Máster Universitario en

Biodiversidad en Áreas Tropicales y su Conservación

POPULATION STRUCTURE AND GENETIC VARIABILITY OF LIVESTOCK DOG BREEDS IN THE IBERIAN PENINSULA, AZORES, CANARY AND BALEARIC ISLANDS

Trabajo Fin de Máster presentado por:

Belén Villegas Giménez

Para optar al título de Máster

Tutor: José Ignacio Doadrio, Ph. D., Annie Machordom Ph. D.

Museo Nacional de Ciencias Naturales - Biodiversidad y biología evolutiva

doadrio@mncn.csic.es; annie@mncn.csic.es

Curso académico 2017 – 2019

AGRADECIMIENTOS

Al Programa de Máster Oficial “Biodiversidad en Áreas Tropicales y su Conservación”, impartido en la ciudad de Madrid, y a sus instituciones auspiciantes: el Consejo Superior de Investigaciones Científicas y la Universidad Internacional Menéndez Pelayo. Incluyo aquí a compañeros, profesores y equipo organizador, que junto con su disponibilidad han aportado una formación fundamental y los medios para diseñar este trabajo. Mención especial al director del Máster, Javier Diéguez Uribeondo, por la dedicación y seguimiento en lo profesional, pero, sobre todo, en lo personal estos dos años.

A mis tutores, José Ignacio Doadrio Villarejo y Annie Machordom Barbé, por darme su apoyo desde el primer momento y por adentrarme en el mundo de la investigación y la conservación a través de la genética, ayudándome en la comprensión de sus herramientas y metodologías. Agradezco su buen criterio, su apoyo, ilusión y paciencia. A la gente del Museo de Ciencias Naturales de Madrid, y particularmente, infinitas gracias a Lourdes Alcaraz Gacituaga y Silvia Perea Aranda ya que sin ayuda, habría sido imposible sacar este trabajo adelante. Así como a las organizaciones caninas y todos los propietarios que han colaborado con el estudio.

A mi familia, amigos y amigas, que han seguido a mi lado durante todo este tiempo, aunque os lo intento decir, no es comparable a lo agradecido que estoy de contar con todas vosotras y vosotros en mi vida.

ABSTRACT

Dogs have been helping in the work of the farms since the domestication of livestock animals. However, Iberian livestock dog breeds are still poorly known. In this work we use microsatellites to study the relationship between breeds and/or racial groups of Iberian dogs to know their population structure and genetic variability. To do so, over 1,200 livestock guarding (LGD) and herding (LHD) breeds will be genotyped for 18 microsatellite markers. We estimated the molecular diversity indices for each breed and analysed livestock dogs' population structure using the Bayesian clustering procedure. We found higher genetic variability values in LGD than in LHD. Structure analysis showed that the subdivision of 7 groups in LGD is the most likely subdivision agreeing with the breeds, and Pyrenees mountain dog and Pyrenean mastiff had the lowest genetic variability. Within LHD non-Iberian breeds Czech Wolf Dog and German Shepherd Dog had lowest genetic variability and the highest variability values were the Portuguese island dogs. In the LHD K = 9 it was the best, but it did not distinguish all races. Some unrecognized Iberian races separated even before the recognized non-Iberian. These results reflect the importance of transhumance in favouring the genetic flow between distant populations that form the same race. This is the most complete genetic study with microsatellites of Iberian livestock dogs to date. It shows that breeds should be viewed as historic dynamic populations, and it is necessary to continue the study of Iberian livestock guarding and herding dogs and decipher their genetic characteristics in order to protect their low populations.

Key Words (7): population genetics, genetic diversity, genetic structure, microsatellites, guarding dogs, herding dogs, Iberian Peninsula.

BACKGROUND

The wolf was the first species domesticated by man and its genetic study has aroused scientific interest by the high phenotypic and behavioural variability acquired since the period of their original domestication, given rise in numerous dog breeds. This has taken place in recent years, to different genetic studies including the sequencing of complete genomes (Lindblad-Toh et al. 2005). The origin of the domestication is in a continuous debate and has interest not only since historical and cultural aspects, but because of the importance of dating the first bottleneck that leads to the *Linkage disequilibrium* observed in the dog and that should help us explain the particularities of its genome and evolution (Wayne and Ostrander 2007).

Different results on the origin of dog have depended on genetic markers and breeds employed, as well as of the origin and age of the samples. Work done with mitochondrial DNA has suggested that domesticated dogs originated in East Asia, due to their greater diversity of haplotypes, about 15,000YBP (Savolainen et al. 2002); 16,300YBP (Pang et al. 2009) or 33,000 YBP (Wang et al. 2016). But when in the mitochondrial analyses are included archaeological samples and different wolves populations, one or two distinct origins are supported, but ever emerge an European origin dated between 20,000-40,000 YBP (Botigué et al. 2017; Frantz et al. 2016) The European origin is in agreement with the archeological register that postulates as the oldest reliable fossils European dogs between 15,000 and 16,000 YBP (Benecke 1987; Janssens et al. 2018). When genetic markers are Single Nucleotide Polymorphisms (SNPs), a dual origin is proposed in Europe and the Middle East (vonHoldt et al. 2010), an origin in Central Asia (Shannon et al. 2015) or an origin in East Asia and Europe (Parker et al. 2017). However, a debate exists on the origin of the current dog breeds. Most of them originated about 200 years ago and that resulted in

second historical bottleneck for each breed. Some authors establish a single Asiatic origin (Ollivier et al. 2018) while others indicate the existence of breeds of European origin (Parker et al. 2017).

Different markers and breeds have been studied to tackle the relationships of the current dog breeds (Pertoldi et al. 2013; Pires et al. 2006; Suárez et al. 2013). The most extensive and understandable work on relationships between breeds is that carried out by Parker et al. (2017), which included 161 breeds in their phylogenetic analysis. In the former work, only two Iberian breeds of Ibizan hound and Portuguese water dog were studied. However, the results of a recent study on the mitochondrial DNA of Iberian Mesolithic dogs, dated between 7,903-7,570 YBP, has shown the interest of studying Iberian dogs (Pires et al. 2019). In Iberian Mesolithic dogs, the named Haplotype A is the most common as in current European dog breeds; nevertheless it is a very rare haplotype in other non-Iberian archaeological deposits (Pires et al. 2019). Thus it was postulated that the origin of Haplotype A is from Asia and arrived in Europe during the Neolithic migrations, replacing the old European haplotypes and erasing the actual genetic structure of European dogs (Ollivier et al. 2018).

However, there is no scientific or historical data that could imply contacts between Iberian and Asia populations before the Neolithic. This involves that the European wolves in the Mesolithic already had Haplotype A, and that this genetic marker has been transferred to the Iberian dogs by recruitment in local or independent domestication processes (Pires et al. 2019). In addition, in Spain it exists one remain of a dog, dated for Paleolithic about 16,000 YBP that is the oldest dog remain adequated dated, but whose genetic data are not available (Altuna and Mariezkurrena 2017; de-la-Rúa et al. 2019; Janssens et al. 2018).

In Spain and Portugal the Fédération Cynologique Internationale (FCI) has recognized

twelve and eight breeds respectively. The Portuguese Kennel Club (PKC) recognizes eleven breeds and the Spanish Royal Canine Society (SRCS) recognizes 27 breeds. Including the breeds in regional catalogues, and local breeds not officially recognized, the number of dog breeds in Spain and Portugal is about 45.

In this work we study 24 breeds of shepherd dogs of Spain and Portugal. Within these breeds two functional and morphological groups exist: Livestock Herding Dogs (LHD) that are smaller and used by shepherds to handle flock, while Livestock Guardian Dogs (LGD) are bigger and do not receive instructions from the shepherds. The study of both groups of shepherd dogs makes it possible to compare two groups with different selection history. LGD are described with precision since roman period (Columella 1954). Since the creation of Mesta (pastoral organization) in XIII century, the care and selection of these dogs in Spain were ruled until XVIII century. Different histories on LGD in Iberian Peninsula have been postulated since the historical and epic literature. However, genetic data found two clades of LGD: one for central European breeds and another for Mediterranean ones (Parker et al. 2017). In Spain, only the Spanish mastiff and the Pyrenean mastiff are officially recognized, but many LGD individuals spread throughout the territory with only a functional selection. In contrast, no description of LHD appears in photographs or literature in Spain until the 19th century. LHD could be more modern and include diversified races in different clades, but no studies are available.

The study of different Iberian dog breeds with populations subjected to different selection pressures and with different times of formation should provide information on how these parameters influence the genetic variability and structure of dog populations. But also the study of shepherd dogs should help us understand the relationship between shepherd communities, especially nomadic and transhumant throughout history, which leave few archaeological traces of their presence due to the precariousness of their lives and utensils

(Toscano et al. 1998).

Microsatellites have been widely used to study population variability and structure (Bruford and Wayne 1993). In this work we use microsatellites to study the relationship between breeds and/or racial groups of Iberian dogs to know their population structure and genetic variability.

MATERIAL AND METHODS

Blood or saliva samples were collected from 1,248 individuals from 27 breeds. Of them, 24 are autochthones from Spain and Portugal. For LGD, we studied a total of 719 samples of 7 Iberian breeds (Table 1). Within LGD, we take a larger number of samples of Spanish Mastiff due to their larger population size, with around 25,000 individuals spread across much of the Spanish territory. Of a total of 461 studied Spanish mastiffs, 321 were working dogs, including 108 transhumant and 213 non-transhumant herds throughout the Iberian Peninsula. These samples were collected from 115 farms and 15 provinces. In the north and central northwest of Spain, 225 Spanish mastiffs are used to protect goats, sheep or cows from the wolf, and 82 live in territory also inhabited by bear. In south and central northeast of Spain, 96 mastiffs protect livestock from mesopredators such as fox, lynx or Egyptian mongoose (Table 2).

The 140 samples of non-working dogs are officially registered dogs in SRCS and belong to 41 different breeders and Kennels names. Non-working dogs were selected to remove nearby kinships from a genetic database of 229 dogs.

We divided the Spanish mastiff in two groups: working and non-working dogs. Within working dogs we distinguished 17 groups by transhumant and non-transhumant dogs and by the different regions where they develop their work, to study the presence of structure within Spanish mastiff, the breed more spread in the farms of Spain (Figure 2).

For LHD we collected 529 samples of 17 breeds from Spain and Portugal (Table 1). Also two breeds, also widely extended by Iberian farmers, as Border collie and German shepherd dog were studied. Czech wolf dog was analysed for possible crossings with Spanish Wolfhound. Mainly the samples come of working dogs, except the Portuguese breeds. From Leonese shepherd dog we studied samples from two groups. One belongs to

one single breeder, but eliminating close kinships, whose founding group was established in 1992. The other was working dogs from 7 different farms. We expanded the number of samples to 75 for the Spanish wolfhound, a breed not recognized by official canine administrations. The Can of Chira is another breed not recognized but we were not able to obtain more than 19 samples of 9 farmers. The original distribution of the breeds is shown in Figure 1.

1. Handling of samples and DNA extraction

Blood samples were preserved in EDTA until DNA isolation using the Qiagen DNeasy® Blood & Tissue kit. DNA from saliva samples, preserved in Oragene-Animal transport and purification tubes (DNAgenotek), was isolated using the Oragene-Animal DNA collection kit (DNAgenotek). The DNA was extracted, following the standard manufacturers' protocols adapting the Oragene-Animal DNA protocol for saliva samples. DNA amplification was tested by electrophoresis using 2% agarose gels with SYBR® Safe (Life Technologies). Negative controls were included throughout the entire process to monitoring potential DNA contaminations. We use 18 microsatellite markers [AHTK211 (CFA26), CXX279 (CFA22), REN169O18 (CFA29), INU055 (CFA10), REN54P11 (CFA18), INRA21 (CFA21), AHT137 (CFA11), REN169D01 (CFA14), AHTH260 (CFA16), AHTK253 (CFA23), INU005 (CFA33), INU030 CFA13), FH2848 (CFA2), AHT121 (CFA13), FH2054 (CFA12), REN162C04 (CFA7), AHTH171 (CFA6) and REN247M23 (CFA15)] (Table 3), recommended by the International Society of Animal Genetics (ISAG), using the Canine Genotypes Panel 1.1 Kit (Thermo Scientific) following the manufacturer's instructions. Positive amplifications were genotyped for all individuals in the Sequencing and Genetic Diagnostic service - SECUGEN. Microsatellite alleles were processed using GENEMAPPER v 4.0 (ABI, Foster City, CA, USA) and manually checked.

2. Study of genetic diversity

a. Genetic variability

The molecular diversity indices for each breed were estimated. The number of alleles, polymorphism, observed and expected heterozygosity, allelic richness and F_{IS} (Kirby 1975) were calculated using the GENETIX 4.2 software (Belkhir 2004). Statistical significance of F_{IS} was tested using $\alpha = 0.05$ for 10,000 permutations in GENETIX, and sequential Bonferroni corrections with FSTAT software (Goudet 2002).

b. Genetic differentiation

The ARLEQUIN vs. 3.5.2.2 (Excoffier and Lischer 2010) software was used to estimate genetic differentiation of the populations through the F-statistics, by pairwise mean F_{ST} (Wright 1949) using 10,000 permutations. To assess the relative contribution of genetic variation to structure within and between breed's populations we performed several analyses of molecular variance (AMOVA) testing for genetic structure between groups of populations and among groups according to their function, similarity and geographic distribution.

Statistical significance of F_{ST} was tested using $\alpha = 0.05$ for 10,000 permutations, using sequential Bonferroni corrections. Populations Hardy-Weinberg equilibrium were was also tested with 1,000,000 steps in Markov chain and Bonferroni corrections were used.

c. Population structure

To analyse the livestock dogs' population structure, the Bayesian clustering procedure was implemented through the software STRUCTURE vs. 2.3.4 (Hubisz et al. 2009). We grouped guardian and herding dogs into sub datasets and according to their genetic similarity. Within LGDs, for Spanish Mastiff we established 18 populations based on their geographical distribution and if they were transhumant or non-transhumant. We conducted

5 independent runs for each value of the putative number of populations between 1 and n, where n is the number of breeds in each LGD sub dataset, using 100.000 *burn-in* period followed by 5,000,000 MCMC repetitions and using 10,000 *burn-in* period followed by 500,000 MCMC repetitions for LHD. Structure Harvester (Earl 2012) and CLUMPAK (Kopelman et al. 2015) were used to truncate the different runs of each K and to display the graphical output for the best K for each analysis, respectively. A second analysis was implemented for the LGD group, considering the first biggest split showed by the K=2 runs, that is eliminating the seven first populations.

Finally, the genetic variance between the obtained groups and individuals was confronted in a Factorial Correspondence Analysis (FCA), which was performed using GENETIX. Populations with a large number of individuals were randomly reduced through the R program, to minimize a possible size sample effect.

To understand the origin of the Spanish Mastiff, a distance analysis was performed using the index of DENEI72 and the NJ method clustering (Saitou and Nei 1987) with the software PAST 3.04 (Hammer et al. 2001). The populations of Spanish Mastiffs were also compared to other breeds of LGDs. A representative number of individuals from each population were used. Populations with a large number of individuals were randomly reduced through the R program (a minimum of 12 and a maximum of 28).

RESULTS

1. Genetic variability

Among LGD a total of 24 populations belong to 7 breeds were established and between 14 (Palencia non-transhumant) and 140 (Spanish Mastiff non-working group) individuals were analysed within each population. Among LHD we analysed 20 different breeds between 10 (Garafian Sheepdog and Portuguese Water dog) and 75 (Spanish Wolfhound) dogs. We found higher genetic variability values in LGD than in LHD.

Livestock Guardian Dogs

In LGD the consanguinity values were mostly positive because the expected heterozygosity was greater than that observed, except in the Pyrenean Mastiff and the Pyrenees mountain where the observed was greater than expected. In general, the values were close to 0 and after the Bonferroni correction, only Non-transhumant of León and Non-transhumant of Cádiz were significant.

The lowest genetic variability observed was found in the Pyrenees mountain dog (H_{OBS} = 0.651; H_{EXP} = 0.637; Avg = 4.94) and Pyrenean mastiff (H_{OBS} = 0.607; H_{EXP} = 0.59; Avg = 4.94). But low values were also found in transhumant Cáceres (H_{OBS} = 0.693; H_{EXP} = 0.672; Avg = 5.06) and Zamora (H_{OBS} = 0.697; H_{EXP} = 0.691; Avg = 5.61). The other races of LGD showed similar variability values, especially the mastiffs of Leon and breeders. Curiously the working and non-working dogs had similar values of F_{IS} in Spanish mastiff (Table 4).

Livestock Herding dogs

In HD we found F_{IS} values were near to 0 and negative consanguinity values in the Castilian-Manchego Sheepdog, Can of Chira, Portuguese Water dog, Terceira Cattle dog and the Saint Miguel Cattle dog. After the Bonferroni correction, values from Spanish

Wolfhound, Portuguese Sheepdog, Manchego Sheepdog, Spanish Water dog, Basque Shepherd, Garafian Sheepdog and Majorca Shepherd dog were significant.

The lowest genetic variability observed was found in non-Iberian breeds Czech Wolfhound (H_{OBS} = 0.523; H_{EXP} = 0.526: Avg = 3.61) and German Shepherd dog (H_{OBS} = 0.580; H_{EXP} = 0.608; Avg = 5.17). The highest value were the Portuguese island dogs Terceira Cattle dog (H_{OBS} = 0.724; H_{EXP} = 0.699: Avg = 6.00) and Saint Miguel Cattle dog (H_{OBS} = 0.711; H_{EXP} = 0.664: Avg = 4.67) (Table 5).

2. Genetic structure

The F_{ST} values among populations were higher in HD that ranged from 0.0487 to 0.3012, compared to LGD (0.0113 to 0.1791). Likewise, AMOVA results showed that most of the variation for all comparisons was found within populations (> 84%) for both groups of dogs. Among the populations, the HD exhibited a higher percentage of variation (> 12%) with respect to LGD comparisons (5%) (Tables 6 and 7).

Livestock Guardian Dogs

F_{ST} values for LGD were significant in all combinations for the two Pyrenean breeds (Pyrenean Mountain dog and Pyrenean Mastiff). The group of non-working Spanish mastiff and all Portuguese breeds also had significant values of F_{ST} for many pairs of combinations. An exception is Transmontano Mastiff, a recently recognized breed for PKC. Clustering with the genetic distances of Nei (1972) resulted in relations with very little support between the branches. Only a group formed by the Pyrenean Mountain dog, Pyrenean Mastiff and Estrela Mountain dog (Tree not shown) was well supported (99%).

The AMOVA results showed that in LGD most of the variation was explained within populations (92-97%). There were low values of variance between groups of non-working Spanish Mastiff dogs and group of Leonnese transhumant dogs (0,1%). Highest values of

variance were to explain the variation observed between comparisons of Zamora Spanish mastiff and other groups (Table 8).

Structure analysis based on Bayesian methods (STRUCTURE) showed that the subdivision of 7 groups of the analysed samples is the most likely subdivision (Figure 3A), according to the method of Evanno et al (2005) (K = 7 Mean (LnProb) = -39654.080, Mean = 0.995). At the first possible partition (K=2) breeds were mostly clustered by Spanish Mastiff non-working group and the populations from north-western of Spain; and the rest of Spanish Mastiff groups and non-Spanish Mastiff breeds splitting apart. The Pyrenean Mastiff and Pyrenean Mountain dog split in K=3 associated. In K=5 Pyrenean Mountain dog, Pyrenean Mastiff, Estrela Mountain dog, Rafeiro of Alentejo and north western Spanish Mastiff are grouped in different clusters. To K=7 all LGD are well differentiated except Transmontano Mastiff and different groups of Spanish Mastiff from mainly central and southern Spain. These dogs seem have genetic signatures of other Iberian LGD breeds (Figure 4, see Annex Figure 1).

In addition, in a second analysis by removing the groups of Spanish Mastiff well separated in first cluster from K = 2, result in six groups (K = 6 Mean (LnProb) = -23546.480, Mean = 0.991) (Figure 3B). In K = 2 the Pyrenean Mastiff and Pyrenean Mountain dog were separated in a different cluster. To K = 6 all the breeds were differentiated in independent clusters with Transmontano Mastiff associated to most of the Spanish Mastiff of central and southern Spain. This group of Transmontano Mastiff and Spanish Mastiff have individuals with genetic traits of Rafeiro of Alentejo and Castro Laboreiro breeds (Figure 5; see Annex Figure 2).

To eliminate possible biases due to different sample size, we realized a Factorial Analysis of Correspondence (FCA). The four principal axes of a FCA explain most of the variance (Table 10; Figure 7). Axes I and IV and II and IV shown a distribution of the samples very

similar to Bayesian analysis. Axis I and IV separated the three more morphologically different breeds (Pyrenean Mountain dog, Pyrenean Mastiff and Castro Laboreiro). Spanish Mastiff no-working group and groups of north-western Spain can also be seen as a group, but less separate, probably due to the correction of the number of samples. No-working dogs were closer to the Transhumant group than other Spanish mastiff groups. Regarding Rafeiro of Alentejo, it also separated from other groups, but this did not happen with Transmontano Mastiff and Estrela Mountain dogs, which are within the groups of south and central Spanish Mastiffs. However, Estrela Mountain dog was separated when we analysed the II and IV axes. To the II and IV axes all the breeds are corrected clustered except Transmontano Mastiff that was within of Spanish Mastiff groups.

Livestock Herding Dogs

Within each group, the global F_{ST} among LHD populations were values ranging from 0.0487 to 0.3012 between the Basque Shepherd and Saint Miguel Cattle dog; and from Czech Wolfhound and Terceira Cattle dog comparisons, respectively. These pairwise F_{ST} comparisons among breeds within LHD populations were statistically supported, shown a high level of differentiation (Table 7).

Among LHDs populations, AMOVA showed canary breeds and water dog breeds exhibited 12.86 and 12.88% of variation respectively, but in general, almost no variation (<4%) was found between groups (Table 8).

To determine if each LHD breed form a breed-specific cluster, we investigated the patterns of genetic structure using Bayesian methods in Structure software. The subdivision in K=9 was identified as the second most likely number of populations in the dataset using the delta K method (Mean (LnProb) = -23546.480, Mean = 0.991) (Figure 3C).

To the first partition K = 2 Leonese shepherd and wolfhounds dogs (Spanish Wolfhound, German Shepherd, Czech Wolfhound, Galician Palleiro Sheepdog and Herreño Shepherd dog) were grouped apart from other breeds. At the second partition (K=3) the Leonese Shepherd seems that the dogs of the only one known breeder was fully assigned to its own private cluster, which means that the dogs of the same breed from different farmers was exhibiting admixture with the other group breeds (Figure 6).

Border Collie with Portuguese Sheepdog and Can of Chira split associated in K=4 and the Villano Bulldog Cattle dog, which became different cluster in K=5. The best K (K=9 minor) was achieved assigning independently the following breeds: Leonese Shepherd dog, Border Collie, Can of Chira and Villano Bulldog Cattle dog. Dogs belong to Spanish Wolfhound were assigned together to Herreño Shepherd dog and Galician Palleiro Sheepdog. German Shepherd dog and Czech Wolfhound not were different groups to K=9.

The samples of Portuguese Sheepdog and those of the Manchego Sheepdog were assigned to different breeds. The Castilian-Manchego Sheepdog was assigned to Saint Miguel Cattle dog and Basque Shepherd while Catalan Sheepdog, Portuguese and Spanish Water dogs and Majorca Cattle dog were together. The individuals of Terceira Cattle dog were assigned next to Garafian Sheepdog until K=17, where Saint Miguel Cattle dog split too. Galician Palleiro Sheepdog split apart in K=10, as same as Catalan Sheepdog in K=13 (see Annex, Figure 3).

The four main axes of an ACF explain most of the variance (Table 11; Figure 8). Axis I and II differentiated mainly three groups of races: wolfhounds, Garafian and the rest of shepherds. The different breeds of wolfhounds were well separated, especially the Czech Wolfhound. In axes I and III, the Leonese shepherd dog and cattle dogs, Villano Bulldog and Saint Miguel, were separated from the other breeds.

DISCUSSION

Dogs have accompanied Iberian villages since the Palaeolithic (Altuna and Mariezkurrena 2017). First they have been used as guardians against predators, increasing in size. After, larger herds resulted in the selection of LHD to help manage them. It can be observed in dogs of the Peninsula from the Neolithic to the Iron Age and in different types of dogs in archaeological sites from Roman Era (Pires et al. 2018; Toscano et al. 1998). Some Iberian breeds underwent a strict selection, but others have a very small population census due to the abandonment of agriculture. This biological and cultural diversity limits the study of how different selection pressures affect the genome, or historical and cultural aspects we know from literature, but need to be tested.

In general, in the Iberian LGD breeds, genetic diversity was high and higher than in LHD. The Pyrenean Mastiff (H_{EXP} = 0.590) and the mountain of the Pyrenees (H_{EXP} = 0.637) are an exception that suffered a sharp population decline associated with wolf extinction in the Pyrenees. The Pyrenean Mountain dog was used again when the wolf came back, especially in the Alps (Landry et al. 1999), but it still has a low genetic diversity. The Pyrenean Mastiff is currently a breed as a pet and has been artificially crossed with other breeds producing the higher observed heterozygosity. Castro Laboreiro had low diversity genetic due to have never done large transhumance and has been more isolated in a small region in Portugal (Geraldes 1996). As expected, we found significant positive inbreeding coefficients for non-working Spanish Mastiff group, due to controlled and closed breeding, but also in three different working mastiff groups (Transhumant and non-transhumant from León and Non-transhumant of Cádiz), probably because of their isolation and smaller population and also controlled breeding. However, transhumance causes Spanish mastiffs populations of the northern and southern peninsular to be in contact as happens in Italy with LGD (Talenti et al. 2018). Although in general showed genetic heterogeneity, within

mastiffs and with Transmontano Mastiff and Rafeiro of Alentejo, Amova show almost no variance among geographically remote populations and it may be due to contact in the border between Spain and Portugal or a more functional and relaxed selection. Nevertheless, Rafeiro of Alentejo had a high degree of inbreeding, probably because our samples come from official breeders. Non-working group were closer to Leonese dogs and can confirm that the modern selection will start from León's mastiff dogs because they are the largest. The traditional transhumant movements of livestock through the *Via de la Plata* and main León ravines explain the relationship between dogs of León, Valladolid, Segovia and Salamanca, avoiding the creation of independent races unlike in Italy and Portugal (Talenti et al. 2018).

For LGD, all the Portuguese and Pyrenean breeds were differentiated with all the analysis and the Pyrenees were the closest genetically. These breeds have been linked to the group that includes Mediterranean and other ancient LGD breeds (Parker et al. 2017; Talenti et al. 2018). The other Iberian LGD cannot be attributed to central LGD or Mediterranean because not central Mastiffs were included in our analyses.

The LHD had lower genetic variability and higher population structure because this breeds suffer higher selection pressure because inbreeding in current times (Pedersen et al. 2013). LHD had similar genetic variability values regardless of being peninsular or island breeds. The German shepherd and the Czech wolfhound had the lowest values, possibly for a lower foundational number of specimens (Randi and Lucchini 2002). In the case of the German shepherd, we made a huge effort to collect the samples from different Spanish bloodlines, so the low variability cannot be attributed to high consanguinity.

All our population structure analysis group wolfhound was well differentiated. The German shepherd is related to American races as ancestor (Parker et al. 2017), but in the

Iberian wolfhounds there is more genetic variability and diversity. So a new line of research is opened on a possible Iberian ancestor of the American races come from the Iberian races born with the discovery of America. This must be solved with a larger number of samples and molecular markers. The Leonese pastor was also well differentiated but only for Kennel Dogs. Those come from farmers have been mixed with other breeds of shepherds. The Garafian shepherd was well differentiated from other breeds when we corrected the bias produced by the small number of specimens analysed. The Can of Chira, an unofficially recognized breed, showed a genetic uniformity and a greater separation than other officially recognized breeds. Surprisingly, the Portuguese Water dog that had been considered the rarest dog in the world (Molinari 1989), had high variability values (H_{EXP} = 0.620) disagreeing with Pires et al. (2006). On the contrary, Spanish water dog had lower values than expected (H_{EXP} = 0.702) and did not form its own cluster as a breed, and the same happened with Basque Sherpherd dog. For Azores island breeds we found higher values that expected in Terceira Cattle dog and Saint Miguel Cattle dog. It is supposed that these breeds were brought to the islands by boat and their high differentiation values reflect their isolation (Canicultura. 2004; Paschoud and Triquet 2007). In Canarian islands, Garafian Shepdog and Lobito Herreño had F_{ST} > 10% with almost all breeds, and Garafian Shepdog show lower levels than expected. These two breeds could be affected by sample bias, so we consider increasing the number of individuals. Likewise, in Balears islands, Majorca Cattle dog had F_{ST} > 10% for almost all the breeds, whereas it showed lower heterocigosity (H_{OBS} = 0.674; H_{EXP} = 0.709) and significant consanguinity values. Finally, Villano Bulldog Cattle dog had a high differentiation degree and split as a breed, so it should be studied.

CONCLUSIONS

1. LGD have more genetic diversity than LHD; due to LHD have been under higher selection pressure because inbreeding in current times.
2. Within of LGD Pyrenean Mountain dog and Pyrenean Mastiff had the lowest genetic diversity as consequence of their loss of functionality to protect livestock of wolf.
3. Within LHG, German shepherd dog, a breed with high selection pressure since the XIX century, had the lowest genetic diversity levels along with Czech wolfhound, a breed with a very low founding number and raised in consanguinity.
4. All the breeds of LGD were identified by the analyses of structuration, except the Transmontano Mastiff, a recently recognized breed in Portugal.
5. Spanish Mastiff was distributed throughout the centre and northwest of Spain. South and east populations were grouped with Trasmontano Mastiff and seem to have characteristics of many of the Iberian LGD breeds.
6. Transhumance that is still in force in north-western and central Spain, in addition to the presence of wolves, seems to be a source of genetic flow that makes it difficult to divide mastiff populations in different breeds, such as in Italy and Portugal.
7. Non-working Spanish Mastiff group seems to be close to the mastiffs of León since individuals for the commercial breed were probably extracted.
8. 8. Within LHD the population structure recognized 9 entities that group the 20 races studied.
9. The wolfhounds group was identified with the two foreign species German Shepherd dog and Czech wolfhound, always differentiated from the Spanish group

consisting of the Galician Palleiro Sheepdog, Herreño Shepherd dog and Spanish Wolfhound.

10. Within LHD structure, some breeds recognized by the FCI or regional official entities do not differ in a cluster, whereas others that are not officially recognized, such as Can of Chira, were well assigned to their own group in the analysis of the structure, so it should be studied more in depth.
11. Leonese Shepherd dog was an enigmatic race that was not related to other functional groups, which is endangered by crosses with other breeds in the field.
12. Canarian races differ in analyses sensitive to population size. We propose to expand the number of faces and include the Canarian breed of Majorero Sherperd dog.
13. Cattle dogs were a group identified by structure analysis and its breeds form particular clusters. Villano Bulldog Cattle dog is not officially recognised and it is the first to separate.
14. Water dogs are recognised breeds that seem to have been crossed several times and structural analysis do not support. In addition, Portuguese water dog has been considered a rare breed but we found high diversity levels.
15. Dog breeds should be viewed as dynamic populations, modified by admixture, introgression, and genetic isolation. It is necessary to continue the study of Iberian livestock guarding and herding dogs and decipher their genetic characteristics in order to protect their low populations.

REFERENCES

Altuna J, Mariezkurrena K (2017) Origenes y evolución de la domesticación en el País Vasco - Iconografía europea de animales domésticos. Centro del Patrimonio Cultural Vasco Gobierno Vasco

Belkhir K (2004) 1996-2004 GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations. http://www univmontp2 fr/~ genetix/genetix/genetix htm

Benecke N (1987) Studies on early dog remains from Northern Europe. Journal of Archaeological Science 14:31-49

Botigué LR et al. (2017) Ancient European dog genomes reveal continuity since the Early Neolithic. Nature Communications 8:16082 doi:10.1038/ncomms16082

https://www.nature.com/articles/ncomms16082 - supplementary-information

Breen M et al. (2001) Chromosome-Specific Single-Locus FISH Probes Allow Anchorage of an 1800-Marker Integrated Radiation-Hybrid/Linkage Map of the Domestic Dog Genome to All Chromosomes. Genome Res 11:1784-1795

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939-943 doi:https://doi.org/10.1016/0959-437X(93)90017-J

Canicultura. CPd (2004) Barbado da Terceira.

Columella (1954) On Agriculture vol II. Loeb Classical Library 407 edn. Harvard University Press, Cambridge, MA

de-la-Rúa C, Altuna J, Mariezkurrena K, Hervella M (2019) Domesticación del perro en Europa. Contribución del yacimiento de Eralla (Zestoa, Gipuzcoa). Paper presented at the XV Reunión Nacional del Cuaternario, Bilbao, 1-5 July

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method vol 4, vA.2 July 2014 edn. Springer. doi:10.1007/s12686-011-9548-7

Excoffier L, 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

Francisco LV, Langsten AA, Mellersh CS, Neal CL, Ostrander EA (1996) A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mamm Genome 7:359-362 doi:10.1007/s003359900104

Frantz LAF et al. (2016) Genomic and archaeological evidence suggest a dual origin of domestic dogs. Science 352:1228-1231

Geraldes A (1996) Brandas e inverneiras: particularidades do sistema agro-pastoril" crastejo".

Goudet J (2002) Fstat software. Institute of Ecology BB, UNIL

Guyon R et al. (2003) A 1-Mb resolution radiation hybrid map of the canine genome. Proceedings of the National Academy of Sciences 100:5296 doi:10.1073/pnas.0831002100

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontologia electronica 4:9

Holmes NG, Dickens HF, Parker HL, Binns MM, Mellersh CS, Sampson J (1995) Eighteen canine microsatellites. Anim Genet 26:132a-133 doi:10.1111/j.1365-2052.1995.tb02659.x

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular ecology resources 9:1322-1332

Janssens L, Giemsch L, Schmitz R, Street M, Van Dongen S, Crombé P (2018) A new look at an old dog: Bonn-Oberkassel reconsidered. Journal of Archaeological Science 92:126-138

Kirby GC (1975) Heterozygote frequencies in small subpopulations. Theor Popul Biol 8:31-48 doi:https://doi.org/10.1016/0040-5809(75)90037-4

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15:1179-1191 doi:10.1111/1755-0998.12387

Landry J-M, Olsson P, Siegenthaler A, Jackson P, Farrell A (1999) The use of guard dogs in the Swiss Alps: a first analysis. KORA, Koordinierte Forschungsprojekte zur Erhaltung und zum Management der …,

Lindblad-Toh K et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803

Lingaas F et al. (1997) Towards construction of a canine linkage map: Establishment of 16 linkage groups. Mamm Genome 8:218-221 doi:10.1007/s003359900393

Mariat D, Kessler JL, Vaiman D, Panthier JJ (1996) Polymorphism characterization of five canine microsatellites. Anim Genet 27:434-435 doi:10.1111/j.1365-2052.1996.tb00514.x

Molinari C (1989) El Cao de Aguas Portugues. Um perro pescador vol 29. SpainEdiciones Guau, SA, Madrid,

Nei M (1972) Genetic distance between populations. The American Naturalist 106:283-292

Ollivier M et al. (2018) Dogs accompanied humans during the Neolithic expansion into Europe. Biol Lett 14:20180286 doi:10.1098/rsbl.2018.0286

Pang J-F et al. (2009) mtDNA Data Indicate a Single Origin for Dogs South of Yangtze River, Less Than 16,300 Years Ago, from Numerous Wolves. Mol Biol Evol 26:2849-2864 doi:10.1093/molbev/msp195

Parker HG, Dreger DL, Rimbault M, Davis BW, Mullen AB, Carpintero-Ramirez G, Ostrander EA (2017) Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. Cell reports 19:697-708

Paschoud JM, Triquet R (2007) Fédération Cynologique Internationale. Saint Miguel Cattle dog. vol 340.

Pedersen N, Liu H, Theilen G, Sacks B (2013) The effects of dog breed development on genetic diversity and the relative influences of performance and conformation breeding. J Anim Breed Genet 130:236-248 doi:10.1111/jbg.12017

Pertoldi C et al. (2013) Characterization of the genetic profile of five Danish dog breeds1. J Anim Sci 91:5122-5127 doi:10.2527/jas.2013-6617

Pires AE et al. (2019) The curious case of the Mesolithic Iberian dogs: An archaeogenetic study. Journal of Archaeological Science 105:116-129 doi:https://doi.org/10.1016/j.jas.2019.03.002

Pires AE et al. (2018) Roman dogs from the Iberian Peninsula and the Maghreb–A glimpse into their morphology and genetics. Quaternary International 471:132-146

Pires AE, Ouragh L, Kalboussi M, Matos J, Petrucci-Fonseca F, Bruford MW (2006) Mitochondrial DNA Sequence Variation in Portuguese Native Dog Breeds: Diversity and Phylogenetic Affinities. J Hered 97:318-330 doi:10.1093/jhered/esl006

Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (Canis lupus) populations by Bayesian admixture analyses of microsatellite variation. Conserv Genet 3:29-43

Richman M, Mellersh CS, André C, Galibert F, Ostrander EA (2001) Characterization of a minimal screening set of 172 microsatellite markers for genome-wide screens of the canine genome. J Biochem Biophys Methods 47:137-149 doi:https://doi.org/10.1016/S0165-022X(00)00160-3

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425 doi:10.1093/oxfordjournals.molbev.a040454

Savolainen P, Zhang Y-p, Luo J, Lundeberg J, Leitner T (2002) Genetic Evidence for an East Asian Origin of Domestic Dogs. Science 298:1610 doi:10.1126/science.1073906

Shannon LM et al. (2015) Genetic structure in village dogs reveals a Central Asian domestication origin. Proceedings of the National Academy of Sciences 112:13639 doi:10.1073/pnas.1516215112

Suárez NM, Betancor E, Fregel R, Pestano J (2013) Genetic characterization, at the mitochondrial and nuclear DNA levels, of five Canary Island dog breeds. Anim Genet 44:432-441 doi:10.1111/age.12024

Talenti A et al. (2018) Studies of modern Italian dog populations reveal multiple patterns for domestic breed evolution. Ecology and evolution 8:2911-2925

Toscano LGV, Serrano MLC, Córdoba de Oya B (1998) El origen de los mastines ibéricos. Complutum 9:117-135

vonHoldt BM et al. (2010) Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464:898 doi:10.1038/nature08837

https://www.nature.com/articles/nature08837 - supplementary-information

Wang G-D et al. (2016) Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res 26:21 doi:10.1038/cr.2015.147

https://www.nature.com/articles/cr2015147 - supplementary-information

Wayne RK, Ostrander EA (2007) Lessons learned from the dog genome. Trends Genet 23:557-567

Wright S (1949) THE GENETICAL STRUCTURE OF POPULATIONS. Annals of Eugenics 15:323-354 doi:10.1111/j.1469-1809.1949.tb02451.x

TABLES

Table 1: The breed or population sampled (Group), number of individuals (N) and level of recognition of the twenty-seven dog populations that will be used in this study. The breeds, if not recognized internationally by the International Cynological Federation (FCI), are recognized by the Portuguese Kennel Club (CPC), Spanish Kennel Club (RSCE), or yet not recognized by any entity (–).

	Group	Abreviation	Recognition	N
Livestock guarding dogs	Spanish Mastiff	SM	FCI	461
	Transmontano Mastiff	TM	PKC	55
	Castro Laboreiro dog	CLD	FCI	27
	Rafeiro of Alentejo	RA	FCI	44
	Estrela Mountain dog	EMD	FCI	58
	Pyrenean Mastiff	PM	FCI	41
	Pyrenean Mountain dog	PMD	FCI	33
Livestock herding dogs	Leonese Shepherd dog	LSD	RSCE	50
	Border Collie	BC	FCI	25
	Spanish Wolfhound	SW	–	75
	German Shepherd dog	GSD	FCI	28
	Czech Wolfhound	CW	FCI	18
	Herreño Shepherd dog	HSD	RSCE	13
	Galician Palleiro Sheepdog	GPS	–	22
	Portuguese Sheepdog	PS	FCI	25
	Manchego Shepherd dog	MSD	–	30
	Castilian-Manchego Sheepdog	CMS	RSCE	58
	Can of Chira	CC	–	19
	Catalan Sheepdog	CS	FCI	20
	Portuguese Water dog	PWD	FCI	10
	Terceira Cattle dog	TCD	PKC	22
	Spanish Water dog	SWD	FCI	22
	Basque Shepherd dog	BSD	RSCE	45
	Garafian Sheepdog	GS	RSCE	10
	Majorca Cattle dog	MCD	FCI	26
	Villano Bulldog Cattle dog	VBCD	–	23
	Saint Miguel Cattle dog	SMCD	FCI	18

Table 2: Origin of the Spanish Mastiff samples.

Spanish Mastiff	N	Transhumant	Non-Transhumant	Wolf	Bear
Non-working Spanish Mastiff	140				
Transhumant of León	69	X		X	X
Non-transhumant of León	66		X	X	
Transhumant of Soria	3	X			
Non-transhumant of Segovia	5		X	X	

Non-transhumant of Valladolid	6		X	X	
Transhumant of Cáceres	13	X		X	X
Non-transhumant of Palencia	14		X	X	
Non-transhumant of Salamanca	4		X	X	
Non-transhumant of Caceres	14		X		
Non-transhumant of Zamora	27		X	X	
Transhumant of Zamora	19	X		X	
Transhumant of Sierra Morena	4	X			
Non-transhumant of Jaén	7		X		
Non-transhumant of Badajoz	27		X		
Non-transhumant of Cádiz	28		X		
Non-transhumant of Huelva-Doñana	8		X		
Non-transhumant of Granada	5		X		
Total	461	108	213	225	82

Table 3: Microsatellites analysed, repeat motifs, allelic ranges and references.

Name	Repeat	Allelic Range	Reference
AHT121	Di	74-118	(Holmes et al. 1995)
AHT137	Di	128-160	(Holmes et al. 1995)
AHTh171	Di	216-240	(Breen et al. 2001)
AHTh260	Di	230-254	(Breen et al. 2001)
AHTk211	Di	82-98	(Lingaas et al. 1997)
AHTk253	Di	280-300	(Lingaas et al. 1997)
AME	-	174–218	(Holmes et al. 1995)
CXX279	Di	109–13	(Richman et al. 2001)
FH2054	Tetra	121-181	(Francisco et al. 1996)
FH2848	Di	224-244	(Breen et al. 2001)
INRA21	Di	87-103	(Mariat et al. 1996)
INU005	Di	104-136	(Holmes et al. 1995)
INU030	Di	136-156	(Holmes et al. 1995)
INU055	Di	196-210	(Holmes et al. 1995)
REN162C04	Di	189-215	(Guyon et al. 2003)
REN169D01	Di	192-220	(Guyon et al. 2003)
REN169O18	Di	145-171	(Guyon et al. 2003)
REN247M23	Di	263-283	(Guyon et al. 2003)
REN54P11	Di	223-245	(Guyon et al. 2003)

Table 4: Genetic diversity indices for Livestock guarding dog populations using 18 microsatellites. Populations groups, number of individuals (N), expected heterozygosity (H_{EXP}), observed heterozygosity (H_{OBS}), polymorphism (P), average number of alleles (Avg), F-coefficient of inbreeding (F_{IS}) and G_{IS} are listed. Values for some populations are missing due to their low sample size. Non-significant G_{IS} values. (*) Significant F_{IS} values based on a test of 10,000 permutations. (**) Significant F_{IS} values supported by Bonferroni.

Populations	N	H_{EXP}	H_{OBS}	P (0.99)	Avg	F_{IS}	G_{IS}
Non-working Spanish Mastiff	140	0.702	0.684	1.00	6.33	0.029*	0.029
Transhumant of León	69	0.703	0.685	1.00	6.67	0.033*	0.033
Non-transhumant of León	66	0.733	0.702	1.00	7.44	0.050**	0.050
Transhumant of Cáceres	13	0.672	0.693	1.00	5.06	0.009	0.009
Non-transhumant of Palencia	14	0.723	0.734	1.00	5.50	0.024	0.024
Non-transhumant of Caceres	14	0.712	0.729	1.00	5.83	0.014	0.014
Non-transhumant of Zamora	27	0.753	0.758	1.00	7.11	0.012	0.012
Transhumant of Zamora	19	0.691	0.697	1.00	5.61	0.018	0.018
Non-transhumant of Badajoz	27	0.718	0.707	1.00	7.00	0.034	0.034
Non-transhumant of Cádiz	28	0.759	0.707	1.00	6.83	0.088**	0.088
Transmontano Mastiff	55	0.751	0.739	1.00	7.78	0.025	0.025
Castro Laboreiro dog	27	0.716	0.682	1.00	6.89	0.066*	0.066
Rafeiro of Alentejo	44	0.757	0.736	1.00	7.39	0.040*	0.040
Estrela Mountain dog	58	0.717	0.705	1.00	7.17	0.026	0.026
Pyrennean Mastiff	41	0.590	0.607	1.00	4.94	-0.016	-0.016
Pyrenean Mountain dog	33	0.637	0.651	1.00	4.94	-0.007	-0.007

Table 5: Genetic diversity indices for Herding dog populations using 18 microsatellites. Populations groups, number of individuals (N), expected heterozygosity (H_{EXP}), observed heterozygosity (H_{OBS}), number of polymorphic loci (P) and average number of alleles (Avg), F-coefficient of inbreeding (F_{IS}) and G_{IS} are listed. Values for some populations are missing due to their low sample size. Non-significant G_{IS} values. (*) Significant F_{IS} values based on a test of 10,000 permutations. (**) Significant F_{IS} values supported by Bonferroni.

Populations	N	H_{EXP}	H_{OBS}	P (0.99)	Avg	F_{IS}	G_{IS}
Leonese Shepherd dog	50	0.667	0.665	1.00	6.33	0.014	0.014
Border Collie	25	0.654	0.630	1.00	5.83	0.056*	0.056
Spanish Wolfhound	75	0.705	0.679	1.00	7.39	0.045**	0.045
German Shepherd dog	28	0.608	0.580	1.00	5.17	0.065	0.065
Czech Wolfhound	18	0.526	0.523	1.00	3.61	0.036	0.036
Herreño Shepherd dog	13	0.622	0.642	1.00	4.11	0.009	0.009
Galician Palleiro Sheepdog	22	0.692	0.677	1.00	6.17	0.046	0.046
Portuguese Sheepdog	25	0.686	0.642	1.00	5.33	0.084**	0.084
Manchego Sheepdog	30	0.749	0.709	1.00	7.56	0.070**	0.068
Castilian-Manchego Sheepdog	28	0.646	0.673	1.00	5.72	-0.024	-0.024

Can of Chira	19	0.643	0.686	1.00	5.06	-0.038	-0.038
Catalan Sheepdog	20	0.676	0.662	1.00	5.33	0.047	0.047
Portuguese Water dog	10	0.620	0.667	1.00	4.39	-0.022	-0.022
Terceira Cattle dog	22	0.699	0.724	1.00	6.00	-0.012	-0.012
Spanish Water dog	22	0.702	0.633	1.00	5.67	0.122**	0.122
Basque Shepherd	45	0.717	0.678	1.00	6.89	0.065**	0.065
Garafian Sheepdog	10	0.605	0.541	1.00	4.78	0.165**	0.165
Majorca Shepherd dog	26	0.709	0.674	1.00	6.39	0.068**	0.068
Villano Bulldog Cattle dog	23	0.593	0.566	1.00	5.06	0.069**	0.069
Saint Miguel Cattle dog	18	0.664	0.711	1.00	4.67	-0.043	-0.043

Table 6: Genetic differentiation F_{ST} values between LGDs populations for 18 microsatellites. Comparisons statistically supported by Bonferroni corrections in bold.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
1	0																							
2	**0.0113**	0																						
3	**0.0130**	**0.0147**	0																					
4	0.0301	0.0270	0.0163	0																				
5	0.0392	0.0284	0.0167	0.0736	0																			
6	0.0208	0.0320	0.0219	0.0710	0.0274	0																		
7	**0.0411**	**0.0364**	0.0150	0.0236	0.0537	0.0424	0																	
8	**0.0325**	**0.0295**	0.0154	0.0104	0.0259	0.0282	0.0193	0																
9	**0.0637**	**0.0550**	0.0382	0.0399	0.0336	0.0708	0.0392	0.0175	0															
10	**0.0628**	**0.0477**	**0.0272**	0.0671	0.0517	0.0561	0.0265	0.0280	0.0305	0														
11	**0.0435**	**0.0347**	**0.0227**	0.0171	0.0358	0.0203	0.0204	0.0042	0.0082	0.0158	0													
12	**0.0564**	**0.0592**	**0.0335**	0.0452	0.0666	0.0368	0.0272	0.0400	0.0331	**0.0480**	**0.0239**	0												
13	0.0441	0.0408	0.0271	0.0444	0.0321	0.0007	0.0327	0.0223	0.0213	0.0146	0.0082	0.0312	0											
14	**0.0609**	**0.0612**	0.0374	0.0542	0.0330	0.0391	0.0414	0.0258	0.0424	0.0350	0.0328	0.0583	0.0004	0										
15	**0.0524**	**0.0477**	**0.0248**	0.0149	0.0266	0.0376	**0.0290**	0.0072	0.0166	**0.0327**	**0.0192**	**0.0355**	0.0127	0.0241	0									
16	**0.0484**	**0.0386**	**0.0275**	0.0499	0.0165	**0.0478**	**0.0436**	0.0253	0.0114	**0.0355**	0.0146	**0.0479**	0.0250	0.0294	**0.0188**	0								
17	**0.0608**	**0.0412**	0.0296	0.0188	0.0308	0.0359	0.0297	0.0070	0.0086	0.0361	0.0122	0.0411	0.0168	0.0343	0.0100	0.0164	0							
18	0.0274	0.0271	-0.0041	0.0354	-0.0005	0.0152	0.0124	0.0132	0.0092	0.0086	0.0080	0.0157	0.0026	0.0040	0.0140	0.0126	0.0160	0						
19	**0.0483**	**0.0438**	**0.0261**	0.0258	0.0308	0.0259	0.0271	0.0197	0.0159	**0.0295**	0.0105	**0.0279**	-0.0029	0.0295	**0.0169**	**0.0204**	0.0152	0.0123	0					
20	**0.1139**	**0.1051**	**0.0864**	0.1086	**0.1069**	**0.0946**	**0.0850**	**0.0828**	0.0546	**0.0639**	**0.0565**	**0.0768**	0.0558	**0.0892**	**0.0714**	**0.0630**	**0.0701**	0.0550	**0.0644**	0				
21	**0.0682**	**0.0577**	**0.0407**	0.0585	0.0487	**0.0573**	**0.0496**	**0.0369**	0.0209	**0.0356**	**0.0254**	**0.0357**	0.0332	0.0391	**0.0359**	**0.0250**	0.0138	0.0182	**0.0377**	**0.0674**	0			
22	**0.0890**	**0.0852**	**0.0721**	0.0838	**0.0808**	**0.0638**	**0.0721**	**0.0592**	0.0447	**0.0638**	**0.0466**	**0.0665**	0.0531	**0.0466**	**0.0455**	**0.0573**	**0.0466**	0.0517	**0.0499**	**0.0895**	**0.0557**	0		
23	**0.1549**	**0.1580**	**0.1237**	**0.1791**	**0.1772**	**0.1553**	**0.1353**	**0.1253**	**0.1488**	**0.1361**	**0.1261**	**0.1275**	**0.1163**	**0.1426**	**0.1098**	**0.1378**	**0.1532**	**0.1275**	**0.1117**	**0.1533**	**0.1500**	**0.1528**	0	
24	**0.0957**	**0.0892**	**0.0771**	**0.0979**	**0.0613**	**0.0976**	**0.0736**	**0.0818**	**0.0521**	**0.0714**	**0.0632**	**0.0962**	**0.0743**	**0.0716**	**0.0686**	**0.0704**	**0.0673**	**0.0539**	**0.0694**	**0.0931**	**0.0705**	**0.0865**	**0.1605**	**0**

Table 7: Genetic differentiation F_{ST} values between LHDs populations for 18 microsatellites. All were significant values supported by Bonferroni corrections.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	0																			
2	0.1597	0																		
3	0.1139	0.1105	0																	
4	0.1785	0.1889	0.0736	0																
5	0.2492	0.2635	0.1627	0.1562	0															
6	0.1612	0.1063	0.0755	0.1386	0.2235	0														
7	0.1421	0.1639	0.0650	0.0935	0.2004	0.1420	0													
8	0.1486	0.1112	0.0842	0.1672	0.2174	0.1104	0.1480	0												
9	0.1853	0.1979	0.1085	0.1905	0.2988	0.1582	0.1354	0.1412	0											
10	0.1044	0.1526	0.0840	0.1628	0.2320	0.1253	0.1167	0.1044	0.1330	0										
11	0.1390	0.1272	0.0857	0.1519	0.2255	0.1174	0.1077	0.0831	0.0962	0.0730	0									
12	0.1697	0.1890	0.1186	0.1787	0.2415	0.1703	0.1458	0.1681	0.2259	0.1233	0.1412	0								
13	0.1595	0.1257	0.0927	0.1508	0.2169	0.1155	0.1160	0.1032	0.1312	0.0928	0.0721	0.1249	0							
14	0.2032	0.1859	0.1611	0.2379	0.3012	0.1857	0.2128	0.1654	0.2173	0.1437	0.1325	0.2197	0.1406	0						
15	0.1751	0.1490	0.1375	0.2054	0.2873	0.1710	0.1613	0.1379	0.1601	0.1183	0.0854	0.1868	0.0964	0.1531	0					
16	0.0732	0.1036	0.0620	0.1350	0.1907	0.1084	0.0801	0.0817	0.1074	0.0587	0.0526	0.1121	0.0816	0.1369	0.0994	0				
17	0.1683	0.1332	0.1122	0.1702	0.2622	0.1330	0.1568	0.1561	0.1774	0.1144	0.1145	0.1604	0.1078	0.1924	0.1282	0.0878	0			
18	0.1238	0.1595	0.1192	0.1771	0.2615	0.1501	0.1451	0.1312	0.1719	0.1385	0.1102	0.2091	0.1525	0.2423	0.1646	0.0757	0.1348	0		
19	0.1535	0.1326	0.0838	0.1721	0.2376	0.0972	0.1548	0.1202	0.1277	0.0959	0.0929	0.1351	0.0952	0.1703	0.1523	0.0747	0.0992	0.1401	0	
20	0.1118	0.0995	0.0590	0.1346	0.1990	0.0853	0.0979	0.1020	0.1276	0.0833	0.0776	0.1390	0.0945	0.1547	0.1150	0.0487	0.0722	0.0877	0.0706	0

Table 8: Analysis of Molecular Variance for LGD: 1. For the populations' groups Spanish mastiffs and not Spanish mastiffs; 2. For the populations' groups one and two from Bayesian procedure; 3. For the populations' groups one, two, three, four, five and six from Bayesian procedure; 4. For the populations' groups one, two, three, four, five, six and seven from Bayesian procedure; 5. For the populations' groups Working and Non-working mastiffs; 6. for the populations' groups Non-Transhumant and Transhumant mastiffs; 7. For the populations' groups Non Transhumant and León Transhumant mastiffs, Zamora Transhumant mastiffs, Sierra Morena Transhumant mastiffs and Cáceres Transhumant mastiffs; 8. For the populations' groups Zamora Transhumant mastiffs, Zamora non-transhumant Shelves and the group of the rest of populations mastiffs; 9. for the populations' groups from the North and South peninsular mastiffs; 10. For the populations' groups North mastiffs, Zamora mastiffs and the group of the rest of populations mastiffs; 11. For the populations' groups Transmontano Mastiff, South mastiffs, and Zamora mastiffs; 12. For the populations' groups Castro Laboreiro dog, South mastiffs, and Zamora mastiffs; 13. For the populations' groups Rafeiro of Alentejo, South mastiffs, and Zamora mastiffs; 14. For the populations' groups Estrela Mountain dog, South mastiffs, and Zamora mastiffs.

LGD Structure tested	Variance (%)	F_{ST}	*p* Value
1. Spanish vs. no-Spanish		0.07013	0.000±0.000
Among groups	1.58		
Among populations within groups	5.43		
Within populations	92.99		
2. One vs. two from Bayesian		0.07267	0.000±0.000
Among groups	2.26		
Among populations within groups	5.01		
Within populations	92.73		
3. One to six from Bayesian		0.07388	0.000±0.000
Among groups	4.95		
Among populations within groups	2.44		
Within populations	92.61		
4. One to seven from Bayesian		0.07384	0.000±0.000
Among groups	5.49		
Among populations within groups	1.89		
Within populations	92.62		
5. Exposition vs. no-Exposition		0.03269	0.000±0.000
Among groups	0.35		
Among populations within groups	2.92		
Within populations	96.73		
6. Non-Transhumant vs. Transhumant		0.02947	0.000±0.000
Among groups	0.11		
Among populations within groups	2.84		
Within populations	97.05		
7. Non-transhumant vs. geographical transhumant groups		0.03288	0.000±0.000
Among groups	0.96		
Among populations within groups	2.33		
Within populations	96.71		
8. Zamora transhumant vs. Zamora non-transhumant vs. Rest		0.03120	0.000±0.000
Among groups	0.33		
Among populations within groups	2.79		
Within populations	96.88		
9. North vs. South		0.03301	0.000±0.000
Among groups	0.73		

Among populations within groups	2.57		
Within populations	96.7		
10. North vs. Zamora vs. Rest		0.03193	0.000±0.000
Among groups	1.02		
Among populations within groups	2.18		
Within populations	96.81		
11. Transmontano Mastiff vs. South vs. Zamora		0.02078	0.000±0.000
Among groups	1.02		
Among populations within groups	2.18		
Within populations	96.81		
12. Castro Laboreiro dog vs. South vs. Zamora		0.04561	0.000±0.000
Among groups	2.62		
Among populations within groups	1.94		
Within populations	95.44		
13. Rafeiro of Alentejo vs. South vs. Zamora		0.02806	0.000±0.000
Among groups	0.91		
Among populations within groups	1.89		
Within populations	97.19		
14. Estrela Mountain dog vs. South vs. Zamora		0.04306	0.000±0.000
Among groups	2.34		
Among populations within groups	1.96		
Within populations	95.69		

Table 9: Analysis of Molecular Variance for LHD; 1. For the populations' groups herding dogs and cattle dogs; 2. For the populations' groups herding dogs and Cattle dogs including Majorca Cattle dog; 3. For the populations' groups herding dogs and canary breeds; 4. For the populations' groups herding dogs and wolfhound breeds; 5. For the populations' groups herding dogs and peninsula wolfhound breeds; 6. For the populations' groups herding dogs and water-like breeds; 7. For the populations' groups herding dogs and water dog breeds; 8. For the populations' groups herding dogs and sheepdogs' breeds, for the microsatellites.

LHD Structure tested	**Variance (%)**	**F_{ST}**	***p* Value**
1. Herding vs. peninsular cattle dogs		0.15819	0.000±0.000
Among groups	3.98		
Among populations within groups	11.84		
Within populations	84.18		
2. Herding vs. Cattle dogs		0.14685	0.000±0.000
Among groups	2.71		
Among populations within groups	11.98		
Within populations	85.32		
3. Herding vs. Canary breeds		0.12963	0.000±0.000
Among groups	0.10		
Among populations within groups	12.86		
Within populations	87.04		
4. Herding vs. wolfhound breeds		0.13913	0.000±0.000
Among groups	2.13		
Among populations within groups	11.79		

Within populations	86.09		
5. Herding vs. peninsular wolfhound breeds		0.13965	0.000±0.000
Among groups	2.15		
Among populations within groups	11.82		
Within populations	86.03		
6. Herding vs. water-like breeds		0.12950	0.000±0.000
Among groups	0.09		
Among populations within groups	12.86		
Within populations	87.05		
7. Herding vs. water dog breeds		0.12907	0.000±0.000
Among groups	0.03		
Among populations within groups	12.88		
Within populations	87.09		
8. Herding vs. Sheepdogs		0.13183	0.000±0.000
Among groups	0.59		
Among populations within groups	12.59		
Within populations	86.82		

Table 10: Table of the Factorial Correspondence Analysis LGD. The first two principal axes of a FCA explain 14.05% and 11.93% for axes 1 and 2.

Factor LGD	Val. propres	(%)	Accumulated
1	0.09128	14.05	14.05
2	0.07746	11.93	25.98
3	0.06811	10.49	36.47
4	0.05881	9.05	45.52

Table 11: Table of the Factorial Correspondence Analysis LHD. The first two principal axes of a FCA explain 14.05% and 11.93% for axes 1 and 2.

Factor LHD	Val. propres	(%)	Accumulated
1	0.13872	12.66	12.66
2	0.10793	9.85	22.51
3	0.09514	8.68	31.19
4	0.07929	7.23	38.42

FIGURES

Figure 1

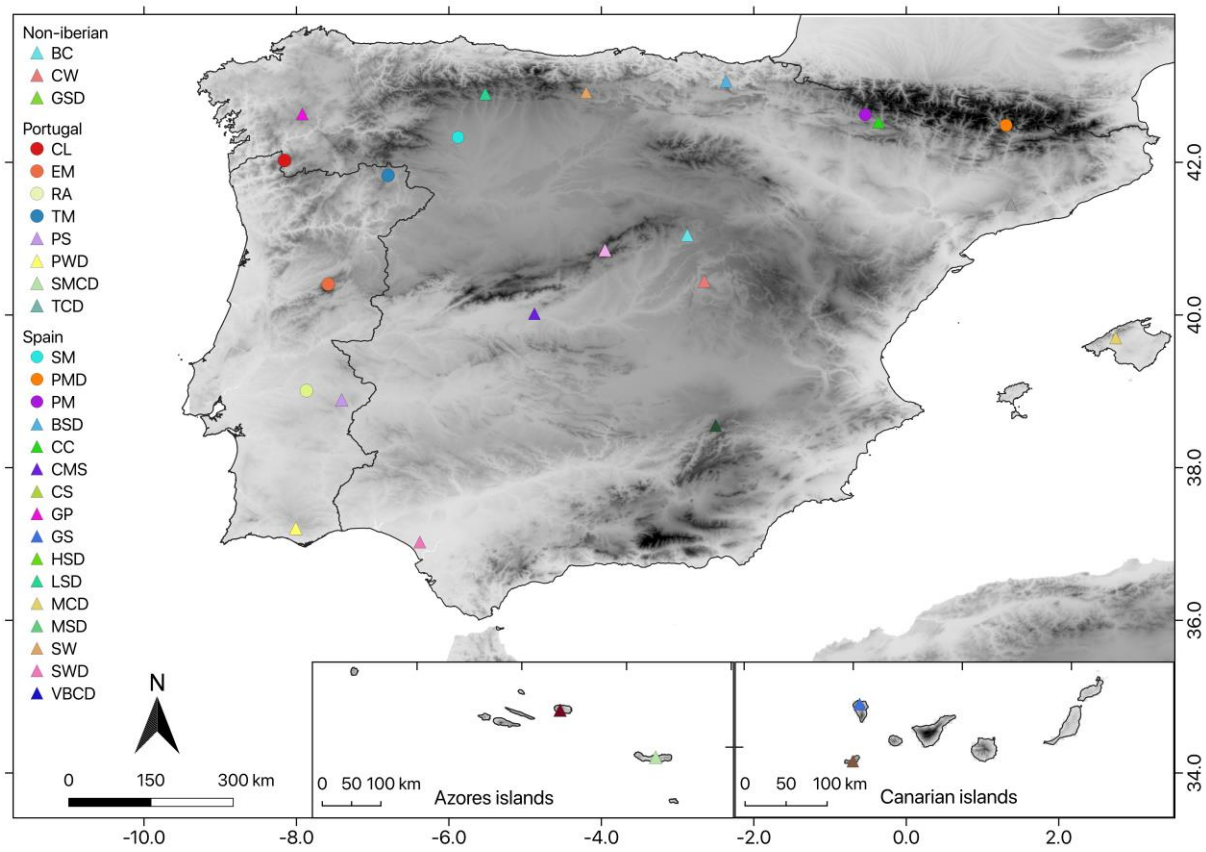


Figure 2

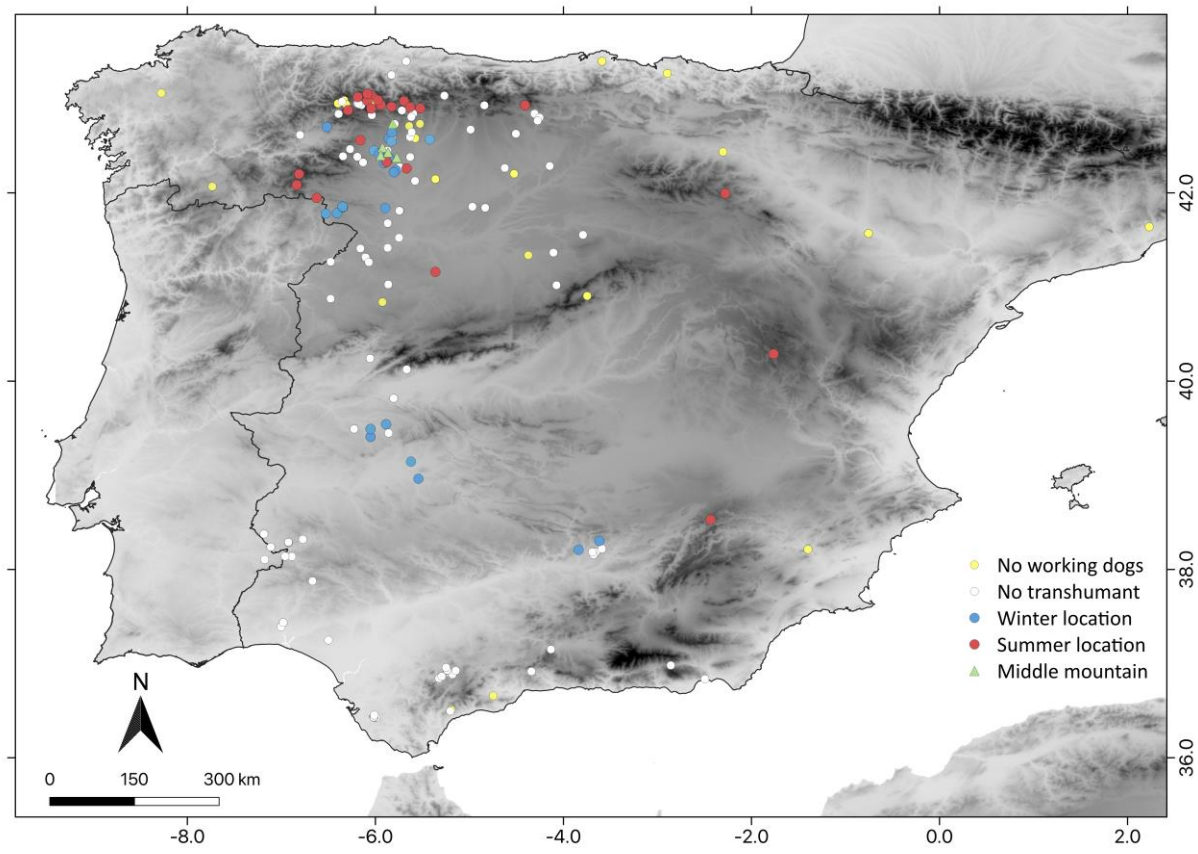

No working dogs
No transhumant
Winter location
Summer location
Middle mountain
N
0
150
300 km
-8.0
-6.0
-4.0
-2.0
0.0
2.0
42.0
40.0
38.0
36.0

Figure 3

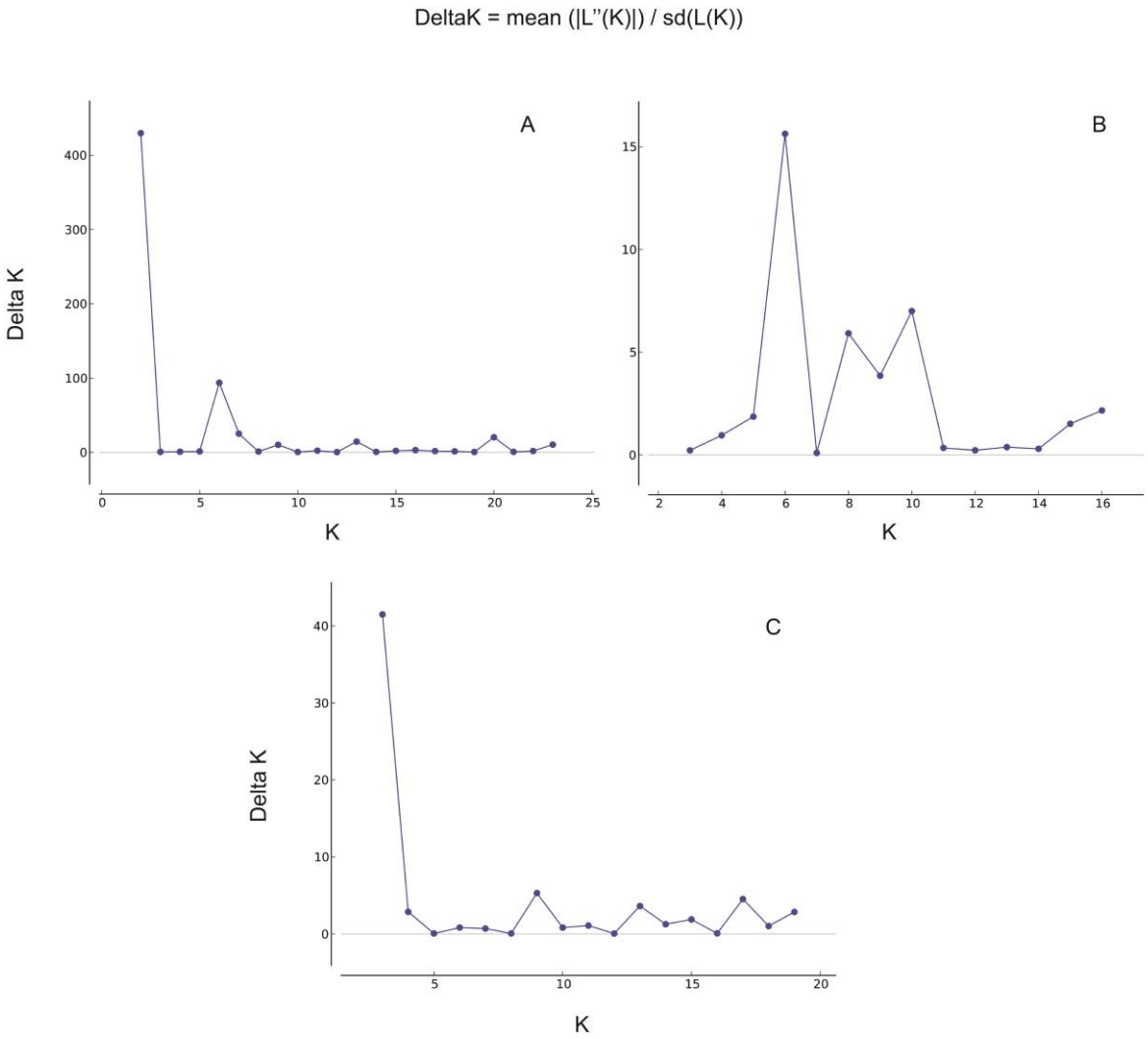


Figure 4

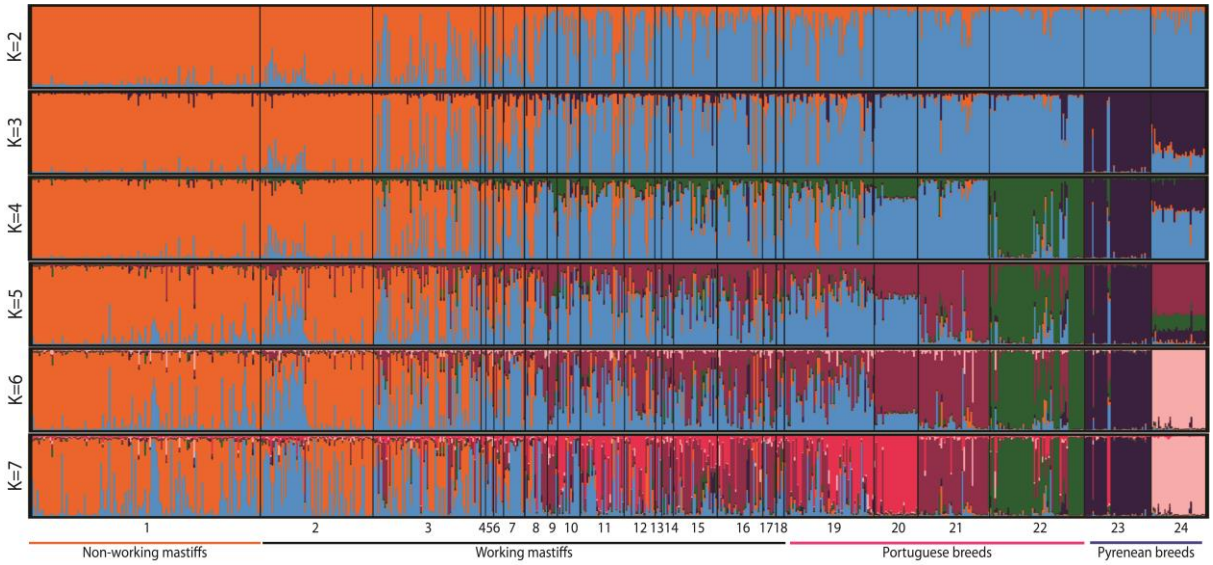


Figure 5

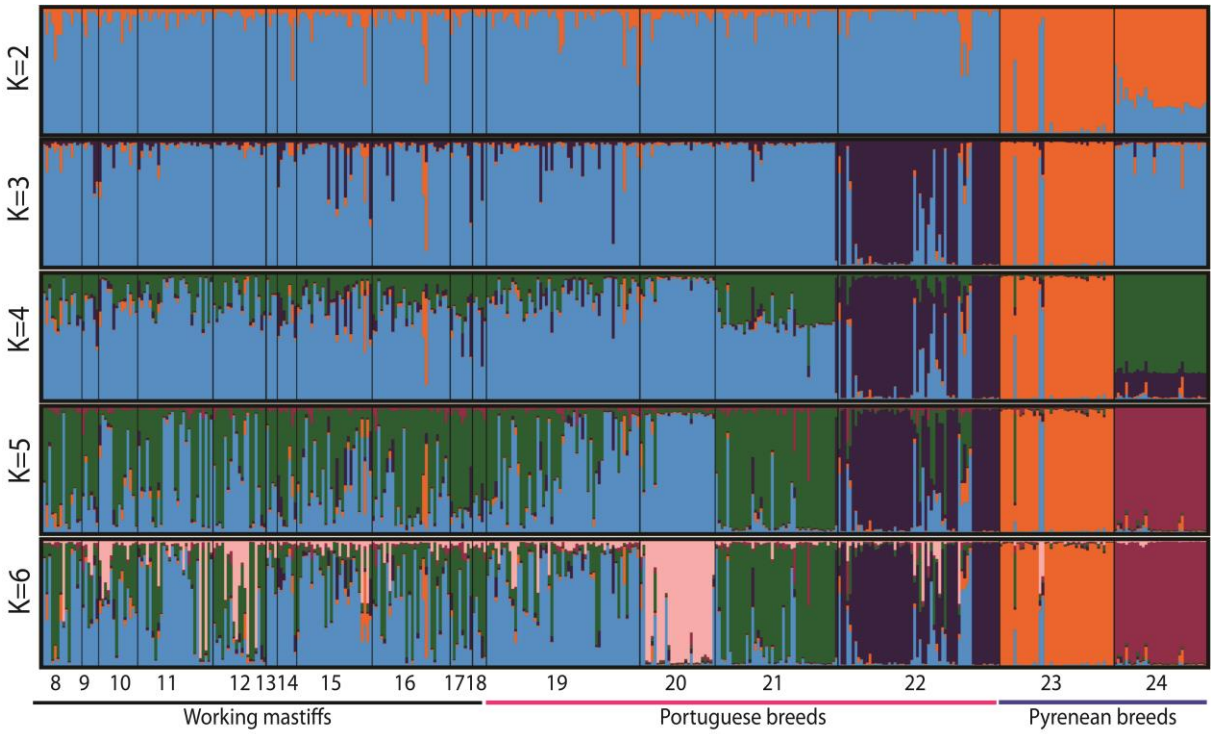


Figure 6

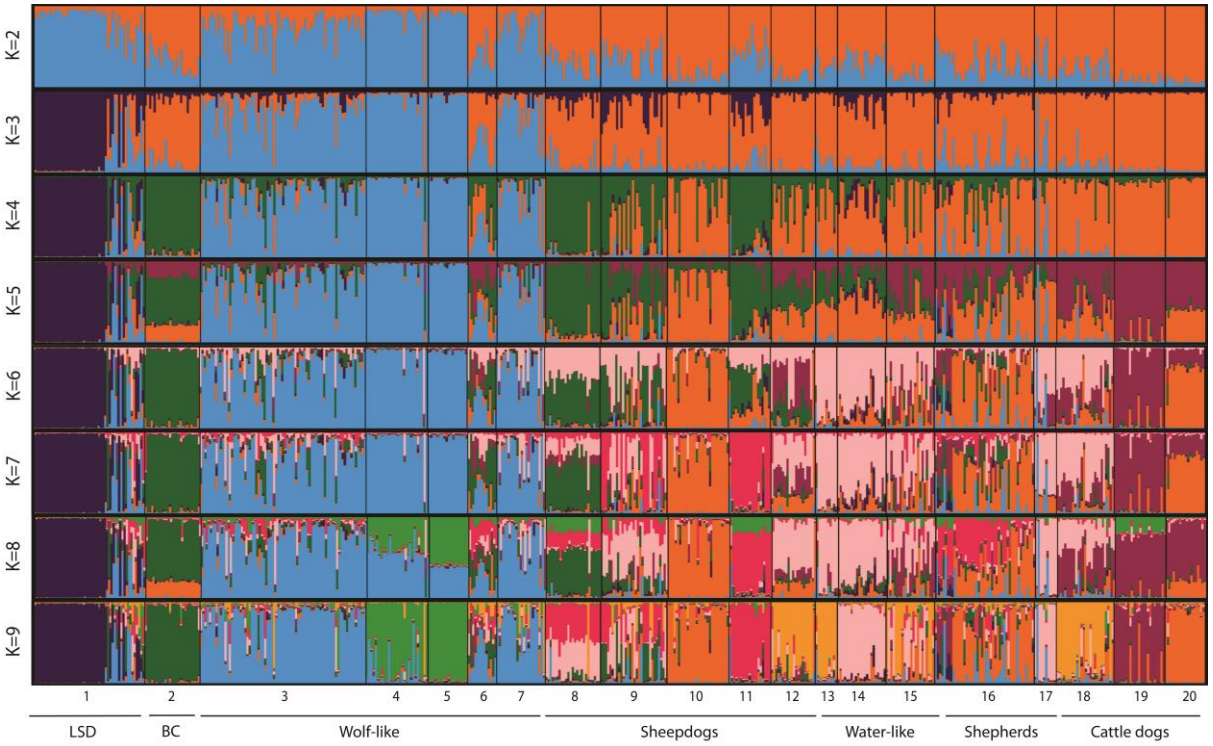
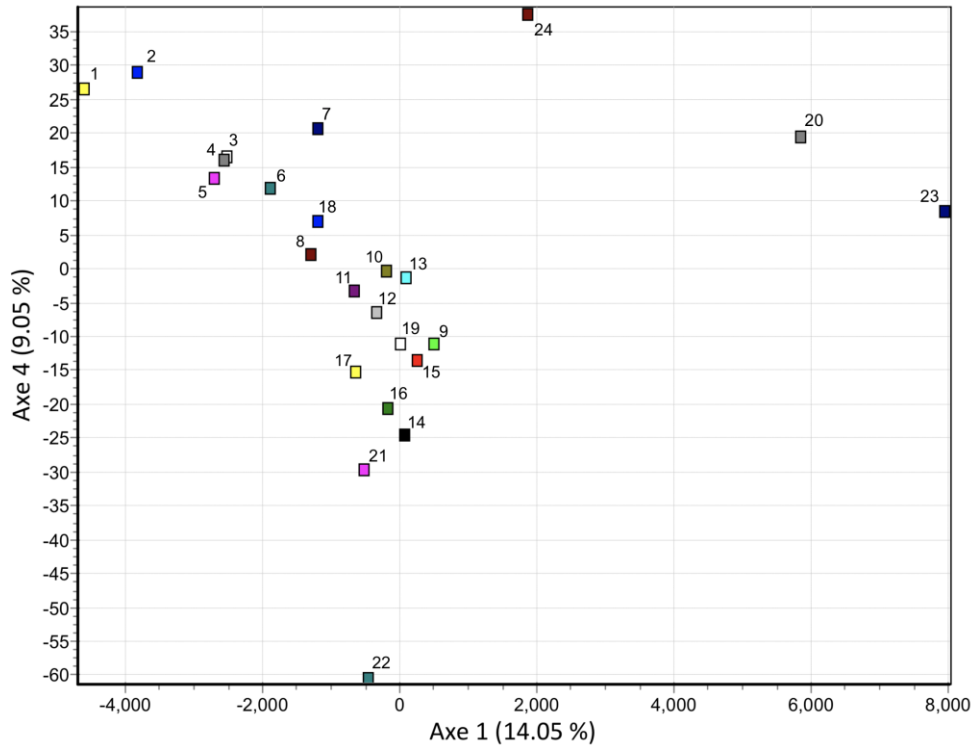


Figure 7


Axe 4 (9.05 %)
Axe 1 (14.05 %)
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
-55
-60
-4,000
-2,000
0
2,000
4,000
6,000
8,000
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

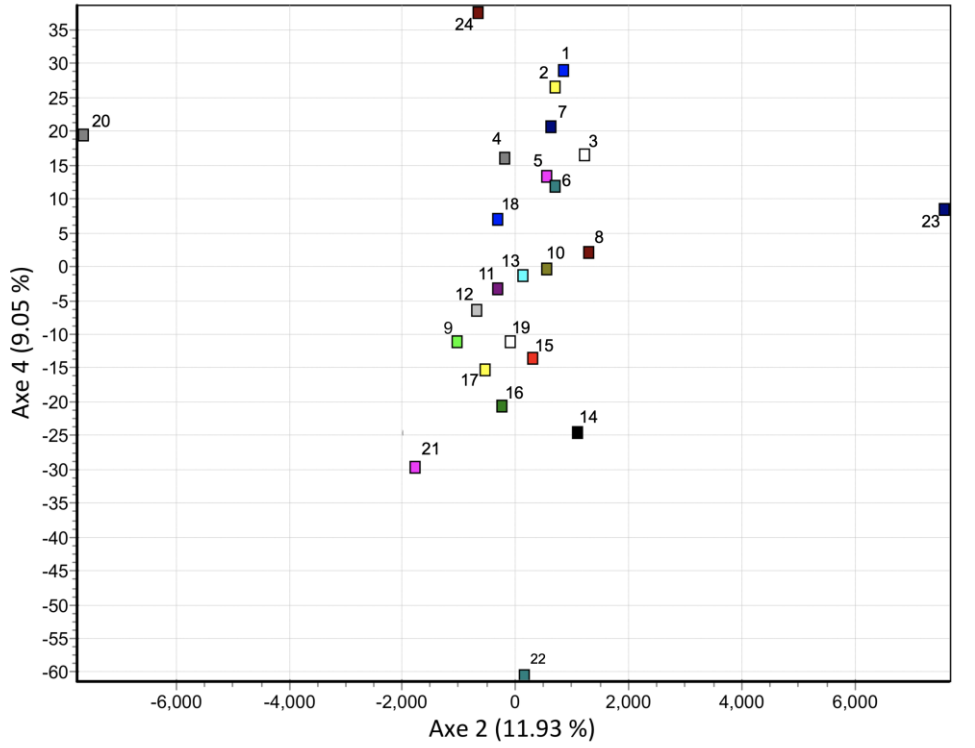
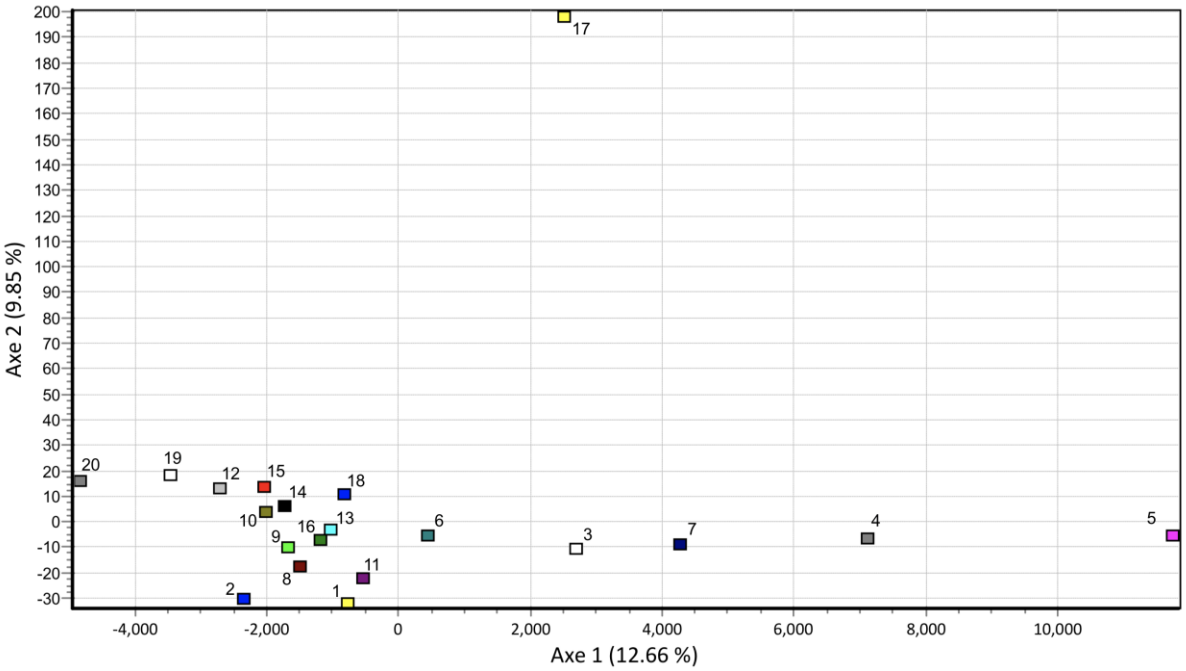

Axe 4 (9.05 %)
Axe 2 (11.93 %)
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
-55
-60
-6,000
-4,000
-2,000
0
2,000
4,000
6,000
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Figure 8


Axe 2 (9.85 %)
Axe 1 (12.66 %)
17
19
20
12
15
14
18
10
13
16
9
6
3
7
4
5
11
8
2
1

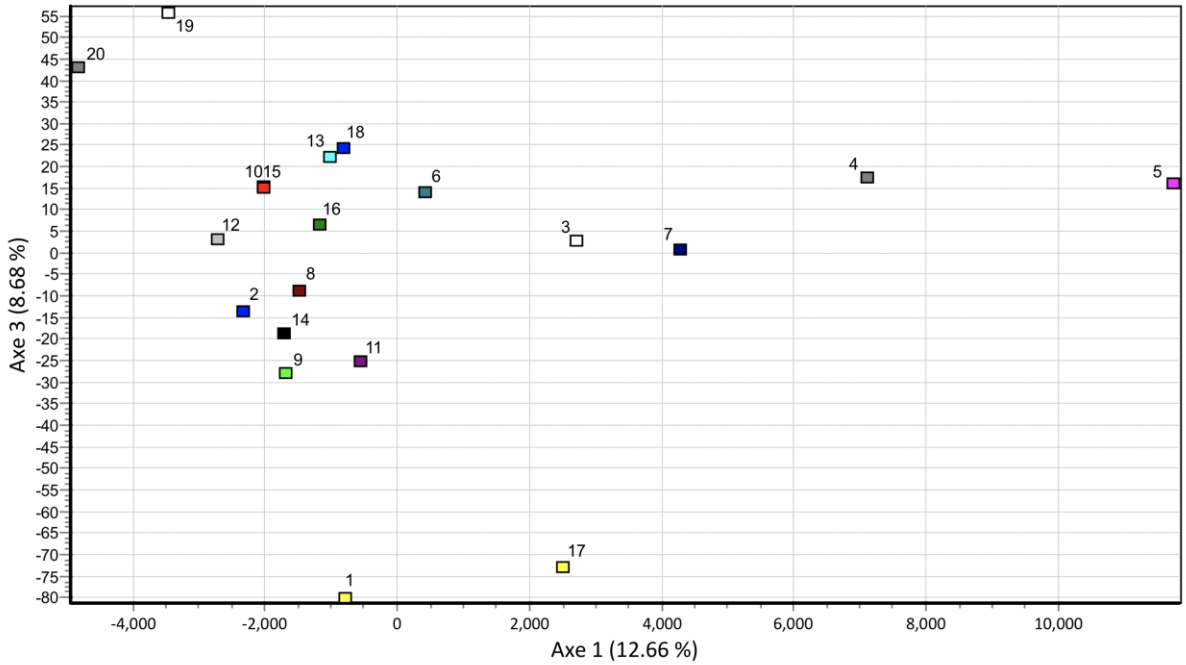

Axe 3 (8.68 %)
Axe 1 (12.66 %)
19
20
13
18
10
15
6
4
5
16
12
3
7
8
2
14
11
9
17
1

FIGURE LEGENDS

Figure 1: Location of the breeds. The circles are LGD: Spanish Mastiff (SM); Transmontano Mastiff (TM); Castro Laboreiro dog (CLD); Rafeiro of Alentejo (RA); Estrela Mountain dog (EMD); Pyrenean Mastiff (PM); Pyrenean Mountain dog (PMD). The triangles are HGD: Leonese Shepherd dog (LSD); Border Collie (BC); Spanish Wolfhound (SW); German Shepherd dog (GSD); Czech Wolfhound (CW); Herreño Shepherd dog (HSD); Galician Palleiro Sheepdog (GPS); Portuguese Sheepdog (PS); Manchego Shepherd dog (MSD); Castilian-Manchego Sheepdog (CMS); Can of Chira (CC); Catalan Sheepdog (CS); Portuguese Water dog (PWD); Terceira Cattle dog (TCD); Spanish Water dog (SWD); Basque Shepherd dog (BSD); Garafian Sheepdog (GS); Majorca Cattle dog (MCD); Villano Bulldog Cattle dog (VBCD); Saint Miguel Cattle dog (SMCD).

Figure 2: Location for Spanish Mastiff samples.

Figure 3: DeltaK graphs from analysis using STRUCTURE. (A) For the first LGD dataset the highest probability by Evanno method is achieved for K=2 and K= 6 for 24 populations. (B) For the second LGD dataset the highest probability by Evanno method is achieved for K= 6 for 17 populations. (C) For the LHD dataset the highest probability by Evanno method is achieved for K= 2 and K= 9.

Figure 4: Proportion of membership of 719 individuals from 24 mastiff populations from K=2 to K=7, as calculated by Structure software. Multiple runs for each K were concatenated using ClumpaK, and distruct was used to generate images.

Figure 5: Proportion of membership of 424 individuals from 17 mastiff populations from K=2 to 6, as calculated by Structure software. Multiple runs for each K were concatenated using ClumpaK, and distruct was used to generate images.

Figure 6: Proportion of membership of 529 individuals from 20 herding dogs populations from K=2 to 9, and including K = 9 cluster, as calculated by Structure software. Multiple runs for each K were concatenated using ClumpaK, and distruct was used to generate image

Figure 7: Scatter plot of the Factorial Correspondence Analysis. Each colour represents a dog breed population. The first two principal axes of a FCA explain 14.05% and 11.93% for axes 1 (top) and 2 (button). Inset: distribution of LGDs in the same FCA for axes 1 and 2 versus axe 4.

Figure 8: Scatter plot of the Factorial Correspondence Analysis. Each colour represents a dog breed population. The first principal axe of a FCA explains 12.66%. Inset: distribution of LHDs in the same FCA for axe 1 versus axes 2 and 3.

Annex

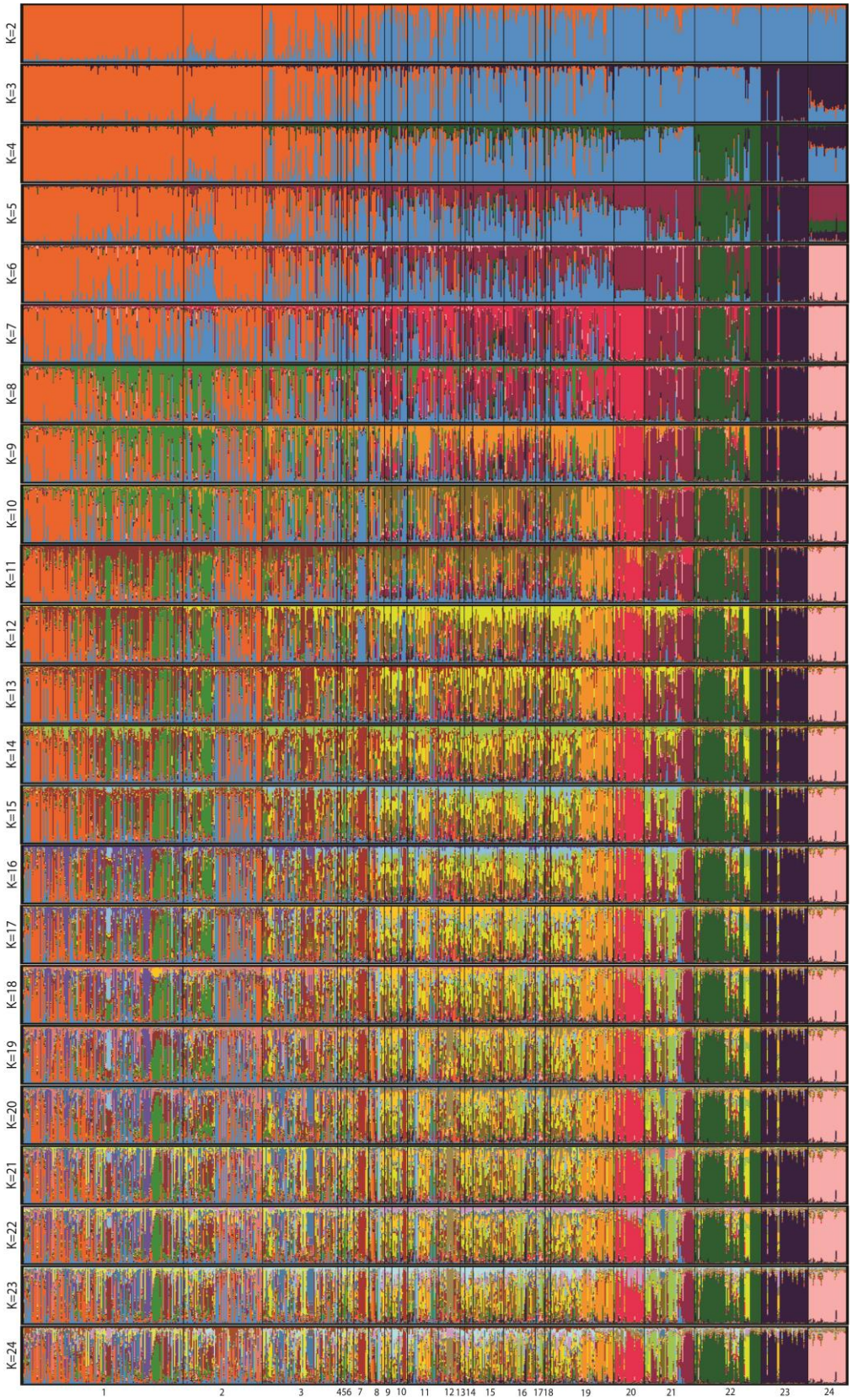


Figura 1: Proportion of membership of 719 individuals from 24 mastiff populations from K = 2 to K = 24.

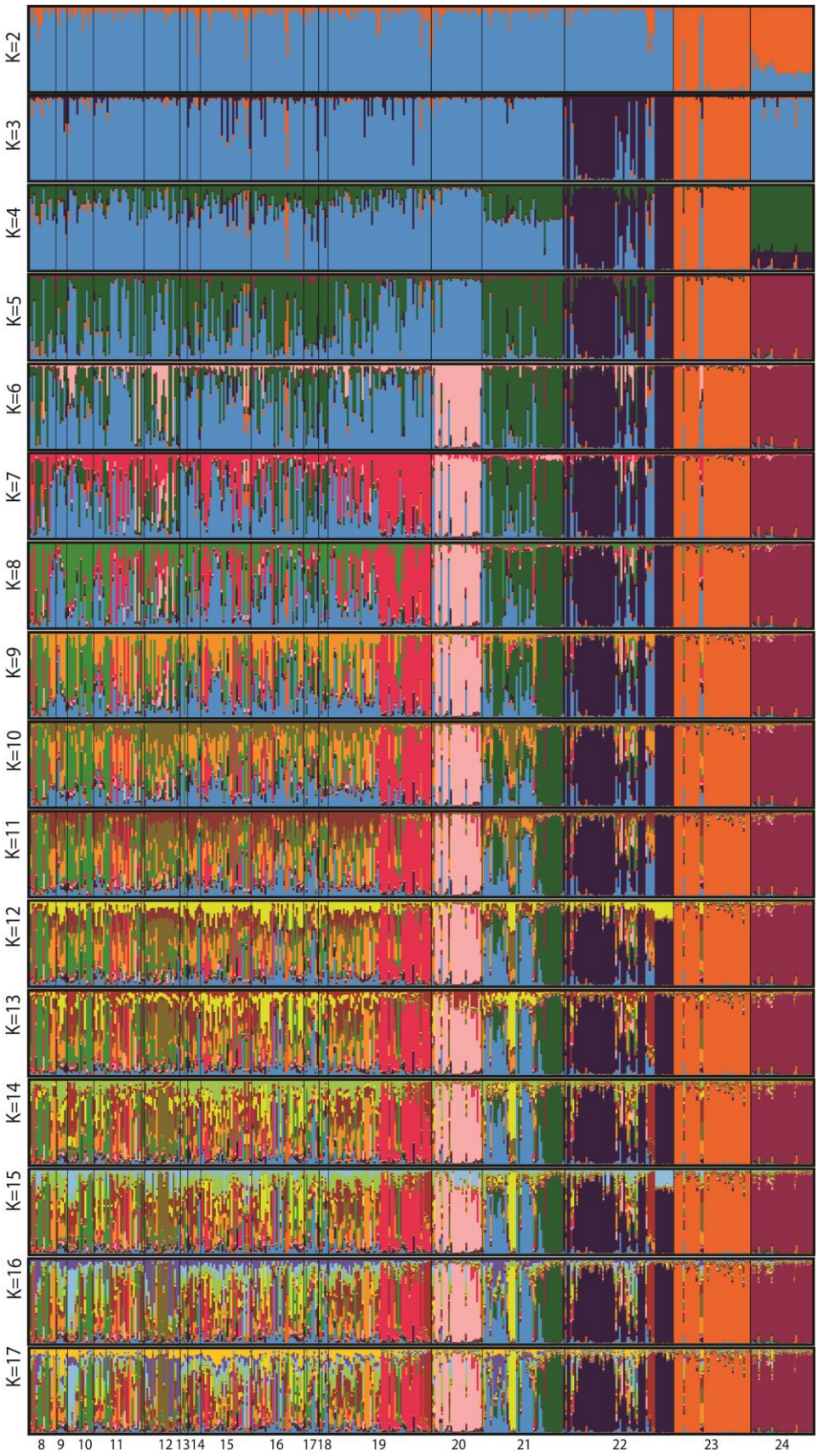


Figura 2: Proportion of membership of 719 individuals from 17 mastiff populations (8 - 24) from K = 2 to K = 17.

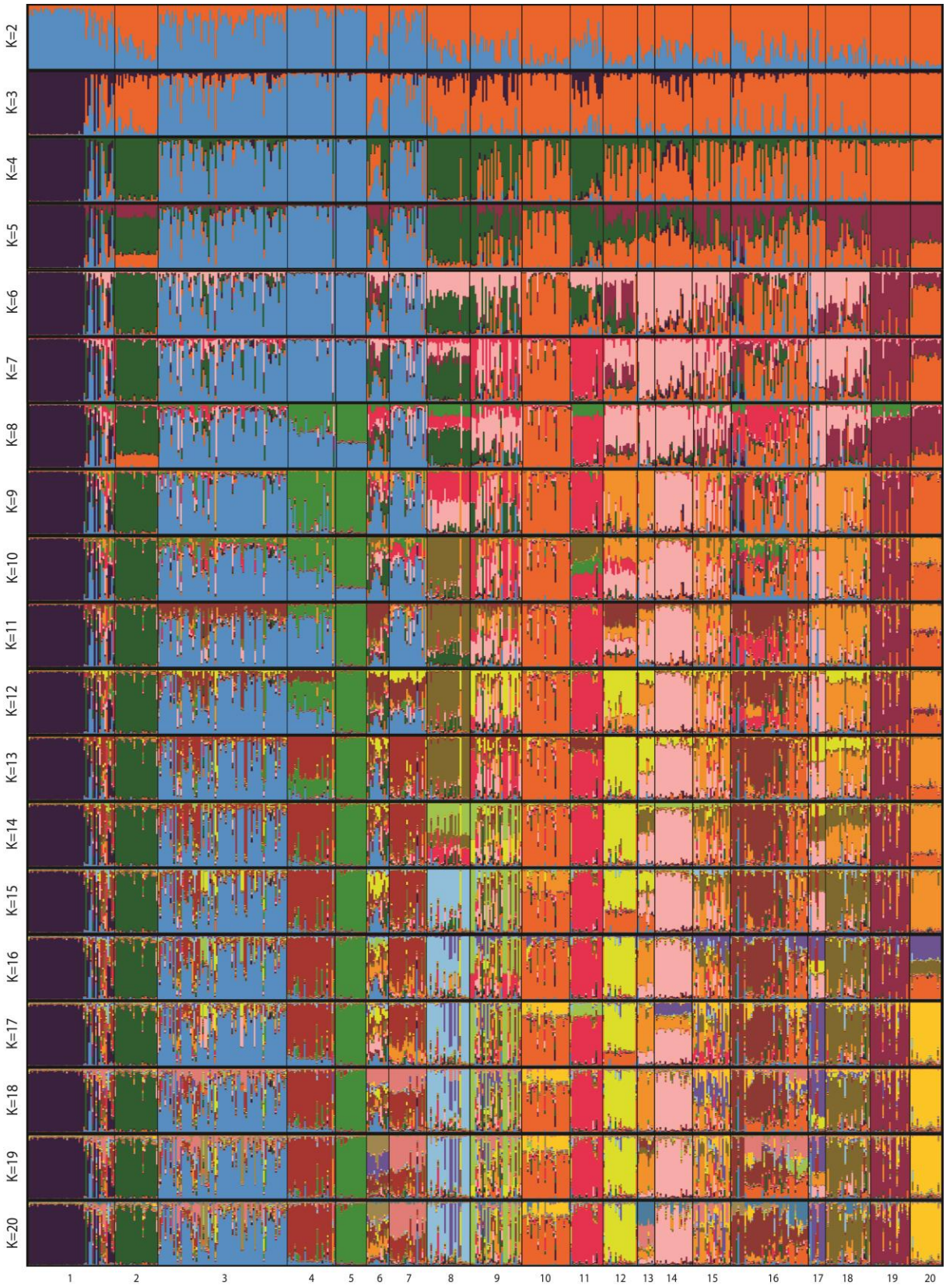


Figure 3: Proportion of membership of 529 individuals from 20 herding dogs populations from K = 2 to K = 20.