimal K.

A graphic representation of the estimated membership coefficients to the clusters for each individual, obtained running structure setting K from 2 to 5, is shown in Figure 1. Each individual is represented by a single vertical line broken into K coloured segments, whose lengths are proportional to each of the K inferred clusters.

For K = 5 most of the individuals could be unambiguously assigned to the five breed clusters. Table 4 demonstrates the proportion of membership of the five breeds in each of the five clusters. Individuals of Piemontese and Marchigiana show the highest level of genetic admixture and could not be differentiated for K values between 2 to 5. Differentiation within the Turkish Grey breed is first observed for K = 2, with 0.23% of the individuals assigned to a separate cluster before differences at the breed level are detected. Raising the K value to 5, the Turkish Grey individuals remain grouped in two distinct populations (Figure 1).

Table 1 *F* Statistics per breed over all loci following Weir & Cockerham (1996)

Breed	*Fst*	*Fit*	*Fis*
Turkish Grey	0.3235	0.3975	0.1861
Hungarian Grey	0.3596	0.3431	−0.0480
Maremmana	0.3795	0.3541	−0.0718
Marchigiana	0.3231	0.3617	0.1066
Piemontese	0.3763	0.3688	−0.0202

Table 2 Pairwise *Fst* between breeds estimated as in Weir & Cockerham (1996)

	Turkish Grey	Hungarian Grey	Maremmana	Marchigiana
Hungarian Grey	0.118			
Maremmana	0.107	0.124		
Marchigiana	0.100	0.107	0.081	
Piemontese	0.109	0.119	0.094	0.081

Table 3 Nei (1972) standard diversity index

OTU	Turkish Grey	Hungarian Grey	Maremmana	Marchigiana
Hungarian Grey	0.064			
Maremmana	0.063	0.066		
Marchigiana	0.070	0.059	0.051	
Piemontese	0.062	0.069	0.056	0.057

Discussion

The SNPs used in this study are not a random sample. However, they do not show to be under selection and are fairly scattered among all chromosomes (with exception of 23, 24 and 27; Table S1). A selection bias may be present because of the small size of the animal panel used for SNPs discovery. However, this is a limit common to many studies, even using larger SNP panels (The Bovine HapMap Consortium 2009). Our panel was composed by eight individuals belonging to different breeds which are deemed to

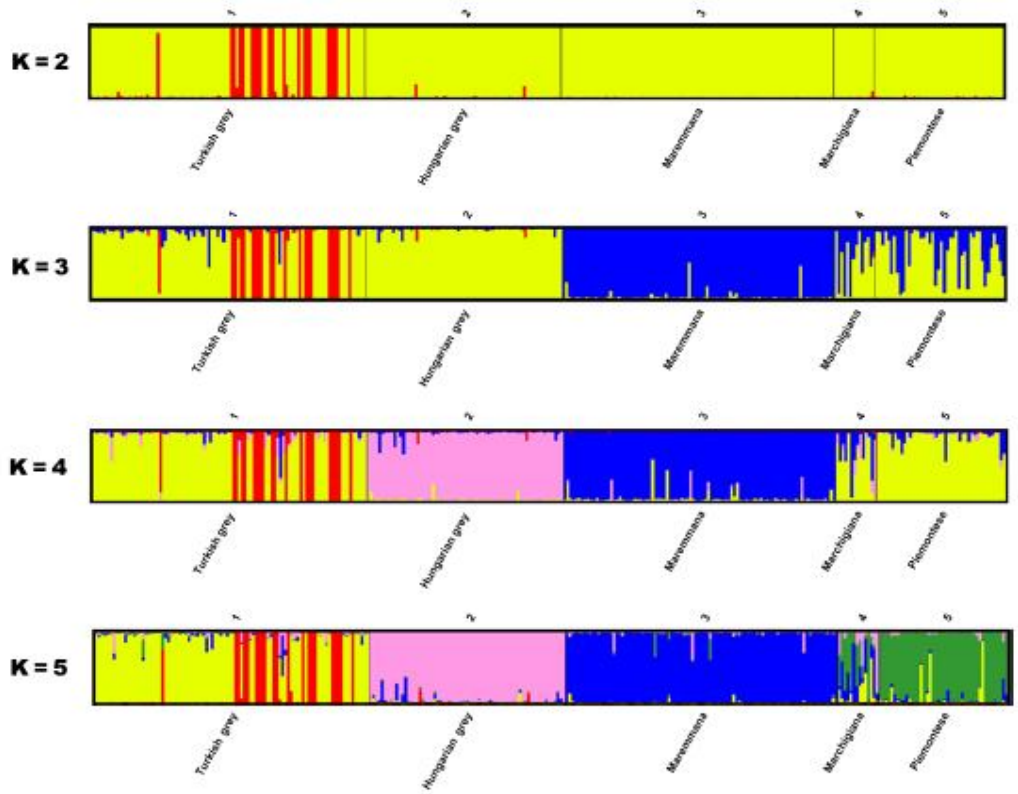


Figure 1 Estimated membership coefficients of each individual to the model-based clusters obtained running structure setting K from 2 to 5. Each individual is represented by a single vertical line broken into K coloured segments, whose lengths are proportional to each of the K inferred clusters. The predefined groups (breeds) are separated by black lines.

Table 4 Proportion of assignment of individuals of the five breeds in each of the five inferred clusters obtained from STRUCTURE analysis

Predefined populations	Inferred clusters				
	1	2	3	4	5
Turkish Grey	0.229	0.706	0.027	0.026	0.012
Hungarian Grey	0.008	0.013	0.947	0.025	0.006
Maremmana	0.003	0.012	0.021	0.949	0.015
Marchigiana	0.008	0.149	0.168	0.283	0.392
Piemontese	0.004	0.074	0.017	0.031	0.873

diverge since many generations (Williams *et al.* 2009) and therefore should include a substantial amount of species variation. Within population variance estimate (*Fis*) was positive in Turkish Grey breed, indicating a significant level of inbreeding. A value suggesting moderate inbreeding is also observed in Marchigiana. The other three breeds showed values compatible with random mating. Therefore we can conclude that mating strategies used in Maremmana and Hungarian Grey during the recovery were appropriate, while a different mating strategy could be suggested at least in one of the Turkish Grey populations.

The estimate of Nei's genetic distances show that Maremmana is more distant from Hungarian Grey (0.66) than from Turkish Grey (0.63), despite the recent admixture with the former breed. Maremmana is closer to Marchigiana (0.51) than to either Hungarian Grey or Turkish Grey. Maximum genetic distances were observed between Turkish Grey and Marchigiana (0.70) and Hungarian Grey and Piemontese (0.62). Data are supported also by pairwise *Fst*, showing that Maremmana is closer to Turkish

Grey (0.107) than to Hungarian Grey (0.124). The high *Fst* values observed in Hungarian Grey (Table 1) may reflect human selection, a long-time isolation of the breed, or a limited number of founders. The last two hypotheses seem more compatible with the history of the breed.

From the *Structure* analysis, already with a K value of 2 about 1/4 of the Turkish Grey individuals were assigned to a separate group, remaining apart from the other breeds. This may be attributable to the high level of inbreeding in the population analysed (*Fis* 0.19). It must be noticed that Turkish Grey cattle samples were collected from two distinct populations: the first raised by breeders of Enez district, the second belonging to the Bandirma district of Balikesir. Raising the K value Maremmana (K = 3), then Hungarian Grey (K = 4) and finally Marchigiana and Piemontese together (K = 5) are assigned to separate groups. It can be hypothesized that the different conservation strategies used in Enez district and in Bandirma district have very different effects on the genetic asset of the populations. Individuals of Piemontese and Marchigiana show the highest level of genetic admixture, revealing that the genotype distributions of these two breeds are more similar than those of the other breeds examined. This is not surprising, in that the two breeds, beside having a different history, underwent specific selection for beef production.

From this analysis, despite their similar morphology, Hungarian Grey and Maremmana are clearly identified as genetically distinct breeds. This could be attributed to either a different origin of the breeds or a consequence of the recent history, that led to the selection and probably fixation of genes. The two breeds were found differentiated on the basis of their allelic frequencies in a previous study (Valentini *et al.* 2006), and this confirms also the morphological differences reported by Maróti-Agóts *et al.* (2005).

As for the Turkish Grey, we observed interesting differences between the population raised by breeders of Enez district, where *in situ* conservation studies are practised, and belonging to the Bandirma district of Balikesir, where *ex situ* conservation studies are practised 400 km far away from the original grey cattle raising area, in the Agriculture Research farm of Ministry of Agriculture. Turkey is very close to cattle domestication centre (Edwards *et al.* 2007); therefore, a higher differentiation is expected because the time for drift and distance to domestication bottleneck and this could explain why one of the two population results genetically differentiated for any value of K tested. Besides being raised far from the original raising area of Grey cattle, the *ex-situ* herds of Bandirma are subject to legal and financial limitation, i.e. farmers are not able to buy new members, and this leads to increment of inbreeding and to genetic drift. On the contrary, private farmers, raising the *in situ* herds for conservation purposes, avoid inbreeding because of more flexibility in management decisions. Moreover, the Enez herds were established 5 years after those of Bandirma, and this could represent another source of difference.

In conclusion, this study demonstrates that morphology and anecdotic accounts might be deceiving in describing a population, as it was in the case of Maremmana and Hungarian Grey. Molecular data are very suggestive, and the observed genetic differentiation of the breeds may deserve more investigation. Further analysis could help in tracing an unbiased picture of past events and provide the basis for a rational exploitation of livestock, suggesting appropriate cross-breeding plans based on genetic distance, or breeding strategies that include the population structure.

Acknowledgements

The authors thank Guido Pezzali for its collaboration in Castelporziano sampling, Paolo Ciorba for technical assistance, Flora Jane Dause for help with the language and two anonymous reviewers for helpful comments. Marchigiana and Piemontese data were obtained under EU-GemQual QLRT CT2000-0147 project.

References

Akey J.M., Zhang G., Zhang K., Jin L., Shriver M.D. (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Res.*, **12**, 1805–1814.

Bartosiewicz L. (1997) The Hungarian Grey cattle: a traditional European breed. *Information*, **21**, 49–60.

Beaumont M.A., Balding D.J. (2004) Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.*, **13**, 969–980.

Beaumont M.A., Nichols R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B*, **263**, 1619–1626.

Bodó I., Gera I., Koppány G. (1996) The Hungarian Grey Cattle Breed. Association of the Hungarian Grey Cattle Breeders, Budapest, Hungary, pp. 128.

Edwards C.J., Bollongino R., Scheu A., Chamberlain A., Tresset A., Vigne J.D., Baird J.F., Larson G., Ho S.Y., Heupink T.H., Shapiro B., Freeman A.R., Thomas M.G., Arbogast R.M., Arndt B., Bartosiewicz L.,

Benecke N., Budja M., Chaix L., Choyke A.M., Coqueugniot E., Döhle H.J., Göldner H., Hartz S., Helmer D., Herzig B., Hongo H., Mashkour M., Ozdogan M., Pucher E., Roth G., Schade-Lindig S., Schmölcke U., Schulting R.J., Stephan E., Uerpmann H.P., Vörös I., Voytek B., Bradley D.G., Burger J. (2007) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc. Biol. Sci.*, **274**, 1377–1385.

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

Goudet J. (1995) *FSTAT* (version 1.2): A Computer Program to Calculate F-Statistics. *J. Hered.*, **86**, 485–48.

Goudet J. (2001) *FSTAT*, a Program to Estimate and Test Gene Diversities and Fixation Indices (version 2.9.3). (available at: http://www2.unil.ch/popgen/softwares/fstat.htm; last accessed 25 April 2010).

Guo S.W., Thompson E.A. (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 359.

Kelley J.L., Madeoy J., Calhoun J.C., Swanson W., Akey J.M. (2006) Genomic signatures of positive selection in humans and the limits of outlier approaches. *Genome Res.*, **16**, 980–989.

Liu K., Muse S. (2001). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics 2005* **21**(9), 2128–2129. (available at: http://statgen.ncsu.edu/powermarker; last accessed 25 April 2010).

Luikart G., England P.R., Tallmon D., Jordan S., Taberlet P. (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat. Rev. Genet.*, **4**, 981–994.

Maróti-Agóts Á., Bodó I., Valentini A., Verny M., Jávorka L., Gera I. (2005) Comparison of Body Measurements of Hungarian Grey and Maremman Cattle Breeds. Proceedings of the 4th World Italian Beef Cattle Congress, Italy, April 29th–May 1st 2005, 223–228.

Mason I.L. (1996) A World Dictionary of Livestock Breeds, Types and Varieties, 4th edn. C.A.B International, Wallingford (UK), 273 pp.

Negrini R., Nicoloso L., Crepaldi P., Milanesi E., Colli L., Chegdani F., Pariset L., Dunner S., Leveziel H., Williams J.L., Ajmone Marsan P. (2008) Assessing SNP markers for assigning individuals to cattle populations. *Anim. Genet.*, **40**, 18–26.

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

Pariset L., Cappuccio I., Joost S., D'Andrea M., Marletta D., Ajmone Marsan P., Valentini A., The ECONOGENE Consortium (2006a) Characterization of single nucleotide polymorphisms in sheep and their variation as evidence of selection. *Anim. Genet.*, **37**, 290–292.

Pariset L., Cappuccio I., Ajmone Marsan P., Dunner S., Luikart G., England P.R., Obexer-Ruff G., Peter C., Marletta D., Pilla F., Valentini A. (2006b) Assessment of population structure by single nucleotide polymorphisms (SNPs) in goat breeds. *J. Chromatogr. B*, **833**, 117–120.

Pariset L., Joost S., Ajmone Marsan P., Valentini A. (2009) Landscape genomics and biased *Fst* approaches reveal single nucleotide polymorphisms under selection in goat breeds of North-East Mediterranean. *BMC Genet.*, **10**, 7.

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

Raymond M., Rousset F. (1995) *Genepop* (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.*, **86**, 248–249.

Rosenberg N.A. (2004) *Distruct*: a program for the graphical display of population structure. *Mol. Ecol. Notes*, **4**, 137–138.

Scherf B.D. (2000) World Watch List for Domestic Animal Diversity, 3rd edn. Food and Agriculture Organization of the United Nations, Rome.

Soysal M.İ., Kok S. (2006) The last survivors of Grey cattle's whose resisting not to be extincted. A case study of characteristics and sustainability of traditional system of native Grey cattle breed. In: Olaizola A., Boutonnet J.P., Bernues A. (eds.), Mediterranean livestock production: uncertainties and opportunities. CIHEAM, Zaragoza, 441 p. (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 78). ISBN 2-85352-374-8. 2. Seminar of the Scientific-Professional Network on Mediterranean Livestock Farming (RME), 2006/05/18-20, Zaragoza, (Spain).

Soysal M.I., Tuna Y.T., Gurcan E.K., Ozkan E. (2004) Farms in turkey: sustainable development in the preservation of animal genetic resources in turkey and in the world. *Trakia J. Sci.*, **2**, 47–53.

The Bovine HapMap Consortium (2009) Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. *Science*, **324**, 528–532.

Valentini A., Marchitelli C., Crisà A., Pariset L., Bodò I., Marò ti-Agò ts A., Nardone A. (2006) Stima della diversità genetica nella popolazione bovina maremmana di Castelporziano e nei bovini Grigi Ungheresi. In: "il sistema ambientale della tenuta presidenziale di Castelporziano. Ricerche sulla complessità di un ecosistema forestale costiero mediterraneo, II serie, Segretariato generale della presidenza della repubblica italiana. Accademia Nazionale delle Scienze detta dei Quaranta "Scritti e Documenti" XXXVII.

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

Williams J.L., Dunner S., Valentini A., Mazza R., Amarger V., Checa M.L., Crisa' A., Razzaq N., Delourme D.,

Grandjean F., Marchitelli C., Garcia D., Perez Gomez R., Negrini R., Ajmone Marsan P., Leveziel H. (2009) Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. *Anim. Genet.*, **40**, 486–491.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Major Allele frequency, Expected and Observed Heterozygosity, Polimorfic information content (PIC), *Fis* (*f*) values for the genotyped SNPs, Exact p-value in HW test of Guo and Thompson (1992). Chromosome (BTA) and chromosomal location relative to Btau 4.0 genome sequence are those reported by Williams *et al.* (2009).

Table S2 SNPs used for haplotype reconstruction (SNPs), Major Allele frequency, Expected and Observed Heterozygosity, Polymorphic information content (PIC), *Fis* (*f*), Exact p-value in HW test of Guo and Thompson (1992).

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

2.3. Genetic Diversity of Sheep Breeds from Albania, Greece, and Italy Assessed by Mitochondrial DNA and Nuclear Polymorphisms (SNPs)

Research Article
TheScientificWorldJOURNAL (2011) 11, 1641–1659
ISSN 1537-744X; doi:10.1100/2011/186342

TheScientificWorldJOURNAL
www.thescientificworld.com

Genetic Diversity of Sheep Breeds from Albania, Greece, and Italy Assessed by Mitochondrial DNA and Nuclear Polymorphisms (SNPs)

Lorraine Pariset,[1] Marco Mariotti,[1] Maria Gargani,[1] Stephane Joost,[2] Riccardo Negrini,[3] Trinidad Perez,[4] Michael Bruford,[4] Paolo Ajmone Marsan,[3] and Alessio Valentini[1]

[1] *Department for Innovation in Biological, Agro-Food and Forest Systems, Tuscia University, 01100 Viterbo, Italy*
[2] *Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland*
[3] *Istituto di Zootecnica, Università Cattolica del Sacro Cuore, 29122 Piacenza, Italy*
[4] *School of Biosciences, University of Wales, Cardiff CF10 3NS, UK*

Received 12 April 2011; Revised 30 July 2011; Accepted 8 August 2011

Academic Editor: Dirk-Jan de Koning

We employed mtDNA and nuclear SNPs to investigate the genetic diversity of sheep breeds of three countries of the Mediterranean basin: Albania, Greece, and Italy. In total, 154 unique mtDNA haplotypes were detected by means of D-loop sequence analysis. The major nucleotide diversity was observed in Albania. We identified haplogroups, A, B, and C in Albanian and Greek samples, while Italian individuals clustered in groups A and B. In general, the data show a pattern reflecting old migrations that occurred in postneolithic and historical times. PCA analysis on SNP data differentiated breeds with good correspondence to geographical locations. This could reflect geographical isolation, selection operated by local sheep farmers, and different flock management and breed admixture that occurred in the last centuries.

KEYWORDS: mtDNA, sheep, SNPs, Mediterranean, domestication

Correspondence should be addressed to Lorraine Pariset, pariset@unitus.it

1. INTRODUCTION

The earliest archaeozoological evidence of domestic sheep comes from a restricted area of south-western Asia: modern Iran, Turkey, and Cyprus [1]. A pioneering genetic study examining the karyotypes of the various species of extant wild sheep [2, 3] showed that domestic sheep derive from the Asiatic mouflon (*Ovis orientalis*) of Anatolia, western Iran, and southwest Iran. Afterwards, a probable migration of the Neolithic farmers occurred out of the Near East and across Europe following two main routes, through the continental heartland up the Danube valley or along the Mediterranean coast [4, 5] crossing the sea to the major islands. Archaeological data and radiocarbon dates on seeds or bones provide support for an earlier arrival in Western Europe via the Mediterranean route rather than the "Danubian" route [6].

Both archaeozoological evidence and genetic evidence indicate that the domestication of wild sheep occurred 8000–9000 years ago. The first appearance of the remains of domestic sheep in the western part of Mediterranean Europe, dating approximately 5400 BC, is believed to reflect a rapid spread by sea [7, 8]. The Mediterranean Sea also had a key role in the history of livestock in postneolithic times, when peoples like Phoenicians, Greeks, Romans, and Berbers probably introduced new species of animals and new breeds of livestock into southwest Europe by sea. Some settlers may have improved local livestock by importing stock from overseas [8], which explained the unexpectedly high within-breed diversity in domestic goats [9, 10], the differential cattle migration along the Mediterranean coast [11], and the close genetic relationship between Tuscan and Near Eastern cattle breeds [12]. The role of the Mediterranean Sea as a natural corridor connecting the South West Europe to the Near East and North Africa is particularly plausible for domestic sheep and goats that were adaptable to various environments and easy to transport due to their size [8]. Subsequently, sheep breeds developed after selective breeding for desirable traits (wool, milk, and meat production) and environmental tolerance. Since domestication, sheep have established a wide geographic range due to their adaptability to poor diets and extreme climatic conditions as well as their manageable size.

The genetic history of sheep has been investigated using three major sources of genomic variation: autosomes, Y chromosome, and mitochondrial genome. Analysis of the nonrecombining region of the Y chromosome has revealed patterns of male-mediated introgression during breed development [13, 14]. Recent surveys have tested collections of animals from southern and northern Europe [15] or Europe and the Middle East [16] using microsatellites and enabled the analysis of genetic partitioning at a continental scale. Interestingly, southern European breeds displayed increased genetic diversity and decreased genetic differentiation compared to their northern European counterparts. This is consistent with the expectation that genetic diversity will be maintained high in populations close to the centre of domestication but decreases with increasing geographic distances. Kijas et al. [17] used a SNP panel to analyse sheep nuclear genome, providing the indication that breeds cluster into large groups based on geographic origin and that SNPs can successfully identify population substructures within individual breeds.

A recent study on retrovirus integrations [18] has provided additional information on the introduction of sheep into Europe, indicating an early arrival of the primitive sheep populations (European mouflons, North-Atlantic Island breeds) and a subsequent advent of wool producing sheep.

However, most of the information about history and domestication of the species have been gathered using mtDNA. The existence of multiple mtDNA lineages and their admixture within breeds [8, 19–22] could be due to multiple domestication events and subsequent human selection or introgression by domestic and wild species.

Mitochondrial DNA analyses in sheep identified an increasing number of maternal lineages: two [23–25], three [20, 26], and then five [22]. The main haplogroups A and B are both found in Asia, while B dominates in Europe. Haplogroup C has been found in Portugal, Turkey, the Caucasus, and China [7]. Haplogroup D, present in Rumanian Karachai and Caucasian animals, is possibly related to the haplogroup A. In contrast to taurine cattle, the sheep haplogroups hardly correlate with geography.

Because of their mode of inheritance, mitochondrial markers are more likely to lead to biased estimates of species phylogeny [27]. Combining nuclear and mitochondrial markers may help in avoiding

TABLE 1: Country of origin, breeds, and acronyms used in computations.

Country	Breed	Acronym
Albania	Bardhoka	BAR
	Ruda	RUD
	Shkordane	SHK
Greece	Kalarritiko	KAL
	Orino	ORI
	Pilioritiko	PIL
	Kefalleneas	KEF
	Lesvos	LES
	Kymi	KIM
	Karagouniko	KAR
	Skopelos	SKO
	Anogeiano	ANO
	Sfakia	SFA
Italy	Bergamasca	BER
	Delle Langhe	LAN
	Laticauda	LAT
	Altamurana	ALT
	Gentile di Puglia	GDP

this problem. Nuclear genome evolves five-to-ten times slower than mtDNA; it is contributed by both parents and its variability is less affected by demographic forces such as bottleneck. Therefore, nuclear markers can detect more recent genetic events that influence the extant divergence of domestic breeds. Several studies have demonstrated that the combination of nuclear and mtDNA markers can increase the information obtained [27–30]. The use of both markers might provide a more accurate and comprehensive understanding of a species' history [31]. SNP markers could help in understanding the recent evolutionary history of domestic animals [10, 32].

We aimed at investigating the geographic distribution of the genetic diversity of sheep breeds in Albania, Greece, and Italy and to gather information on the migration history of the species. To accomplish that, we employed sequence data from the mitochondrial D-loop and 27 nuclear loci (SNPs).

2. MATERIALS AND METHODS

2.1. Sampling and DNA Extraction

We focused on sheep breeds of Albania, Greece, and Italy. Samples of the European mouflon were also included. About twenty unrelated samples per breed were selected. Three animals per flock from 11 farms spread over the traditional rearing area were sampled. A total of 313 animals from 18 sheep breeds were analyzed. Breeds, acronyms used, and country of origin of each breed are reported in Table 1. Part of the samples were obtained from a previous project (Econogene, http://www.econogene.eu/). Blood samples were collected in EDTA tubes and frozen at -20°C until extraction. Genomic DNA was isolated using standard procedures, checked for DNA quality on agarose gel and quantified using a DTX microplate reader (Beckman Coulter) after staining with PicoGreen (Invitrogen).

2.2. Amplification and Sequencing of the Mitochondrial D-Loop

To amplify the partial D-loop of 721 bp, primers described by Tapio et al. [7] were used from 15,541 to 16,261 of the complete sequence described by Hiendleder et al. [33] available in GenBank (NC_001941.1).

Polymerase chain reaction (PCR) was performed in a total volume of 50 μL containing 20 ng of genomic DNA, 40 pMol of each primer (Sigma-Aldrich), 200 μM dNTPs, 5X PCR buffer, and 5 units of *Taq* DNA polymerase (Promega) on a PCR Thermo Cycler (MJ Research). A 5 minutes denaturation step at 95°C was followed by 14 cycles of denaturation at 95°C for 30 sec, annealing for 30 sec starting at 62°C and decreasing 0.5°C per cycle and extension at 72°C for 120 sec, then by 20 cycles of denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec and extension at 72°C for 120 sec; the final extension step was carried out at 72°C for 5 minutes. PCR products were purified through ExoSap-IT (USB Corporation) to remove residual primers and dNTPs and used as templates for forward and reverse sequencing reactions.

Sequencing was performed using the primers described by Tapio et al. [7] with a CEQ 8800 sequencer using DTCS QuickStart Kit and purifying with Agencourt CleanSEQ 96 (Beckman Coulter), according to the manufacturer's instructions. After the optimization of the sequencing protocol, sequencing was outsourced to Macrogen (http://www.macrogen.com/). The sequences of D-loop were submitted to GenBank (accession numbers: JN184789–JN184999).

2.3. Mitochondrial Sequence Analysis

A fragment of 435 bp, running from 15,541 to 16,261 bp (NC_0019041.1), was selected excluding a central region rich in tandem repeats (from 15,644 to 15,932 bp). mtDNA variations were identified on a total of 313 sequences of 18 breeds analyzed and aligned with BioEdit software [34]. DnaSP 5.00 software [35] was used to calculate haplotype, sequence variation, average number of nucleotide differences (D), and average number of nucleotide substitutions (Dxy) per site between breeds. Neighbour-joining tree for all haplotypes was constructed using Mega version 5 [36]. Analysis of molecular variance (AMOVA) was performed with Arlequin version 3.11 [37]. Sequences of the same D-loop fragment in wild sheep, published by Hiendleder et al. [33], were obtained from GenBank, *Ovis vignei arkal* (AY091489.1), *Ovis vignei bocharíensis* (AY091490.1, AY091491.1, and AF039580.1), *Ovis ammon collium* (AY091492.1), *Ovis ammon nigrimontana* (AY091493.1 and AY091494.1), and used as outgroups in phylogenetic analysis.

Geographic distribution of eigenvectors was performed to investigate population genetic differences on the basis of their geographic distances. This approach permitted the generation of a synthetic configuration of locations based on the pairwise genetic distances that matched the real geographic configuration. Principal component analysis (PCA) scores for the first two components, obtained using Nei's 1973 genetic distance, were plotted on a geographic map. As breeds are scattered among several farms, a virtual geographic entity representing the centroid of each breed on geographic maps was created using WGS84 geographical coordinates [38]. For a given component, it is a measure of the variance accounted for by that component. On thematic maps produced with the geographic information system (GIS) Manifold software package (Manifold System, Version 7, Manifold Net Ltd., Carson City, USA, http://www.manifold.net/), all breeds are thus represented according to a geometric distribution (see Figures 3(a) and 3(b)). Breeds showing high eigenvectors contribute sensibly to the explanation of the variance related to the component displayed. Classes were elaborated on the basis of the criterion of the natural breaks (Jenks optimization method). This algorithm reduces the variance within classes and maximizes the variance between classes. Colour classes were chosen in order to support the distinction between the different categories of behaviours observed: green: positive contribution; yellow: intermediary values; red: negative contribution to the component displayed.

2.4. Nuclear Polymorphism Analysis

The same 313 sheep belonging to 18 breeds sequenced at D-loop were genotyped with 37 previously described SNPs [39]. SNP ascertainment bias was minimised by sequencing target DNA in at least 8 individuals from different populations. Large-scale genotyping of all animals was performed by outsourcing to a commercial genotyping company (http://www.Kbioscience.co.uk/).

TABLE 2: Sample size per country (n), number of haplotypes observed (Haplotypes), number of polymorphic sites, mean number of pairwise differences among sequences (pairwise diff.), haplotype diversity (h), and nucleotide diversity (π) are shown.

Country	n	Haplotypes	Polymorphic sites	Pairwise diff.	h	π
Italy	93	62	58	4.180	0.978	0.01007
Greece	167	83	73	5.934	0.934	0.01469
Albania	53	37	57	8.704	0.979	0.02107

Allele frequencies, Nei's estimation of observed and expected heterozygosities (Ho and He, resp.), were calculated using Fstat 2.93 [40]. Weir and Cockerham's [41] estimates of F_{is} per population, F_{st} per locus, and population pairs were calculated for each locus using Genalex 4.0 [42]. The same software was used to test deviations from Hardy-Weinberg equilibrium (HWE) for each locus and population and for locus over all populations; test for conformity with HWE expectations was assessed by calculating the Chi-squared value.

Correlation between geographic and Nei's 1973 pairwise genetic distances was tested using Mantel tests (999 permutation) implemented in Genalex 4.0 software [42].

A PCA was performed on the covariance matrix of SNP frequency data to investigate spatial patterns of genetic variation using GENETIX software [43].

Nei [44] and Reynolds [45] genetic distances between population pairs were calculated using PowerMarker v3.25 [46].

Geographic distribution of eigenvectors was performed as described above using pairwise genetic distances [47] calculated on the basis of the selected SNP markers.

3. RESULTS

3.1. Mitochondrial Haplotypes

Ninety-three polymorphic sites and 154 haplotypes were identified from 313 sequences. Relatively high haplotype diversity was found in all three sampled geographic regions; the largest nucleotide diversity is present in Albania (0.02107) while the highest number of haplotypes observed is recorded in Greece (83) (Table 2).

The average number of nucleotide differences and the average number of nucleotide substitutions per site were used to calculate the genetic distance between breeds. The lowest distance was observed between Laticauda and Anogeiano (D: 2.357—Dxy: 0.006), while the highest distance was observed between Bardhoka and Kymi (D: 12.450—Dxy: 0.03) (Table 3).

AMOVA revealed that mitochondrial diversity is mainly distributed within breeds (95.04%) and only in part among regions (0.90%); low variability was also found among breeds/within regions (4.06%) (Table 4).

3.2. Phylogenetic Analysis and Haplogroups

The NJ tree obtained from mtDNA haplotypes and wild sheep sequences, used as out-group, revealed three of the five haplogroups described in the literature: A, B, and C (Figure 1). Haplogroup B is the most frequent among the analyzed samples (89%), while A and C are less common (8% and 3%, resp.). Greek and Albanian breeds are present in all three haplogroups, while Italian breeds are present only in haplogroups B and A (Table 5). This is shown also in Figure 2, representing the percentage of occurrence of each haplogroup in Albania, Greece, and Italy.

TABLE 3: Average number of nucleotide differences, D (below), and average number of nucleotide substitutions per site between populations, Dxy [48] (above).

	BER	ALT	LAT	LAN	GDP	BAR	RUD	SHK	KAL	ORI	PIL	KEF	LES	KIM	KAR	SKO	ANO	SFA
BER	*	0.009	0.009	0.012	0.011	0.018	0.011	0.021	0.009	0.017	0.009	0.018	0.018	0.025	0.012	0.011	0.008	0.009
ALT	3.908	*	0.007	0.011	0.009	0.017	0.009	0.02	0.007	0.016	0.008	0.016	0.017	0.024	0.011	0.009	0.006	0.008
LAT	3.728	2.93	*	0.01	0.009	0.016	0.009	0.02	0.007	0.015	0.007	0.016	0.017	0.025	0.01	0.008	0.006	0.007
LAN	5.079	4.425	4.279	*	0.012	0.019	0.013	0.022	0.01	0.018	0.01	0.019	0.019	0.026	0.013	0.012	0.01	0.01
GDP	4.632	3.906	3.756	5.134	*	0.018	0.011	0.021	0.009	0.017	0.01	0.017	0.018	0.024	0.012	0.011	0.009	0.009
BAR	7.447	6.933	6.856	8.107	7.579	*	0.018	0.026	0.016	0.023	0.017	0.023	0.024	0.03	0.02	0.018	0.016	0.016
RUD	4.612	3.95	3.7	5.245	4.711	7.647	*	0.021	0.009	0.017	0.01	0.018	0.018	0.025	0.012	0.011	0.008	0.009
SHK	8.778	8.272	8.345	9.269	8.696	10.567	8.789	*	0.02	0.025	0.02	0.025	0.025	0.029	0.023	0.021	0.02	0.02
KAL	3.712	3	2.708	4.183	3.754	6.741	3.778	8.241	*	0.015	0.007	0.016	0.017	0.024	0.01	0.008	0.006	0.007
ORI	7.108	6.488	6.395	7.536	7.132	9.4	7.122	10.395	6.352	*	0.016	0.023	0.023	0.029	0.018	0.017	0.015	0.015
PIL	3.91	3.289	3.011	4.5	4.021	7.049	4.067	8.419	2.978	6.556	*	0.017	0.017	0.025	0.011	0.009	0.007	0.007
KEF	7.349	6.861	6.832	7.853	7.319	9.804	7.475	10.306	6.757	9.486	6.971	*	0.022	0.026	0.019	0.018	0.016	0.016
LES	7.515	7.042	6.941	8.053	7.507	10.1	7.588	10.396	6.896	9.576	7.125	9.367	*	0.027	0.019	0.018	0.016	0.016
KIM	10.426	10.083	10.276	10.994	10.211	12.45	10.438	11.931	10.194	12.028	10.383	11.141	11.375	*	0.026	0.026	0.024	0.024
KAR	5.144	4.478	4.197	5.595	5.139	8.217	5.185	9.378	4.294	7.65	4.503	7.894	7.975	10.975	*	0.012	0.01	0.01
SKO	4.497	3.852	3.538	4.989	4.544	7.585	4.611	8.883	3.472	7.102	3.719	7.455	7.618	10.813	4.936	*	0.008	0.009
ANO	3.514	2.678	2.357	4.053	3.526	6.614	3.505	8.123	2.439	6.132	2.712	6.655	6.73	10.079	4.003	3.301	*	0.006
SFA	3.768	3.164	2.878	4.266	3.801	6.698	3.834	8.015	2.906	6.266	3.182	6.658	6.77	9.757	4.284	3.596	2.643	*

TABLE 4: Hierarchical analysis of molecular variance (AMOVA) with 10,000 permutations.

Source of variation	Variation (%)	Fixation indices [49]	*P* value
Among regions	0.90	FSC: 0.0495	<0.01
Among breeds/within regions	4.06	F_{st}: 0.04960	<0.01
Within breeds	95.04	FCT: 0.00903	<0.5

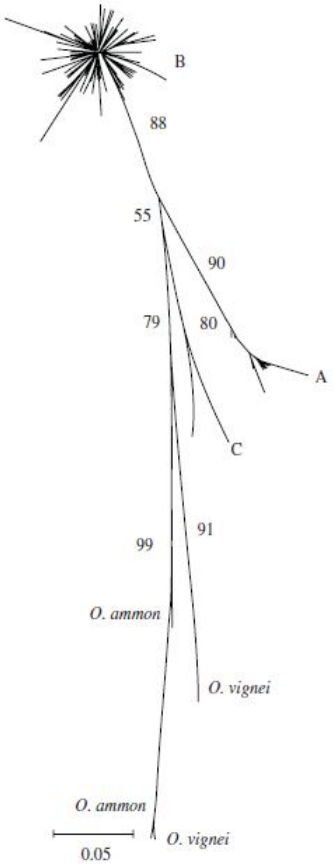


FIGURE 1: Neighbour-joining tree based on the 154 mtDNA haplotypes, showing differences among haplogroups and wild sheep (outgroup). Numbers indicate the percentage bootstrap support (10,000 resamplings).

3.3. SNP Analysis

A total of 37 SNPs identified as polymorphic on eighteen sheep breeds selected throughout Europe [39] were applied to genotype 313 individuals of three Albanian, ten Greek, and five Italian sheep breeds. After removing those found monomorphic in the selected breeds, 27 SNPs were used for the analysis.

The frequencies of the major alleles ranged from 0.99 for the locus *LEP1* to 0.538 for the locus *IL2_1*. Except for *CAST_1*, *LEP1*, *LEP2*, *GDF8*, and *PRNP_1*, which show frequencies of the rare alleles of 0.035,

TABLE 5: Sequences analysed, number of variable sites (# Var), number of haplotypes (Hap), and percentage of haplotypes in the haplogroups A, B, or C per breed.

Breed	Sequences	# Var	Hap	A	B	C
BAR	15	43	14	0.07	0.79	0.14
RUD	20	35	15	0.07	0.93	0
SHK	18	42	11	0.09	0.73	0.18
ANO	19	12	7	0	1	0
KAL	18	18	11	0	1	0
KAR	20	38	14	0.07	0.93	0
KEF	16	31	11	0.18	0.82	0
KIM	8	27	6	0.33	0.77	0
LES	16	38	13	0.15	0.85	0
ORI	18	43	13	0.08	0.84	0.08
PIL	15	17	11	0	1	0
SFA	19	15	11	0	1	0
SKO	18	27	13	0	1	0
ALT	18	19	11	0	1	0
BER	17	30	12	0.08	0.92	0
GDP	19	32	16	0.06	0.94	0
LAN	20	41	17	0.06	0.94	0
LAT	19	17	13	0	1	0

0.010, 0.020, 0.021, and 0.036, respectively, the remaining SNPs have rare allele frequency of greater than 5%. Observed heterozygosity of all loci ranged from 0.019 (*LEP1*) to 0.489 (*IL2_1*), with a mean of 0.250. Expected heterozygosity of the loci ranged from 0.018 (*LEP1*) to 0.474 (*IL2_1*), with a mean of 0.259.

Significant deviations from Hardy-Weinberg equilibrium (HWE) for each locus and population (P value < 0.01) were observed for nine loci. Locus *BMPR* was not in HWE in Skopelos breed. Locus *ACVR2B_2* was not in HWE in Bardhoka population, and locus *SERPINA3* was not in HWE in Karagouniko population. Loci *ZP2* and *KRTAP6* were not in HWE in Altamurana. *PRNP_1* was not in HWE in four Italian breeds: Bergamasca, Delle Langhe, Gentile di Puglia, and Laticauda. *CSN3* was not in HWE in Ruda, Skopelos, and Bergamasca populations; *KRT1* in Delle Langhe, Laticauda, and Karagouniko; *GHR* in Karagouniko and Laticauda.

F_{is} per population ranged from -0.102 (Shkordane) to 0.276 (Altamurana) with a mean of 0.033 (P value $= 0.00010$ on 9720 randomisations).

The analysis of population differentiation revealed an overall F_{st} of 0.048, that is, 4.8% of allelic variation accounted across breeds and 95.2% within breeds. Weir and Cockerham's [41] estimate of F_{st} per locus ranged from 0.0132 (*IL2_2*) to 0.098 (*DESMIN_2*), with a mean of 0.047.

Both Nei's [44] and Reynolds' [45] genetic distances were calculated (Table 6). The distances range from 0.022 (Ruda-Karagouniko) to 0.253 (Sfakia-Gentile di Puglia) using Reynolds' distances and from 0.008 (Ruda-Karagouniko) to 0.096 (Gentile di Puglia-Anogeiano) using Nei's distances. Both indices indicate Ruda and Karagouniko as the breeds with the minimum pairwise distance and show the maximum distances between the Italian breed Gentile di Puglia and Anogeiano, Karagouniko, Kefalleneas, Kymi, Lesvos, Pilioritiko, Sfakia, Skopelos, and Bergamasca breeds. The Mantel test showed correspondence between geographic and genetic distances with a $P = 0.04$.

TABLE 6: Pairwise population matrix of Nei's [44], below, and Reynolds' [45], above, genetic distances. Breed codes are as in Table 1.

	BAR	RUD	SHK	ANO	KAL	KAR	KEF	KIM	LES	ORI	PIL	SFA	SKO	ALT	BER	GDP	LAN	LAT
BAR	*	0.060	0.046	0.065	0.052	0.054	0.055	0.104	0.110	0.117	0.083	0.061	0.070	0.107	0.103	0.167	0.121	0.055
RUD	0.018	*	0.036	0.072	0.082	0.022	0.046	0.067	0.059	0.148	0.104	0.062	0.062	0.079	0.049	0.203	0.100	0.051
SHK	0.017	0.009	*	0.072	0.072	0.039	0.058	0.078	0.087	0.112	0.063	0.080	0.042	0.101	0.062	0.174	0.123	0.058
ANO	0.019	0.029	0.025	*	0.086	0.055	0.075	0.109	0.096	0.147	0.103	0.045	0.098	0.128	0.085	0.247	0.122	0.073
KAL	0.014	0.019	0.023	0.023	*	0.078	0.071	0.127	0.124	0.076	0.062	0.107	0.112	0.142	0.104	0.159	0.119	0.092
KAR	0.013	0.008	0.009	0.020	0.013	*	0.047	0.071	0.072	0.142	0.103	0.057	0.060	0.087	0.048	0.216	0.102	0.054
KEF	0.009	0.015	0.014	0.025	0.013	0.014	*	0.095	0.091	0.141	0.106	0.074	0.076	0.089	0.091	0.248	0.124	0.058
KIM	0.034	0.024	0.025	0.039	0.033	0.021	0.035	*	0.081	0.154	0.134	0.107	0.115	0.102	0.100	0.241	0.157	0.075
LES	0.029	0.017	0.025	0.029	0.028	0.021	0.030	0.028	*	0.180	0.148	0.095	0.110	0.116	0.086	0.249	0.134	0.070
ORI	0.016	0.018	0.018	0.017	0.010	0.009	0.011	0.027	0.027	*	0.090	0.169	0.161	0.182	0.168	0.152	0.163	0.140
PIL	0.024	0.020	0.013	0.021	0.025	0.012	0.020	0.028	0.025	0.021	*	0.124	0.104	0.161	0.125	0.152	0.161	0.107
SFA	0.014	0.026	0.025	0.015	0.026	0.020	0.024	0.040	0.028	0.020	0.021	*	0.098	0.140	0.097	0.253	0.122	0.067
SKO	0.022	0.021	0.013	0.040	0.034	0.020	0.021	0.039	0.037	0.031	0.021	0.034	*	0.128	0.083	0.232	0.152	0.098
ALT	0.031	0.024	0.031	0.047	0.041	0.026	0.025	0.037	0.040	0.030	0.042	0.051	0.041	*	0.119	0.226	0.132	0.073
BER	0.039	0.021	0.021	0.035	0.028	0.019	0.036	0.042	0.032	0.026	0.024	0.040	0.029	0.049	*	0.233	0.070	0.083
GDP	0.077	0.062	0.077	0.096	0.086	0.068	0.079	0.089	0.083	0.076	0.083	0.081	0.094	0.067	0.082	*	0.212	0.173
LAN	0.053	0.045	0.058	0.054	0.047	0.044	0.053	0.076	0.061	0.037	0.058	0.051	0.069	0.057	0.032	0.078	*	0.099
LAT	0.017	0.017	0.019	0.025	0.025	0.016	0.017	0.030	0.024	0.022	0.023	0.021	0.034	0.026	0.035	0.064	0.048	*

3.4. PCA on SNPs

Genetic relationships were also explored by means of PCA. To examine the overall pattern of population differentiation, we considered the first three axes, which cumulatively explained 48.87% of the total inertia contained in the data set (Figure 4). From PCA, it can be seen that the some breeds are quite differentiated, with good correspondence to geographical locations, even if SNPs were few (4 of the 5 Italian breeds are well separated from the main cluster). Particularly, a differentiation between northern Italian (Bergamasca and Delle Langhe) and southern Italian (Gentile di Puglia, Altamurana, and Laticauda) breeds can be seen. The projection of loci in the space formed by the first three principal components (data not shown) shows that the differentiation of outlying breeds is caused by a small number of SNPs: the Delle Langhe population is mainly affected by alleles in the *MSTN* gene along the second component.; the Skopelos breed position in the graph is affected by alleles in the *PRP* gene and the Laticauda breed (lesser) differentiation is mainly due to alleles in the *CALPA* and *LEP* genes.

PCA scores calculated on mtDNA marker for the first 2 components were plotted on a geographic map, using the centroids of the sampling area of each breed (Figures 3(a) and 3(b)). The highest eigenvector contribution (coloured in green) was observed for Albanian and Greek breeds, as expected, even if four breeds show unexpectedly low diversity.

As for mtDNA analyses, the results of the PCA based on SNPs data were also used for making inferences about population genetic differences on the basis of their geographic distances. PCA scores for the first two principal components were plotted on a geographic map, using the centroids of the sampling area of each breed (Figures 3(c) and 3(d)). The line separating the map in two regions shows the isoline for an eigenvalue of 0. RUDA breed is an exception showing an eigenvalue above 0 but located in an area where all other breeds show lower eigenvalues. In accordance with the domestication history of the species, genetic diversity was higher in south-eastern populations than in north-western populations. The first component of Figure 3(c), in fact, shows a regular loss of genetic diversity towards North West.

4. DISCUSSION

Agriculture arose mainly within the distribution range of the wild ancestors of the most valuable livestock, such as the Fertile Crescent of southwest Asia, where early farmers were able to outcompete local hunter-gatherers. Once livestock slowly spread northwest across Europe, farming also shifted northwest from the Fertile Crescent to areas where farming had never arisen independently—first to Greece, then to Italy, and finally to northwest Europe [50]. Therefore, today the most productive farming zones do not correspond to those most productive in the past. Then, founder effect, genetic drift, and natural or artificial selection led to the formation of distinct breeds or varieties [51].

We therefore focused on the analysis of sheep of three countries aligned on this route, to evidence signs of migration. Geographical isolation, natural and artificial selection for physical or productive characters, genetic drift, mutations, and interpopulation gene flows have altered gene frequencies over many generations. The genetic diversity within and across breeds and species forms the basis of our current animal genetic resources for food production and other purposes.

The nature of the markers used for the analysis can affect the detection of geographical structuring, as suggested by Naderi et al. [52]. In fact, mtDNA informativeness is limited because it does not detect male-mediated gene flow and does not predict the nuclear genomic diversity [53]. Moreover, results may be affected by phenomena such as homoplasy, incomplete lineage sorting, effective population sizes, and sex-biased dispersal [27]. By combining markers with different modes of inheritance and rates of evolution this bias can be minimized [54].

Our mtDNA analysis shows higher levels of sheep nucleotide diversity in the South-East, which is congruent with data reported in the literature [22] and congruent with the proximity to the domestication centre. This is confirmed by eigenvector analysis, which showed high contribution to variance by Albanian

FIGURE 2: Geographical distribution of haplogroups, circles are proportional to the number of samples. (green: haplogroup A; purple: haplogroup B; red: haplogroup C).

and Greek breeds, even if four breeds show unexpectedly low diversity. However, this behaviour can be explained by recent isolation or selection for some traits that reduced the overall genetic diversity through bottleneck.

Very high haplotypes diversity was found in all three regions analysed (greater than 0.9), in agreement with previously published works on Portuguese breeds [9], Indian breeds [55], and Balkan sheep [56]. The major mitochondrial variation is distributed within breeds (95.04%), while it is lower among regions (0.90%) and among breeds within regions (4.06%). Phylogenetic methods were employed to examine the evolutionary history of the 18 breeds. Neighbour-joining and median-joining network revealed three of the described haplogroups, A, B, and C. The mouflon shares a haplotype with domestic sheep, as previously reported by Hiendleder et al. [33].

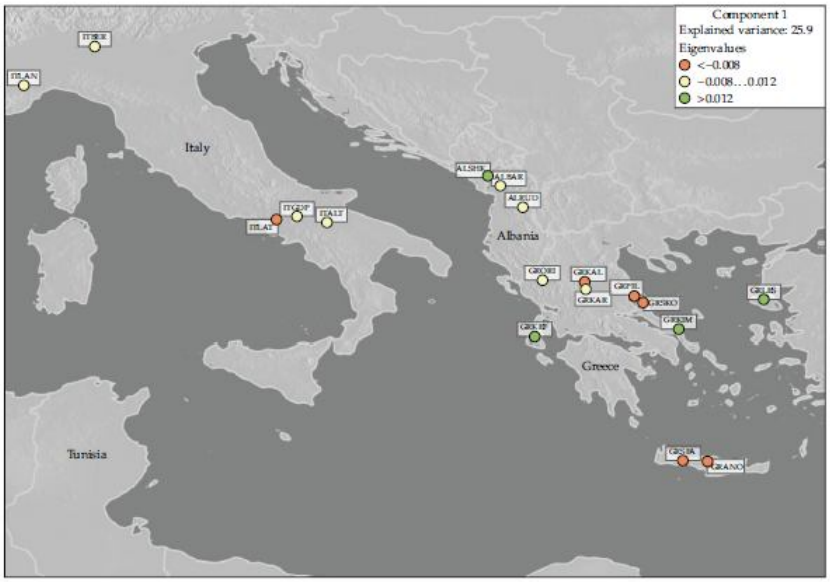
The SNP analysis revealed a rare allele frequency <5% for *LEP1* and *LEP2* loci, in agreement with those observed on a different European breed panel by Pariset et al. [39]. Observed and expected mean heterozygosity also showed similar values to those reported in the same paper. Expected heterozygosity values, which can indicate response to selection, are higher than observed heterozygosity values (Hs 0.063, 0.07, and 0.042; Ho 0.052, 0.06, and 0.038 in *CALPA*, *PRNP-1*, and *GDF8*, resp.).

Among the breeds tested, Altamurana showed the highest F_{is} value suggesting the inbreeding in this population.

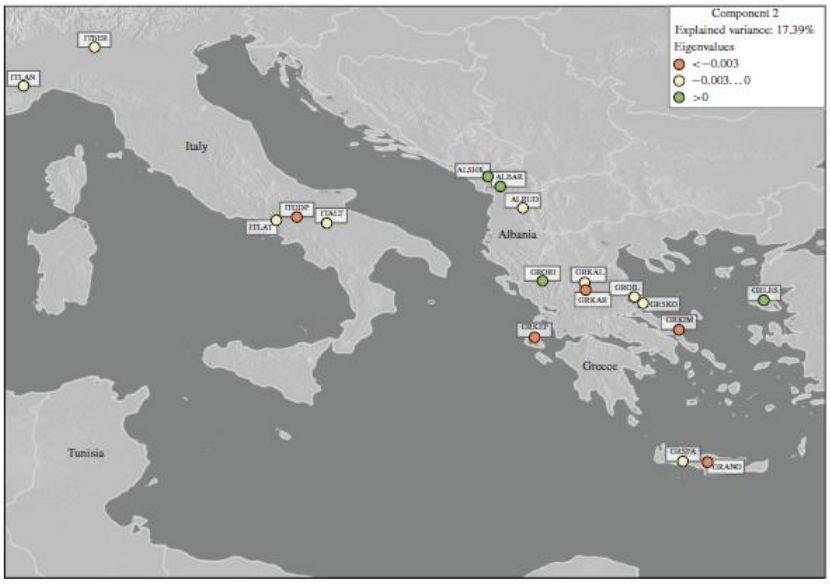
Regarding the phylogeographic structure we found that the 95.2% of variation occurred within breeds indicating the weak phylogeographic structure in sheep. These data are consistent with those previously published by Kijas et al. using a different SNP panel [17]. Sheep generally do not have a strong geographic structure and show a high genetic variability within breeds.

Anyway, Mantel test analysis using SNPs revealed a correlation between genetic and geographic distance. The possibility to assess the presence of a geographic component in genetic diversity using SNPs was already reported in previous studies on sheep [39] and goats [10, 57, 58].

In the PCA, the breeds appear differentiated with 48.9% of the variance explained by the first three principal components. Also this analysis shows a good correspondence to geographical locations: the breeds remaining separated by the main group are all Italian. PCA supports therefore a westward route to Italy that could indicate that transport of animals made by sea as already proposed for cattle [11, 12] and goats [8, 59]. This is particularly plausible because small sized species as sheep are easy to transport during migration

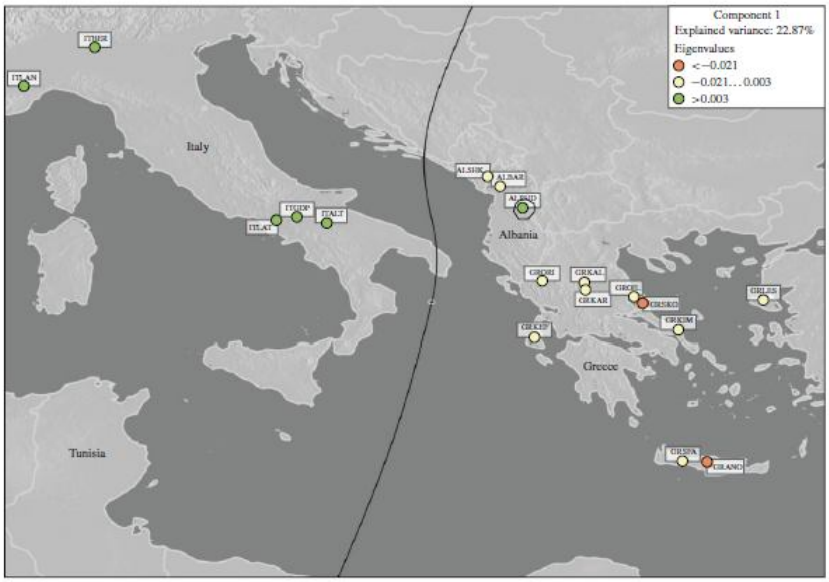


(a)

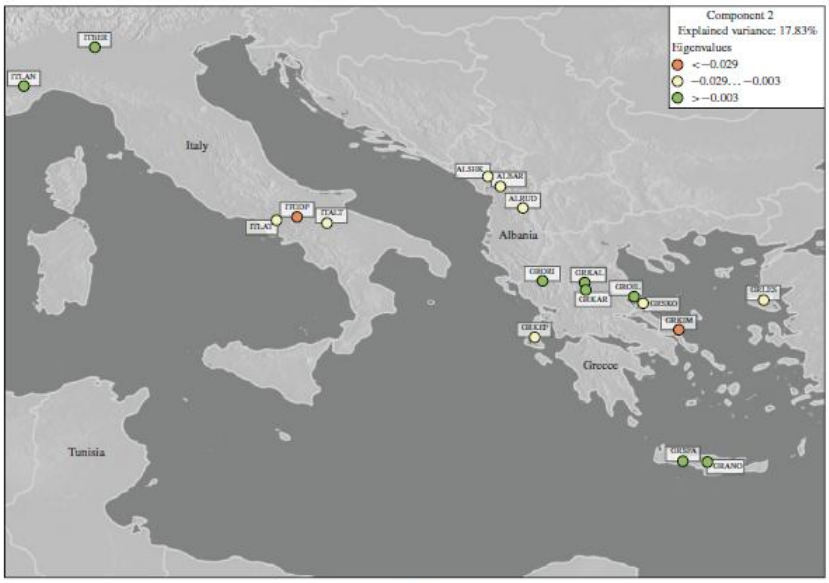


(b)

FIGURE 3: Continued.



(c)



(d)

FIGURE 3: First (a) and second (b) components of eigenvectors spatial distribution calculated on mtDNA marker and first (c) and second (d) components calculated on SNPs markers. Background image is GTOPO30, a global digital elevation model (DEM) with a horizontal grid spacing of 30 arc sec-onds (approximately 1 kilometer) produced by the U.S. Geological Survey's Center for Earth Resources Observation and Science (EROS). The line in the map (c) shows the isoline for an eigenvalue of 0. It appears on panel C only because this is the only geographical configuration for which a limit is so obvious. The circle around the RUDA breed shows an isoline for the eigenvalue of 0.03.

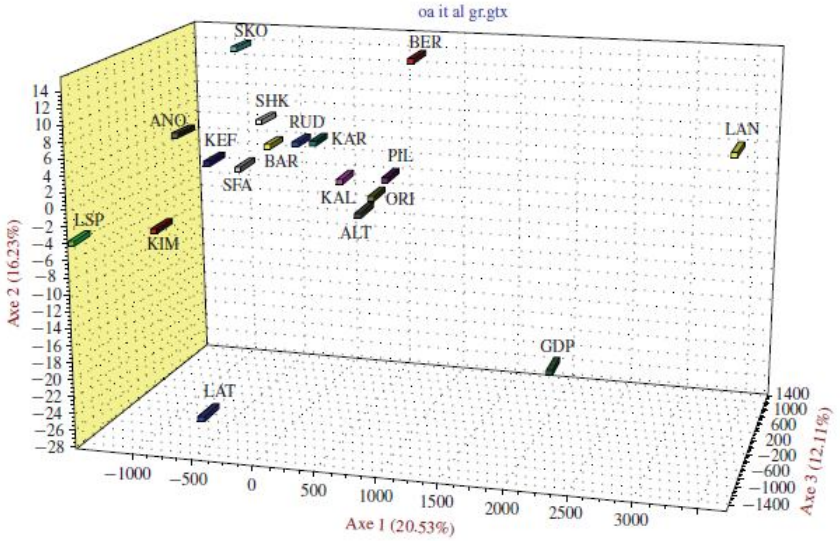


FIGURE 4: PCA analysis of the breeds assessed by SNP markers. The first three axes, cumulatively explaining 48.87% of the total inertia, were considered.

and commercial trade and can adapt easily to various environments [53, 60, 61]. A similar decrease in genetic diversity as well as an increase in the level of differentiation at the breed level from South-East to North-West in European sheep breeds, supporting the hypothesis of livestock migration from the Middle East towards western and northern Europe, was found by Peter et al. [16] and Lawson Handley et al. [15], using other nuclear markers.

The formulation of the modern breed concept during mid-1800s has caused remarkable changes in the livestock sector: large-scale production expanded and its application to breeding and husbandry practices led to the formation of well-defined breeds that were exposed to intense anthropogenic selection. The differentiation of three breeds observed using PCA analysis could be related to a recent selection, which appears to be linked to *CALPA* (Laticauda), *PRNP_1* (Skopelos), and *GDF8* (Delle Langhe) SNP markers. Gentile di Puglia breed seems influenced by both *CALPA* and *GDF8*. In particular, these two genes have an effect on conformation and therefore are an easy target of selection. Other SNPs related to meat traits were found potentially under the effect of selection and apparently not associated with production attitude of the breeds [58].

5. CONCLUSIONS

We employed mtDNA and nuclear SNPs to investigate the genetic diversity of sheep breeds of three countries of the Mediterranean basin: Albania, Greece, and Italy. Our results showed significant genetic differentiation among the sheep breeds, supported by mtDNA and by SNP. The differentiation identified by nuclear markers could indicate a reduced gene flow due to geographical isolation, associated with different flocks management, or an effect of the introduction of different stocks centuries ago (cf. Figure 3(d), showing the 2nd dimension geographic distribution of eigenvectors). In general, D-loop sequence analysis shows a pattern reflecting migrations that occurred in postneolithic and historical times, with the most divergent mtDNA lineage occurring in the southern breeds, as shown in Figure 2 and Table 5. PCA on

SNP data differentiated breeds with good correspondence to geographical locations. It is interesting that the correlation between genetic and geographic distances revealed using nuclear SNPs was not confirmed by mtDNA, for which Mantel test was not significant. Our results seem to indicate a better correlation between geographic distances and autosomal markers.

APPENDIX

Members of the Econogene Consortium

Abo-Shehada Mahamoud, Parasitology Resesarch Laboratory, Faculty of Veterinary Medicine, Jordan University of Science and Technology, Ibid, Jordan; Ajmone Marsan Paolo, Istituto di Zootecnica, Università Cattolica del S. Cuore, Piacenza, Italy; Al Tarrayrah Jamil, Department of Animal Production, Aristotle Faculty of Agriculture, University of Thessaloniki, Thessaloniki, Greece; Angiolillo Antonella, Dipartimento di Scienze Animali Vegetali e dell'Ambiente, Università del Molise, Campobasso, Italy; Baret Philip, Department of Animal Genetics, Catholic University of Louvain, Louvain-la-Neuve, Belgium; Baumung Roswitha, University of Natural Resources and Applied Life Sciences, Vienna, Vienna, Austria; Beja-Pereira Albano, Laboratoire de Biologie des Populations d'Altitude, Université Joseph Fourier et Centre National de la Recherche Scientifique, Grenoble, France; Bertaglia Marco, Department of Food Economics and Consumption Studies, University of Kiel, Kiel, Germany; Bordonaro Salvatore, Dipartimento di Scienze Agronomiche e delle Produzioni Animali, Università degli Studi di Catania, Catania, Italy; Bruford Mike, School of Biosciences, Cardiff University, Cardiff, UK; Caloz Régis, Laboratoire de Systémes d'Information Géographique (LASIG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; Canali Gabriele, Istituto di Economia Agro-Alimentare, Università Cattolica del S. Cuore, Piacenza, Italy; Canon Javier, Universidad Computense de Madrid, Madrid, Spain; Cappuccio Irene, Dipartimento di Produzioni Animali, Università della Tuscia, Viterbo, Italy; Carta Antonello, Istituto Zootecnico e Caseario per la Sardegna, Sassari, Italy; Cicogna Mario, Istituto di Zootecnia Generale, Università degli Studi di Milano, Milano; Italy, Crepaldi Paola, Istituto di Zootecnia Generale, Università degli Studi di Milano, Milano, Italy; Dalamitra Stella, Biological Sciences, University of East Anglia, Norwich, UK; Daniela Krugmann, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Dobi Petrit, Department of Animal Production, Faculty of Agriculture, Tirana, Albania; Dominik Popielarczyk, Department of Sheep and Goat Breeding, Warsaw Agricultural University, Warsaw, Poland; Dunner Susana, Universidad Computense de Madrid, Madrid, Spain; D'Urso Giuseppe, Dipartimento di Scienze Agronomiche e delle Produzioni Animali, Università degli Studi di Catania, Catania, Italy; El-Barody M. A. A., Animal Production Department, Faculty of Agriculture, Minia University, Minia, Egypt; England Phillip, Laboratoire de Biologie des Populations d'Altitude, Université Joseph Fourier et CNRS, Grenoble, France; Erhardt Georg, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Ertugrul Okan, Veterinary Faculty, Ankara University, Ankara, Turkey; Eva-Maria Prinzenberg, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Eveline Ibeagha-Awemu, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Ewa Strzelec, Department of Sheep and Goat Breeding, Warsaw Agricultural University, Warsaw, Poland; Fadlaoui Aziz, Department of Animal Genetics, Catholic University of Louvain, Louvain-la-Neuve, Belgium; Fornarelli Francesca, Istituto di Zootecnia Generale, Università degli Studi di Milano, Milano, Italy; Garcia David, Universidad Computense de Madrid, Madrid, Spain; Georgoudis Andreas, Department of Animal Production, Faculty of Agriculture, Aristotle University of Thessaloniki, Thessaloniki, Greece; Gesine Lühken, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Giovenzana Stefano, Istituto di Zootecnia Generale, Università degli Studi di Milano, Milano, Italy; Gutscher Katja, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Hewitt Godfrey, Biological Sciences, University of East Anglia, Norwich, UK; Hoda Anila, Department of Animal Production, Faculty of Agriculture,

Tirana, Albania; Horst Brandt, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Istvan Anton, Research Institute for Animal Breeding and Nutrition, Department of Genetics, Herceghalom, Hungary; Juma Gabriela, School of Biosciences, Cardiff University, Cardiff, UK; Joost Stéphane, Laboratoire de Systèmes d'Information Géographique (LASIG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; Jones Sam, The Sheep Trust, Biology Area-8, University of York, York, England; Karetsou Katerina, Department of Animal Production, Faculty of Agriculture, Aristotle University of Thessaloniki, Thessaloniki, Greece; Kliambas Georgios, Department of Animal Production, Faculty of Agriculture, Aristotle University of Thessaloniki, Thessaloniki, Greece; Koban Evren, Department of Biology, Middle East Technical University, Ankara, Turkey; Kutita Olga, Department of Animal Production, Faculty of Agriculture, Aristotle University of Thessaloniki, Thessaloniki, Greece; Lazlo Fesus, Research Institute for Animal Breeding and Nutrition, Department of Genetics, Herceghalom, Hungary; Lenstra Johannes A., Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; Ligda Christina, National Agricultural Research Foundation, Thessaloniki, Greece; Lipsky Shirin, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Luikart Gordon, Laboratoire de Biologie des Populations d'Altitude, Université Joseph Fourier et CNRS, Grenoble, France; Marie-Louise Glowatzki, Institut für Genetik Ernährung und Haltung von Haustieren, Universität Bern, Bern, Switzerland; Marilli Marta, Istituto di Zootecnia Generale, Università degli Studi di Milano, Milano, Italy; Marletta Donata, Dipartimento di Scienze Agronomiche e delle Produzioni Animali, Università degli Studi di Catania, Catania, Italy; Milanesi Elisabetta, Istituto di Zootecnica, Università Cattolica del S. Cuore, Piacenza, Italy; Negrini Riccardo, Istituto di Zootecnica, Università Cattolica del S. Cuore, Piacenza, Italy; Nijman, Isaäc J., Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; Obexer-Ruff Gabriela, Institut für Genetik Ernährung und Haltung von Haustieren, Universität Bern, Bern, Switzerland; Papachristoforou Christos, Agricultural Research Institute, Nicosia, Cyprus; Pariset Lorraine, Dipartimento di Produzioni Animali, Università della Tuscia, Viterbo, Italy; Pellecchia Marco, Istituto di Zootecnica, Università Cattolica del S. Cuore, Piacenza, Italy; Peter Christina, Institut für Tierzucht und Haustiergenetik, Justus Liebig-Universität of Gießen, Gießen, Germany; Perez Trinidad, School of Biosciences, Cardiff University, Cardiff, UK; Pietrolà Emilio, Dipartimento di Scienze Animali, Vegetali e dell'Ambiente, Università del Molise, Campobasso, Italy; Pilla Fabio, Dipartimento di Scienze Animali, Vegetali e dell'Ambiente, Università del Molise, Campobasso, Italy; Roman Niznikowski, Department of Sheep and Goat Breeding, Warsaw Agricultural University, Warsaw, Poland; Roosen Jutta, Department of Food Economics and Consumption Studies, University of Kiel, Kiel, Germany; Scarpa Riccardo, Environment Department, University of York, York, UK; Sechi Tiziana, Istituto Zootecnico e Caseario per la Sardegna, Sassari, Italy; Taberlet Pierre, Laboratoire d'Ecologie Alpine (LECA), Université Joseph Fourier, Grenoble, France; Taylor Martin, Biological Sciences, University of East Anglia, Norwich, UK; Togan Inci, Department of Biology, Middle East Technical University, Ankara, Turkey; Trommetter Michel, Institut de recherché Agronomique, Unité d'Économie et Sociologie Rurales, Grenoble, France; Valentini Alessio, Dipartimento di Produzioni Animali, Università della Tuscia, Viterbo, Italy; Van Cann, Lisette M., Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; Vlaic Augustin, Department of Animal Genetics, Faculty of Zootechnics, University of Cluj-Napoca, Cluj-Napoca, Romania; Wiskin Louise, Biological Sciences, University of East Anglia, Norwich, UK; Zundel Stephanie, Laboratoire d'Ecologie Alpine (LECA), Université Joseph Fourier, Grenoble, France.

ACKNOWLEDGMENTS

This work has been supported by a PRIN 2007 (prot. 2007XRBAEN_003) of Italian Ministry of University and Research. Part of the sampling was operated in the frame of the EU ECONOGENE project (QLK5-CT2001-02461). The content of the publication does not necessarily represent the views of the commission or its services. The authors wish to thank Gabriella Porcai and Paolo Ciorba for technical assistance.

REFERENCES

[1] J. D. Vigne, I. Carrère, and J. Guilaine, "Unstable status of early domestic ungulates in the Near East: the example of Shillourokambos (Cyprus, IX-VIIIth millennia Cal. B.C.)," *Bulletin de Correspondance Hellenique*, vol. 43, supplement, pp. 239–251, 2003.

[2] C. F. Nadler, K. V. Korobitsina, R. S. Hoffmann, and N. N. Vorontsov, "Cytogenetic differentiation, geographic distribution, and domestication in Palaearctic sheep (*Ovis*)," *Zeitschrift für Säugetierkunde*, vol. 38, pp. 109–125, 1973.

[3] T. D. Bunch, W. C. Foote, and J. J. Spillett, "Translocations of acrocentric chromosomes and their implications in the evolution of sheep (*Ovis*)," *Cytogenetics and Cell Genetics*, vol. 17, no. 3, pp. 122–136, 1976.

[4] K. Dobney and G. Larson, "Genetics and animal domestication: new windows on an elusive process," *Journal of Zoology*, vol. 269, no. 2, pp. 261–271, 2006.

[5] M. A. Zeder, "Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 33, pp. 11597–11604, 2008.

[6] J. Zilhão, "Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 24, pp. 14180–14185, 2001.

[7] M. Tapio, N. Marzanov, M. Ozerov et al., "Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas," *Molecular Biology and Evolution*, vol. 23, no. 9, pp. 1776–1783, 2006.

[8] F. Pereira, S. J. M. Davis, L. Pereira, B. McEvoy, D. G. Bradley, and A. Amorim, "Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry," *Molecular Biology and Evolution*, vol. 23, no. 7, pp. 1420–1426, 2006.

[9] F. Pereira, L. Pereira, B. Van Asch, D. G. Bradley, and A. Amorim, "The mtDNA catalogue of all Portuguese autochthonous goat (*Capra hircus*) breeds: high diversity of female lineages at the western fringe of European distribution," *Molecular Ecology*, vol. 14, no. 8, pp. 2313–2318, 2005.

[10] L. Pariset, A. Cuteri, C. Ligda, P. Ajmone-Marsan, and A. Valentini, "Geographical patterning of sixteen goat breeds from Italy, Albania and Greece assessed by Single Nucleotide Polymorphisms," *BMC Ecology*, vol. 9, article 20, 2009.

[11] T. Cymbron, A. R. Freeman, M. I. Malheiro, J. D. Vigne, and D. G. Bradley, "Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations," *Proceedings of the Royal Society B*, vol. 272, no. 1574, pp. 1837–1843, 2005.

[12] M. Pellecchia, R. Negrini, L. Colli et al., "The mystery of Etruscan origins: novel clues from *Bos taurus* mitochondrial DNA," *Proceedings of the Royal Society B*, vol. 274, no. 1614, pp. 1175–1179, 2007.

[13] J. R. S. Meadows, O. Hanotte, C. Drögemüller et al., "Globally dispersed Y chromosomal haplotypes in wild and domestic sheep," *Animal Genetics*, vol. 37, no. 5, pp. 444–453, 2006.

[14] J. R. S. Meadows and J. W. Kijas, "Re-sequencing regions of the ovine Y chromosome in domestic and wild sheep reveals novel paternal haplotypes," *Animal Genetics*, vol. 40, no. 1, pp. 119–123, 2009.

[15] L. J. Lawson Handley, K. Byrne, F. Santucci et al., "Genetic structure of European sheep breeds," *Heredity*, vol. 99, no. 6, pp. 620–631, 2007.

[16] C. Peter, M. Bruford, T. Perez, S. Dalamitra, G. Hewitt, and G. Erhardt, "Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds," *Animal Genetics*, vol. 38, no. 1, pp. 37–44, 2007.

[17] J. W. Kijas, D. Townley, B. P. Dalrymple et al., "A genome wide survey of SNP variation reveals the genetic structure of sheep breeds," *PLoS ONE*, vol. 4, no. 3, Article ID e4668, 2009.

[18] B. Chessa, F. Pereira, F. Arnaud et al., "Revealing the history of sheep domestication using retrovirus integrations," *Science*, vol. 324, no. 5926, pp. 532–536, 2009.

[19] S. Y. Chen, Z. Y. Duan, T. Sha, J. Xiangyu, S. F. Wu, and Y. P. Zhang, "Origin, genetic diversity, and population structure of Chinese domestic sheep," *Gene*, vol. 376, no. 2, pp. 216–223, 2006.

[20] S. Pedrosa, M. Uzun, J. J. Arranz, B. Gutiérrez-Gil, F. San Primitivo, and Y. Bayón, "Evidence of three maternal lineages in near eastern sheep supporting multiple domestication events," *Proceedings of the Royal Society B*, vol. 272, no. 1577, pp. 2211–2217, 2005.

[21] J. R. S. Meadows, K. Li, J. Kantanen et al., "Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe," *Journal of Heredity*, vol. 96, no. 5, pp. 494–501, 2005.

[22] J. R. S. Meadows, I. Cemal, O. Karaca, E. Gootwine, and J. W. Kijas, "Five ovine mitochondrial lineages identified from sheep breeds of the near east," *Genetics*, vol. 175, no. 3, pp. 1371–1379, 2007.
[23] N. J. Wood and S. H. Phua, "Variation in the control region sequence of the sheep mitochondrial genome," *Animal Genetics*, vol. 27, no. 1, pp. 25–33, 1996.
[24] S. Hiendleder, H. Lewalski, R. Wassmuth, and A. Janke, "The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype," *Journal of Molecular Evolution*, vol. 47, no. 4, pp. 441–448, 1998.
[25] S. Hiendleder, B. Kaupe, R. Wassmuth, and A. Janke, "Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies," *Proceedings of the Royal Society B*, vol. 269, no. 1494, pp. 893–904, 2002.
[26] J. Guo, L. X. Du, Y. H. Ma et al., "A novel maternal lineage revealed in sheep (*Ovis aries*)," *Animal Genetics*, vol. 36, no. 4, pp. 331–336, 2005.
[27] J. Flanders, G. Jones, P. Benda et al., "Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum*: contrasting results from mitochondrial and microsatellite data," *Molecular Ecology*, vol. 18, no. 2, pp. 306–318, 2009.
[28] G. Heckel, R. Burri, S. Fink, J. F. Desmet, and L. Excoffier, "Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*," *Evolution*, vol. 59, no. 10, pp. 2231–2242, 2005.
[29] J. Bryja, C. Smith, A. Konečný, and M. Reichard, "Range-wide population genetic structure of the European bitterling (*Rhodeus amarus*) based on microsatellite and mitochondrial DNA analysis," *Molecular Ecology*, vol. 19, no. 21, pp. 4708–4722, 2010.
[30] M. M. Tollefsrud, J. H. Sønstebø, C. Brochmann, O. Johnsen, T. Skrøppa, and G. G. Vendramin, "Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European Picea abies," *Heredity*, vol. 102, no. 6, pp. 549–562, 2009.
[31] J. W. O. Ballard and M. C. Whitlock, "The incomplete natural history of mitochondria," *Molecular Ecology*, vol. 13, no. 4, pp. 729–744, 2004.
[32] G. L. Gonçalves, G. R. P. Moreira, T. R. O. Freitas, D. Hepp, D. T. Passos, and T. A. Weimer, "Mitochondrial and nuclear DNA analyses reveal population differentiation in Brazilian Creole sheep," *Animal Genetics*, vol. 41, no. 3, pp. 308–310, 2010.
[33] S. Hiendleder, K. Mainz, Y. Plante, and H. Lewalski, "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," *Journal of Heredity*, vol. 89, no. 2, pp. 113–120, 1998.
[34] T. A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symposium Series*, vol. 41, pp. 95–98, 1999.
[35] P. Librado and J. Rozas, "DnaSP v5: a software for comprehensive analysis of DNA polymorphism data," *Bioinformatics*, vol. 25, no. 11, pp. 1451–1452, 2009.
[36] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," *Molecular Biology and Evolution*. In press.
[37] L. Excoffier, G. Laval, and S. Schneider, "Arlequin ver. 3.0: an integrated software package for population genetics data analysis," *Evolutionary Bioinformatics Online*, vol. 1, pp. 47–50, 2005.
[38] S. Joost, L. Colli, P. V. Baret et al., "Integrating geo-referenced multiscale and multidisciplinary data for the management of biodiversity in livestock genetic resources," *Animal Genetics*, vol. 41, no. 1, pp. 47–63, 2010.
[39] L. Pariset, I. Cappuccio, P. Ajmone-Marsan et al., "Characterization of 37 breed-specific single-nucleotide polymorphisms in sheep," *Journal of Heredity*, vol. 97, no. 5, pp. 531–534, 2006.
[40] J. Goudet, "FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)," 2001, http://www2.unil.ch/popgen/softwares/fstat.htm.
[41] B. S. Weir and C. C. Cockerham, "Estimating F-statistics for the analysis of population structure," *Evolution*, vol. 38, no. 6, pp. 1358–1370, 1984.
[42] R. Peakall and P. E. Smouse, "GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research," *Molecular Ecology Notes*, vol. 6, no. 1, pp. 288–295, 2006.
[43] K. Belkhir, P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme, (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier, France.

[44] M. Nei, "Genetic distance between populations," *American Naturalist*, vol. 106, no. 949, p. 283, 1972.
[45] J. Reynolds, B. S. Weir, and C. C. Cockerham, "Estimation of the coancestry coefficient: basis for a short-term genetic distance," *Genetics*, vol. 105, no. 3, pp. 767–779, 1983.
[46] K. Liu and S. V. Muse, "PowerMaker: an integrated analysis environment for genetic maker analysis," *Bioinformatics*, vol. 21, no. 9, pp. 2128–2129, 2005.
[47] M. Nei, "Analysis of gene diversity in subdivided populations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, no. 12, pp. 3321–3323, 1973.
[48] M. Nei, *Molecular Evolutionary Genetics*, Columbia University Press, New York, NY, USA, 1987.
[49] M. Slatkin and R. R. Hudson, "Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations," *Genetics*, vol. 129, no. 2, pp. 555–562, 1991.
[50] J. Diamond, "Evolution, consequences and future of plant and animal domestication," *Nature*, vol. 418, no. 6898, pp. 700–707, 2002.
[51] C. Berthouly, B. Bed'Hom, M. Tixier-Boichard et al., "Using molecular and multivariate methods to study the genetic diversity on local European and Asian chickens breeds," *Animal Genetics*, vol. 39, pp. 121–129, 2008.
[52] S. Naderi, H. R. Rezaei, P. Taberlet et al., "Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity," *PLoS ONE*, vol. 2, no. 10, Article ID e1012, 2007.
[53] M. W. Bruford, D. G. Bradley, and G. Luikart, "DNA markers reveal the complexity of livestock domestication," *Nature Reviews Genetics*, vol. 4, no. 11, pp. 900–910, 2003.
[54] G. M. Hewitt, "Genetic consequences of climatic oscillations in the Quaternary," *Philosophical Transactions of the Royal Society B*, vol. 359, no. 1442, pp. 183–195, 2004.
[55] V. C. Pardeshi, N. Y. Kadoo, M. N. Sainani, J. R. S. Meadows, J. W. Kijas, and V. S. Gupta, "Mitochondrial haplotypes reveal a strong genetic structure for three Indian sheep breeds," *Animal Genetics*, vol. 38, no. 5, pp. 460–466, 2007.
[56] M. Ćinkulov, Z. Popovski, K. Porcu et al., "Genetic diversity and structure of the West Balkan Pramenka sheep types as revealed by microsatellite and mitochondrial DNA analysis," *Journal of Animal Breeding and Genetics*, vol. 125, no. 6, pp. 417–426, 2008.
[57] I. Cappuccio, L. Pariset, P. Ajmone-Marsan et al., "Allele frequencies and diversity parameters of 27 single nucleotide polymorphisms within and across goat breeds," *Molecular Ecology Notes*, vol. 6, no. 4, pp. 992–997, 2006.
[58] L. Pariset, I. Cappuccio, S. Joost et al., "Characterization of single nucleotide polymorphisms (SNPs) in sheep and their variation as an evidence of selection," *Animal Genetics*, vol. 37, pp. 290–292, 2006.
[59] J. Cañón, D. García, M. A. García-Atance et al., "Geographical partitioning of goat diversity in Europe and the Middle East," *Animal Genetics*, vol. 37, no. 4, pp. 327–334, 2006.
[60] G. Luikart, L. Gielly, L. Excoffier, J. D. Vigne, J. Bouvet, and P. Taberlet, "Multiple maternal origins and weak phylogeographic structure in domestic goats," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 10, pp. 5927–5932, 2001.
[61] H. Fernández, S. Hughes, J. D. Vigne et al., "Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 42, pp. 15375–15379, 2006.

This article should be cited as follows:

Lorraine Pariset, Marco Mariotti, Maria Gargani, Stephane Joost, Riccardo Negrini, Trinidad Perez, Michael Bruford, Paolo Ajmone Marsan, and Alessio Valentini, "Genetic Diversity of Sheep Breeds from Albania, Greece, and Italy Assessed by Mitochondrial DNA and Nuclear Polymorphisms (SNPs)," *TheScientificWorldJOURNAL*, vol. 11, pp. 1641–1659, 2011.

3. Conclusion

After the domestication and during the colonization of Europe, part of cattle and sheep biodiversity was lost, due to genetic drift and bottlenecks that have acted during the migration and subsequent expansion of populations. Also, a series of secondary migrations have accompanied human migrations in history. Subsequently geographical isolation that limited gene flow between groups of animals and environmental and human-induced selection, resulted in populations with different characteristics, adapted to extreme environments, and specific capacities for the production of milk, meat or fiber. Over the last two centuries many of these populations have been gradually standardized for several morphological and productive traits, more subject to selective pressure and isolated from the reproductive point of view. This process led to the formation of the breeds we know today and the fragmentation of the genetic variability of indigenous populations.

Through the analysis of the molecular markers it possible to identify the genetic variability within and between modern breeds and this is crucial for decision-making on the overall theme of conservation.

In the article “Polymorphisms within the Toll-Like Receptor (TLR)-2, -4 and -6 Genes in Cattle”, nucleotide sequences of bovine TLR2, TLR4 and TLR6 genes were screened to identify novel SNPs that can be used in studies of cattle resistance to diseases. Toll-like Receptors play an important role in the recognition of components of pathogens (PAMPs, Pathogen Associated Molecular Patterns) and subsequent activation of the innate immune response, which then leads to development of adaptive immune responses. In total eight SNPs were identified, three in TLR2, three in TLR4 and two in TLR6 and were deposited in NCBI dbSNP. These three genes are very important because they could be involved in immune response against various bovine diseases. In fact, TLR2 and TLR6 are critical in the immune response against Gram positive bacteria, TLR4 against Gram negative bacteria and virus. The polymorphisms in TLRs may reduce the ability of the protein to recognise ligands, particularly a non-synonymous SNP identified in TLR6 gene is very interesting because it causes an amino acid substitution (Asp↔Asn).

Furthermore, the SNPs characterization was performed by analysing a conspicuous number of individuals from 16 European breeds, and the main statistics were calculated. Even if from our analysis the SNPs do not appear located in loci under selection, a deviation of three SNPs from Hardy Weinberg equilibrium was observed. We hypothesize that some of the polymorphisms were fixated many generations ago within breed and the coalescent model could not be powerful enough

to reveal selection events so far in the past. It would be interesting to apply a more powerful model to confirm the absence of selection in the SNPs and their suitableness as neutral markers.
Also, distance based phylogenetic analysis was used to describe the relationships between breeds regarding the investigated TLRs. The lowest distance values are observed between Charolais and Asturiana de los Valles (0.002), while the highest distance is observed between Highlands and Jersey (0.117). Furthermore, the Jersey breed results very distant from all the other breeds of Great Britain, confirming the results obtained by AFLP and suggesting isolation within the Jersey island as the major cause of distinctiveness (Negrini et al., 2007). Indeed, Nei distances show that the highest genetic diversity observed in the geographically isolated breeds: it is suggestive (Figure 2) that the breeds of Great Britain (Aberdeen Angus, Highlands, South Devon and Jersey), using the analysed polymorphisms, are distributed accordingly to their geographic provenience.

In the paper "Relationships between Podolic cattle breeds assessed by single nucleotide polymorphisms (SNPs) genotyping", the relationship among the analysed Podolic breeds was examined and it was checked whether they genetic state reflect their history. Podolic cattle include a very ancient group of breeds, considered to be straight descendant from the Auroch (*Bos primigenius*). Podolic breeds are present in various European areas, and many of them are seriously endangered of extinction. Hungarian Grey, Maremmana and Turkish Grey belong to the same Podolic group of cattle and show similar external conformation. These breeds recently underwent a similar demographic reduction. After World War II, in Hungarian Grey inbreeding was avoided by using a rotational mating scheme based on six local Hungarian Grey sires, two imported sires of the same breed and three Maremmana sires introduced during the early 1970s. All three of these breeds face risks for their future survival because of inbreeding, indiscriminate crossbreeding and substitution with cosmopolitan more productive breeds.
Ninety-nine SNPs, selected from a previously characterized panel of 701 SNPs in candidate genes for meat quality (Williams et al., 2009), were genotyped on individuals belonging to the three Podolic breeds as well as on individuals of two Italian beef breeds, Marchigiana and Piemontese, which have different selection and migration histories, and were chosen as example of unrelated breeds. The SNPs used in this study, being located in genes involved in meat traits, are not a random sample. However, they do not show to be under selection and are fairly scattered among all chromosomes (with exception of 23, 24 and 27).
Within population variance estimate (Fis) was positive in Turkish Grey breed, indicating a significant level of inbreeding. A value suggesting moderate inbreeding is also observed in Marchigiana. The other three breeds showed values compatible with random mating. Therefore we

can conclude that mating strategies used in Maremmana and Hungarian Grey during the recovery were appropriate, while a different mating strategy could be suggested at least in one of the Turkish Grey populations. The estimate of Nei's genetic distances show that Maremmana is more distant from Hungarian Grey (0.66) than from Turkish Grey (0.63), despite the recent admixture with the former breed. Maremmana is closer to Marchigiana (0.51) than to either Hungarian Grey or Turkish Grey. Maximum genetic distances were observed between Turkish Grey and Marchigiana (0.70) and Hungarian Grey and Piemontese (0.62). Data are supported also by pairwise Fst, showing that Maremmana is closer to Turkish Grey (0.107) than to Hungarian Grey (0.124). The high Fst values observed in Hungarian Grey may reflect human selection, a long-time isolation of the breed, or a limited number of founders. The last two hypotheses seem more compatible with the history of the breed.

From this analysis, despite their similar morphology, Hungarian Grey and Maremmana are clearly identified as genetically distinct breeds. This could be attributed to either a different origin of the breeds or a consequence of the recent history, that led to the selection and probably fixation of genes. The two breeds were found differentiated on the basis of their allelic frequencies in a previous study (Valentini et al. 2006), and this confirms also the morphological differences reported by Maròti-Agòts et al. (2005). As for the Turkish Grey, we observed interesting differences between the population raised by breeders of Enez district, where in situ conservation studies are practised, and belonging to the Bandirma district of Balikesir, where ex situ conservation studies are practised 400 km far away from the original grey cattle raising area, in the Agriculture Research farm of Ministry of Agriculture. Turkey is very close to cattle domestication centre (Edwards et al. 2007); therefore, a higher differentiation is expected because the time for drift and distance to domestication bottleneck and this could explain why one of the two population results genetically differentiated for any value of K tested. Besides being raised far from the original raising area of Grey cattle, the ex situ herds of Bandirma are subject to legal and financial limitation, i.e. farmers are not able to buy new members, and this leads to increment of inbreeding and to genetic drift. On the contrary, private farmers, raising the in situ herds for conservation purposes, avoid inbreeding because of more flexibility in management decisions. Moreover, the Enez herds were established 5 years after those of Bandirma, and this could represent another source of difference.

In conclusion, this study demonstrates that morphology and anecdotic accounts might be deceiving in describing a population, as it was in the case of Maremmana and Hungarian Grey. Molecular data are very suggestive, and the observed genetic differentiation of the breeds may deserve more investigation. Further analysis could help in tracing an unbiased picture of past events and provide

the basis for a rational exploitation of livestock, suggesting appropriate cross-breeding plans based on genetic distance, or breeding strategies that include the population structure.

In the paper "Genetic Diversity of Sheep Breeds from Albania, Greece, and Italy Assessed by Mitochondrial DNA and Nuclear Polymorphisms (SNPs)", we employed mtDNA and nuclear SNPs to investigate the genetic diversity of sheep breeds of three countries of the Mediterranean basin: Albania, Greece, and Italy. 154 mtDNA haplotypes were detected by means of D-loop sequence analysis in 313 animals from 18 sheep breeds and relatively high haplotype diversity was found in all three sampled geographic regions. Phylogenetic analysis revealed haplogroups A, B and C in Albanian and Greek samples, while Italian individuals clustered in haplogroups A and B.
Furthermore, the same sheep sequenced at D-loop were genotyped with 37 previously described SNPs (Pariset et al., 2006). After removing those found monomorphic in the selected breeds, 27 SNPs were used for the analysis.
The nature of the markers used for the analysis can affect the detection of geographical structuring, as suggested by Naderi et al., 2007. In fact, mtDNA informativeness is limited because it does not detect male-mediated gene flow and does not predict the nuclear genomic diversity (Bruford et al., 2003). By combining markers with different modes of inheritance and rates of evolution this bias can be minimized (Hewitt 2004). Our mtDNA analysis shows higher levels of sheep nucleotide diversity in the South-East, which is congruent with data reported in the literature (Meadows et al., 2007) and congruent with the proximity to the domestication centre. This is confirmed by eigenvector analysis, which showed high contribution to variance by Albanian and Greek breeds, even if four breeds show unexpectedly low diversity. However, this behaviour can be explained by recent isolation or selection for some traits that reduced the overall genetic diversity through bottleneck. The SNP analysis revealed a rare allele frequency <5% for LEP1 and LEP2 loci, in agreement with those observed on a different European breed panel by Pariset et al., 2006. Observed and expected mean heterozygosity also showed similar values to those reported in the same paper. Expected heterozygosity values, which can indicate response to selection, are higher than observed heterozygosity values (Hs 0.063, 0.07, and 0.042; Ho 0.052, 0.06, and 0.038 in CALPA, PRNP-1, and GDF8, resp.). Among the breeds tested, Altamurana showed the highest Fis value suggesting the inbreeding in this population. Regarding the phylogeographic structure we found that the 95.2% of variation occurred within breeds indicating the weak phylogeographic structure in sheep. These data are consistent with those previously published by Kijas et al., 2009 using a different SNP panel. Sheep generally do not have a strong geographic structure and show a high genetic variability within breeds.

In the PCA, the breeds appear differentiated with 48.9% of the variance explained by the first three principal components. Also this analysis shows a good correspondence to geographical locations: the breeds remaining separated by the main group are all Italian. PCA supports therefore a westward route to Italy that could indicate that transport of animals made by sea as already proposed for cattle (Pellecchia et al., 2007), and goats (Cañón at al., 2006). This is particularly plausible because small sized species as sheep are easy to transport during migration and commercial trade and can adapt easily to various environments. A similar decrease in genetic diversity as well as an increase in the level of differentiation at the breed level from South-East to North-West in European sheep breeds, supporting the hypothesis of livestock migration from the Middle East towards western and northern Europe, was found by Peter et al., 2007, and Lawson Handley et al., 2007, using other nuclear markers.

To finish, D-loop sequence analysis shows a pattern reflecting migrations that occurred in post Neolithic and historical times, with the most divergent mtDNA lineage occurring in the southern breeds. PCA on SNP data differentiated breeds with good correspondence to geographical locations. It is interesting that the correlation between genetic and geographic distances revealed using nuclear SNPs was not confirmed by mtDNA, for which Mantel test was not significant. Our results seem to indicate a better correlation between geographic distances and autosomal markers.

To conclude, the study of molecular markers in livestock is very important for the future. In fact, the biodiversity conservation of farm animals is universally recognized as a challenge for the next years. The genetic diversity is a key resource to allow adaptation of the species to rapidly changing climatic conditions, for the survival of the species to the attack of new pathogens, to address the improvement genetic selection towards new goals and ultimately to ensure adequate food for future generations. The local breeds, in addition to being an invaluable reservoir of genetic variability, also play a social-ecological and cultural role relevant to land conservation, rural communities and their traditions

4. References

1. Achilli A, Bonfiglio S, Olivieri A, Malusa` A, Pala M, et al. (2009) The Multifaceted Origin of Taurine Cattle Reflected by the Mitochondrial Genome. PLoS ONE 4(6): e5753. doi:10.1371/journal.pone.0005753

2. Bollongino R, Elsner J, Vigne J-D, Burger J (2008) Y-SNPs Do Not Indicate Hybridisation between European Aurochs and Domestic Cattle. PLoS ONE 3(10): e3418. doi:10.1371/journal.pone.0003418

3. Bonfiglio S, Achilli A, Olivieri A, Negrini R, Colli L, et al. (2010) The Enigmatic Origin of Bovine mtDNA Haplogroup R: Sporadic Interbreeding or an Independent Event of Bos primigenius Domestication in Italy? PLoS ONE 5(12): e15760. doi:10.1371/journal.pone.0015760

4. Bruford MW, Bradley DG and Luikart G (2003) Genetic analysis reveals complexity of livestock domestication. Nature Rev. Genet. 4: 900-910.

5. Cañón, J., García, D., García-Atance, M.A., Obexer-Ruff, G., Lenstra, J.A., Ajmone-Marsan, P., andDunner, S. (2006) Geographical partitioning of goat diversity in Europe and the Middle East. Anim. Genet. **37**(4), 327-334.

6. Edwards C.J., Bollongino R., Scheu A., Chamberlain A., Tresset A., Vigne J.D., Baird J.F., Larson G., Ho S.Y., Heupink T.H., Shapiro B., Freeman A.R., Thomas M.G., Arbogast R.M., Arndt B., Bartosiewicz L., Benecke N., Budja M., Chaix L., Choyke A.M., Coqueugniot E., Do¨ hle H.J., Go¨ ldner H., Hartz S., Helmer D., Herzig B., Hongo H., Mashkour M., Ozdogan M., Pucher E., Roth G., Schade-Lindig S., Schmo¨ lcke U., Schulting R.J., Stephan E., Uerpmann H.P., Vo¨ ro¨ s I., Voytek B., Bradley D.G., Burger J. (2007) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. Proc. Biol. Sci., 274, 1377–1385.

7. 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 36:331-6

8. Hewitt G. M., "Genetic consequences of climatic oscillations in the Quaternary," Philosophical Transactions of the Royal Society B, vol. 359, no. 1442, pp. 183–195, 2004.

9. 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.

10. Hiendleder S, Mainz K, Plante Y, Lewalski H (1998a) 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 89:113-20

11. Hiendleder S., Lewalski H., Wassmuth R., and Janke A. (1998) The complete mitochondrial DNA sequence of the Domestic Sheep (Ovis aries) and comparison with the other major ovine haplotypes. Journal of Molecular Evolution 47 (4), 441-448.

12. Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J.F., McGrath, A., Wilson, P., Ingersoll, RG., McCulloch, R., McWilliam, S., Tang, D., McEwan, J., Cockett, N., Oddy, V.H., Nicholas, F.W., Raadsma, H., Barris, W., Bishop, S.S., Coltman, D., Crawford, A., Eggen, A., Erhardt, G., Forage, R., Hanotte, O., Hunt, P., Jianlin, H., Li, K., Marsan, P.A., Miller, J.E., Pemberton, J., and Schibler, L. (2009) A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS ONE **4(3)**, e4668.

13. Landegren U., Nilsson M. and Kwok P.Y. Reading bits of genetic information: methods for single nucleotide polymorphism analysis. *Genome Research* 1998, Vol.8:769-776;

14. Larson G., Ranran Liu, Xingbo Zhao, Jing Yuan, Dorian Fuller, Loukas Barton, Keith Dobney, Qipeng Fan, Zhiliang Gu, Xiao-Hui Liu, Yunbing Luo, Peng Lv, Leif Andersson, and Ning Li. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. PNAS 107: 7686-7691.

15. Lawson Handley, L.J., Byrne, K., Santucci, F., Townsend, S., Taylor, M., Bruford, M.W., and Hewitt, G.M. (2007) Genetic structure of European sheep breeds. Heredity **99**, 620-631

16. Maròti-Agòts A´ ., Bodò I., Valentini A., Verny M., Jàvorka L., Gera I. (2005) Comparison of Body Measurements of Hungarian Grey and Maremman Cattle Breeds. Proceedings of the 4th World Italian Beef Cattle Congress, Italy, April 29th–May 1st 2005, 223–228.

17. Meadows JRS, (2008). Re-sequencing regions of the ovine Y chromosome in domestic and wild sheep reveals novel paternal haplotypes. Anim Genet 40: 119–123.

18. Meadows JRS, Cemal I, Karaca O, Gootwine E, Kijas JW (2007). Five ovine mitochondrial lineages identified from sheep breeds of the near East. Genetics 175: 1371–1379.

19. Meadows JRS, Hiendleder S, and Kijas JW (2011). Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. Heredity **106**, 700-706 (April 2011) doi:10.1038/hdy.2010.122

20. Mullis K.B. and Faloona F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 1987, Vol.155:335-50.

21. Naderi, S., Rezaei, H.R., Taberlet, P., Zundel, S., Rafat, S.A., Naghash, H.R., El-Barody, M.A.A., Ertugrul, O., and Pompanon, F. (2007) Large-Scale Mitochondrial DNA Analysis of the Domestic Goat Reveals Six Haplogroups with High Diversity. PLoS ONE **2**(10), e1012.

22. Negrini, R.; Nijman, I.J.; Milanesi, E.; Moazami-Goudarzi, K.; Williams, J.L.; Erhardt, G.; Dunner, S.; Rodellar, C.; Valentini, A.; Bradley, D.G.; Olsaker, I.; Kantanen, J.; Ajmone-Marsan, P.; Lenstra J.A. European Cattle Genetic Diversity Consortium. Differentiation of European cattle by AFLP fingerprinting. Anim. Genet. 2007, 38, 60-66.

23. Pariset L., I. Cappuccio, P. Ajmone-Marsan et al., "Characterization of 37 breed-specific single-nucleotide polymorphisms in sheep," *Journal of Heredity*, vol. 97, no. 5, pp. 531–534, 2006.

24. Pedrosa S, Uzun M, Arranz JJ, Gutie´rrez-Gil B, San Primitivo F, Bayo´n Y. 2005. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. Proc R Soc Lond B Biol Sci 272:2211-7

25. Pellecchia M., R. Negrini, L. Colli et al., "The mystery of Etruscan origins: novel clues from Bos Taurus mitochondrial DNA," Proceedings of the Royal Society B, vol. 274, no. 1614, pp. 1175–1179, 2007.

26. Peter, C., Bruford, M.W., Perez, T., Dalamitra, S., Hewitt, G.M., Erhardt, G., and Econogene Consortium (2007) Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. Anim. Genet. **38**, 37-44.

27. Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. & Martin, W. Genetics and geography of wild cereal domestication in the near east. Nature Rev. Genet. 3, 429–441 (2002).

28. Shen R., Fan J.B., Campbell D., Chang W., Chen J., Doucet D., Yeakley J., Bibikova M., Wickham Garcia E., McBride C., Steemers F., Garcia F., Kermani B. G., Gunderson K. and Oliphant A. High throughput SNP genotyping on universal bead arrays. *Mutation Research* 2005, Vol.573:70–82.

29. Taberlet P., A. Valentini, H.R. Rezaei, S. Naderi, F. Pompanon, R. Negrini, P. Ajmone-Marsan, Are cattle, sheep, and goats endangered species? Mol. Ecol. 17 (2008) 275–284.

30. Taberlet P., E. Coisac, J. Pansu, F. Pompanon, Conservation genetics of cattle, sheep and goats. C. R. Biologies 334 (2011) 247-254.

31. Valentini A., Marchitelli C., Crisa` A., Pariset L., Bodo` I., Maròti-Agòts A., Nardone A. (2006) Stima della diversita` genetica nella popolazione bovina maremmana di Castelporziano e nei bovini Grigi Ungheresi. In: ''il sistema ambientale della tenuta presidenziale di Castelporziano. Ricerche sulla complessita` di un ecosistema forestale costiero mediterraneo, II serie, Segretariato generale della presidenza della repubblica italiana. Accademia Nazionale delle Scienze detta dei Quaranta ''Scritti e Documenti'' XXXVII.

32. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, et al. (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92: 16–24.

33. Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. and Zabeau M. AFLP: a new technique for DNA fingerprinting. *Nucleic Acid Research* 1995, Vol.23(21):4407-4414

34. Weber J.L. and May P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal Of Human Genetics* 1989, Vol.44(3):388-396.

35. Williams J.A., Scott I.M., Atkin A.L., Brook W.J., Russell M.A., Bell J.B. Genetic and molecular analysis of vgU and vgW: two dominant vg alleles associated with gene fusions in Drosophila. *Genetics* 1990, Vol.125: 833-844

36. Williams J.L., Dunner S., Valentini A., Mazza R., Amarger V., Checa M.L., Crisa‘ A., Razzaq N., Delourme D., Grandjean F., Marchitelli C., Garcıa D., Perez Gomez, R., Negrini R., Ajmone Marsan P., Leveziel H. (2009) Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. Anim. Genet., 40, 486–491.

37. Wood NS, Phua SH. 1996. Variation in the control region sequence of the sheep mitochondrial genome. Anim Genet 27:25-33