Animal, page 1 of 8 © The Animal Consortium 2018
doi:10.1017/S1751731117003573

animal

Investigating the population structure and genetic differentiation of livestock guard dog breeds

D. Bigi[1a], S.P. Marelli[2a], L. Liotta[3], S. Frattini[2], A. Talenti[2], G. Pagnacco[2], M. Polli[2†] and P. Crepaldi[2]

[1]*Department of Agricultural and Food Sciences (DISTAL), University of Bologna, 40127 Bologna, Italy;* [2]*Department of Veterinary Medicine, University of Milan, 20133 Milan, Italy;* [3]*Department of Veterinary Science, University of Messina, 98168 Messina, Italy*

(Received 20 June 2017; Accepted 27 November 2017)

Livestock guarding dogs are a valuable adjunct to the pastoral community. Having been traditionally selected for their working ability, they fulfil their function with minimal interaction or command from their human owners. In this study, the population structure and the genetic differentiation of three Italian livestock guardian breeds (Sila's Dog, Maremma and Abruzzese Sheepdog and Mannara's Dog) and three functionally and physically similar breeds (Cane Corso, Central Asian Shepherd Dog and Caucasian Shepherd Dog), totalling 179 dogs unrelated at the second generation, were investigated with 18 autosomal microsatellite markers. Values for the number of alleles per locus, observed and expected heterozygosity, Hardy–Weinberg Equilibrium, F *stats, Nei's and Reynold's genetic distances, clustering and sub-population formation abilities and individual genetic structures were calculated. Our results show clear breed differentiation, whereby all the considered breeds show reasonable genetic variability despite small population sizes and variable selection schemes. These results provide meaningful data to stakeholders in specific breed and environmental conservation programmes.*

Keywords: dog, livestock guardian breeds, breed differentiation, population structure, genetic variability

Implications

The work will provide essential information about the genetic structure and uniqueness of some important Italian dog breeds. These populations play a pivotal role for the pastoral community, helping them in the management of the livestock resources on field. These data will help the improvement of the current and future breed conservation programmes.

Introduction

The dog was the first animal to be domesticated (Vonholdt *et al.*, 2010) and has been successfully employed as a model organism for the study of mammalian size, shape, behaviour and naturally occurring disease risk studies (Vilà *et al.*, 1997). Domestic dog breed differentiation traces back to the 19th century (Gagliardi *et al.*, 2011), where breeds can be defined as intraspecific groups showing uniform morphological and behavioural characteristics developed by artificial selection (Irion *et al.*, 2003). Livestock guarding dogs are an adjunct to the pastoral community, selected to fill the role of protecting agricultural flocks in the absence of direction or oversight by a human shepherd. They are traditionally dogs with a large body size, and display early juvenile motor patterns that persist longer during development than in other breeds, whereas also characterized by a weakened stereotypical behavioural sequence in adults (Coppinger and Schneider, 1995). In these breeds many regional phenotypic groups can be identified, with a primary driver of selection corresponding to coat colour (Coppinger and Schneider, 1995). Italy has a long tradition in the sheep industry and is home to many different livestock guardian dog populations devised to succeed in varied conditions. Some of these breeds are officially recognized by Fédération Cynologique Internationale (FCI), whereas others are locally identified landraces (Table 1). The long-standing practice of transhumance, the seasonal movement of agricultural flocks between grazing lands along traditional routes, allowed a continuous genetic flow of genes among and between guardian dog breeds working in overlapping areas (Ceh and Dovc, 2014).

Autosomal microsatellites are a well-known effective and powerful tool to investigate genetic structure and diversity of dog breeds as already proved by many authors (Parker, 2012; Bigi *et al.*, 2015; Sechi *et al.*, 2016). Likewise, gene flow within breeds and detection of breed hybridization have been

a These two authors contributed equally to this work.
† E-mail: michele.polli@unimi.it

Table 1 *Information about the six dog breeds: Code, Fédération Cynologique Internationale (FCI) classification, height at withers according to official standard, coat type, coat colour, in the notes the links within the breeds and the inclusion in the research panel are reported*

Breeds	Breed code	FCI classification	Height at withers	Coat	Colour	Notes
Central Asia Shepherd Dog	CAD	Group 2 Pinscher and Schnauzer type, Molossoid Breeds, Swiss Mountain- and Cattle Dogs. Section 2.2. Molossoid Breeds, Mountain type. Without working trial	Males: minimum 70 cm; females: minimum 65 cm	Abundant, straight coarse and with well-developed undercoat	Any, except genetic blue and genetic brown in any combination and black mantel on tan.	Probable crosses with local breeds. Specific behavioural characteristic
Caucasian Shepherd dog	CSD	Group 2 Pinscher and Schnauzer type- Molossian and Swiss Mountain and Cattle Dogs. Section 2.2 Molossian/Mountain type. Without working trial.	Males: desirable height 72 to 75 cm; minimum 68 cm; females: desirable height 67 to 70 cm; minimum 64 cm	Straight, coarse, stand-off coat with well-developed undercoat. The length of guard coat as well as the undercoat should not be less than 5 cm	Any solid colour, piebald or spotted colour. Except for solid black; diluted black or black in any combination or genetic blue or liver brown colour	Probable crosses with local breeds. Specific behavioural characteristic
Italian Corso Dog	COR	Group 2 Pinscher and Schnauzer, Molossian and Swiss Mountain- and Cattledogs. Section 2 Molossoide breeds, Mastiff type. With working trial	Males: 64 to 68 cm. Females: 60 to 64 cm. With a tolerance of 2 cm, more or less taller	Short, shiny, very dense with a slight undercoat of vitreous texture	Black, lead-grey, slate-grey, light grey, light fawn; dark fawn and stag red; dark wheat colour (etc.)	Overlapping diffusion areas. Traditionally farm dog working with livestock too
Maremma and the Abruzzes Sheepdog	MSD	Group 1 Sheepdogs and cattle dogs except Swiss cattledogs. Section 1 Sheepdogs. Without working trial	Males: 65 to 73 cm. Females: 60 to 68 cm.	Very well furnished. Hair long, rather harsh to the touch, close to straight horsehair	Solid white. Shades of ivory, pale orange or lemon is tolerated but only in certain limits	Probable crosses with local breeds. Livestock guarding, overlapping diffusion areas
Mannara Dog	MAN	Group 2 Pinscher and Schnauzer type, Molossoid Breeds, Swiss Mountain- and Cattle Dogs. Section 2.2. Molossoid Breeds, Mountain type. Without working trial	Males: minimum 65 cm. Females: minimum 59 cm	Hair long , dense and close, no skin should be visible, presence of undercoat	Fawn:from cream to mahogany tan, solid black, brindle, light brown, black and tan, particolour with white	Local breed
Sila Dog	SIL	Group 1 Sheepdogs and cattle dogs except Swiss cattledogs. Section 1 Sheepdogs. Without working trial	Males: 65 to 73 cm. Females: 60 to 68 cm	Abundant, long, coarse, flattening on the body, light waving admitted	Black and tan, black and silver, wolf grey and tan, silver wolf grey and tan, solid black, brindle in combination with other colours too	Local breed

successfully demonstrated with microsatellites (Kopaliani *et al*., 2014). The aim of this study is to analyse the population structure and genetic differentiation of Italian livestock guardian breeds (Maremma and Abruzzese Sheepdog (MSD); Mannara's Dog (MAN); Sila's Dog (SIL)) and selected breeds that share original regions and working areas (Cane Corso (COR)) or morphological and behavioural characteristics (Central Asian Shepherd Dog (CAD); Caucasian Shepherd Dog (CSD)).

There is an increasing worldwide interest in identifying native genetic resources to be considered as important targets for conservation projects (Lee *et al*., 2014). The three Italian livestock guardian breeds in the present work should be considered genetic resources with high cultural value due to their ancient origin, strictly linked to well-established traditional sheep and goat farming practices in marginal areas, which are now disappearing in favour of highly productive farms based on intensive rearing conditions. We describe the genetic makeup of the studied breeds. Through characterizing these local flock guardian dogs as distinct regionally and culturally relevant breeds, we will provide breeders and conservationists with valuable genetic data for enhanced preservation.

Material and methods

Sampling and DNA extraction

A total of 179 blood samples from 59 MAN, 22 MSD, 36 COR, 26 SIL, 24 CAD and 12 CSD dogs were collected by Italian Kennel Club (ENCI) partner laboratories for the official pedigree certification scheme and field inspections according to European rules for animal welfare (Council of Europe, 1986). Three-generation pedigrees of dogs belonging to FCI officially recognized breeds (COR, MSD, CAD, CSD) and supplementary studbook registered breeds (MAN, SIL) were obtained from the online Italian Kennel Club studbooks. All sampled dogs were unrelated for at least two generations. DNA was extracted from whole EDTA blood or DNA sample cards (Vet Kard with FTA System®; Prion Diagnostica srl, Rho, MI, Italy), using IllustraTM 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 18 microsatellite markers mapping on 17 different autosomes (*AHT121* (map location on CFA13), *AHT137* (CFA11), *AHTH171* (CFA6), *AHTH260* (CFA16), *AHTK211* (CFA26), *AHTK253* (CFA23), *CXX279* (CFA22), *FH2054* (CFA12), *FH2848* (CFA2), *INRA21* (CFA21), *INU005* (CFA33), *INU030* (CFA12), *INU055* (CFA10), *REN162C04* (CFA7), *REN169D01* (CFA14), *REN169O18* (CFA29), *REN247M23* (CFA15) and *REN54P11* (CFA18)). Gender was confirmed using an amelogenin marker (*AMELY* and *AMELX*) that was included in the Canine Genotypes TM Panel 1.1, F-860S/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 Companion Animals of the International Society for Animal Genetics (Budowle *et al*., 2005). Polymerase chain reaction amplification and electrophoresis (ABI Prism® 310 Genetic Analyzer; Thermo Fisher Scientific, Waltham, MA, USA) were carried out following the manufacturer's protocols. Individual genotypes were determined using the software Genescan® 3.7 (ABI).

Data analysis

The software Arlequin v.3.5 (Excoffier and Lischer, 2010) was used to calculate the number of alleles per locus, observed heterozygosity (H_O), and expected heterozygosity (H_E) and test for Hardy–Weinberg equilibrium (HWE). Arlequin v.3.5 was also used to compute Wright's *F* statistics, where F_{IS} is the inbreeding coefficient of the individual relative to the breed, F_{IT} is the inbreeding coefficient of the individual relative to the total population, and F_{ST} is the differentiation between breeds (Weir and Cockerham, 1984). In addition we, calculated the pairwise R_{st} index, which provides more adequate measures of differentiation based on microsatellites markers (Slatkin, 1995). Multiple comparisons were corrected to statistical Bonferroni significance (Rice, 1989) and tested using 1000 permutations. To evaluate the significance of genetic differentiation between populations an analysis of molecular variance (AMOVA; Excoffier and Lischer, 2010) using the exact test implemented in ARLEQUIN software was performed. Allelic richness (AR) was computed in FSTAT to reduce the effect of a different sample size (Goudet, 2001).

Genetic divergence among breeds was estimated by Nei's standard genetic distance (*D*) (Nei, 1972) and Reynolds' unweighted distances (D_R) (Reynolds *et al*., 1983), using Genedist (Brzustowski, 2001) software. The Phylip v.3.65 package (Felsenstein, 1989) was used to construct phenograms based on neighbour-joining analysis of Nei's and Reynolds's distances. Support for the tree nodes was assessed by 1000 bootstraps of gene frequencies using Seqboot and compiled with Consense, both functions of Phylip v.3.65 (Felsenstein, 1989).

Animals were clustered on the basis of their genotypes using a Bayesian approach in Structure 2.3.4 (Pritchard *et al*., 2000), using the population as prior information in the analysis. Using a model with admixture and correlated allele frequencies, we performed ten independent runs for each value of the putative number of subpopulations (K) between 1 and 10, with a burn-in period of 50 000 followed by 500 000 MCMC repetitions. The most likely number of clusters in our sample was determined using ΔK as described by Evanno *et al*. (2005). CLUMPP (Jakobsson and Rosenberg, 2007) was used to match the data from the multiple runs for each K, and DISTRUCT (Rosenberg, 2003) was used to display the results. Main splits were analysed to detect the relationships among lineages.

To visualize the genetic structure of individuals, a principal components analysis (PCA) based on inter-individual genetic distance in GenAlex 6 (Peakall and Smouse, 2012) was also conducted. Genetic distances were calculated as described in the GenAlex 6 guide for co-dominant data.

Results

Genetic variability

All samples were successfully genotyped with a total of 185 alleles detected across the 18 microsatellite loci analysed. All of the dog breeds were polymorphic across the microsatellite panel. The average number of alleles per locus was 9.1 and the number of alleles ranged between seven (*AHTK211* and *INRA21*) and 17 (*AHTH121*) (Table 2). The private alleles observed in the studied breeds are reported in the Supplementary Table S1. The majority of the private alleles in the six breeds were below 0.05 in frequency; allele 174 of locus FH2054 in COR had a frequency of 0.157.

Average estimates of H_O and H_E over all loci and breeds were 0.72 and 0.72 (Table 3), with the lowest values detected in COR ($H_E = 0.68$; $H_O = 0.65$) and the highest in CAD ($H_E = 0.76$; $H_O = 0.74$). Among the six dog breeds considered, the lowest value of AR was found in the COR (5.11) and the highest in the CAD (6.85). The HWE exact test revealed 21 significant deviations (19.4%) between H_E and H_O after Bonferroni correction. H_O was significantly higher than H_E in three cases; in all other cases, H_E was higher than H_O (Table 3).

The overall F_{IS} value among all loci was significantly higher than zero after Bonferroni correction in the MSD (0.10) indicating heterozygote deficiency in this breed.

A total of 38 private alleles were recorded. All the populations, except for CSD, showed at least one private allele, with the CAD having the highest number of private alleles ($N_{PA} = 16$).

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

Locus	N_A	N_{PA}	F_{IS}	F_{ST}	F_{IT}
AHTH121	17	6	0.044	0.057***	0.098***
AHT137	13	1	0.023	0.093***	0.113***
AHTH171	11		0.058	0.077***	0.131***
AHTH260	9	1	0.057	0.158***	0.206***
AHTK211	7		−0.014	0.078***	0.065
AHTK253	8	2	−0.004	0.074***	0.070
CXX279	12	6	0.095*	0.124***	0.208***
FH2054	11	3	0.031	0.078***	0.107***
FH2848	10	2	0.043	0.102***	0.141***
INRA21	7		0.013	0.043***	0.056
INU005	9	1	−0.050	0.049***	0.001
INU030	8	2	−0.003	0.069***	0.065
INU055	9	2	0.005	0.142***	0.146**
REN162C04	11	4	−0.015	0.144***	0.132**
REN169D01	12	2	−0.038	0.051***	0.015
REN169018	9		0.031	0.024*	0.055
REN247M23	10	3	0.057	0.122***	0.173***
REN54P11	12	3	0.099**	0.113***	0.201***
Average	9.1	2.1	0.024*	0.089***	0.011***

N_A = number of alleles; N_{PA} = number of private alleles; F_{IS} and F_{IT} = measurements of the deviation from Hardy–Weinberg proportions within populations and in the total population, respectively; F_{ST} is the genetic differentiation over subpopulations.
* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Relationships among breeds

Genetic differentiation among breeds was highly significant ($P < 0.001$) for all loci, as revealed by AMOVA. The average F_{ST} values indicate that ~9% of the total genetic variation was explained by breed differences, with the remaining 91% corresponding to differences among individuals. A significant excess of homozygotes across all breeds was found for *CXX279* and *REN54P11* ($P_{CXX279} = 0.018$; $P_{REN54P11} = 0.006$). Even if these excess of homozygosity could be also related to the microsatellites technique, we decided to keep them in the subsequent analyses.

All pairwise F_{ST} values were significantly different from zero ($P < 0.001$), ranging from 0.058 between the MSD and the SIL to 0.166 between the COR and the CSD (Table 4). R_{ST} statistic gave us similar average values, although ranging from 0.032 to 0.139 (Table 5). The rooted bootstrapped neighbour-joining phylogeny with Nei's *D* produced a dendrogram (Figure 1a) with CAD and CSD adjacent (89% confidence).On a separate branch are the MSD and SIL monophyletic clades (61% confidence). Finally, the COR branch is outside of this primary clade with confidence levels below 50%. The D_R analysis produced the same phylogenetic arrangement of breeds, with comparable confidence values (Figure 1b).

The first three axes of the principle coordinate analysis account for 62.69% of the total genetic variance between individual dogs (Figure 2). Coordinate 1 (24.95%) separates the COR from the flock guardian breeds. The second coordinate (20.71%) separates the MAN from the remaining five breeds. Coordinate 3 (17.03%) isolates the SIL from the remaining breeds.

Clustering using Bayesian approaches was performed on the entire data set with an increasing number of inferred clusters from $K = 2$ to $K = 10$ and produced consistent results. ΔK analysis showed the primary peak when $K = 3$, a secondary peak was obtained at $K = 7$ (Supplementary Figure S1). Therefore, we analysed the sub-sequential splits from two to eight to identify the strict relationships among breeds (Figure 3). The initial division of $K = 2$ identifies a signature primarily in COR but also present in MAN. The $K = 3$ division defines a strong COR signature and a new signature in most, but not all, MAN. When $K = 5$, the SIL presents as a unique grouping. For $K = 7$, the CAD, COR, MSD and SIL appear to have predominantly breed-specific signatures, whereas substructure is apparent within the MAN and CSD breeds.

Discussion

A clear differentiation of the six studied breeds has been shown at the nuclear DNA level, with all 18 microsatellites genotyping as polymorphic in the analysed breeds. The average number of alleles ($N = 9.1$) corresponds to that calculated by Suárez and colleagues in a 2013 investigation of the genetic makeup of Canary Island dog breeds, both FCI recognized and local. The number of private alleles shows the difference between breeds showing similar results to

Table 3 *Basic information and genetic diversity parameters of six dog breeds*

Breeds	Breed code	Reference population size	Sample size	N_{PA}	*AR*	H_O	H_E	F_{IS}	HWE deviations[1]
Central Asia Shepherd Dog	CAD	7890	24	16	6.85	0.74	0.76	0.04	−3
Caucasian Shepherd Dog	CSD	2967	12	0	5.19	0.68	0.70	0.04	−3
Cane Corso	COR	32 717	36	4	5.11	0.68	0.65	0.01	+1/−2
Mannara's Dog	MAN	83	59	5	5.63	0.75	0.73	−0.01	+2/−4
Maremma Sheep Dog	MSD	12 000	22	5	5.75	0.69	0.75	0.10**	−3
Sila's Dog	SIL	134	26	8	5.37	0.73	0.71	−0.01	−3

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; CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.
Reference Population Size: ENCI studbook entries: 2007–2016.
[1]Negative and positive values indicate the number of loci showing heterozygote deficiency and excess, respectively.
**$P < 0.01$.

Table 4 F_{ST} *estimates*[1] *as a measure of genetic distance between dog breeds*

	CAD	CSD	COR	MAN	MSD	SIL
CAD	–					
CSD	0.063	–				
COR	0.109	0.166	–			
MAN	0.047	0.103	0.102	–		
MSD	0.054	0.102	0.098	0.075	–	
SIL	0.080	0.122	0.125	0.078	0.058	–

CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.
All pairwise F_{ST} differences were significantly larger than 0 ($P < 0.001$).
[1]F_{ST} estimates calculated as described in Weir and Cockerham (1984).

Table 5 R_{ST} *estimates*[1] *as a measure of genetic distance between dog breeds*

	CAD	CSD	COR	MAN	MSD	SIL
CAD	–					
CSD	0.052	–				
COR	0.061	0.102	–			
MAN	0.041	0.081	0.100	–		
MSD	0.036	0.069	0.070	0.032	–	
SIL	0.108	0.088	0.139	0.086	0.084	–

CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.
[1]R_{ST} estimates calculated as described in Slatkin (1995). All pairwise R_{ST} differences were significantly larger than 0 ($P < 0.01$).

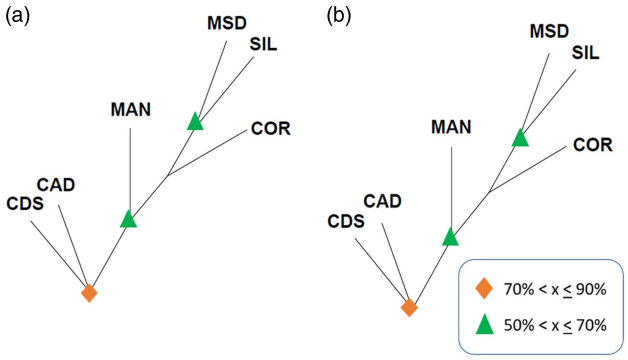

Figure 1 (colour online) Neighbour-joining trees of the six dog populations built from (a) Nei's standard genetic distance (*D*) (b) Reynolds unweighted distances (*DR*). Symbols indicate the ratio of 1000 bootstrap replications across loci. CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.

those presented by Irion *et al.* (2005) about Bali street dog originality. The *F* statistics of these Italian livestock dogs revealed a broad variation among loci and breeds. A high across-breed average F_{ST} value indicates greater divergence of breeds within the selected population. A lower across-breed average F_{ST} value would imply that the included breeds are more closely related to each other. The genetic divergence of autosomal microsatellite markers in the present Italian flock guardian breeds resulted in an across-breed average F_{ST} of 0.09, a comparable value to that found in pointing ($F_{ST} = 0.11$; Parra *et al.* 2008), Spanish ($F_{ST} = 0.10$; Jordana *et al.* 1992) and Asian ($F_{ST} = 0.15$; Kim *et al.* 2001) dog breeds. R_{ST} statistics showed similar average values and patterns. A recent analysis of the genetic makeup of Italian shepherd dog breeds, of which the MSD was also included, showed higher average F_{ST} value of 0.20 (Bigi *et al.*, 2015). Amongst the livestock guardian breeds of the present study, only the MSD showed a significant F_{IS} ($P \leqslant 0.01$), consistent with previously reported data (Bigi *et al.*, 2015). This indicates that the MSD probably represents the most uniform of the considered breeds, with the most ancient official recognition and the highest population size.

The breed-to-breed F_{ST} values estimate pairwise genetic differentiation. In regard to the livestock guardian breeds of the present study, all pairwise F_{ST} values were highly significant ($P \leq 0.001$). These results highlight the significant differences between the studied breeds. The lowest F_{ST} value occurred between MAN and CAD (0.047) revealing a probable past shared genetic history. On the other hand, the highest F_{ST} (0.166) was recorded between COR and CSD, confirming the historical expectation of these two breeds having developed independently.

The present breeds show high H_O values, ranging from 0.68 to 0.75. The observed values are closer to those reported by Pedersen *et al.* (2012) for village dogs and working gundog breeds ($H_O = 0.63$ to 0.80) than for breeds selected

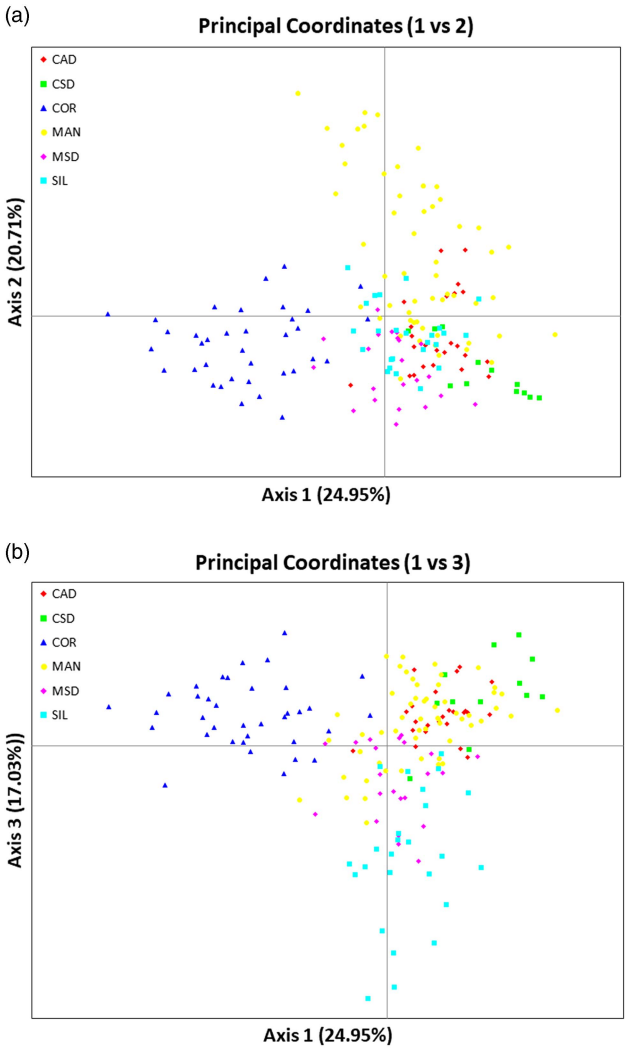

Figure 2 (colour online) (a) Principal component analysis of allele frequencies for six dog populations. The projection is shown on the first two axes. (b) The same samples as in (a) were plotted using dimension coordinates 2 *v.* 3. CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.

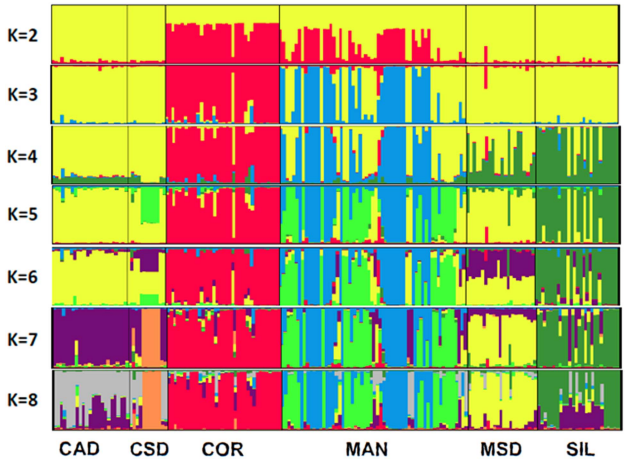

Figure 3 (colour online) Proportion of membership of 179 individuals from six dog breeds for $K = 2$ to 8, as calculated by Structure software. Multiple runs for each K were concatenated using Clumpp, and Distruct was used to generate images. Each individual, represented as a vertical line is partitioned into coloured segments whose length is proportional to the individual's probability of assignment (Q) to the Kth group. CAD = Central Asia Shepherd Dog, CSD = Caucasian Shepherd Dog, COR = Corso, MAN = Mannara Dog, MSD = Maremma Sheep Dog, SIL = Sila Dog.

for conformation events ($H_O = 0.43$ to 0.57). Previous studies, using comparable numbers of microsatellite markers, of working lines of CAS and CSD reported a mean H_O of 0.78 (Kopaliani *et al.*, 2014), whereas the mean H_O of 12 modern breeds was found to be 0.67 (Streitberger *et al.*, 2012). The results published by Ceh and Dovc (2014) regarding flock guardian breeds from the Balkans are likewise comparable (H_O 0.59 to 0.76). However, the Norwegian Lundehund, a very rare breed characterized by strong genetic erosion and high inbreeding level, has an H_O of 0.05 (Pfahler and Distl, 2013).

The breeds utilized in this study continue to be raised locally to fulfil their intended guardian purpose, therefore, selection has not been substantially influenced by scores obtained through canine conformation competitions. All of the study dogs were, however, registered as purebred animals with ENCI, either as a recognized breed or through preliminary studbook recording, both of which require a minimum of three purebred generations of unique individuals. However, as international recognition of breed status dates to between 1956 (MSD) and 2004 (COR) for each of the four internationally recognized breeds, admixture between similar breeds of common function and geography before those dates is likely. High heterozygosity could be considered a characterizing trait in working breeds, sheep guardians in particular, who are selected mainly for working attitude and natural morphology, limiting inbreeding coefficients and the common sire effect. Breed clubs and associations are interested in limiting inbreeding levels increase even though no specific rules have been officialised by the Italian Kennel Club, furthermore a high effective population size of the founder groups could be supposed. Even though the six breeds have very different population sizes (Table 3), they all maintain high heterozygosity levels. The CAD breed shows the highest number of private alleles and AR, consistent with previously published data (Kopaliani *et al.*, 2014). This reflects its ancient origin close to one of the main domestication hubs (Vonholdt *et al.*, 2010).

The genetic differentiation of the six breeds is shown in Figures 1a and b which report consistent results, a clear cluster separating the CAD and the CSD breeds of Western Asian origin from the Italian breeds, underlining their unique genetic composition (Ceh and Dovc, 2014; Kopaliani *et al.*, 2014). The four Italian breeds considered maintain their phylogenetic differentiation regardless of analysis involving genetic distance or allele frequency.

The first coordinate of the PCA (Figure 2a and b) confirms our hypothesis that the COR is particularly divergent from the flock guardian breeds. However, coordinates two and three isolate the MAN and SIL, respectively. The CAD and the CSD, once again, consistently cluster together, connoting some shared genetic

component. In similar analyses, the Perro da Presa Canario, a breed with shared morphological and behavioural phenotypes to the COR, was likewise shown to differentiate from sheepdog breeds of the Canary Islands (Suárez *et al*., 2013).

The population structure analyses show clear breed differentiation that accompanies the inheritance of phenotypical traits where natural and/or artificial breed barriers are present. In this way, the similarity within a breed is higher than between breeds even if dogs belonging to different breeds were living in overlapping areas and were sharing the same function and working ability (Suárez *et al*., 2013). The first breed to differentiate from the others ($K = 2$) is the COR, again reproducing the pattern of breed relatedness reported by Suárez *et al*. (2013). The COR is an internationally recognized breed selected for conformation and working attributes. Conversely, SIL and MAN, that separate through STRUCTURE analysis at $K = 4$ and $K = 5$, are breeds which are still closely linked to their traditional agricultural roles and for which selection is heavily weighted for behaviour rather than appearance. The highest ΔK was calculated to be at $K = 3$, even though at $K = 5$ all the Italian livestock guard dog breeds are differentiated. Our results are comparable with those obtained by Ceh and Dovc (2014) in Balkan breeds underlining the strong commonalities linking together the livestock guarding dog breeds throughout Europe.

In the present study, the genetic characteristics of six livestock guardian breeds were analysed to show how all the considered breeds are genetically different. None of the considered breeds, even if characterized by small population size, seems to be victim of phenotypic selection leading to genetic drift, loss of alleles or high inbreeding values. Reported results can assist stakeholders in improving conservation programmes tailored to the specific needs and characteristics of each breed and livestock production system. The diffusion of these breeds should be strongly encouraged to prevent not only the extinction of highly historical and culturally valuable Italian livestock farming dog breeds, but also the abandoning of marginal area farming.

Acknowledgements

The authors are grateful to Elaine A. Ostrander and Dayna L. Dreger for their editing and suggestions to this manuscript. The authors are also grateful to the Italian Kennel Club (ENCI) and VETOGENE for the information provided, to the Mannara breeders club and, finally, to Isabella Biafora and the Sila breeders club.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731117003573

References

Bigi D, Marelli SP, Randi E and Polli M 2015. Genetic characterization of four native Italian shepherd dog breeds and analysis of their relationship to cosmopolitan dog breeds using microsatellite markers. Animal 9, 1921–1928.

Brzustowski J 2001. GeneDist: population to population genetic distance calculator. Retrieved on 25 May 2017 from http://www2.biology.ualberta.ca/jbrzusto/GeneDist.php.

Budowle B, Garofano P, Hellman A, Ketchum M, 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.

Ceh E and Dovc P 2014. Population structure and genetic differentiation of livestock guard dog breeds from the Western Balkans. Journal of Animal Breeding and Genetics 131, 313–325.

Coppinger R and Schneider R 1995. Evolution of working dogs. In: The domestic dog: its evolution, behaviour and interactions with people. Cambridge University Press, Cambridge, UK. pp. 21–47.

Council of Europe 1986. European convention for the protection of vertebrate animals used for experimental and other scientific purposes CETS 123. Strasbourg. Retrieved on 18 December 2015 from http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/123.

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 Lischer HEL 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.

Felsenstein J 1989. PHYLIP – Phylogeny Inference Package (Version 3.2). pp. 164–166. Retrieved on 28 November 2014 from http://evolution.genetics.washington.edu/phylip/.

Gagliardi R, Llambí S, García C and Arruga MV 2011. Microsatellite characterization of Cimarron Uruguayo dogs. Genetics and Molecular Biology 34, 165–168.

Goudet J 2001. FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2.9.3. Retrieved on 25 May 2017 from https://www2.unil.ch/popgen/softwares/fstat.htm.

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.

Irion DN, Schaffer AL, Grant S, Wilton AN and Pedersen NC 2005. Genetic variation analysis of the bali street dog using microsatellites. BMC Genetics 6, 6.

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.

Jordana J, Piedrafita J, Sanchez A and Puig P 1992. Comparative F statistics analysis of the genetic structure of ten Spanish dog breeds. Journal of Heredity 83, 367–374.

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.

Kopaliani N, Shakarashvili M, Gurielidze Z, Qurkhuli T and Tarkhnishvili D 2014. Gene flow between wolf and shepherd dog populations in Georgia (Caucasus). Journal of Heredity 105, 345–353.

Lee E-W, Choi S-K and Cho G-J 2014. Molecular Genetic Diversity of the Gyeongju Donggyeong dog in Korea. Journal of Veterinary Medical Science 76, 1359–1365.

Nei M 1972. Genetic distance between populations. The American Naturalist 106, 283–292.

Parker HG 2012. Genomic analyses of modern dog breeds. Mammalian Genome 23, 19–27.

Parra D, Méndez S, Cañón J and Dunner S 2008. Genetic differentiation in pointing dog breeds inferred from microsatellites and mitochondrial DNA sequence. Animal Genetics 39, 1–7.

Peakall R and Smouse PE 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research –an update. Bioinformatics 28, 2537–2539.

Pedersen N, Liu H, Theilen G and Sacks B 2012. 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.

Pfahler S and Distl O 2013. A massive reduction of the genetic diversity in the Lundehund. Animal Genetics 45, 154.

Pritchard JK, Stephens M and Donnelly P 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Reynolds J, Weir BS and Cockerham CC 1983. Estimation of 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.

Rosenberg NA 2003. Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137–138.

Sechi S, Polli M, Marelli S, Talenti A, Crepaldi P, Fiore F, Spissu N, Dreger DL, Zedda M, Dimauro C, Ostrander EA, Di Cerbo A and Cocco R 2016. Fonni's dog: morphological and genetic characteristics for a breed standard definition. Italian Journal of Animal Science 16, 22–30.

Slatkin M 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457–462.

Streitberger K, Schweizer M, Kropatsch R, Dekomien G, Distl O, Fischer MS, Epplen JT and Hertwig ST 2012. Rapid genetic diversification within dog breeds as evidenced by a case study on Schnauzers. Animal Genetics 43, 577–586.

Suárez NM, Betancor E, Fregel R and Pestano J 2013. Genetic characterization, at the mitochondrial and nuclear DNA levels, of five Canary Island dog breeds. Animal Genetics 44, 432–441.

Vilà C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J and Wayne RK 1997. Multiple and ancient origins of the domestic dog. Science 276, 1687–1689.

Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkahloun A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, Ding Z-L, Zhang Y-P, Bustamante CD, Ostrander EA, Novembre J and Wayne RK 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464, 898–902.

Weir BS and Cockerham CC 1984. Estimating F-Statistics for the analysis of population structure. Evolution 38, 1358.