Characterization of biodiversity in six goat breeds reared in Southern Italy by means of microsatellite and SNP markers

A. Criscione[1], D. Marletta[1], T. Ådnøy[2], S. Bordonaro[1], A. M. Guastella[1], S. Lien[2], G. D'Urso[1]

[1] Dipartimento di Scienze Agronomiche, Agrochimiche e delle Produzioni Animali. Università di Catania, Italy

[2] Centre for Integrative Genetics and Department of Animal and Aquacultural Sciences. Norwegian University of Life Sciences. Ås, Norway

Corresponding author: Andrea Criscione Dipartimento di Scienze Agronomiche, Agrochimiche e delle Produzioni Animali. Università di Catania. Via Valdisavoia 5, 95123 Catania, Italy - Tel.: +39 095 234481 - Fax: +39 095 234345 - E-mail: andreacriscione@tiscali.it

ABSTRACT: An integrated analysis, using 20 microsatellite markers and 32 SNP markers belonging to the casein cluster has been carried out on 174 goats from 6 local goat breeds and populations from Southern Italy. Microsatellite markers provided 216 alleles (10.8 per *locus*; from 6.6 to 8.2 per breed).
The average expected heterozygosity (H_e) was 0.732. *Fis* value (0.148) indicated a general heterozygosity deficiency. A high number of intragenic haplotypes (56) have been detected at casein loci, 25 at α_{S1} casein (CSN1S1), 12 at β casein (CSN2), 8 at α_{S2} casein (CSN1S2), 11 at κ casein (CSN3), when 1% frequency was required for each breed. The breeds with higher production and management level, fixed the lowest number of combinations at casein *loci* (28). Molecular data have been used to calculate genetic distances and in clustering analysis.

Key words: Goats, Biodiversity, Microsatellites, SNP.

INTRODUCTION – Goat breeding in Southern Italy is still exploiting marginal agricultural resources. Local breeds and populations are scarcely managed: selective breeding is often absent and gene flow is almost continuous between populations reared in the same area.
Microsatellites and SNP markers have proved useful in assessing genetic diversity of goat populations (Canon *et al.*, 2006; Pariset *et al.*, 2006). An integrated analysis of these markers is needed in order to better characterize and exploit goat genetic resources of Southern Italy.

MATERIAL AND METHODS – Genomic DNA has been extracted from blood collected from 174 unrelated goats (31 Girgentana, 30 Argentata dell'Etna, 30 Maltese, 22 Rossa Mediterranea, 30 Messinese, 31 Aspromonte). A set of 20 microsatellites markers (CSRD0247, SRCRSP0023, INRA0023, ILSTS0087, McM0527, OarFCB0020, SRCRSP0005, SRCRSP0008, MAF65, SRCRSP0024, PZ963, HSC, ILSTS019, INRA0005, INRA063, ILSTS005, MAF70, SRCRSP0006, MAF209, TGLA53) has been amplified in three PCR multiplex reactions and ran on a automated DNA sequencer ABI PRISM 377 equipped with GeneScan Analysis ver. 3.1.2 and Genotyper ver. 2.5 softwares (Applied Biosystems). Moreover the 250 kb region spanning over the casein genes has been also characterized at 32 polymorphic SNP (12 at CSN1S1, 7 at CSN2, 4 at CSN1S2 and 9 at CSN3 *locus*) by means of MassArray System (Sequenom).
FastPHASE 1.1 software (Scheet and Stephens, 2006) was used to construct haplotypes, within each *locus*, from SNP genotypes. Haplotypes with a frequency of less than 1% were omitted from the dataset. The linkage disequilibrium (LD) was assessed by HAPLOVIEW program (Barrett, 2005) using r^2 statistic. Estimated haplotype frequencies have been analysed by DISPAN (Ota, 1993). Principal statistical parameters of genetic diversity and F-

statistics (Weir and Cockerham, 1984), have been calculated with Microsatellites analyser (Dieringer and Schlötterer, 2002), Excel Microsatellite Toolkit (Park, 2001) and FSTAT (Goudet, 2001) programs.
Da genetic distance (Nei *et al.*, 1983), was computed using MICROSAT (Minch *et al.*, 1995) and DISPAN programs. Phylogeny trees were obtained from genetic distance matrices using Phylip 3.66 software package (Felsenstein, 2005) and SplitsTree4 4.6 software (Huson and Briant, 2006). STRUCTURE 2.0 (Pritchard *et al.* 2000a) was implemented to assess the maximum likelihood number of clusters in the dataset, independent from breed affiliation.

RESULTS AND CONCLUSION – All microsatellites markers were polymorphic in each breed: a total of 216 alleles have been observed across all 20 *loci* (10.8 per *locus*). MAF209 *locus* showed the lowest PIC value (0.299) while SRCRSP0023 *locus* the highest one (0.881).
Allelic richness ranged from 2.9 (MAF209) to 12.2 (SRCRSP0023) with a mean value of 7.5. The effective number of alleles ranged from 3.4 in Girgentana to 4.1 in Argentata. The expected heterozygosity (H_e) over all population varied between 0.325 and 0.893 with a mean value of 0.751. Genetic differentiation among breeds (*Fst*) accounted for 3%, hence 97% of the genetic variation is among individuals. *Fis* value was 0.148 indicating a general heterozygosity deficiency. Girgentana goat breed presented the lowest variability (6.6 alleles/*locus*, H_e=0.706) in the dataset, while Argentata appeared the most variable population (8.2 alleles/*locus*, H_e=0.753).
Average H_e was 0.732. SNP markers highlighted two monomorphic sites at CSN2 *locus*, in Girgentana and Rossa Mediterranea goat; at those sites the other breeds showed frequencies always lower than 8.1% for the same allele. CSN1S2 *locus*, despite the lowest number of SNP markers, appeared to be the most informative according to the differentiation index *Gst* (0.058), the gene diversity (H_t=0.818) and the low internal linkage evaluated by r^2.
In the whole sample (174 goats) 33 haplotypes have been detected (12 CSN1S1, 8 CSN2, 6 CSN1S2, 7 CSN3). A higher number of haplotypes (56) resulted from the analysis of separate breeds (25 CSN1S1, 12 CSN2, 8 CSN1S2, 11 CSN3) when 1% frequency was required for each breed. For each *locus* the two most frequent haplotypes accounted for more than 50%, except for Messinese and Rossa at CSN1S2 *locus*.
Breeds with higher production and management level, fixed a lower number of combinations at casein *loci*: the lowest number of haplotypes (28) was observed in Girgentana, Maltese and Rossa.
The Aspromonte and Argentata goat showed the highest variability with 34 and 33 haplotypes respectively. LD was not equally spread over the chromosome segment; CSN1S1 and CSN3 *loci* highlighted the highest r^2 values. At microsatellites markers, Maltese and Girgentana goat were the most divergent breeds whilst Argentata and Messinese appeared to be the closest ones (Table 1).

Table 1. Da distance matrices estimated by microsatellites (below diagonal) and SNP markers (above diagonal).

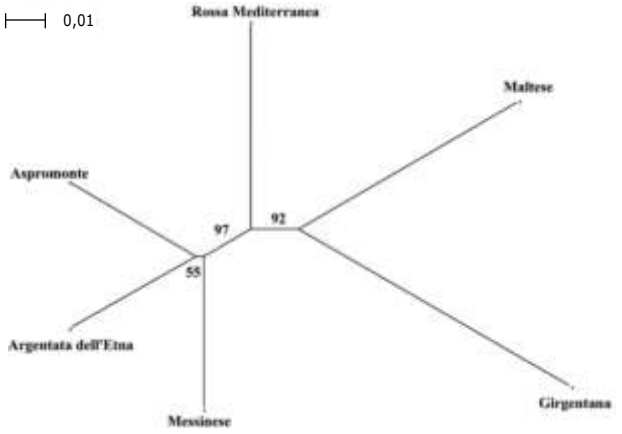
	Messinese	Argentata	Aspromonte	Maltese	Rossa	Girgentana
Messinese		0.118	0.117	0.113	0.096	0.144
Argentata dell'Etna	0.072		0.081	0.139	0.112	0.150
Aspromonte	0.074	0.075		0.176	0.155	0.189
Maltese	0.138	0.126	0.124		0.092	0.174
Rossa Mediterranea	0.106	0.116	0.095	0.148		0.161
Girgentana	0.139	0.118	0.126	0.182	0.166	

Maltese and Girgentana were also the only two breeds to cluster significantly for each k value used in the Bayesian model approach. *Da* distance estimated at SnPs was the highest and the lowest in Aspromonte-Girgentana, and Aspromonte-Argentata, respectively (Table 1). UPGMA tree in figure 1 was drawn from *Da* distance matrix obtained analysing jointly all the molecular markers (microsatellites and SNP) (Table 2).
The dendrogram clearly clustered two groups, one including the two breeds (Girgentana and Maltese) the other one formed by Aspromonte, Argentata and Messinese populations. Outcomes provided by each class of markers here used, were partly complementary.
Microsatellites markers are of interest for assessing relationships among breeds. However the integration of neutral and specific markers, related with production traits, may allow to better characterize, taking into account the selective breeding pressure, the genetic structure of animal resources.

Table 2. Da distance matrix estimated by microsatellites and SNP markers.

	Messinese	Argentata	Aspromonte	Maltese	Rossa	Girgentana
Messinese						
Argentata dell'Etna	0.080					
Aspromonte	0.081	0.076				
Maltese	0.133	0.128	0.133			
Rossa Mediterranea	0.105	0.115	0.105	0.138		
Girgentana	0.139	0.123	0.137	0.180	0.165	

Figure 1. UPGMA tree drawn from *Da* distance estimated by microsatellites and SNP markers.



REFERENCES – Barrett JC, Fry B., Maller J., Daly M.J. 2005 Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263-65. **Cañon** J., Garcia D., Garcia-Atance M.A., Obexer-Ruff G., Lenstra J.A., Ajmone-Marsan P., Dunner S., ECONOGENE Consortium 2006 Geographical partitioning of goat diversity in Europe and the Middle East *Anim. Genet.* 37, 327–34. **Dieringer** D. & Schlötterer C. 2002 Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* 3, 167-69. **Felsenstein** J. 2005 PHYLIP ver 3.6. *Department of Genome Sciences, Washington University, Seattle http://evolution.genetics.washington.edu/phylip/getme.html.* **Goudet** J. 2001 FSTAT, a program to estimate and test gene diversities and fixation indices (ver. 2.9.3) *http://www.unil.ch/izea/softwares/fstat.html.* **Huson D.H.,** Bryant D. 2006 Application of phylogenetic networks in evolutionary studies *Mol. Biol. Evol.* 23, 254–67 www.splitstree.org. **Minch** E., Ruiz-Linares A., Goldstein D., Feldman M., Cavalli-Sforza L.L. 1995 Microsat (ver 1.4d): a computer program for calculating various statistics on microsatellite allele data *Stanford University Medical Center, Stanford. http://human.stanford.edu/microsat/microsat.html.* **Nei** M., Tajima F., Tateno Y. 1983 Accuracy of estimated phylogenetic trees from molecular data *J. Mol. Evol.* 19, 153-70. **Ota** T., 1993 Dispan: Genetic Distance and Phylogenetic Analysis *Pennsylvania St. Un., University Park, P.A. ftp://ftp.bio.indiana.edu/molbio/ibmpc.* **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., ECONOGENE Consortium 2006 Assessment of population structure by single nucleotide polymorphisms (SnPs) in goat breeds *J. Chromat. B* 833, 117–20. **Park** S.D.E. 2001 Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection. *Ph.D. thesis*, University of Dublin. **Pritchard** J.K., Stephens M., Donnely P. 2000a Inference of population structure using multilocus genotype data. *Genet.* 155, 945-49. **Scheet** P., Stephens M. 2006 A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase *Am. J. Hum. Genet.* 78, 629–44. **Weir** B.S., Cockerham C.C. 1984 Estimating F-statistics for the analysis of population structure *Evol.* 38, 1358-70.