

# Genetic characterization of four native Italian shepherd dog breeds and analysis of their relationship to cosmopolitan dog breeds using microsatellite markers

D. Bigi<sup>1†</sup>, S. P. Marelli<sup>2</sup>, E. Randi<sup>3,4</sup> and M. Polli<sup>2</sup>

<sup>1</sup>Department of Agricultural and Food Science and Technology (DISTAL), University of Bologna, Via F.lli Rosselli 107, 42123 Reggio Emilia, Italy; <sup>2</sup>Dipartimento di Scienze Veterinarie e Sanità Pubblica, Università degli Studi di Milano, via Celoria 10, 20133 Milano, Italy; <sup>3</sup>Laboratorio di Genetica, Istituto Superiore per la Protezione e Ricerca Ambientale, Ozzano Emilia, Bologna, Italy; <sup>4</sup>Department 18/ Section of Environmental Engineering, Aalborg University, Sohngårdsholmsvej 57, 9000, Aalborg, Denmark

(Received 12 December 2014; Accepted 10 July 2015; First published online 6 August 2015)

---

*Very little research into genetic diversity of Italian native dog breeds has been carried out so far. In this study we aimed to estimate and compare the genetic diversity of four native Italian shepherd dog breeds: the Maremma, Bergamasco, Lupino del Gigante and Oropa shepherds. Therefore, some cosmopolitan dog breeds, which have been widely raised in Italy for a long time past, have also been considered to check possible influence of these dog populations on the Italian autochthonous breeds considered here. A total of 212 individuals, belonging to 10 different dog breeds, were sampled and genotyped using 18 autosomal microsatellite loci. We analyzed the genetic diversity of these breeds, within breed diversity, breed relationship and population structure. The 10 breeds considered in this study were clearly genetically differentiated from each other, regardless of current population sizes and the onset of separate breeding history. The level of genetic diversity explained 20% of the total genetic variation. The level of  $H_E$  found here is in agreement with that found by other studies. The native Italian breeds showed generally higher genetic diversity compared with the long established, well-defined cosmopolitan dog breeds. As the Border Collie seems closer to the Italian breeds than the other cosmopolitan shepherd dogs considered here, a possible utilization of this breed to improve working performance in Italian traditional working shepherd dogs cannot be ignored. The data and information found here can be utilized in the organization of conservation programs planned to reduce inbreeding and to minimize loss of genetic variability.*

---

**Keywords:** genetic diversity, population genetic structure, domestic dog, shepherd dog, canine

## Implications

In this study we aimed to estimate and compare the genetic diversity of four native Italian shepherd dog breeds and their relationship to cosmopolitan dog breeds reared in Italy. As these Italian autochthonous breeds are involved in sheep- and cattle-farming in mountainous regions, they play a role in the economic tissue of marginal and disadvantaged areas, it is important to preserve their breeding. These data and information found here can be utilized in the organization of conservation programs planned to reduce inbreeding and to minimize loss of genetic variability.

## Introduction

The dog is probably the most phenotypically diverse vertebrate species, thanks to its long history of domestication. Most of

the various pure-breed dogs are morphologically clearly distinct and they also differ in behavior, physical properties and specific inherited diseases. Based on available breed histories, the majority of existing dog breeds were developed in the 19th century. At present, more than 400 canine breeds are registered worldwide by the Federation Cynologique Internationale (FCI, World Canine Organisation) and other federations such as the American and British Kennel Clubs (Parker *et al.*, 2004; Streitberger *et al.*, 2011).

To understand the effect of breed structure on the genomic makeup of the dog, molecular markers can be put to use in both small and large data sets. Autosomal microsatellites have been utilized to study genetic diversity in several dog breeds, primarily for the purposes of determining the power of exclusion for parentage applications (Altet *et al.*, 2001; DeNise *et al.*, 2004). Autosomal microsatellites have also been utilized to study genetic diversity in several dog breeds. In dogs, microsatellite typing, combined with phylogenetic

---

† E-mail: daniele.biggi@unibo.it

analysis and modern genetic clustering methods, allows the definition of related groups of breeds and that genetic relatedness among breeds often correlates with morphological similarity and shared geographic origin (Kim *et al.*, 2001; Irion *et al.*, 2003; Parker *et al.*, 2004; Wade, 2011; Ciampolini *et al.*, 2013; Mellanby *et al.*, 2013). However, these studies have focused mainly on cosmopolitan breeds, while peripheral breeds have been studied less.

Effective management of farm animal resources requires comprehensive knowledge of the breeds' characteristics including data on population size and structure, geographical distribution, and within and between breed genetic diversity. Information regarding genetic structure in non-cosmopolitan dog breeds and local populations is still lacking (Pires *et al.*, 2009). No previous studies about genetic diversity of Italian native dog breeds have been carried out so far, except for the breed Bracco Italiano (Ciampolini *et al.*, 2011; Cecchi *et al.*, 2013), the Italian Greyhound (Pedersen *et al.* 2013) and a research into the dog breeds raised in France (Leroy *et al.*, 2009) that included two Italian autochthonous breeds (Italian Corso Dog and Romagna Water Dog). The aim of this study was to conduct a comparative analysis of the genetic diversity of four native shepherd dog populations in order to investigate the extent of genetic variation characterizing these native dogs, based on microsatellite loci. Moreover we decided to analyze the possible influence of some cosmopolitan dog breeds reared in Italy on these native breeds. At present, 16 Italian dog breeds are officially recognized by the FCI. This study focused on two of these autochthonous breeds: the Maremma Sheepdog is a breed of livestock guard dog indigenous to central Italy, particularly to Abruzzo and the Maremma region of Tuscany and Lazio. It is a large size dog with a solid, muscular build, a thick white coat, a large head and a black nose. It has been used for centuries by Italian shepherds to guard sheep from wolves. As in Italy, in the last few decades, the number of wolves has been rapidly increasing, the use of this dog breed for guarding sheep is growing; the Bergamasco Shepherd originated in the Italian Alps near Bergamo, where it was traditionally used as a herding dog. Today it is largely a show and performance dog. It is a medium size dog, well-proportioned and harmonious, having a rustic appearance. The breed's most distinctive feature is the unusual felted coat, a normal and healthy characteristic of the breed. Moreover we also considered two not officially recognized Italian native dog breeds: the Lupino del Gigante Shepherd is a medium-sized herding dog, it has a morphological wolf-like appearance in terms of formation, movement, coat, but smaller in size with variable wolf gray, red, black, white and blue merle coats. It originated in the Apennine Mountains near the city of Reggio Emilia. The use of this shepherd dog to herd and guard the flocks dates back centuries, but recently the strong reduction of sheep-breeding in the area has led to a decline in the numbers of the breed which is now reduced to about 300 specimens; the Oropa Shepherd has been used for centuries in the Biella region of the Western Italian Alps to herd indigenous cattle and sheep.

Similar types are found throughout the Alps, in Eastern Italy, Switzerland, and possibly in France as well. The coat color of this medium-size dog breed, characterized by wolf like morphology, is usually merle with black and tan markings, black or different shades of brown (Marelli *et al.*, 2009). We also considered three cosmopolitan shepherd dog breeds (Belgian Shepherd, Border Collie and German Shepherd Dog), which have been widely bred in Italy for a long time past, to check possible influence of these dog populations on the Italian autochthonous breeds considered here. We also included the Siberian Husky and the Alaskan Malamute, the two most common spitz-like breeds in Italy, due to a close resemblance of some Lupino Del Gigante specimens to spitz-like breeds; in particular regarding coat texture and undercoat density, head shape and ear shape, thickness and carriage. For the same reason the Czechoslovakian Wolfdog, a quite popular breed in Italy (1024 new entries in 2013 ENCI studbooks), was considered as well, being this hybrid breed very similar morphologically to Lupino del Gigante. Using microsatellite data, we analyzed the genetic diversity of these breeds, within breed diversity, breed relationship, population structure. The four Italian shepherd dogs considered here, because of their small population size are endangered. In particular, based on their current number of specimens and following the Food and Agriculture Organization classification, the two breeds Lupino del Gigante and Oropa are at risk of extinction. Therefore, the conservation of these breeds is urgently needed.

## Material and methods

### *Sampling and DNA extraction*

A total of 212 samples from 22 Bergamasco Shepherds (BER), 26 Belgian Shepherds (BEL), 22 Maremma Sheepdogs (MSD), 15 Siberian Huskies (SHU), 12 Alaskan Malamutes (ALM), 24 Czechoslovakian Wolfdogs (CWD), 24 German Shepherd Dogs (GSD), 22 Border Collies (BCO), 19 Oropa Shepherds (ORO), 26 Lupino del Gigante Shepherds (LGS), were collected from shows according to Italian Kennel Club pedigree certification scheme and field inspections. We recorded the following individual dog's information: ID number (microchip), breed, age, gender, health status. Pedigrees of the dogs belonging to FCI certified breeds were downloaded from the Italian Kennel Club studbooks. We used genealogical registrations for the uncertified dog breeds Oropa and Lupino del Gigante Shepherds. All sampled dogs were unrelated in the second generation. DNA was extracted from whole EDTA blood or DNA sample cards (Vet Kard with FTA System<sup>®</sup>, Prion Diagnostica srl), using Illustra<sup>™</sup> blood genomicPrep Mini Spin Kit (GE Healthcare UK Limited, Amersham Place, Little Chalfont, Buckinghamshire, UK), according to the manufacturer's protocols.

### *Microsatellite genotyping*

Samples were genotyped using eighteen microsatellite markers mapping on 17 different autosomes [AHTK211 (map location

on *CFA26*, *CXX279* (CFA22), *REN169018* (CFA29), *INU055* (CFA10), *REN54P11* (CFA18), *INRA21* (CFA21), *AHT137* (CFA11), *REN169D01* (CFA14), *AHTH260* (CFA16), *AHTK253* (CFA23), *INU005* (CFA33), *INU030* (CFA12), *FH2848* (CFA2), *AHT121* (CFA13), *FH2054* (CFA12), *REN162C04* (CFA7), *AHTH171* (CFA6) and *REN247M23* (CFA15)]. Gender was confirmed using an amelogenin marker (AMELY and AMELX) that was included in the Canine Genotypes TM Panel 1.1, F-860 S/L microsatellite panel (Finnzymes Diagnostics; Finnzymes OY, Keilaranta 16 A, 02150 Espoo, Finland). These microsatellites are included in the panel of loci by Applied Genetics Committee of Companying Animals of the International Society for Animal Genetics (ISAG) (Budowle *et al.*, 2005). PCR amplification and electrophoresis (ABI Prism® 310 Genetic Analyzer - Applied Biosystems) were carried out following the manufacturer's protocols.

Individual genotypes were determined using the software genescan® 3.7 (ABI).

#### Data analysis

The software Arlequin version 3.5 (Excoffier and Lisher, 2010) was used to: (i) calculate the number of alleles per locus and the observed ( $H_O$ ) and expected heterozygosity ( $H_E$ ); (ii) test for Hardy–Weinberg equilibrium (HWE). For multiple comparisons, the statistical significance levels were Bonferroni-corrected (Rice, 1989); (iii) compute Wright's  $F_{IS}$ ,  $F_{IT}$ , and  $F_{ST}$  fixation index (Wright, 1965; Weir and Cockerham, 1984) and their statistical significance was tested using 1000 permutations and to evaluate the significance of genetic differentiation between populations with an analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) (iv) compute a matrix of Reynolds unweighted distances,  $D_R$ , between breeds (Reynolds *et al.*, 1983).

Fstat software (Goudet, 2001) was used to estimate the presence and frequency of private alleles and to calculate the allelic richness ( $AR$ ) standardized for variation in sample size.

The software Splitstree version 4.11.3 (Huson and Bryant, 2006) was used to draw a Neighbor-net starting from Reynolds's genetic distances ( $D_R$ ).

Dog genotypes were clustered using a Bayesian approach to infer populations by the Structure version 2.2 computer program (Pritchard *et al.*, 2000; Falush *et al.*, 2003). Using a model with admixture and correlated allele frequencies, we made 10 independent runs for each value of the putative number of subpopulations ( $K$ ) between 1 and 20, with a burn-in period of 20 000 followed by 100 000 MCMC repetitions. The most likely number of clusters in our sample was determined using the method of Evanno *et al.* (2005). Clump (Jacobsson and Rosenberg, 2007) was used to concatenate the data from the multiple runs for each  $K$ , and Distruct (Rosenberg, 2004) was used to graphically display the results.

To visualize the genetic structure of individuals, a principal coordinates analysis based on inter-individual genetic distance in GenAlex 6 (Peakall and Smouse, 2006) was also conducted. Genetic distances were calculated as described in the GenAlex 6 guide for co-dominant data.

#### Results

A total number of 195 alleles were detected across the 18 microsatellite loci analyzed which were polymorphic in all the dog breeds. The average number of alleles per locus was 10.8 and the number of alleles ranged between 7 (*INU055* and *REN247M23*) and 17 (*REN169018*).

Observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosities, allelic richness ( $AR$ ), and  $F_{IS}$  for each dog breed are presented in Table 1.

Among the 10 dog breeds considered, the lowest value of  $AR$  was found in Belgian Shepherd (3.74) and the highest in Maremma Sheepdog (5.75). Also for the microsatellite loci tested across populations, the lowest observed and expected heterozygosities were observed in the Belgian Shepherd, whereas the highest were found in the Maremma Sheepdog.

The HWE exact test revealed 31 significant deviations (17.2%) between expected ( $H_E$ ) and observed ( $H_O$ ) heterozygosity after

**Table 1** Basic information and genetic diversity parameters of 10 dog breeds

Breed	Breed code	Reference	Sample size	$N_{PA}$	$AR$	$H_O$	$H_E$	$F_{IS}$	HWE deviations <sup>1</sup>
		population size							
Alaskan Malamute	ALM	6963	12	3	4.94	0.57	0.64	0.12*	+2/–4
Belgian Shepherd	BEL	3830	26	5	3.74	0.53	0.55	0.04	–2
Bergamasco Shepherd	BER	1928	22	3	5.52	0.66	0.71	0.08*	–5
Border Collie	BCO	36 718	22	5	5.34	0.68	0.70	0.02	–2
Czechoslovakian Wolfdog	CWD	12 142	24	4	3.99	0.57	0.56	–0.01	–1
German Shepherd Dog	GSD	28 9865	24	6	3.91	0.56	0.57	0.02	+1/–3
Lupino del Gigante Shepherd	LGS	300	26	12	5.01	0.64	0.68	0.06*	
Maremma Sheep Dog	MSD	12 352	22	8	5.75	0.69	0.77	0.11***	–4
Oropa Shepherd	ORO	250	19	1	4.63	0.70	0.62	–0.14***	+4/–1
Siberian Husky	SHU	8637	15	1	3.91	0.61	0.62	0.02	–2

$N_{PA}$  = number of private alleles;  $AR$  = allelic richness;  $H_E$  = expected heterozygosity;  $H_O$  = observed heterozygosity;  $F_{IS}$  = within-breed inbreeding coefficient; HWE deviations = number of loci departing from Hardy–Weinberg equilibrium expectations.

<sup>1</sup>Negative and positive values indicate the number of loci showing heterozygote deficiency and excess, respectively.

\* $P < 0.05$ ; \*\*\* $P < 0.001$ .

Bonferroni correction.  $H_0$  was significantly higher than  $H_E$  in seven cases; in all other cases,  $H_E$  was higher than  $H_0$ .

The overall  $F_{IS}$  value among all loci was significantly higher than zero, after Bonferroni correction, in five breeds (0.12 in the Alaskan Malamute, 0.08 in the Bergamasco Shepherd, 0.06 in Lupino del Gigante Shepherd, 0.10 in the Maremma Sheepdog and  $-0.14$  in the Oropa Shepherd), indicating heterozygote deficiency in ALA, BER, LGS and MSD and a significant level of heterozygote excess in ORO.

Forty-eight private alleles were identified. At least one private allele was identified in each of the analyzed

populations with the Lupino del Gigante Shepherd having the highest number of private alleles ( $N_{PA} = 12$ ), while Oropa Shepherd and Siberian Husky have the lowest ( $N_{PA} = 1$ ).

F statistics are shown in Table 2. Genetic differentiation among breeds was very significant: the average  $F_{ST}$  values indicate that around 20% of the total genetic variation was explained by breed differences, with the remaining 80% corresponding to differences among individuals. Genetic differentiation among breeds was highly significant ( $P < 0.001$ ) for all loci. A significant excess of homozygotes across all breeds were found for *CXX279*, *AHTK121*, *AHTK253* and *REN54P11*. On average, breeds had a 3.4% ( $P < 0.001$ ) deficit of heterozygotes, whereas the total population had a 22.5% ( $P < 0.001$ ) deficit of heterozygotes, although some of the markers showed heterozygosity excess (*INRA21*, *INU030*, *INU055*, *INU005* and *AHTH137*).

AMOVA revealed that 20% of the observed total variance occurred among varieties and among groups whereas 80% was explained by differences within individuals and among individuals within varieties (data not shown).

Pairwise  $F_{ST}$  values correlate with cluster analyses, ranging from 0.069 between the Maremma Sheepdog and the Border Collie specimens to 0.362 between the Belgian Shepherd and the Czechoslovakian Wolfdogs (Table 3). All  $F_{ST}$  values were significantly different from zero ( $P < 0.001$ ).

The principal coordinates analysis (Figure 1), both the first and the second axis, which accounted for 31.34% and for 17.74% of the total variance, separated the Belgian Shepherd from the other dog breeds. The German Shepherd and the Czechoslovakian Wolfdog grouped together and are clearly separated from the other breeds. All the other breeds form a single big group in the center of the graph even if the Alaskan Malamute and the Siberian Husky form a subgroup.

The Neighbor-net, reconstructed from Reynolds's distances, is reported in Figure 2. The German Shepherd and the Czechoslovakian Wolfdog clearly clustered together. This result is expected on the basis of known CWD breed history, as this dog is a hybrid breed of German Shepherd

**Table 2** Characterization of the 18 analysed microsatellite loci in 10 dog breeds

Locus	$N_A$	$N_{PA}$	$F_{IS}$	$F_{ST}$	$F_{IT}$
<i>CXX279</i>	11	5	0.133**	0.117***	0.234***
<i>AHTH211</i>	8	2	0.031	0.266***	0.289***
<i>AHTH121</i>	16	4	0.156***	0.277***	0.390***
<i>INRA21</i>	9	1	-0.057	0.157***	0.109***
<i>AHTH171</i>	14	5	0.057	0.225***	0.269***
<i>AHTK253</i>	10	1	0.139**	0.221***	0.330***
<i>FH2054</i>	12	2	0.033	0.128***	0.157***
<i>REN54P11</i>	11	1	0.090*	0.163***	0.238***
<i>AHTH260</i>	10		0.027	0.181***	0.203***
<i>FH2848</i>	9	1	0.040	0.122***	0.157***
<i>INU030</i>	8	1	-0.038	0.209***	0.179***
<i>INU055</i>	7		-0.025	0.122***	0.100***
<i>AHT137</i>	14	5	-0.012	0.131***	0.121***
<i>REN162C04</i>	9	2	0.060	0.134***	0.187***
<i>REN169D01</i>	16	6	0.023	0.218***	0.236***
<i>REN247M23</i>	7	1	0.045	0.379***	0.407***
<i>REN169018</i>	17	8	0.013	0.276***	0.285***
<i>INU005</i>	11	3	-0.059	0.210***	0.164***
Average	11.1	2.7	0.034***	0.196***	0.225***

$N_A$  = number of alleles;  $N_{PA}$  = number of private alleles;  $F_{IS}$  and  $F_{IT}$  are measurements of the deviation from Hardy-Weinberg proportions within populations and in the total population, respectively;  $F_{ST}$  is the genetic differentiation over subpopulations.

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

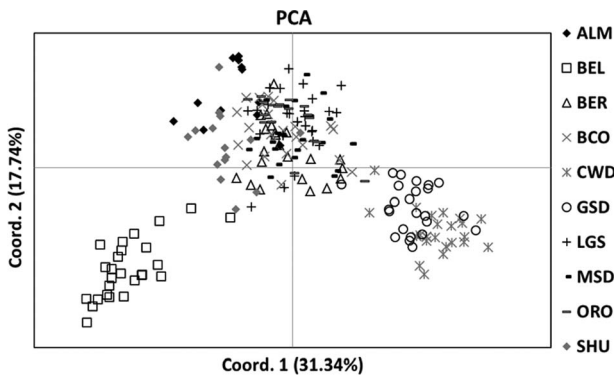
**Table 3**  $F_{ST}$  estimates<sup>1</sup> as a measure of genetic distance between dog breeds

	ALM	BEL	BER	BCO	CWD	GSD	LGS	MSD	ORO	SHU
ALM	–									
BEL	0.269	–								
BER	0.171	0.230	–							
BCO	0.156	0.238	0.114	–						
CWD	0.296	0.362	0.180	0.234	–					
GSD	0.282	0.333	0.207	0.190	0.146	–				
LGS	0.176	0.244	0.145	0.143	0.219	0.176	–			
MSD	0.113	0.211	0.093	0.069	0.177	0.134	0.106	–		
ORO	0.177	0.249	0.152	0.134	0.242	0.186	0.113	0.089	–	
SHU	0.184	0.264	0.180	0.224	0.307	0.271	0.204	0.140	0.225	–

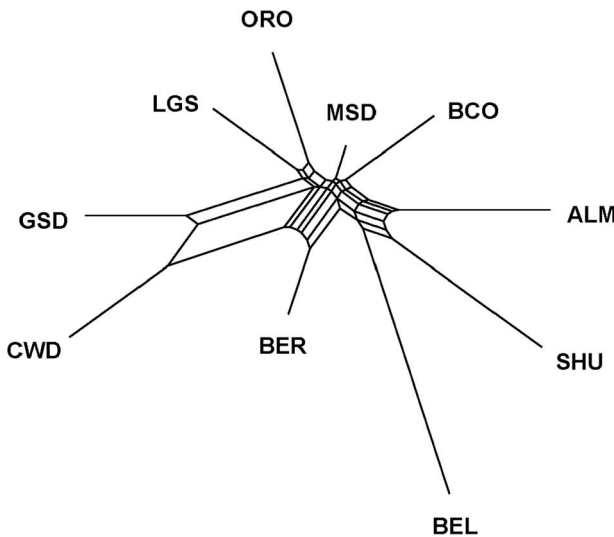
BER = Bergamasco Shepherd; BEL = Belgian Shepherd; MSD = Maremma Sheepdog; SHU = Siberian Husky; ALM = Alaskan Malamute; CWD = Czechoslovakian Wolfdog; GSD = German Shepherd Dog; BCO = Border Collie; ORO = Oropa Shepherd; LGS = Lupino del Gigante Shepherd.

All pairwise  $F_{ST}$  differences were significantly larger than 0 ( $P < 0.001$ ).

<sup>1</sup> $F_{ST}$  estimates calculated as described in Weir and Cockerham (1984).



**Figure 1** Principal coordinates analysis plot for ten dog populations. BER = Bergamasco Shepherd; BEL = Belgian Shepherd; MSD = Maremma Sheepdog; SHU = Siberian Husky; ALM = Alaskan Malamute; CWD = Czechoslovakian Wolfdog; GSD = German Shepherd Dog; BCO = Border Collie; ORO = Oropa Shepherd; LGS = Lupino del Gigante Shepherd.

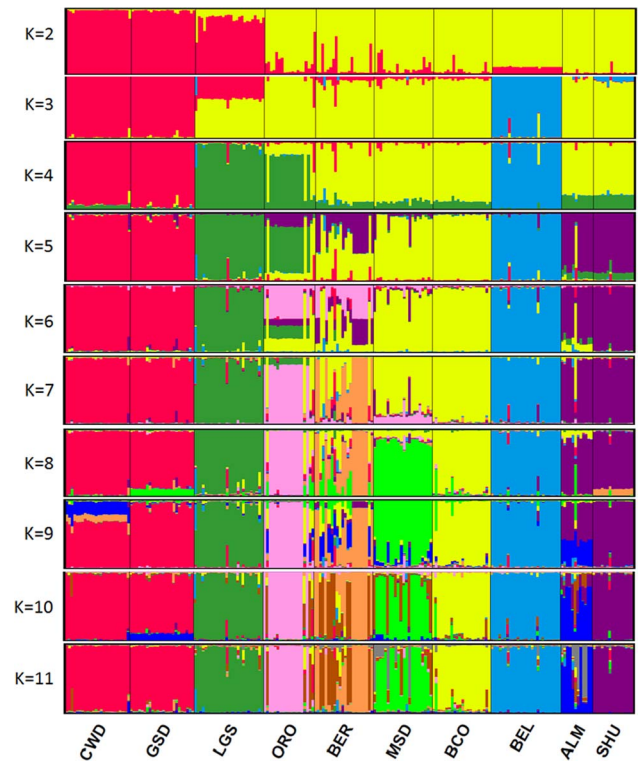


**Figure 2** Neighbor-net based on Reynolds distance matrix between populations. BER = Bergamasco Shepherd; BEL = Belgian Shepherd; MSD = Maremma Sheepdog; SHU = Siberian Husky; ALM = Alaskan Malamute; CWD = Czechoslovakian Wolfdog; GSD = German Shepherd Dog; BCO = Border Collie; ORO = Oropa Shepherd; LGS = Lupino del Gigante Shepherd.

Dog × Carpathian wolf founders (Randi *et al.*, 2014). The two Italian breeds Lupino del Gigante Shepherd and Oropa Shepherd grouped together as well as the two breeds Maremma Sheepdog and the Border Collie. Another cluster is formed by the Alaskan Malamute and the Siberian Husky.

Clustering using Bayesian approaches was performed on the entire data set with an increasing number of inferred clusters from  $K = 2$  to 20 and produced consistent results (Figure 3). The software structure identified  $K = 11$  the most probable number of genetic clusters in our sample and the likelihood scores increased for each  $K$  up to  $K = 11$  ( $\ln = 10\,399.2$ ).

Assignment isolated the Belgian breed from the other breeds as early as three groups, and this group maintained its integrity through the analysis. At the same time,



**Figure 3** Proportion of membership of 212 individuals from 10 dog populations for  $K = 2$  to 11, as calculated by Structure software. Multiple runs for each  $K$  were concatenated using clumpp, and distruct was used to generate images. BER = Bergamasco Shepherd; BEL = Belgian Shepherd; MSD = Maremma Sheepdog; SHU = Siberian Husky; ALM = Alaskan Malamute; CWD = Czechoslovakian Wolfdog; GSD = German Shepherd Dog; BCO = Border Collie; ORO = Oropa Shepherd; LGS = Lupino del Gigante Shepherd.

Czechoslovakian Wolfdog and German Shepherd shared the same cluster and only when  $K$  increased to 16 they separated in two distinct clusters (data not shown). For  $K = 5$ , there was a clear separation of Alaskan Malamute and Siberian Husky from the other populations. When  $K$  increased to 6, the Lupino del Gigante Shepherd separated from the other clusters. When  $K$  became 7, Oropa Shepherd also formed a separated cluster, while only for  $K = 8$  Maremmano Sheepdog and Border Collie separated in two distinct clusters. As  $K$  increased to 10, Alaskan Malamute and Siberian Husky separated into single clusters, while Bergamasco Shepherd seems to be split in two different clusters.

### Discussion

The 10 breeds considered in this study were clearly genetically differentiated from each other, regardless of current population sizes and the onset of separate breeding history (Table 2). Our results are in agreement with findings of previous studies based on larger number of breeds which showed that artificial selection and breed management have determined clear genetic distinctions among breeds (Parker *et al.*, 2004 and 2007). The among-breeds genetic diversity

explained 20% of the total genetic variation, which is quite similar to the results of other dog studies, where values ranged from 18% to 27% (Irion *et al.*, 2003; Koskinen, 2003; Parker *et al.*, 2004). The observed levels of differentiation among dog breeds are higher than values that were found in other domestic species like horses (11%), donkeys (11%), bovines (7%) and pigs (14%) (Canon *et al.*, 2001; Boitard *et al.*, 2010; Bigi and Perrotta, 2012; Colli *et al.*, 2012). Most of the different pure-breed dogs are morphologically distinct and also differ in behavior and physical properties, so the higher level of differentiation in dogs is probably due to the lower utilization of crossbreeding in this species than with other species of domestic animals.

The level of heterozygosity found here (average  $H_E = 0.64$  among dog breeds) is in agreement with that found by other studies (e.g. Irion *et al.*, 2003, 0.39 to 0.76; Bjornerfeldt *et al.*, 2008, 0.40 to 0.77). The native Italian breeds showed generally higher genetic diversity ( $H_E$  ranging from 0.62 to 0.77) compared with the long established, well-defined cosmopolitan dog breeds. Ciampolini *et al.* (2011) found a similar  $H_E$  value (0.63) in the Bracco Italiano breed. The main reason of these findings may be the highly variable gene pool for the Italian breeds and the intense artificial selection, inbreeding and low number of founders for cosmopolitan breeds. Moreover many cosmopolitan breeds, selected originally for performance are now bred mainly for conformation. As reported by Pedersen *et al.* (2013), selection for conformation induces high inbreeding, while selection for performance is more appropriate to maintain diversity. Performance traits are much less heritable than the phenotypic traits. Conformation characteristics such as size, coat color and body shape usually depend on single or small groups of genes and can be easily fixed by inbreeding. On the other hand, performance traits are genetically complex and heritability is highly variable depending on the particular trait.

The highest  $H_E$  values and the highest  $AR$  were recorded in the Maremma Sheepdog, but then a significant ( $P < 0.001$ )  $F_{IS}$  indicated a certain level of inbreeding. MSD is characterized by a good number of puppies entered every year in the Italian Studbook (780; year 2013) with the presence of numerous breeding kennels, so the high level of  $H_E$  in this breed could be due to population consistency and the considerable number of breeders. Also the Bergamasco Shepherd showed a significant ( $P < 0.05$ )  $F_{IS}$  value. A finding like this could have several explanations other than inbreeding, for example the inclusion of closely related individuals (although this influence was minimized by our sampling strategy), hidden population structure or the widespread use of a small number of preferred sires. Both the breeds, MSD and BER, are selected to obtain typical, show winning, specimens leading to an increase in inbreeding levels. Also the Lupino del Gigante Shepherd showed a moderate level of inbreeding probably due to the small population size. In contrast, the negative significant ( $P < 0.001$ )  $F_{IS}$  value found in Oropa Shepherd, in spite of the small population size, indicated in this breed an excess of

heterozygosity, likely due to a documented crossbreeding with other shepherd breeds, in the last decades. Moreover ORO is characterized by a variegated area of origin being present on the plain, in the hills and in very high mountain valleys. It can herd both sheep and cows. These factors lead to slightly different morphological types with different genetic assets.

In several microsatellites, a total of 21 putative private alleles occurred only as a single allele in the whole population (alleles 143 and 157 in ATH137, alleles 80, alleles 114 and 116 in AHTH121, allele 239 in AHTH171, allele 99 in AHTH211, allele 300 in AHTK253, alleles 130 and 242 in CXX279, allele 248 in FH2848, allele 180 in FH2054, alleles 112 and 116 in INU005, allele 164 in INU030, allele 198 and 216 in REN162C04, alleles 156 and 178 in REN169018, allele 166 in REN169D01, allele 274 in REN247M23 and allele 224 in REN54P11). As the incidence of these alleles was so rare, it was not possible to rule out that their identification resulted from a sampling artifact.

The Lupino del Gigante Shepherd exhibited the highest number of private alleles ( $N_{PA} = 12$ ) with a normal level of genetic diversity (as measured by  $H_E$  and  $AR$ ). The high differentiation of this dog population from other breeds is also confirmed by the early segregation in structure, which suggests strong effects of reproductive isolation in the history of the breed.

The Lupino del Gigante Shepherd and the Oropa Shepherd showed a common cluster in the Neighbor-net (Figure 2). The structure analyses partially confirmed this finding, as when  $K = 4$ , 66% of the ORO and 94% of LGS individuals shared the same cluster. The morphological characteristics of these two breeds are similar and both breeds are utilized for the same purpose in two different but not distant geographic areas. A common origin of the two breeds seems probable as in the past, the traditional transhumance created many occasions of exchange between flocks coming from different areas. The Bergamasco Shepherd did not cluster with other breeds in the Neighbor-net, while in structure analysis, this breed is closer to the Oropa Shepherd, when  $K = 6$ . A within-breed substructure was detected when  $K = 10$  as this breed split into two main groups. Both BER and ORO originated in the Alps where pasture areas for sheep are overlapping and so cross mating between breeds is clearly possible. Considering the reduced number of BER breeding kennels, the subgroups differentiation could be done with different foundation stocks and with continuous in-kennel selection, thus leading to genetic drift and differentiation. Despite the close resemblance of some Lupino Del Gigante specimens to spitz-like breeds, ALM and SHU do not show any close genetic relationship with this breed, neither does the CWD. The Maremma Sheepdog groups with the Border Collie in the Network (Figure 2) and also the results of the Bayesian approaches show a common cluster for these two breeds when  $K = 6$ . MSD and BCO have different origins, are quite different morphologically and are also utilized for different purposes, as the MSD is a guard dog and the BCO is a herding dog. However this unexpected finding is

in agreement with the results of another study (Parker *et al.*, 2007), that considered 132 different dog breeds, where the Border Collie clustered together with the Kuvasz. This last breed is a Hungarian dog, that is similar to the Maremma Sheepdog, it is utilized for the same purpose and may have a common ancestor, together with other European dogs, like Tatra from Poland, Sarplaninac from the Balkans and the Pyrenean Mountain Dog (Breber, 1990). In general BCO seems closer to the Italian breeds than the other two cosmopolitan dogs considered here with the exception of LGS, which both in structure and in Neighbor-net analyses seems closer to GSD and to CWD. A possible utilization of BCO to improve working performance in Italian traditional working shepherd dogs cannot be ignored.

To conclude, the genetic characterization, in spite of the small population size, showed relatively high genetic diversity in Italian native shepherd dog breeds considered, which is important to maintain. Therefore, effective and proper breeding management schemes in these dog breeds will be advisable in order to avoid the excessive increase of inbreeding, thus avoiding in significant inbreeding depression and in significant loss of genetic variation. As these dog breeds have an active role in the sheep and cattle farming practiced in mountainous disadvantaged area, it is important to preserve their breeding and diffusion.

## Acknowledgments

Samples were provided by numerous dog breeders. The authors thank Cristian Lelli for the Lupino del Gigante and Dr Luca Sala and 'Amici del Cane Pastore di Oropa' Association for supplying the samples.

## References

- Altet L, Francino O and Sanchez A 2001. Microsatellite polymorphism in closely related dogs. *Journal of Heredity* 92, 276–279.
- Bigi D and Perrotta G 2012. Genetic structure and differentiation of the Italian Catria horse. *Journal of Heredity* 103, 134–139.
- Bjornerfeldt S, Hailer F, Nord M and Vila C 2008. Assortative mating and fragmentation within dog breeds. *BMC Evolutionary Biology* 8, 28.
- Boitard S, Chevalet C, Mercat MJ, Meriaux JC, Sanchez A, Tibau J and Sancristobal M 2010. Genetic variability, structure and assignment of Spanish and French pig populations based on a large sampling. *Animal Genetics* 41, 608–618.
- Breber P 1990. Il Cane da Pastore Maremmano-Abruzzese. Nuova Cinofilia – Editoriale Olimpia, Firenze, Italy.
- Budowle B, Garofano P, Hellman A, Ketchum M, Sree S, Kanthaswamy S, Parson W, Van Haeringen W, Fain S and Broad T 2005. Recommendations for animal DNA forensic and identity testing. *International Journal of Legal Medicine* 119, 295–302.
- Canon J, Alexandrino P, Bessa I, Carleos C, Carretero Y, Dunner S, Ferran N, Garcia D, Jordana J, Laloe D, Pereira A, Sanchez A and Moazami-Goudarzi K 2001. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genetics Selection Evolution* 33, 311–332.
- Cecchi F, Paci G, Spaterna A and Ciampolini R 2013. Genetic variability in Bracco Italiano dog breed assessed by pedigree data. *Italian Journal of Animal Science* 12, 348–352.
- Ciampolini R, Cecchi F, Bramante A, Casetti F and Presciuttini S 2011. Genetic variability of the Bracco Italiano dog breed based on microsatellite polymorphism. *Italian Journal of Animal Science* 10, 267–270.
- Ciampolini R, Cecchi F, Paci G, Policardo C and Spaterna A 2013. Investigation on the genetic variability of the American Pit Bull Terrier dogs belonging to an Italian breeder using microsatellite markers and genealogical data. *Cytology and Genetics* 47, 217–221.
- Colli L, Perrotta G, Negrini R, Bomba L, Bigi D, Zambonelli P, Verini Supplizi A, Liotta L and Ajmone-Marsan P 2012. Detecting population structure and recent demographic history in endangered livestock breeds: the case of the Italian autochthonous donkeys. *Animal Genetics* 44, 69–78.
- DeNise S, Johnston E, Halverson J, Marshall K, Rosenfeld D, McKenna S, Sharp T and Edwards J 2004. Power of exclusion for parentage verification and probability of match for identity in American kennel club breeds using 17 canine microsatellite markers. *Animal Genetics* 35, 14–17.
- Evanno G, Regnaut S and Goudet J 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14, 2611–2620.
- Excoffier L and Lisher HE 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10, 564–567.
- Excoffier L, Smouse P and Quattro J 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.
- Falush D, Stephens M and Pritchard JK 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Goudet J 2001. FSTAT. A program to estimate and test gene diversities and fixation indices. Version 2.9.3.2. Retrieved November 28, 2014, from <http://www.unil.ch/izea/software/fstat.html>
- Huson DH and Bryant D 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23, 254–267.
- Irion DN, Schaffer AL, Famula TR, Eggleston ML, Hughes SS and Pedersen NC 2003. Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *Journal of Heredity* 94, 81–87. *The Veterinary Journal* 196, 92–97.
- Jakobsson M and Rosenberg NA 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806.
- Kim KS, Tanabe Y, Park CK and Ha JH 2001. Genetic variability in East Asian dogs using microsatellite loci analysis. *Journal of Heredity* 92, 398–403.
- Koskinen MT 2003. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. *Animal Genetics* 34, 297–301.
- Leroy G, Verrier E, Meriaux JC and Rognon X 2009. Genetic diversity of dog breeds: within-breed diversity comparing genealogical and molecular data. *Animal Genetics* 40, 323–332.
- Marelli SP, Michelazzi M, Colombo E, Della Noce L and Guidobono Cavalchini L 2009. Phenotypical evaluation of "Cane da Pastore delle Alpi" canine population. In: *Proceedings of the ASPA 18th Congress*, 9 to 12 June, 2009, Palermo, Italy, p. 215.
- Mellanby RJ, Ogden R, Clements DN, French AT, Gowa AG, Powell R, Corcoran B, Schoeman JP and Summers KM 2013. Population structure and genetic heterogeneity in popular dog breeds in the UK. *The Veterinary Journal* 196, 92–97.
- Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA and Kruglyak L 2004. Genetic structure of the purebred domestic dog. *Science* 304, 1160–1164.
- Parker HG, Kukekova AV, Akey DT, Goldstein O, Kirkness EF, Baysac KC, Mosher DS, Aguirre GD, Acland GM and Ostrander EA 2007. Breed relationships facilitate fine-mapping studies: a 7.8-kb deletion cosegregates with Collie eye anomaly across multiple dog breeds. *Genome Research* 17, 1562–1571.
- Peakall R and Smouse PE 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6, 288–295.
- Pedersen N, Liu H, Theilen G and Sacks B 2013. The effects of dog breed development on genetic diversity and the relative influences of performance and conformation breeding. *Journal of Animal Breeding and Genetics* 130, 236–248.
- Pires AE, Amorim IR, Ginja C, Gomes M, Godinho I, Simoes F, Oom M, Petrucci-Fonseca F, Matos J and Bruford MW 2009. Molecular structure in peripheral dog breeds: portuguese native breeds as a case study. *Animal Genetics* 40, 383–392.

## Bigi, Marelli, Randi and Polli

- Pritchard JK, Stephens M and Donnelly P 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Randi E, Hulva P, Fabbri E, Galaverni M, Galov A, Kusak J, Bigi D, Bolfikova BC, Smetanova M and Caniglia R 2014. Multilocus detection of wolf  $\times$  dog hybridization in Italy, and guidelines for marker selection. *PLoS One* 9 (1), e86409.
- Reynolds J, Weir BS and Cockerham CC 1983. Estimation for the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105, 767–779.
- Rice WR 1989. Analyzing tables of statistical tests. *Evolution* 43, 223–225.
- Rosenberg NA 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 4, 137–138.
- Streitberger K, Schweizer M, Kropatsch R, Dekomien G, Distl O, Fischer MS, Epplen JT and Hertwig ST 2011. Rapid genetic diversification within dog breeds as evidenced by a case study on Schnauzers. *Animal Genetics* 43, 577–586.
- Wade CM 2011. Inbreeding and genetic diversity in dogs: results from DNA analysis. *The Veterinary Journal* 189, 183–188.
- Weir BS and Cockerham CC 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- Wright S 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420.