BMC Research Notes

CrossMark

Genome-wide diversity and runs of homozygosity in the "Braque Français, type Pyrénées" dog breed

Salvatore Mastrangelo[1†], Filippo Biscarini[2*†] , Barbara Auzino[6], Marco Ragatzu[3], Andrea Spaterna[4,5] and Roberta Ciampolini[5,6]

Abstract

Objective: Braque Français, type Pyrénées is a French hunting-dog breed whose origin is traced back to old pointing gun-dogs used to assist hunters in finding and retrieving game. This breed is popular in France, but seldom seen elsewhere. Despite the ancient background, the literature on its genetic characterization is surprisingly scarce. A recent study looked into the demography and inbreeding using pedigree records, but there is yet no report on the use of molecular markers in this breed. The aim of this work was to genotype a population of Braque Français, type Pyrénées dogs with the high-density SNP array to study the genomic diversity of the breed.

Results: The average observed (H_O) and expected (H_E) heterozygosity were 0.371 ($\pm$ 0.142) and 0.359 ($\pm$ 0.124). Effective population size (N_e) was 27.5635 runs of homozygosity (ROH) were identified with average length of 2.16 MB. A ROH shared by 75% of the dogs was detected at the beginning of chromosome 22. Inbreeding coefficients from marker genotypes were in the range $F_{IS} = [-0.127, 0.172]$. Inbreeding estimated from ROH (F_{ROH}) had mean 0.112 ($\pm$ 0.023), with range [0.0526, 0.225]. These results show that the Braque Français, type Pyrénées breed is a relatively inbred population, but with still sufficient genetic variability for conservation and genetic improvement.

Keywords: Dog, Braque Français, type Pyrénées, SNP, Genetic diversity, Molecular markers, Inbreeding, Runs of homozygosity, Heterozygosity

Introduction

Genetic variability and structure in domestic animals largely depend on breeders' decisions and practices. In selection, breeding within a closed population is common practice, since it allows to fix the desired characteristics and traits of the best representatives of the breed. However, this mating practice can lead to high rates of inbreeding and associated risks (higher frequency of recessive disorders, inbreeding depression), which are a serious threat especially to small populations and to populations originating from a small number of ancestors [1, 2]. Furthermore, intense directional selection for specialized animal types may result in a reduced genetic basis available to the populations, which leads to a dramatic loss of genetic variability, especially in dogs where mating between close relatives is frequently used [3, 4]. Concerns about the potential effects of inbreeding and reduced diversity on health, functionality and welfare of animals within dog breeds, have led to a call for improved genetic management practices. Hence, managing genetic diversity has become a major focus for dog breeders, herd books and authorities [5]. Studies have shown that there is a loss of the total amount of genetic diversity in modern dog populations [6–8]. In particular, purebred dogs have been intensely selected by resorting, in some lines, to close-breeding where popular dominant sires were repeatedly used for mating, resulting in a reduction of genetic diversity. Traditionally, genealogical data

*Correspondence: filippo.biscarini@ibba.cnr.it
†Salvatore Mastrangelo and Filippo Biscarini contributed equally to this work
[2] CNR-IBBA, Via Bassini 15, 20133 Milano, Italy
Full list of author information is available at the end of the article

BioMed Central

Mastrangelo *et al. BMC Res Notes* (2018) 11:13

Page 2 of 6

has been used to assess genetic diversity in dogs [3–5, 9, 10]. However, the use of genealogical data is limited by the incomplete or inaccurate available pedigree records. Genomics offer novel applications that have great potential to increase our understanding of the genome of domestic animals and to improve the efficiency of conservation and selection programs [11–13]. STR (short-tandem repeats) molecular markers have been used initially to estimate genetic diversity in the absence of pedigree records [9, 10, 14, 15]. More recently, the availability of high density single nucleotide polymorphism (SNP) arrays, has increased the accuracy, the throughput and the cost-effectiveness of genomic analyses for conservation genetics [16, 17]. Indeed, the large numbers of SNPs throughout the genome makes these markers particularly suitable for the detection of genomic regions where a reduction in heterozygosity occurred and offers new opportunities to improve the accuracy of genetic diversity estimates.

Braques Françaises are hunting dogs, originating from a very old type of gun-dog used for pointing the location of game birds to hunters. There are two breeds of Braque Français, both from the south of France: the Braque Français, type Gascogne (larger size) and the Braque Français, type Pyrénées (smaller size). They are popular hunting dogs in France, but are seldom seen elsewhere. The original Braque Français type of pointing dog has existed since the eighteenth century. The first breed club was formed in 1850, and the standards for both breeds were written in 1880. The demography and inbreeding levels of the Braque Français, type Pyrénées breed were estimated using pedigree records [18]. There is yet no report of the genetic characterization of this breed using molecular markers. The aim of this work was to study the genomic diversity of Braque Français, type Pyrénées dogs using data high-density SNP data. Available data, methods and results are hereby presented.

Main text

DNA sampling, genotyping and quality control

Blood samples were collected from 48 Braque Français, type Pyrénées unrelated individual dogs (27 females, 21 males), to capture a representative sample of the within-breed genetic diversity. Genomic DNA (gDNA) was extracted from blood samples through a standard ethanol fractionation with concentrated sodium chloride (6 M NaCl) and sodium dodecyl sulphate (10% SDS). The concentration of DNA was adjusted to 50 ng/μL per sample. All dogs were genotyped with the Illumina CanineHD BeadChip, containing 173,662 SNPs. Genotyping was carried out in the laboratories of "Dipartimento di Scienze Agrarie, Alimentari e Forestali", University of Palermo (Italy). SNP data were filtered to exclude unmapped loci

not assigned to any chromosomes and loci on the sex chromosomes: only SNPs located on the 38 autosomes were considered. Additionally, SNPs with call-rate < 95% and minor allele frequency (MAF) < 5% were removed from the dataset, as well as animals with over 10% missing genotypes.

Genetic diversity parameters, effective population size and runs of homozygosity (ROH)

Basic genetic diversity indices were estimated, including: (i) observed heterozygosity ($H_O = \frac{1}{L} \sum_{l=1}^{L} \left(\frac{n_{AB}}{N} \right)$); (ii) expected heterozygosity ($H_E = \frac{1}{L} \sum_{l=1}^{L} \left(1 - p_l^2 - q_l^2 \right)$); (iii) the inbreeding coefficient based on the difference between observed and expected homozygous genotypes ($F_{IS} = \frac{1}{L} \sum_{l=1}^{L} \left(1 - \frac{H_{Ol}}{H_{El}} \right)$, [19]); and (iv) minor allele frequency ($MAF_l = \frac{n_{B_l}}{2N_l}$). Here: L is the n. of SNP loci; n_{AB} is the number of heterozygous genotypes at each locus l; N is the number of individuals (sample size); p and q are the frequencies of, respectively, the A (major) and B (minor) alleles.

The contemporary effective population size (N_e) was estimated based on linkage disequilibrium (LD; [20]). To reduce the impact of SNP ascertainment bias from linkage between loci, unlinked SNPs were selected based on variance inflation factor (VIF, a measure of multicollinearity in multiple regression) below 2 in 50-SNP sliding window with 5-SNP step. A VIF threshold of 2 is recommended for small sample sizes to avoid removing too many SNPs (http://zzz.bwh.harvard.edu/plink/summary.shtml#prune).

Runs of homozygosity (ROH: segments of continuous homozygous genome, [21]) were detected in each individual dog using the sliding-window based method described in [22, 23]. A 50-SNP long sliding window was used to scan the genome; the proportion of overlapping homozygous windows to call a ROH was 0.05; maximum two missing and one heterozygous SNP were allowed in a run; the minimum number of SNPs to call a ROH (s) was calculated as proposed by Lencz et al. [24] to minimize the number of false positive ROH:

$$s = \frac{\ln \frac{\alpha}{N \cdot L}}{\ln(1 - \overline{H_O})} \tag{1}$$

where α is the tolerated proportion of false positive ROH (0.05 in the present study); N and L are the sample size and number of SNP loci; $\overline{H_O}$ is the average observed heterozygosity across individuals and SNPs. The minimum length of a ROH was set to 1 MB to exclude short ROH arising as a consequence of linkage disequilibrium (LD) [5]. Additionally, a minimum density of one SNP every 50 kB and a maximum 100 kB gap between consecutive SNPs were required to define a ROH.

Mastrangelo *et al. BMC Res Notes* (2018) 11:13

Page 3 of 6

The inbreeding coefficient based on ROH (F_{ROH}) for each animal was estimated as the ratio between the sum of the length of all ROH (L_{ROH}) and the total length of the autosomal genome covered by SNPs on the array ($L_{AUTO} = 2268.83$ MB).

$$F_{ROH} = \frac{\sum L_{ROH}}{L_{AUTO}} \qquad (2)$$

Software

Quality control filtering of genotypic data, ROH detection and the estimation of genetic parameters were all performed using the *PLINK* software package [22]. N_e was estimated using NeEstimator v.2 [25], with unlinked SNP selected using *PLINK*. The *R* programming environment for statistical computing v.3.2.3 [26] was used for data manipulation, summary statistics, preparation of tables and figures. The specific *PLINK* command lines are detailed in Additional file 1.

Results

After filtering for quality (call-rate, genome assembly), the final number of SNPs retained for the analysis was 94,065. All 48 dogs had high quality genotyping and were included in the analysis. The average observed (H_O) and expected (H_E) heterozygosities were 0.371 (± 0.142) and 0.359 (± 0.124), respectively; the average MAF was 0.269 (± 0.132). The H_O and H_E values reported here are comparable with previous studies that examined the genetic diversity in dogs: Pollinger et al. [27] and Pilot et al. [12] estimated average H_O between 0.2 and 0.3 in modern dog breeds [12, 27]. Pertoldi et al. [6] reported H_E estimates in five Danish dog breeds similar to the expected heterozygosity found here in Braque Français, type Pyrénées dogs. The reported MAF values also were consistent with the range found in literature [11, 28]. Recently, Stronen et al. [29] reported lower values of genetic diversity in the endangered Norwegian Lundehund (e.g. $H_O = 0.038$) compared to three reference breeds. In the same breed, a very low MAF (0.033) has been estimated [7]. In comparison, the genetic diversity levels reported here for Braque Français, type Pyrénées dogs seem to indicate that this is not an endangered breed. The 5560 SNPs on the sex chromosomes were not used in this study, because of potentially ambiguous heterozygous SNP calls in males. As an indication, we report basic SNP descriptive statistics in males and females. Monomorphic SNPs on the sex chromosomes were 41.2 and 37.3% in males and females respectively, with MAF 0.145 and 0.151, and missing-rate 0.0301 and 0.0293.

The estimated contemporary effective population size (N_e) in Braque Français, type Pyrénées dogs was 27; this value indicates a potential risk of inbreeding and reduction in genetic diversity. In a recent a study on Bullmastiff dogs, a similar N_e value (29.1) was estimated using the same method [5]. The authors reported that a small N_e is a reflection of population size, unequal use of breeding animals and unequal founder contributions.

A total of 5635 ROH were identified with an average of 117.38 ROH per dog (range between 52 and 230 ROH per animal). Table 1 reports the number of ROH and their average length per length class (0–2, 2–4, 4–8, 8–16 MB). The average number of ROH in Mediterranean dog breeds ranged from 12 (Mastino Abruzzese) to 114 (Saint Bernard) [30]: this positions the Braque Français, type Pyrénées in the higher part of the range. The longest ROH (12.5 MB) was found on chromosome 16. Figure 1 shows the proportion of times (dogs) each SNP falls inside a ROH, plotted against the SNP position along the dog genome. Especially on chromosome 22 there are a number of SNP loci which appear to fall relatively often within a ROH in the Braque Français, type Pyrénées dog breed. Indeed, at the beginning of chromosome 22 (5–10 MB) there is a ROH shard by most of the dogs in the analysed sample (Fig. 2).

Inbreeding coefficients estimated from SNP marker loci as F_{IS} were in the range $[-0.127, 0.172]$: some negative values were obtained, which corresponded to dogs with lower than average homozygosity. Inbreeding estimated from runs of homozygosity (F_{ROH}) had a mean of 0.112 (± 0.023), with a range between 0.0526 to 0.225. In German shepherd dogs, similar F_{ROH} values have been reported [8]. F_{ROH} from 0.061 (Jack Russell Terrier) to 0.151 (Bulldog) was estimated in a panel of canine breeds [5]. The Pearson linear correlation between the two measures of molecular inbreeding (F_{IS} and F_{ROH} was 0.91. Previously, from pedigree records an average genealogical inbreeding coefficient (F, [31]) of 4.3% was estimated [18].
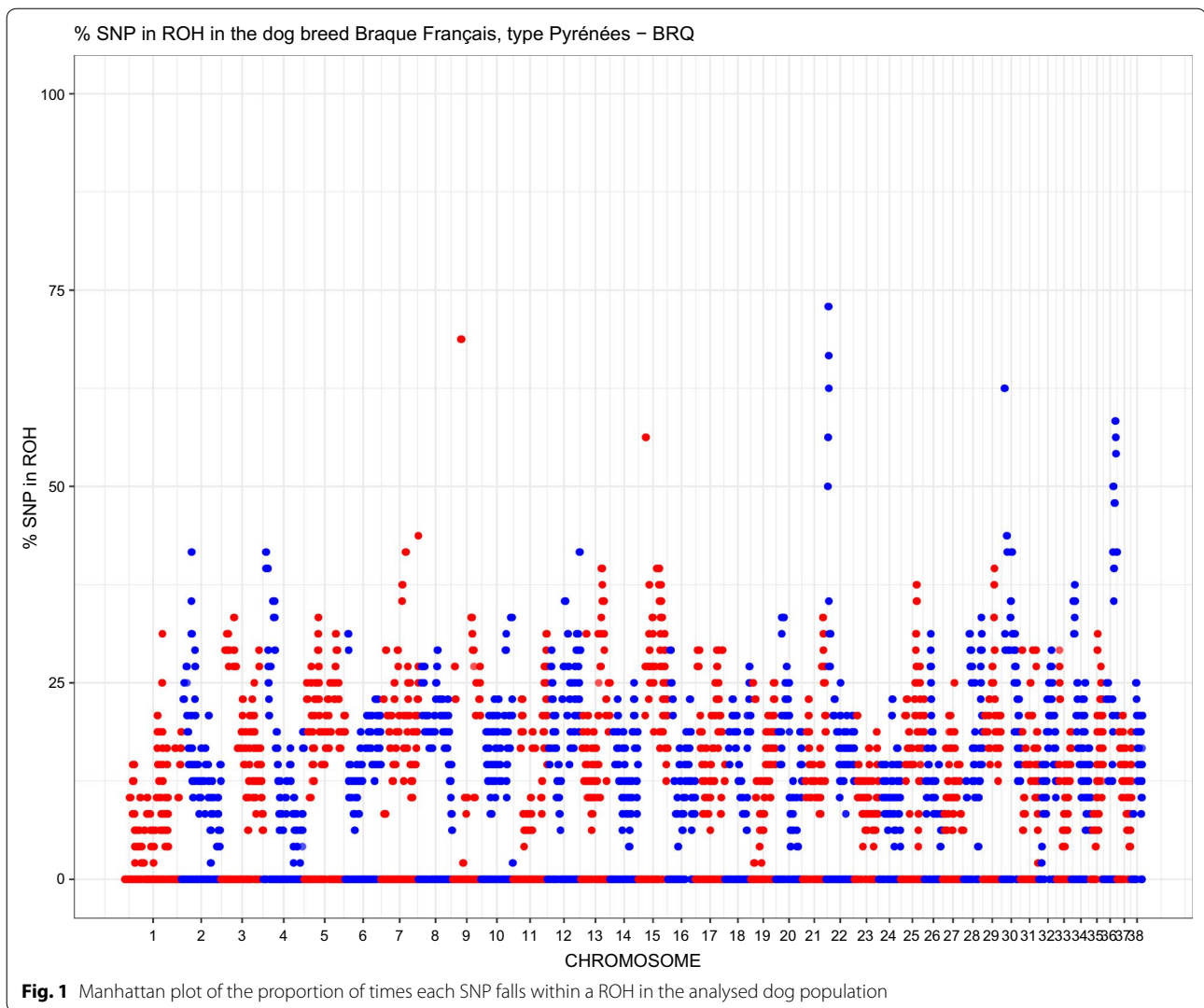
Limitations

The results presented in this note constitute the first report on the genomic characterization of the Braque Français, type Pyrénées dog breed using molecular markers. Previously, the demographics and diversity of this breed had been analysed resorting to pedigree data exclusively. The use of molecular data represents

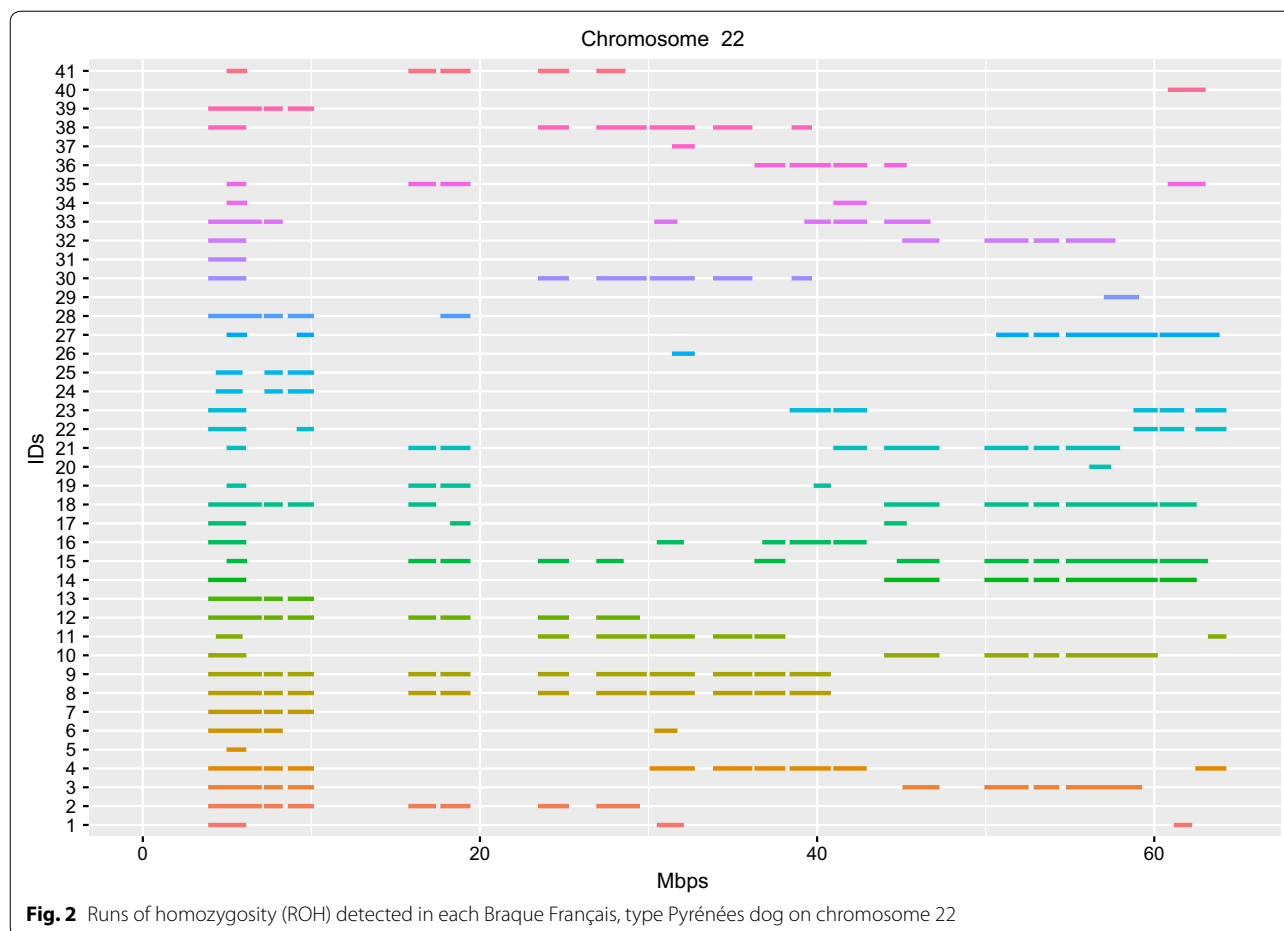**Table 1 ROH counts and average length per class (MBps)**

	Class (MBps)	Group	nRuns	Avg length (MBps)
1	1–2	BRQ	3390	1.43
2	2–4	BRQ	1752	2.67
3	4–8	BRQ	463	5.16
4	8–16	BRQ	29	9.22

BRQ Braque Français, type Pyrénées

Mastrangelo *et al. BMC Res Notes* (2018) 11:13

Page 4 of 6

Fig. 1 Manhattan plot of the proportion of times each SNP falls within a ROH in the analysed dog population

an important step towards a more accurate description of the genome of the Braque Français, type Pyrénées dogs. This improved genomic information will prove to be very useful for the management of the breed, both in the perspective of conservation and in that of breeding and selection. However, this is but a preliminary study, that aimed at providing essential initial information to be used subsequently for a larger and more comprehensive studies. Genomic data on a panel of relevant dog breeds are currently being acquired, to be used for a large-scale genome-wide characterization of dog breeds and the genetic distances and phylogenetic relationships between them. Additionally, phenotypes on the hunting ability

and morphology of Braque Français, type Pyrénées dogs are being collected: together with the already available genotypic data, these phenotypes will be used in genome-wide association studies (GWAS) and similar approaches (e.g. see [32]) to detect SNP loci and regions of the genome that play a role in relevant phenotypes for the breed. As an illustration, the strong ROH signal detected on chromosome 22 may be associated to hunting ability or other phenotypic characteristics of the Braque Français, type Pyrénées dog breed. Furthermore, genomic data can be used to improve the accuracy of pedigree records, thereby allowing for more meaningful comparisons between genealogic and molecular inbreeding. All

Mastrangelo *et al. BMC Res Notes* (2018) 11:13

Page 5 of 6

Fig. 2 Runs of homozygosity (ROH) detected in each Braque Français, type Pyrénées dog on chromosome 22

together, the results presented here provide an interesting example of the use of molecular markers to understand the genetic background and history of small canine breeds like the Braque Français, type Pyrénées.

Additional file

Additional file 1. Plink command lines. File with the Plink command lines used to: (i) edit the SNP data; (ii) select unlinked SNP loci for the estimation of N_e; (iii) detect runs of homozygosity (ROH).

Abbreviations
SNP: single nucleotide polymorphisms; ROH: runs of homozygosity.

Authors' contributions
FB, BA and SM performed all statistical analyses. FB wrote most of the manuscript. RC, SM, MR and AS contributed largely to the writing of the paper. RC sampled the animals and generated the SNP genotype data. All authors read and approved the final manuscript.

Author details
¹ Dipartimento di Scienze Agrarie e Forestali, Università di Palermo, Palermo, Italy. ² CNR-IBBA, Via Bassini 15, 20133 Milano, Italy. ³ Club Italiano Braque Français Type Pyrénées, Capalbio, GR, Italy. ⁴ Scuola di Scienze Mediche Veterinarie, University of Camerino, Matelica, MC, Italy. ⁵ Centro Interuniversitario di Ricerca e di Consulenza sulla Genetica e la Clinica del cane, Matelica, MC, Italy. ⁶ Dipartimento di Scienze Veterinarie, Università di Pisa, V.le delle Piagge 2, 56124 Pisa, Italy.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
The genotypic data used in this preliminary work will be publicly available after publication of results from the main study including Braque Français, type Pyrénées and all other breeds selected for comparison. In the meantime, data may be requested directly to the corresponding author.

Consent for publication
Not applicable.

Ethics approval and consent to participate
No experimental studies were conducted on the animals. Animal records were provided by "ENCI" (Ente Nazionale Cinofilia Italiana), the institution that officially manages data for all dog breeds registered in Italy. Blood samples for genotyping were collected by official veterinarians. "Club Italiano Bracco

Mastrangelo *et al. BMC Res Notes* (2018) 11:13

Page 6 of 6

Francese", whose president is coauthor of the paper, consented for this study to access the blood samples.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Taberlet P, Valentini A, Rezaei H, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P. Are cattle, sheep, and goats endangered species? Mol Ecol. 2008;17(1):275–84.
2. Kristensen TN, Sørensen AC. Inbreeding-lessons from animal breeding, evolutionary biology and conservation genetics. Anim Sci. 2005;80(2):121–33.
3. Leroy G, Rognon X, Varlet A, Joffrin C, Verrier E. Genetic variability in french dog breeds assessed by pedigree data. J Anim Breed Genet. 2006;123(1):1–9.
4. Cecchi F, Paci G, Spaterna A, Ciampolini R. Genetic variability in bracco italiano dog breed assessed by pedigree data. Ital J Anim Sci. 2013;12(3):54.
5. Mortlock S-A, Khatkar MS, Williamson P. Comparative analysis of genome diversity in bullmastiff dogs. PLoS ONE. 2016;11(1):0147941.
6. Pertoldi C, Kristensen TN, Loeschcke V, Berg P, Praebel A, Stronen AV, Proschowsky HF, Fredholm M. Characterization of the genetic profile of five danish dog breeds. J Anim Sci. 2013;91(11):5122–7.
7. Pfahler S, Distl O. A massive reduction of the genetic diversity in the lundehund. Anim Genet. 2014;45(1):154.
8. Mortlock S-A, Booth R, Mazrier H, Khatkar MS, Williamson P. Visualization of genome diversity in german shepherd dogs. Bioinform Biol Insights. 2015;9(Suppl 2):37.
9. Leroy G, Verrier E, Meriaux J, Rognon X. Genetic diversity of dog breeds: within-breed diversity comparing genealogical and molecular data. Anim Genet. 2009;40(3):323–32.
10. Ciampolini R, Cecchi F, Paci G, Policardo C, Spaterna A. Investigation on the genetic variability of the american pit bull terrier dogs belonging to an italian breeder using microsatellite markers and genealogical data. Cytol Genet. 2013;47(4):217–21.
11. Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Pielberg GR, Sigurdsson S, Fall T, Seppälä EH, Hansen MS, Lawley CT, et al. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. PLoS Genet. 2011;7(10):1002316.
12. Pilot M, Malewski T, Moura AE, Grzybowski T, Oleński K, Kamiński S, Fadel FR, Alagaili AN, Mohammed OB, Bogdanowicz W. Diversifying selection between pure-breed and free-breeding dogs inferred from genome-wide snp analysis. Genes Genomes Genet. 2016;6(8):2285–98.
13. Dreger DL, Rimbault M, Davis BW, Bhatnagar A, Parker HG, Ostrander EA. Whole genome sequence, snp chips and pedigree structure: building demographic profiles in domestic dog breeds to optimize genetic trait mapping. Dis Models Mech. 2016. https://doi.org/10.1242/dmm.027037.
14. Ciampolini R, Cecchi F, Bramante A, Casetti F, Presciuttini S. Genetic variability of the bracco italiano dog breed based on microsatellite polimorphism. Ital J Anim Sci. 2011;10(4):59.
15. Bigi D, Marelli S, Randi E, Polli M. Genetic characterization of four native italian shepherd dog breeds and analysis of their relationship to cosmopolitan dog breeds using microsatellite markers. Animal. 2015;9(12):1921–8.
16. Allendorf FW, Hohenlohe PA, Luikart G. Genomics and the future of conservation genetics. Nat Rev Genet. 2010;11(10):697.
17. Bruford MW, Ginja C, Hoffmann I, Joost S, Orozco-terWengel P, Alberto FJ, Amaral AJ, Barbato M, Biscarini F, Colli L, et al. Prospects and challenges for the conservation of farm animal genomic resources, 2015–2025. Front Genet. 2015;6:314.
18. Cecchi F, Paci G, Spaterna A, Ragatzu M, Ciampolini R. Demographic approach on the study of genetic parameters in the dog Braque Français type Pyrénées italian population. Ital J Anim Sci. 2016;15(1):30–6.
19. Wright S. The interpretation of population structure by f-statistics with special regard to systems of mating. Evolution. 1965;19(3):395–420.
20. Jones A, Ovenden J, Wang Y. Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. Heredity. 2016;117(4):217–23.
21. McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, et al. Runs of homozygosity in European populations. Am J Hum Genet. 2008;83(3):359–72.
22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
23. Bjelland D, Weigel K, Vukasinovic N, Nkrumah J. Evaluation of inbreeding depression in holstein cattle using whole-genome snp markers and alternative measures of genomic inbreeding. J Dairy Sci. 2013;96(7):4697–706.
24. Lencz T, Lambert C, DeRosse P, Burdick KE, Morgan TV, Kane JM, Kucherlapati R, Malhotra AK. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. Proc Natl Acad Sci. 2007;104(50):19942–7.
25. Do C, Waples RS, Peel D, Macbeth G, Tillett BJ, Ovenden JR. Neestimator v2: re-implementation of software for the estimation of contemporary effective population size (ne) from genetic data. Mol Ecol Resour. 2014;14(1):209–14.
26. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2017. https://www.R-project.org/.
27. Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, et al. Genome-wide snp and haplotype analyses reveal a rich history underlying dog domestication. Nature. 2010;464(7290):898.
28. Parker HG, Dreger DL, Rimbault M, Davis BW, Mullen AB, Carpintero-Ramirez G, Ostrander EA. Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. Cell Rep. 2017;19(4):697–708.
29. Stronen AV, Salmela E, Baldursdóttir BK, Berg P, Espelien IS, Järvi K, Jensen H, Kristensen TN, Melis C, Manenti T, et al. Genetic rescue of an endangered domestic animal through outcrossing with closely related breeds: a case study of the norwegian lundehund. PLoS ONE. 2017;12(6):0177429.
30. Dreger DL, Davis BW, Cocco R, Sechi S, Di Cerbo A, Parker HG, Polli M, Marelli SP, Crepaldi P, Ostrander EA. Studies of the Fonni's dogs from sardinia show commonalities between development of pure breeds and population isolates. Genetics. 2016;204:737–55.
31. Wright S. Coefficients of inbreeding and relationship. Am Nat. 1922;56(645):330–8.
32. Biscarini F, Biffani S, Stella A. Más allá del gwas: alternativas para localizar qtls. arXiv preprint arXiv:1504.03802. 2015.