

Research article

Open Access

Geographical patterning of sixteen goat breeds from Italy, Albania and Greece assessed by Single Nucleotide Polymorphisms

Lorraine Pariset*¹, Antonella Cuteri¹, Christina Ligda², Paolo Ajmone-Marsan³, Alessio Valentini¹ and ECONOGENE Consortium⁴

Address: ¹Dipartimento di Produzioni Animali, Università degli Studi della Tuscia, Viterbo, Italy, ²National Agricultural Research Foundation, Thessaloniki, Greece, ³Istituto di Zootecnica, Università Cattolica del Sacro Cuore, Piacenza, Italy and ⁴ <http://www.econogene.eu>

Email: Lorraine Pariset* - pariset@unitus.it; Antonella Cuteri - anto.cut@libero.it; Christina Ligda - chligda@otenet.gr; Paolo Ajmone-Marsan - paolo.ajmone@unicatt.it; Alessio Valentini - alessio@unitus.it; ECONOGENE Consortium - paolo.ajmone@unicatt.it

* Corresponding author

Published: 2 September 2009

Received: 17 March 2009

BMC Ecology 2009, 9:20 doi:10.1186/1472-6785-9-20

Accepted: 2 September 2009

This article is available from: <http://www.biomedcentral.com/1472-6785/9/20>

© 2009 Pariset et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: SNP data of goats of three Mediterranean countries were used for population studies and reconstruction of geographical patterning. 496 individuals belonging to Italian, Albanian and Greek breeds were genotyped to assess the basic population parameters.

Results: A total of 26 SNPs were used, for a total of 12,896 genotypes assayed. Statistical analysis revealed that breeds are not so similar in terms of genetic variability, as reported in studies performed using different markers. The Mantel test showed a strongly significant correlation between genetic and geographic distance. Also, PCA analysis revealed that breeds are grouped according to geographical origin, with the exception of the Greek Skopelos breed.

Conclusion: Our data point out that the use of SNP markers to analyze a wider breed sample could help in understanding the recent evolutionary history of domestic goats. We found correlation between genetic diversity and geographic distance. Also PCA analysis shows that the breeds are well differentiated, with good correspondence to geographical locations, thus confirming the correlation between geographical and genetic distances. This suggests that migration history of the species played a pivotal role in the present-day structure of the breeds and a scenario in which coastal routes were easier for migrating in comparison with inland routes. A westward coastal route to Italy through Greece could have led to gene flow along the Northern Mediterranean.

Background

The domestic goat (*Capra hircus*) has been an important livestock resource since its domestication [1-3].

Even now, goats are of great importance in many developing countries, to exploit marginal agricultural resources, and in developed countries, for the production of high

quality products and the achievement of sustainable development of rural areas. However, breeding programs and selection schemes in goats are less advanced than in other livestock.

Archaeological evidence indicates a probable migration of the Neolithic farmers 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 [4-7] 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 [8]. In addition, genetic data revealed a weaker degree of phylogeographic structuring in domestic goats compared to other livestock species [9,10], which probably results from high gene flow at the intercontinental level, suggesting that goats have been extensively transported [2,11]. This led to initial settlement in the Balkans and southern Italy [12]. A decrease of genetic diversity likely occurred during this colonization process in Europe [13].

The Mediterranean Sea had a key role in the history of livestock also in post-Neolithic times, when civilisations like Phoenicians, Greeks, Romans and Berbers probably introduced new species of animals and new breeds of livestock in southwest Europe arriving by sea. Some colonists may have improved local livestock importing stock from overseas [14,15], explaining the unexpectedly high diversity in breeds of domestic goats [16], the differential cattle migration along the Mediterranean coast [14] and the close genetic relationship between Tuscan cattle breeds and Near Eastern breeds [17]. The role of the Mediterranean Sea as a natural corridor connecting the peninsula to the Near East, North Africa, and southern Europe is particularly plausible for domestic sheep and goat, species adaptable to various environments and easy to transport [15].

A relative lack of breed standardization, herdbook breeding, parentage control and rigorous management may have facilitated gene flow between geographically nearby breeds in south-eastern Europe [18]. Gene flow may also have a historical background as goats were actively traded all over the Mediterranean basin during the Phoenician, Greek and Roman periods. However, in some circumstances, gene flow is limited by distance and local management, which reduces the effective population size as a result of genetic isolation.

So far, studies on genetic diversity of goats have focused on Swiss [19] and Asian breeds [20-26]. Only a few were relative to Mediterranean breeds [27-29], but all of these employed mitochondrial markers. A wide-range analysis of goat diversity in Europe and the Middle East has been conducted using microsatellites [18]. The use of microsatellites, the most common method used today to estimate neutral genetic diversity [30,31], presents disadvantages such as null alleles, interpretation difficulty of allele calling and size homoplasy [32]. Single nucleotide polymorphisms (SNPs) are useful markers for estimating

parameters such as population history and inference of relationships [33,34] and they could potentially become the marker of choice in ecological and conservation studies [35,36]. The use of these markers is promoted also by a rapid development of genotyping techniques [32,37,38].

We applied SNP genotyping to the study of goat breeds from Italy, Albania and Greece to analyse genetic diversity of goats in these areas.

Results

Population Genetics

A total of 26 SNPs identified as polymorphic on a panel of European goat breeds [39] were used to genotype 496 individuals belonging to 6 Albanian, 2 Greek and 8 Italian breeds, for a total of 12,896 genotypes assayed. Loci were analysed to identify SNPs under selection in a separated study [40]. Analysis were then performed excluding the three SNPs in two loci identified as outliers, even if the use of all the 26 SNPs did not lead to substantial differences in the results.

Expected heterozygosity of the loci ranged from 0.016 (FABP4) to 0.495 (GHR), with a mean of 0.300. *F_{is}* per population ranged from 0.017 (Liquenasi) to 0.197 (Capore) (Table 1). Observed heterozygosity of the loci determined from SNP frequencies (Table 2) ranged from 0.012 (FABP4) to 0.463 (GHR), with a mean of 0.272.

Table 1: Breeds analysed, country of origin and their sample sizes (N), average observed heterozygosity (H_o), average expected heterozygosity (H_e) and F_{is} calculated per locus per sample and averaged over populations.

| Breed | label | Origin | N | H _o | H _e | F _{is} |
|------------------------------------|-------|---------|----|----------------|----------------|-----------------|
| Argentata dell' Etna ² | ARG | Italy | 31 | 0.280 | 0.322 | 0.131 |
| Bionda dell' Adamello ² | BIO | Italy | 31 | 0.318 | 0.340 | 0.066 |
| Camosciata | CAM | Italy | 31 | 0.273 | 0.330 | 0.173 |
| Capore | CAP | Albania | 31 | 0.221 | 0.275 | 0.197 |
| Dukati | DUK | Albania | 31 | 0.295 | 0.309 | 0.047 |
| Girgentana ^{1,2} | GIR | Italy | 32 | 0.274 | 0.286 | 0.041 |
| Greek goat | GRG | Greece | 31 | 0.293 | 0.327 | 0.104 |
| Grigia molisana ^{1,2} | GMO | Italy | 31 | 0.292 | 0.306 | 0.048 |
| Hasi | HAS | Albania | 31 | 0.272 | 0.307 | 0.115 |
| Liquenasi | LIQ | Albania | 31 | 0.296 | 0.301 | 0.017 |
| Mati | MAT | Albania | 30 | 0.246 | 0.273 | 0.100 |
| Muzhake | MUZ | Albania | 31 | 0.259 | 0.295 | 0.124 |
| Orobica ² | ORO | Italy | 31 | 0.287 | 0.295 | 0.030 |
| Sarda | SAR | Italy | 31 | 0.287 | 0.311 | 0.078 |
| Skopelos2 | SKO | Greece | 31 | 0.177 | 0.201 | 0.122 |
| Valdostana ^{1,2} | VAL | Italy | 31 | 0.293 | 0.308 | 0.047 |
| Mean | | | | 0.273 | 0.299 | 0.090 |

¹ Listed by FAO as endangered or critical <http://fao.org>. ² Number of breeding females below the EU thresholds (<10000). (Commission Regulation (EC), n. 445/2002).

Table 2: Locus, SNP, major allele frequency (m.a.f), observed heterozygosity (Ho), expected heterozygosity (He), Fst values and deviations from Hardy-Weinberg equilibrium (p-value) for each genotyped SNP.

| locus | SNP | m.a.f. | Ho | He | Fst | p-value |
|----------------------------------|-----------|--------|-------|-------|-------|---------|
| Activin receptor IIB | ACVR2B | 0.803 | 0.308 | 0.294 | 0.076 | 0.815 |
| Calpastatin | CAST | 0.948 | 0.096 | 0.087 | 0.124 | 0.742 |
| K-casein | CSN3 | 0.761 | 0.314 | 0.348 | 0.046 | 0.660 |
| CathepsinK | CTSK | 0.945 | 0.098 | 0.102 | 0.029 | 0.997 |
| Desmin | DES | 0.602 | 0.452 | 0.457 | 0.051 | 0.999 |
| MHC class II DQA gene | HLA-DQA_1 | 0.884 | 0.116 | 0.191 | 0.077 | 0.000* |
| MHC class II DQA gene | HLA-DQA_2 | 0.672 | 0.158 | 0.400 | 0.102 | 0.000* |
| MHC class II DRB gene | HLA-DRB | 0.838 | 0.226 | 0.258 | 0.057 | 0.077 |
| Fatty acid-binding protein 4 | FABP4 | 0.992 | 0.012 | 0.016 | 0.009 | 0.096 |
| Fibronectin | FN1 | 0.811 | 0.238 | 0.292 | 0.050 | 0.309 |
| Growth differentiation factor 9B | GDF9 | 0.808 | 0.215 | 0.298 | 0.043 | 0.004* |
| Growth hormone receptor | GHR | 0.514 | 0.463 | 0.495 | 0.011 | 0.794 |
| Interleukin-4 | IL4 | 0.508 | 0.440 | 0.432 | 0.144 | 0.932 |
| Interleukin-2 | IL2_1 | 0.983 | 0.033 | 0.032 | 0.008 | 1.000 |
| Interleukin-2 | IL2_2 | 0.839 | 0.243 | 0.246 | 0.096 | 0.964 |
| Integrin B1 | ITGB1 | 0.637 | 0.422 | 0.452 | 0.025 | 0.947 |
| Beta-lactoglobulin | LGB | 0.757 | 0.338 | 0.341 | 0.078 | 0.998 |
| Melatonin | MTNR1A | 0.558 | 0.453 | 0.451 | 0.093 | 0.865 |
| Myostatin | GDF8 | 0.942 | 0.098 | 0.106 | 0.031 | 0.940 |
| Prion protein | PRNP_1 | 0.624 | 0.422 | 0.439 | 0.070 | 0.965 |
| Prion protein | PRNP_2 | 0.649 | 0.408 | 0.431 | 0.058 | 0.641 |
| Toll-like receptor 4 | TLR4 | 0.571 | 0.441 | 0.475 | 0.033 | 0.231 |
| microsatellite | U80 | 0.838 | 0.254 | 0.262 | 0.037 | 0.996 |
| | Mean | 0.706 | 0.272 | 0.300 | 0.059 | |

*: Highly significant

The frequencies of the major alleles are reported in table 2 and ranged from 0.508 for the locus IL4 to 0.992 for the locus FABP4. Except for IL2-1 and FABP4, showing frequencies of the rare alleles of 0.017 and 0.008, respectively, all other SNPs have a frequency of the rare allele greater than 5%, as observed for the same loci on a different European breeds panel [39]. Also mean observed and expected heterozygosity showed similar values to those reported in the same paper. The frequencies of the major alleles per SNP and per population are presented in Additional file 1. Some populations presented fixed alleles in a number of SNPs. Particularly, Skopelos and Girgentana showed the highest number of fixed alleles (5). SNPs FABP4 and IL2_1 are fixed in 12 and 8 populations, respectively (Additional file 1).

Significant deviations from Hardy-Weinberg equilibrium over all populations (p -value < 0.005) were observed in three loci (Table 2). Significant deviations from Hardy-Weinberg equilibrium (HWE) for each locus and population were observed for five loci. Locus CSN3 was not in HWE in Muzake population. Locus HLA-DQA_1 was not in HWE in Hasi, Skopelos and Argentata dell'Etna populations. HLA-DQA_2 was not in HWE in seven populations, five of which are Italian: Capore, Argentata dell'Etna, Bionda dell'Adamello, Camosciata, Sarda, Valdostana. Locus HLA-DRB is not in HWE in two popula-

tions: Capore and Greek Goat. Locus TLR4 was not in HWE in Camosciata population only.

Significant deviations from linkage equilibrium (p -value < 0.01) were observed for two of the analysed locus pairs: HLA-DQA_2 and GDF9 loci, PRP_1 and PRP_2. Loci have not yet been mapped in the goat; HLA-DQA_2 and GDF9 are located on *Bos taurus* chromosome 23 and 7, respectively; the two SNPs on PRP are 458 bp apart on *B. taurus* chromosome 13 <http://www.ncbi.nlm.nih.gov>. The log-likelihood based test of genotypic differentiation for all populations showed statistically significant values (p -value < 0.01) except for GHR, IL2_1 and FABP4 loci which have very little informativeness.

The analysis revealed an overall F_{ST} of 0.063. That is, 6.3% of allelic variation was accounted across breeds and 93.7% within breeds. Weir and Cockerham's [41] estimate of F_{ST} per locus ranged from 0.008 (IL2_1) to 0.144 (IL4), with a mean of 0.059 (Table 2). Pairwise F_{ST} values are given in Table 3. F_{ST} estimates, between Skopelos and Valdostana (0.19), Orobica (0.19), Camosciata (0.22), Girgentana (0.18) e Argentata dell'Etna (0.15) populations, were higher than all other pairwise comparisons suggesting the Skopelos as the most diverse population.

Table 3: Pairwise F_{ST} below diagonal, Nei's standard distance (Nei 1972) above diagonal.

| | ARG | BIO | CAM | CAP | DUK | GIR | GMO | GRG | HAS | LIQ | MAT | MUZ | ORO | SAR | SKO | VAL |
|-----|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|-------|-------|-------|-------|-------|
| ARG | 0 | 0.033 | 0.080 | 0.036 | 0.075 | 0.063 | 0.045 | 0.038 | 0.043 | 0.035 | 0.061 | 0.037 | 0.103 | 0.018 | 0.167 | 0.085 |
| BIO | 0.013 | 0 | 0.067 | 0.065 | 0.078 | 0.063 | 0.067 | 0.045 | 0.062 | 0.058 | 0.067 | 0.050 | 0.083 | 0.029 | 0.156 | 0.056 |
| CAM | 0.058 | 0.044 | 0 | 0.132 | 0.127 | 0.094 | 0.115 | 0.068 | 0.097 | 0.090 | 0.136 | 0.116 | 0.138 | 0.050 | 0.231 | 0.114 |
| CAP | 0.016 | 0.043 | 0.110 | 0 | 0.068 | 0.064 | 0.052 | 0.027 | 0.032 | 0.033 | 0.036 | 0.023 | 0.116 | 0.045 | 0.106 | 0.107 |
| DUK | 0.056 | 0.059 | 0.108 | 0.048 | 0 | 0.118 | 0.061 | 0.068 | 0.053 | 0.039 | 0.036 | 0.035 | 0.129 | 0.077 | 0.114 | 0.127 |
| GIR | 0.045 | 0.044 | 0.074 | 0.044 | 0.102 | 0 | 0.106 | 0.056 | 0.098 | 0.088 | 0.097 | 0.093 | 0.156 | 0.046 | 0.193 | 0.104 |
| GMO | 0.026 | 0.047 | 0.094 | 0.032 | 0.042 | 0.089 | 0 | 0.051 | 0.060 | 0.027 | 0.071 | 0.054 | 0.127 | 0.053 | 0.106 | 0.117 |
| GRG | 0.019 | 0.025 | 0.046 | 0.006 | 0.050 | 0.037 | 0.032 | 0 | 0.031 | 0.038 | 0.048 | 0.037 | 0.093 | 0.027 | 0.116 | 0.074 |
| HAS | 0.024 | 0.042 | 0.077 | 0.011 | 0.034 | 0.081 | 0.041 | 0.011 | 0 | 0.038 | 0.036 | 0.028 | 0.091 | 0.051 | 0.139 | 0.106 |
| LIQ | 0.016 | 0.038 | 0.068 | 0.012 | 0.020 | 0.071 | 0.009 | 0.019 | 0.018 | 0 | 0.041 | 0.020 | 0.096 | 0.037 | 0.101 | 0.110 |
| MAT | 0.043 | 0.047 | 0.117 | 0.016 | 0.017 | 0.080 | 0.053 | 0.029 | 0.017 | 0.022 | 0 | 0.017 | 0.099 | 0.067 | 0.090 | 0.105 |
| MUZ | 0.018 | 0.029 | 0.096 | 0.002 | 0.015 | 0.075 | 0.035 | 0.018 | 0.008 | 0.001 | -0.003 | 0 | 0.094 | 0.045 | 0.090 | 0.084 |
| ORO | 0.086 | 0.064 | 0.120 | 0.097 | 0.112 | 0.140 | 0.111 | 0.075 | 0.073 | 0.079 | 0.081 | 0.076 | 0 | 0.114 | 0.204 | 0.145 |
| SAR | -0.002 | 0.007 | 0.027 | 0.023 | 0.058 | 0.027 | 0.034 | 0.007 | 0.031 | 0.018 | 0.047 | 0.025 | 0.096 | 0 | 0.139 | 0.086 |
| SKO | 0.150 | 0.142 | 0.221 | 0.088 | 0.098 | 0.180 | 0.090 | 0.100 | 0.123 | 0.085 | 0.073 | 0.072 | 0.189 | 0.121 | 0 | 0.199 |
| VAL | 0.065 | 0.034 | 0.091 | 0.086 | 0.110 | 0.086 | 0.099 | 0.054 | 0.086 | 0.091 | 0.087 | 0.065 | 0.128 | 0.065 | 0.187 | 0 |

The Mantel test showed a strongly significant correlation between genetic and geographic distances (0.40, p-value < 0.001, over 1000 permutations).

PCA analysis

Genetic relationships were also explored by means of principal component analysis. The coefficients of the linear combinations reveal which SNPs most affect the component value. As for the first component, SNP IL4 presents extreme positive and SNP LGB extreme negative values, respectively. Likewise, the second component is mostly affected by the SNPs ACVR2B and MTNR1A, with positive sign, and by the SNPs HLA-DQA_2, IL4 and LGB with negative sign. To examine the overall pattern of population differentiation, PCA was conducted with the first two axes, which cumulatively explained 52% of the total inertia contained in the data set (Figure 1). Breeds are grouped according to geographical origin, with the exception of the Greek Skopelos breed.

Genetic distance

Distance-based phylogenetic analysis was used to describe the relationships between breeds. Table 3 presents the Nei [42] genetic distance relating the 16 goat breeds studied. The lowest distance values are observed between Muzhake and Mati (0.017) and Sarda and Argentata dell'Etna (0.018). The Greek Skopelos breed results the most distant one, very distant from all the 4 breeds of northern Italy and from Argentata dell'Etna and Girgentana.

Discussion

Archaeological evidence showed that two main colonization routes took place in Europe after the initial domestication events in the Fertile Crescent: the Mediterranean route and the Danubian route. Cañon et al. [18], using microsatellites, report a 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 goat breeds, supporting the hypothesis of migration of domestic livestock from the Middle East towards western and northern Europe.

Our results indicate that a highly significant correlation between genetic and geographic distance exists. The pres-

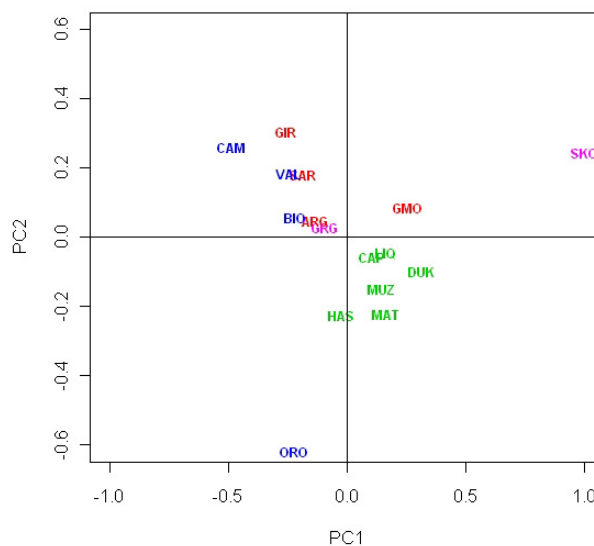


Figure 1
Principal component analysis (PCA) of allele frequencies from twenty three SNP loci genotyped in the sixteen goat breeds. Projection on axis 1 and axis 2, which cumulatively explained 52% of the total inertia contained in the data set. Breeds acronyms as in Table 1. Albanian breeds, green; North Italian breeds, blue; South Italian breeds, red; Greek breeds, pink.

ence of a geographic component in genetic diversity was already reported in breeds of Northern and Southern Italy in a previous study using SNPs [43] and it is confirmed here in a larger area. Such a geographic component is generally not observed when using mitochondrial markers. As reported by Luikart et al. [11] only 10% of the variance assessed by mtDNA is partitioned among continents. This could be due to the nature of the markers used for the analysis, as suggested by Naderi et al. [44]. In fact, mtDNA informativeness is limited because it does not detect male-mediated gene flow and does not predict the nuclear genomic diversity [45]. In the paper by Naderi et al. [44] the breeds cannot be distinguished on the basis of mtDNA, even if authors report that more than 77% of the mtDNA variation is found within breeds, while there is a low regional differentiations of haplotypes. At a regional scale, the lack of geographic structure has been reported using mtDNA in different regions [16,25,28] with the exception of one paper [24].

From PCA analysis the breeds appear well differentiated with 52% of the variance explained by the first two principal components. There is also a good correspondence to geographical locations, thus confirming the correlation between geographical and genetic distances identified by the Mantel test. PCA indicates a westward route to southern Italy through Greece, that may suggest contacts between Albania and Italian peninsula and between Greece and Italian Islands (Sardinia and Sicily). In post-Neolithic times, some colonists may have improved local livestock as well as importing stock from overseas. The transport of animals made by sea has been already proposed for cattle [14,17] and goats [15,18]. The role of the Mediterranean Sea as a natural corridor connecting the Italian peninsula to the Near East, North Africa, and southern Europe is particularly plausible for small sized species, as sheep and goats species adaptable to various environments and easy to transport during human migration and commercial trade [4,11,45].

The Greek Skopelos breed results the most distant one, reflecting the fact that it has been raised only in a island and on the mainland of Magnisia. The distance is not due to inbreeding as $F_{IS} = 0.122$, not the highest value in our breeds (max $F_{IS} = 0.197$ in Capore), but to the lack of admixture with other populations since long time and possibly a natural selection versus local environment. The Skopelos breed is largely differentiated from the other goat populations in Greece, both morphologically and in terms of performance. According to the inhabitants of the Skopelos island, the goat used to live in an uninhabited small island of Northern Sporades, and it was recently domesticated. The breed is also said to have some relationship with the wild goat of the Gioura island, originated from the homonymous island [46]. Breeders, due to

the favourable characteristics of the breed (high prolificacy and high milk production), established a breeders association and applied a genetic improvement programme since 1981. Also the Orobica is very far apart from the other breeds. Again the distance is not attributable to inbreeding ($F_{IS} = 0.030$), but to isolation of this breed in a very secluded valley of Italian Alps.

Among the Italian southern breeds, it is interesting that the lowest distance is seen between Argentata dell'Etna, from Sicily, and Sarda, original from Sardinia. The two islands, although quite far apart, were important trade posts of Phoenician, Punic and Roman traders.

The analysis carried so far excluded SNPs that were proven under selection [40]. If we include these SNPs (CSN1S1_1, CSN1S1_2 and LIPE, [39]) we find that the overall distance pattern remains unchanged but for two breeds of Northern Italy, Bionda and particularly Valdostana, that become closer to Orobica (Figure 2). This is due to casein and LIPE allele frequency that are almost fixed in these breeds for the same allele, while the average allele frequency for the other breeds is 50% (Additional file 1). Caseins have been the first genes to be associated to milk production, characteristics and curding properties [47-49]. It is noteworthy that milk production and cheese making is a primary agricultural activity in North Italy

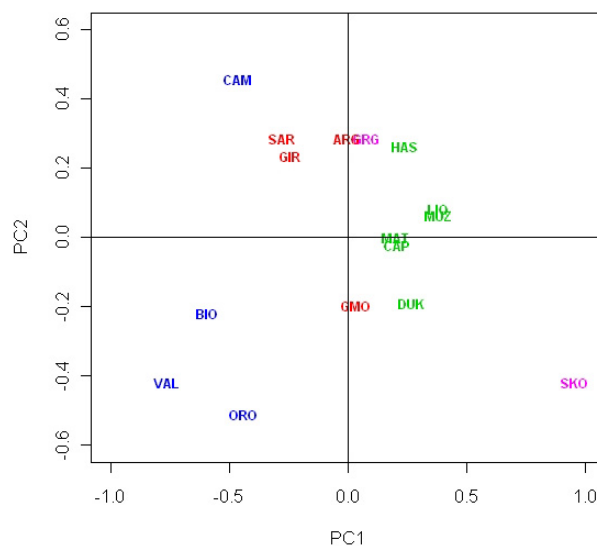


Figure 2
Principal component analysis (PCA) of allele frequencies from twenty six SNP loci genotyped in the sixteen goat breeds. Projection on axis 1 and axis 2, which cumulatively explained 56.8% of the total inertia contained in the data set. Breeds acronyms as in Table 1. Albanian breeds, green; North Italian breeds, blue; South Italian breeds, red; Greek breeds, pink.

since historical times as demonstrated by the high frequency of lactase persistence in humans [50]. We hypothesize that converged selection for caseins and LIPE (an enzyme important for cheese making as well [51]) occurred in Orobica, Valdostana and Bionda dell'Adamello breeds making them "similar" for what concerns their exploitation objectives.

Conclusion

Our data point out that the use of SNP markers to analyze a wider breed sample, until now scarcely employed for genetic population studies in livestock, could help in understanding the recent evolutionary history of domestic goats.

We found correlation between genetic diversity and geographic distance. Also PCA analysis shows that the breeds are well differentiated, with good correspondence to geographical locations, thus confirming the correlation between geographical and genetic distances. This suggests that migration history of the species played a pivotal role in the present-day structure of the breeds. Instead, the limited genetic similarity within main geographical areas suggests that breed differentiation could have occurred in more recent times, after the main migrations.

On the basis of the observed gradient of genetic diversity decreasing from south-east to north-west, and of the signals of the northward dispersal of populations from the

domestication centre, we hypothesize that coastal routes from the domestication centre to Italy through Greece could be a likely explanation for the observed gene flow along the Northern Mediterranean.

Methods

Material

Blood samples of a total of 496 goats, about one third male, were collected in farms spread over the traditional rearing area of each breed (Figure 3). No more than 3 unrelated individuals per flock, from an average of 10 farms per breed, were sampled to reduce the relationship among animals and to increase the breed representativeness. Samples were obtained following the rules of each of the countries involved in sampling. Wherever possible, we used part of samples taken by public veterinaries within national animal health plans. A total of 16 breeds were analysed. The breed names, their acronyms, countries of origin, and the sample sizes are given in Table 1. DNA, extracted by phenol-chlorophorm or commercial kits in the relative sampling laboratory, was tested for quality and concentration by electrophoresis on 0.8% agarose gel, stained with ethidium bromide and compared to a commercial standard.

SNP analysis

SNPs characterization has been described elsewhere [39]. SNP ascertainment bias was minimised by sequencing target DNA in at least 8 individuals from different popula-



Figure 3
Rearing area of each analysed breed.

tions [39]. Large scale genotyping of all animals was performed by outsourcing to a commercial genotyping company <http://www.Kbioscience.co.uk>. Generally, accuracy greater than 99% was achieved. Quality control criteria were adopted (water as negative control, inter plate duplicate testing of a known DNA, intra plate testing of a known DNA). All the SNPs described in [39] were genotyped on our samples. In this investigation 23 SNPs in 19 genes and in one microsatellite were used (Table 2), excluding SNPs showing a frequency >0.05 on our samples.

Data analysis

Allele frequencies were calculated using *FSTAT* 2.93[52]. Observed and expected heterozygosities (H_O and H_E , respectively), Weir and Cockerham's [41] estimate of F_{is} per population, of F_{ST} per locus and population pairs were calculated for each locus using *GENEPOP* 4.0 [53]. The same software was used to test deviations from Hardy-Weinberg equilibrium (HWE) for each locus and population and for locus over all populations using a Markov chain of 100 000 steps and 1000 dememorization steps and to assess deviations from genotypic linkage disequilibrium (LD) for each pair of loci using the same Markov chain parameters as for the HWE test. We performed the Probability-test (the "exact HW test").

Isolation by distance (IBD, [54]) has been assessed by plotting the genetic distance among population pairs as a function of the geographic distance between those pairs to check whether more distant population pairs are more different genetically, providing information on the genetic structure of the populations and [55]. We applied the Mantel test [56,57] implemented in the *ade4* library of *R* 2.6.0 open source software (publicly available at <http://www.r-project.org>) to estimate the correlation between pairwise F_{ST} values and pairwise geographic distance using 1000 replicates to test significance. The matrix of the geographic distance was computed using the coordinates obtained with a global positioning system (GPS).

A principal component analysis (PCA) was performed on the covariance matrix of SNP frequency data to investigate spatial patterns of genetic variation using the *R* 2.6.0 open source software.

Nei [42] genetic distances between populations pairs were calculated to obtain relative estimates of the time that has passed since the populations were established using *POWERMARKER*[58].

Authors' contributions

LP: designed the study, carried out SNP data analysis and drafted the manuscript. AC: completed most of the statistical analysis. CL: provided Greek samples and formulated

parts of the paper. PAM: leader of the Econogene project, provided DNA samples and guidance for the project. AV: participated in developing ideas, in supervision and revision of the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Table S1. Frequency of the Major allele per SNP and per population. Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6785-9-20-S1.doc>]

Acknowledgements

This work was partially supported by the ECONOGENE project, funded by the European Union (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 Flora Jane Dause for help with the language and three anonymous reviewers for helpful comments. Stephane Joost kindly provided figure 3.

Members of the ECONOGENE Consortium are:

Abo-Shehada Mahamoud, Ajmone Marsan Paolo, Al Tarrayrah Jamil, Angiolillo Antonella, Baret Philip, Baumung Roswitha, Beja-Pereira Albano, Bertaglia Marco, Bordonaro Salvatore, Bruford Mike, Caloz Régis, Canali Gabriele, Canon Javier, Cappuccio Irene, Carta Antonello, Cicogna Mario, Crepaldi Paola, Dalamitra Stella, Daniela Krugmann, Dobi Petrit, Dominik Popielarczyk, Dunner Susana, D'Urso Giuseppe, El Barody M. A. A, England Phillip, Erhardt Georg, Ertugrul Okan, Eva-Maria Prinzenberg, Eveline Ibeagha-Awemu, Ewa Strzelec, Fadlaoui Aziz, Fornarelli Francesca, Garcia David, Georgoudis Andreas, Gesine Lühken, Giovenzana Stefano, Gutscher Katja, Hewitt Godfrey, Hoda Anila, Horst Brandt, Istvan Anton, Juma Gabriela, Joost Stéphane, Jones Sam, Karetsou Katerina, Kliambas Georgios, Koban Evren, Kutita Olga, Lazlo Fesus, Lenstra Johannes A, Ligda Christina, Lipsky Shirin, Luikart Gordon, Marie-Louise Glowatzki, Marilli Marta, Marletta Donata, Milanese Elisabetta, Negrini Riccardo, Nijman, Isaäc], Obexer-Ruff Gabriela, Papachristoforou Christos, Pariset Lorraine, Pellecchia Marco, Peter Christina, Perez Trinidad, Pietrolà Emilio, Pilla Fabio, Roman Niznikowski, Roosen Jutta, Scarpa Riccardo, Sechi Tiziana, Taberlet Pierre, Taylor Martin, Togan Inci, Trommetter Michel, Valentini Alessio, Van Cann, Lisette M, Vlaic Augustin, Wiskin Louise, Zundel Stéphane.

References

1. Zeder MA, Hesse B: **The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10,000 years ago.** *Science* 2000, **287(5461)**:2254-2257.
2. MacHugh DE, Bradley DG: **Livestock genetic origins: Goats buck the trend.** *Proceedings of the National Academy of Sciences USA* 2001, **98(10)**:5382-5384.
3. Pringle H: **Neolithic agriculture: The slow birth of agriculture.** *Science* 1998, **282(5393)**:1446.
4. Fernandez H, Hughes S, Vigne JD, Helmer D, Hodgins G, Miquel C, Hanni C, Luikart G, Taberlet P: **Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas.** *Proceedings of the National Academy of Sciences USA* 2006, **103(42)**:15375-15379.

5. Dobney K, Larson G: **Genetics and animal domestication: new windows on an elusive process.** *Journal of Zoology* 2006, **269(2)**:261-271.
6. Luikart G, Fernández H, Mashkour M, England P, Taberlet P, Zeder M, Emshwiller E, Smith B, Bradley D: **Origins and diffusion of domestic goats inferred from DNA markers.** In *Documenting Domestication* Edited by: Zeder MAEE, Smith BD, Bradley DG. Berkeley: Univ of California Press; 2006:294-305.
7. Zeder MA: **Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact.** *Proceedings of the National Academy of Sciences USA* 2008, **105(33)**:11597-11604.
8. Zilhão J: **Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe.** *Proceedings of the National Academy of Sciences USA* 2001, **98(24)**:14180-14185.
9. Bradley DG, MacHugh DE, Cunningham P, Loftus RT: **Mitochondrial diversity and the origin of the African and European cattle.** *proceedings of the National Academy of Sciences USA* 1996, **93**:5131-5135.
10. Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, Robins J, Lowden S, Finlayson H, Brand T, Willerslev E, et al.: **Worldwide Phylogeography of Wild Boar Reveals Multiple Centers of Pig Domestication.** *Science* 2005, **307(5715)**:1618-1621.
11. Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, Taberlet P: **Multiple maternal origins and weak phylogeographic structure in domestic goats.** *Proceedings National Academy of Sciences USA* 2001, **98(10)**:5927-5932.
12. Vigne JD: **The large 'true' Mediterranean islands as a model for the Holocene human impact on the European vertebrate fauna? Recent data and new reflections.** In *The Holocene History of the European Vertebrate Fauna Modern Aspects of Research* Edited by: Benecke N. Rahden/Westf.: Verlag Marie Leidorf GmbH; 1999:295-321.
13. Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P: **Are cattle, sheep, and goats endangered species?** *Molecular Ecology* 2008, **17(1)**:275-284.
14. Cymbron T, Freeman AR, Malheiro IM, Vigne J-D, Bradley DG: **Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations.** *Proceedings of the Royal Society B: Biological Sciences* 2005, **272(1574)**:1837-1843.
15. Pereira F, Davis SJM, Pereira L, McEvoy B, Bradley DG, Amorim A: **Genetic Signatures of a Mediterranean Influence in Iberian Peninsula Sheep Husbandry.** *Molecular Biology and Evolution* 2006, **23(7)**:1420-1426.
16. Pereira F, Pereira L, Van Asch B, Bradley D, A A: **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* 2005, **14(8)**:2313-2318.
17. Pellicchia M, Negrini R, Colli L, Patrini M, Milanese E, Achilli A, Bertorelle G, Cavalli-Sforza LL, Piazza A, Torroni A, et al.: **The mystery of Etruscan origins: novel clues from *Bos taurus* mitochondrial DNA.** *Proceedings of the Royal Society B: Biological Sciences* 2007, **274(1614)**:1175-1179.
18. Cañón J, García D, García-Atance MA, Obexer-Ruff G, Lenstra JA, Ajmone-Marsan P, Dunner S: **Geographical partitioning of goat diversity in Europe and the Middle East.** *Animal Genetics* 2006, **37(4)**:327-334.
19. Saitbekova N, Gaillard C, Obexer-Ruff G, Dolf G: **Genetic diversity in Swiss goat breeds based on microsatellite analysis.** *Animal Genetics* 1999, **30**:36-41.
20. Yang IS, Zhao H, Li K, Peng ZZ, Montgomery GW: **Determination of genetic relationships among five indigenous Chinese goat breeds with six microsatellite markers.** *Animal Genetics* 1999, **30**:452-455.
21. Barker T, Mukherjee M, Selvaraj M: **Genetic variation within and relationships among populations of Asian goats (*Capra hircus*).** *Journal of Animal Breeding and Genetics* 2001, **118(4)**:213-233.
22. Li X-L, Valentini A: **Genetic diversity of Chinese indigenous goat breeds based on microsatellite markers.** *Journal of Animal Breeding and Genetics* 2004, **121(5)**:350-355.
23. Sultana S, Mannen H, Tsuji S: **Mitochondrial DNA diversity of Pakistani goats.** *Animal Genetics* 2003, **34(6)**:417-421.
24. Joshi MB, Rout PK, Mandal AK, Tyler-Smith C, Singh L, Thangaraj K: **Phylogeography and Origin of Indian Domestic Goats.** *Molecular Biology Evolution* 2004, **21(3)**:454-462.
25. Chen S, Su YH, Wu SF, Sha T, Zhang YP: **Mitochondrial diversity and phylogeographic structure of Chinese domestic goats.** *Molecular Phylogenetics and Evolution* 2005, **37**:804-814.
26. Odahara S, Chung H, Choi S, Yu S, Sasazaki S, Mannen H, Park C, Lee J: **Mitochondrial DNA diversity of Korean native goats.** *Asian-Australasian Journal of Animal Sciences* 2006, **19**:482-485.
27. Sardina M, Ballester M, Marmi J, Finocchiaro R, van Kaam J, Portolano B, Folch J: **Phylogenetic analysis of Sicilian goats reveals a new mtDNA lineage.** *Animal Genetics* 2006, **37(4)**:376-378.
28. Amills M, Capote J, Tomas A, Kelly L, Obexer-Ruff G, Angiolillo A, Sanchez A: **Strong phylogeographic relationships among three goat breeds from the Canary Islands.** *Journal of Dairy Research* 2004, **71(03)**:257-262.
29. Azor PJ, Monteagudo LV, Luque M, Tejedor MT, Rodero E, Sierra I, Herrera M, Rodero A, Arruga MV: **Phylogenetic relationships among Spanish goats breeds.** *Animal Genetics* 2005, **36(5)**:423-425.
30. Baumung R, SH , Hoffmann I: **Genetic diversity studies in farm animals - a survey.** *Journal of Animal Breeding and Genetics* 2004, **121**:361-373.
31. Baumung R, Simianer H, Hoffmann I: **Genetic diversity studies in farm animals - a survey.** *Journal of Animal Breeding and Genetics* 2004, **121**:361-373.
32. Schlötterer C: **The evolution of molecular markers - just a matter of fashion?** *Nature Reviews Genetics* 2004, **5**:63-69.
33. Kuhner M, Beerli P, Yamato J, Felsenstein J: **Usefulness of single nucleotide polymorphism data for estimating population parameters.** *Genetics* 2000, **156**:439-447.
34. Glaubitz J, Rhodes OE Jr, Dewoody JA: **Prospects for inferring pairwise relationships with single nucleotide polymorphisms.** *Molecular Ecology* 2003, **12**:1039-1047.
35. Morin PA, Luikart G, Wayne RK, the SNPwg: **SNPs in ecology, evolution and conservation.** *Trends in Ecology and Evolution* 2004, **19(4)**:208-216.
36. Seddon J, Parker H, Ostrander E, Ellegren H: **SNPs in ecological and conservation studies: a test in the Scandinavian wolf population.** *Molecular Ecology* 2005, **14**:503-511.
37. Syvanen A: **Assessing genetic variation: Genotyping single nucleotide polymorphisms.** *Nature Reviews Genetics* 2001, **2**:930-942.
38. Vignal A, Milan D, SanCristobal M, Eggen A: **A review on SNP and other types of molecular markers and their use in animal genetics.** *Genetics Selection Evolution* 2002, **34**:275-305.
39. Cappuccino I, Pariset L, Ajmone-Marsan P, Dunner S, Cortes O, Erhardt G, Lühken G, Gutscher K, Joost S, Nijman IJ, et al.: **Allele frequencies and diversity parameters of 27 single nucleotide polymorphisms within and across goat breeds.** *Molecular Ecology Resources* 2006, **6(4)**:992-997.
40. Pariset L, Joost S, Ajmone Marsan P, Valentini A, Consortium E: **Landscape genomics and biased *FST* approaches reveal Single Nucleotide Polymorphisms under selection in goat breeds of North-East Mediterranean.** *BMC Genetics* 2009, **10(1)**:7.
41. Weir BS, Cockerham CC: **Estimating F-statistics for the analysis of population structure.** *Evolution* 1996, **38**:1358-1370.
42. Nei M: **Genetic Distance between Populations.** *American Naturalist* 1972, **106(949)**:283.
43. Pariset L, Cappuccino I, Ajmone Marsan P, Dunner S, Luikart G, England PR, Obexer-Ruff G, Peter C, Marletta D, Pilla F, et al.: **Assessment of population structure by single nucleotide polymorphisms (SNPs) in goat breeds.** *Journal of Chromatography B* 2006, **833(1)**:117-120.
44. Naderi S, Rezaei H-R, Taberlet P, Zundel Sp, Rafat S-A, Naghash H-R, El-Barody MAA, Ertugrul O, Pompanon Fo: **Large-Scale Mitochondrial DNA Analysis of the Domestic Goat Reveals Six Haplogroups with High Diversity.** *PLoS ONE* 2007, **2(10)**:e1012.
45. Bruford MV, Bradley DG, Luikart G: **DNA markers reveal the complexity of livestock domestication.** *Nature Reviews Genetics* 2003, **3**:900-910.
46. **European Farm Animal Biodiversity System** [<http://efa.bis.tzv.fal.de>]

47. Boyazoglu J, Morand-Fehr P: **Mediterranean dairy sheep and goat products and their quality. A critical review.** *Small Ruminant Research* 2001, **40(1)**:1-11.
48. Caroli A, Chiatti F, Chessa S, Rignanese D, Bolla P, Pagnacco G: **Focusing on the Goat Casein Complex.** *Journal of Dairy Science* 2006, **89(8)**:3178-3187.
49. Sacchi P, Chessa S, Budelli E, Bolla P, Ceriotti G, Soglia D, Rasero R, Cauvin E, Caroli A: **Casein Haplotype Structure in Five Italian Goat Breeds.** *Journal of Dairy Science* 2005, **88(4)**:1561-1568.
50. Anagnostou P, Battaglia C, Coia V, Capelli C, Fabbri C, Pettener D, Destro-Bisol G, Luiselli D: **Tracing the distribution and evolution of lactase persistence in Southern Europe through the study of the T-13910 variant.** *American Journal of Human Biology* 2009, **21(2)**:217-219.
51. Hickey D, Kilcawley K, K N, Beresford TP, Wilkinson MG: **Lipolysis in cheddar cheese made from raw, thermized, and pasteurized milks.** *Journal of Dairy Science* 2007, **90(1)**:47-56.
52. Goudet J: **FSTAT (version 1.2): a computer program to calculate F-statistics.** *Journal of Heredity* 1995, **86**:485-486 [<http://ered.oxfordjournals.org/cgi/gca?allch=&SEARCHID=1&FULLTEXT=goudet&FIRSTINDEX=0&hits=10&RESULTFORMAT=&gca=jhered%3B86%2F6%2F485&allchb=>].
53. Raymond M, Rousset F: **GENEPOP: population genetics software for exact tests and ecumenicism.** *Journal of Heredity* 1995, **86**:248-249.
54. Wright S: **Isolation by distance.** *Genetics* 1943, **28(2)**:114-138.
55. Slatkin M: **Isolation by distance in equilibrium and non-equilibrium populations.** *Evolution* 1993, **47**:264-279.
56. Mantel N: **The detection of disease clustering and a generalized regression approach.** *Cancer Res* 1967, **27**:209-220.
57. Mantel N, Valand R: **A technique of nonparametric multivariate analysis.** *Biometrics* 1970, **26**:547-558.
58. Liu K, Muse S: **Powermarker: new genetic data analysis software.** Version 3.0 edn 2001 [http://statgen.ncsu.edu/brcwebsite/technical_resources_software.php#PowerMarker].

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

