ORIGINAL ARTICLE

Half of 23 Belgian dog breeds has a compromised genetic diversity, as revealed by genealogical and molecular data analysis

K. Wijnrocx, L. François, A. Stinckens, S. Janssens & N. Buys

Department of Biosystems, KU Leuven, Livestock Genetics, 3001 Leuven, Belgium

Summary

The genetic diversity in 23 dog breeds raised in Belgium was investigated using both genealogical analysis and microsatellite markers. Some of these breeds are native breeds, with only small populations maintained. Pedigree and molecular data, obtained from the Belgian kennel club, were used to calculate the inbreeding coefficients, realised effective population size as well as probabilities of gene origin and average observed heterozygosity. Inbreeding coefficients ranged from 0.8 to 44.7% and realised effective population size varied between 3.2 and 829.1, according to the used method and breed. Mean observed heterozygosity ranged from 0.47 to 0.73. Both pedigree and molecular methods reveal low genetic diversity and presence of bottlenecks, especially in native Belgian breeds with small population sizes. Furthermore, principal component analysis on the set of investigated diversity parameters revealed no groups of breeds that could be identified in which similar breeding strategies could be applied to maintain genetic diversity.

Introduction

Selection for specific characteristics in purebred dogs has resulted in a wide diversity of breeds, with more than 400 breeds recognized by the Fédération Cynologique Internationale (FCI). Most dog breeds are closed populations, with no gene flow from outside, and only a small fraction of the dogs are used for reproduction (Calboli *et al.* 2008; Mäki 2010). Since mating between close relatives is frequently used to fixate traits, the expression of inherited defects in purebred dogs has increased, compromising their health and welfare (Leroy *et al.* 2006). This strategy may result in strong bottlenecks within the populations, leading to high rates of inbreeding (Leroy *et al.* 2006). For all these reasons, kennel clubs are more and more interested in parameters that evaluate the genetic variability in order to make decisions about selection and breeding policy.

For the evaluation of the genetic diversity in a population, two approaches are available. The first

approach is the use of genealogical data, which has been available for a long time and has been applied in several dog populations (Lüpke & Distl 2005; Calboli *et al.* 2008; Leroy *et al.* 2009a; Mäki 2010). The main parameters to investigate are the coefficient of inbreeding and coancestry and the associated effective population size (Leroy *et al.* 2009a). However, such approaches have their limitations, mainly due to the limited extent of pedigree knowledge, i.e. the proportion of registered individuals in the history of the breed (Leroy 2011). Registration of animals with limited or no pedigree knowledge of the parents may decrease the overall known pedigree information and may introduce bias in the calculation of parameters, since close relationships between some individuals may not be recognized (Leroy 2011). Furthermore, the existence of pedigree errors will also introduce bias, and this had been reported to be 1–10% (Leroy *et al.* 2012). Complementary, methods based on probabilities of gene origin (measures of genetic contributions of founders, ancestors or founder genomes)

have been applied to detect genetic bottlenecks and effects of genetic drift (Leroy 2011).

The second approach uses genetic markers to evaluate molecular genetic diversity and does not have the limitations of the first approach. This approach has already been applied in dog populations in several studies (Leroy *et al.* 2009a,b). The two main indicators of diversity are heterozygosity and allelic richness. However, this method also has its disadvantages, since it is prone to sampling error mostly due to a small number of genotyped individuals.

The aim of this study was to evaluate and compare the genetic diversity of 23 Belgian dog breeds with different registry sizes, using both methods. We also investigated the possibility to reduce the number of analysed diversity parameters and identify breed groups in which a similar breeding strategy can be applied by using principal component analysis (PCA). By this, we want to simplify the advice for the Belgian studbook and breed clubs in the decisions regarding their breed or breed group, such as use of certain sires or which parameters to investigate.

Materials and methods

Studied breeds

Genealogical and molecular information on 23 breeds were provided by the Belgian studbook KMSH (Koninklijke Maatschappij Sint-Hubertus). Among the breeds studied, all 14 native Belgian breeds with their breed standards kept in Belgium were selected, and these were compared to nine popular breeds of which data were available. This selection was made in consultation with the KMSH. Genealogical data comprised registrations between 1965 and 2013 for all breeds. Molecular data included genotyping results for 19 microsatellite markers (AHT121, AHT137, AHTh171, AHTh260, AHTk211, AHTk253, AMELO-GENIN, CXX279, FH2054, FH2848, INRA21, INU005, INU030, INU055, REN162C04, REN169D001, REN169O18, REN247M23 and REN54P11), recommended for paternity testing by the International Society for Animal Genetics (ISAG), and applied by the KMSH since 2008. Animals retained in this study include all breeding animals in the period 2008–2013, with the exclusion of full sibs.

Genealogical analysis

Genealogical analysis was performed using the PEDIG software (Boichard 2002) and own software routines programmed in SAS 9.3. (SAS Institute Inc., Cary, NC,

USA). For each breed, a reference population was defined as all animals born between 2000 and 2013 (covering at least two generations). Individuals with offspring but no parents in the data were defined as non-inbred and non-related founders in the analyses. The following parameters were calculated for the reference population: number of complete generations equivalent (CGE) (defined as the sum over all generations of the proportion of known ancestors at each generation), coefficient of inbreeding (F) and average coefficient of coancestry (C). Furthermore, the method based on the individual increase in inbreeding coefficients described by Gutiérrez *et al.* (2009) (N_e), as well as the method based on increase in pairwise coancestry (Cervantes *et al.* 2011) (N_{ec}) were used to compute the rate of inbreeding per generation (ΔF), the rate of coancestry (ΔC) and the effective population size (N_e and N_{ec}). Both methods take into account the number of ancestral generations known for each individual, which overcomes difficulties comparing breeds with different pedigree lengths. From a general point of view, for a domestic population, the method based on coancestry is the most appropriate, since it takes into account both differences in pedigree depth and population substructure (Leroy *et al.* 2013). The ratio N_{ec}/N_e was calculated to ascertain the presence of population structure (Cervantes *et al.* 2011).

As an additional method to assess genetic diversity, we used methods based on probabilities of gene origin. We assessed the effective number of founders (f_e) (Lacy 1989), the effective number of ancestors (f_a) (Boichard *et al.* 1997) and the equivalent of founder genomes (N_g) (Lacy 1989) using the PEDIG software (Boichard 2002). In addition the ratios f_e/f (indication of genetic contributions in the population), f_a/f_e (reveals decrease in genetic variation due to impact of genetic bottlenecks on breed populations), and N_g/f_e (describes the impact of genetic drift on the population) were computed (Lacy 1989).

Molecular analysis

For the analysis of microsatellite data, several parameters related to the genetic variability and allelic diversity were calculated: the total number of alleles (N), mean observed heterozygosity (H_o) and mean expected heterozygosity (H_e) using Arlequin (Excoffier *et al.* 2005). Mean allelic richness per population, a measure for the number of alleles, was calculated using FSTAT, using El Mousadik and Petit's rarefaction method (Goudet 1995; El Mousadik & Petit 1996). Weir & Cockerham's (1984) inbreeding coefficient F_{is} was estimated using GENEPOP (Raymond &

Rousset 1995). For each breed, a test for deviation from Hardy–Weinberg equilibrium (HWE) was performed using GENEPOP (Raymond & Rousset 1995).

Principal component analysis

As the number of parameters investigated in the diversity study is high, PCA on the correlation matrix between the parameters was used to simplify the number of investigated traits. Furthermore, we explored the distribution of breeds according to the principal components.

Results

Genealogical analysis

The genealogical parameters calculated for the 23 investigated breeds are shown in Table 1. One of the breeds with the largest registrations per year in Belgium is the German shepherd (mean of 1765 registrations per year in the period 2003–2012), followed by the Malinois shepherd (1151) and the Golden retriever (933). Among breeds with a low number of registrations in the same period are the Phalène (6),

Bichon frisé (9) and the Griffon Belge (11). Pedigree depth, as determined by CGE varied strongly among breeds. Completeness of pedigree was highest in the Malinois shepherd (average CGE equal to 7.1), whereas the Bichon frisé provided the poorest pedigree information (average CGE equal to 2.1). Other breeds with reasonably high CGE were the Bouvier des Flandres (CGE = 6.56), German shepherd (CGE = 6.09) and Schipperke (CGE = 6.09). Breeds with low pedigree completeness comprised mostly populations with a low number of registrations (Laekenois shepherd, Griffon Belge, and Phalène).

Average inbreeding per breed for the reference population ranged from 0.8 (Irish red setter) to 44.7 (Bouvier des Ardennes), whereas average kinship values ranged from 0.5 (Cavalier King Charles spaniel) to 42.2 (Bouvier des Ardennes).

The realised effective population size based on the increase in inbreeding varied between 3.2 (Bouvier des Ardennes) and 829.1 (Australian shepherd). Nine of the analysed populations showed a N_e below 50, namely the Bouvier des Ardennes, Groenendael shepherd, Laekenois shepherd, Tervueren shepherd, Schipperke, Petit Brabançon, Griffon Bruxellois, Bichon frisé, and Phalène. Using the individual

Table 1 Genetic diversity measures based on genealogical data

Breed	Abbrev.	N_ped	NP	CGE	*F*	C	N_e	N_{ec}	N_{ec}/N_e
Australian shepherd	AUST	2494	165	3.44	1.4	0.7	829.1	188.5	0.23
Bichon frisé	BICH	155	9	2.14	10.0	3.5	17.8	18.2	1.02
Bloodhound	HUBR	742	44	3.71	1.4	2.1	108.5	64.8	0.60
Border Collie	BOCO	13 142	875	4.70	1.9	0.8	99.1	284.6	2.87
Bouvier des Ardennes	BARD	281	20	4.87	44.7	42.2	3.2	4.5	1.41
Bouvier des Flandres	BOUV	9452	513	6.56	4.8	1.6	55.8	170.2	3.05
Boxer	BOXR	5851	388	3.91	2.5	1.7	57.6	119.8	2.08
Cavalier King Charles spaniel	CKCS	4783	313	3.56	0.9	0.5	106.6	327.1	3.07
German shepherd	GERM	30 622	1765	6.09	2.1	0.9	119.9	334.9	2.79
Golden retriever	GOLD	14 986	933	4.86	1.6	1.3	123.4	177.9	1.44
Griffon Belge	GBLG	265	11	2.95	1.1	2.3	80.5	34.2	0.43
Griffon Bruxellois	GBXL	654	24	4.17	5.3	3.1	29.3	42.9	1.46
Groenendael shepherd	GROE	2557	172	3.68	2.3	1.3	44.4	149.1	3.36
Irish red setter	ISET	2266	152	3.87	0.8	1.6	187.1	131.3	0.70
Labrador retriever	LABR	13 148	773	4.88	1.9	0.8	106.9	269.6	2.52
Laekenois shepherd	LAEK	559	39	3.10	5.1	5.2	22.1	29.4	1.33
Malinois shepherd	MALI	18 445	1151	7.07	4.8	2.2	61.3	157.2	2.56
Papillon	PAPA	1232	73	3.76	1.9	1.2	63.3	145.6	2.30
Petit Brabançon	BRAB	445	26	3.40	3.8	3.8	21.4	45.5	2.13
Phalène	PHAL	159	6	3.06	1.7	2.3	43.3	26.6	0.61
Rottweiler	ROTT	7247	508	4.40	1.8	1.3	96.4	183.0	1.90
Schipperke	SCHI	1530	82	6.09	7.7	6.0	31.3	39.8	1.27
Tervueren shepherd	TERV	5384	295	6.02	5.0	3.4	48.1	78.3	1.63

Abbrev. = Abbreviation of the breed name, N_ped = number of dogs in the pedigree file, NP = mean number of pups per year (period 2003–2012), CGE = complete generation equivalent, F = average coefficient of inbreeding, C = coefficient of kinship, N_e = effective population size based on individual increase in inbreeding, and N_{ec} = effective population size based on individual increase in coancestry.

increase in coancestry the realised effective size ranged between 4.5 (Bouvier des Ardennes), and 334.9 (German Shepherd). The Australian shepherd, German shepherd, Irish red setter, Golden retriever, Labrador retriever, and Cavalier King Charles spaniel are the only breeds with both N_e and N_{ec} over 100. A high ratio N_{ec}/N_e (larger than 2) was reported in 14 of 23 breeds.

Other indicators for the genetic diversity, namely the probability of gene origin, are shown in Table 2. The number of founders (*f*) ranged from 4 (Bouvier des Ardennes) to 2286 (German shepherd) while the effective number of founders (f_e) ranged from 3 (Bouvier des Ardennes) to 302 (Australian shepherd). A large difference between the number of founders and the effective number of founders was found in some breeds, indicating unbalanced genetic contributions. The effective number of ancestors (f_a) ranged from 3 (Bouvier des Ardennes) to 157 (Cavalier King Charles spaniel) and the equivalent of founder genomes (N_g) ranged from 1.2 (Bouvier des Ardennes) to 78.1 (Cavalier King Charles spaniel). A low f_a/f_e ratio was found to be substantial in the Australian shepherd, Malinois shepherd, Papillon and the Phalène breeds

(0.35; 0.29; 0.36 and 0.36, respectively). In most other breeds, this ratio ranged from 0.40 (Petit Brabançon) to 0.75 (Bichon frisé) but in the Bouvier des Ardennes both effective numbers were almost equal ($f_a/f_e = 0.99$), illustrating the absence of a genetic bottleneck in this breed.

Molecular analysis

Results for the molecular analysis are presented in Table 3. For the 19 autosomal microsatellites a total of 179 alleles were identified, with a mean number of 9.42 alleles per locus. The number of alleles per locus ranged from 6 (AHTk211) to 15 (AHT121). The allelic richness ranged from 2.60 (Bloodhound) to 4.04 (Malinois shepherd). The average expected heterozygosity (H_e) over all loci varied from 0.467

Table 2 Analysis of the probabilities of gene origin for the studied breeds

Breed	*f*	f_e	f_a	N_g	f_e/f	f_a/f_e	N_g/f_e
Australian shepherd	417	302	107	52.0	0.72	0.35	0.17
Bichon frisé	33	13	10	7.2	0.40	0.75	0.55
Bloodhound	124	91	40	18.6	0.74	0.44	0.20
Border Collie	1024	248	115	55.5	0.24	0.47	0.22
Bouvier des Ardennes	4	3	3	1.2	0.78	0.99	0.38
Bouvier des Flandres	509	177	79	31.3	0.35	0.45	0.18
Boxer	573	117	49	28.1	0.20	0.42	0.24
Cavalier King Charles spaniel	708	294	157	78.1	0.41	0.54	0.27
German shepherd	2286	259	121	48.5	0.11	0.47	0.19
Golden retriever	730	139	69	37.5	0.19	0.50	0.27
Griffon Belge	76	53	27	12.4	0.69	0.52	0.23
Griffon Bruxellois	132	50	26	12.6	0.38	0.51	0.25
Groenendael shepherd	272	129	74	34.0	0.47	0.57	0.26
Irish red setter	250	137	61	28.5	0.55	0.44	0.21
Labrador retriever	1011	229	116	55.7	0.23	0.51	0.24
Laekenois shepherd	70	22	16	8.2	0.32	0.72	0.37
Malinois shepherd	645	135	39	21.8	0.21	0.29	0.16
Papillon	239	150	54	26.8	0.63	0.36	0.18
Petit Brabançon	88	44	18	10.0	0.50	0.40	0.23
Phalène	42	32	12	6.5	0.76	0.36	0.20
Rottweiler	598	181	85	37.2	0.30	0.47	0.21
Schipperke	130	31	19	8.0	0.24	0.61	0.25
Tervueren shepherd	439	81	34	14.5	0.18	0.43	0.18

f = number of founders, f_e = effective number of founders, f_a = effective number of ancestors and N_g = number of founder genomes.

Table 3 Molecular genetic diversity measures for the 23 Belgian breeds

Breed	N_{animal}	N	A_r	H_e	H_o	F_{is}	p
Australian shepherd	151	127	3.75	0.658	0.673	−0.023	
Bichon frisé	13	71	3.09	0.581	0.623	−0.077	
Bloodhound	39	63	2.56	0.467	0.462	0.013	
Border Collie	1166	132	3.84	0.665	0.644	0.031	*
Bouvier des Ardennes	22	89	3.80	0.668	0.714	−0.040	
Bouvier des Flandres	446	116	3.54	0.637	0.607	0.047	
Boxer	279	78	2.60	0.488	0.480	0.017	
Cavalier King Charles spaniel	375	77	2.76	0.516	0.505	0.021	
German shepherd	880	113	3.06	0.578	0.562	0.026	
Golden retriever	721	105	3.12	0.564	0.561	0.006	
Griffon Belge	19	68	2.96	0.573	0.590	−0.031	
Griffon Bruxellois	27	74	2.89	0.541	0.552	−0.021	
Groenendael shepherd	151	99	3.02	0.551	0.544	0.013	
Irish red setter	132	89	3.11	0.561	0.555	0.010	
Labrador retriever	672	112	3.34	0.597	0.577	0.034	*
Laekenois shepherd	48	74	2.93	0.565	0.583	−0.031	
Malinois shepherd	1185	139	4.04	0.721	0.698	0.032	*
Papillon	133	100	3.64	0.664	0.674	−0.014	
Petit Brabançon	55	71	2.90	0.506	0.523	−0.034	
Phalène	6	66	3.35	0.636	0.725	−0.155	
Rottweiler	551	89	2.93	0.534	0.536	−0.004	
Schipperke	121	94	3.13	0.568	0.554	0.025	
Tervueren shepherd	378	109	2.98	0.561	0.540	0.038	*

N_{animal} = total number of sampled animals, N = total number of alleles over all loci, A_r = allelic richness, H_o = observed heterozygosity, H_e = expected heterozygosity, and F_{is} = fixation index, p = significance for deviation from Hardy–Weinberg equilibrium.

(Bloodhound) to 0.721 (Malinois shepherd), while the observed heterozygosity (H_o) ranged from 0.462 (Bloodhound) to 0.725 (Phalène). F_{is} values ranged between -0.155 (Papillon) and 0.047 (Bouvier des Flandres), and were significantly different from 0 in all populations but the Groenendael shepherd, German shepherd, Schipperke, Boxer, Rottweiler, Bloodhound, Irish red setter, Golden retriever, Cavalier King Charles spaniel and Papillon.

Principal component analysis

As shown in Figure 1, PCA 1 is composed of the variables number of founders, ancestors and founder genomes, and inbreeding and coancestry, whereas PCA 2 is mainly determined by molecular criteria. The first component accounted for 44% of variation, whereas components 2 and 3 accounted for 31 and 11%, respectively. Measures that were calculated on a similar method are usually grouped together, such as Delta_C, Delta_F, C, and F, and f_e, f_a, f, and N_g on the other hand. The small angle between these measures, represented here as vectors, also indicate a high pairwise correlation between them. Breeds are scattered according to their coordinates on the first three principal components, and no clear clusters of breeds could be identified. Some breeds were clearly separated from the group of breeds by their extreme values for the variables, such as the Bouvier des Ardennes (BARD), which has high inbreeding and coancestry values.

Discussion

The present study analysed pedigree- and molecular data for 23 different dog breeds in Belgium. Some of the analysed breeds are native breeds, with their origins and breed standards located in Belgium and only small populations are maintained. The impact of breeding practices on the genetic diversity and the level of inbreeding are critical and of great interest for the kennel clubs.

The pedigree length and completeness is an important factor to take into account as incomplete pedigree information can result in underestimated inbreeding levels (Shariflou *et al.* 2011). Assessed using CGE, six breeds presented very low values, prompting a cautious interpretation of the population status, as overestimation of number of founders and underestimation of inbreeding are probable (Leroy *et al.* 2006; Shariflou *et al.* 2011). The investigated Belgian populations showed lower CGE's compared to other countries (Leroy *et al.* 2009b; Shariflou *et al.* 2011). For

some breeds, such as the Australian shepherd, a high number of import dogs without pedigree information can explain low CGE values but for other breeds, the reason is unclear.

The effective population size is one of the most important measures of long-term performance of a population, regarding both diversity and inbreeding, and therefore, characterizing the risk status of breeds (FAO, 1998). Unlike N_e, N_{ec} accounts for possible population structure by using the individual increase in coancestry (Cervantes *et al.* 2011). Breeds showing a high N_{ec}/N_e ratio showed a significant amount of substructure, like the Bouvier des Flandres and Golden Retriever, among others. In the Golden Retriever different types ('lines') of dogs are bred (show dogs and retrieving dogs) (Windig & Oldenbroek 2015), which may be a possible explanation for the substructure. In other breeds, such as the Australian shepherd or the Bloodhound, the ratio between N_{ec}/N_e is small, probably indicating active avoidance of mating between relatives. Additionally the use of only a small fraction of purebred animals as reproductive animals results in small N_e values as does an imbalance in founder contributions (Mäki *et al.* 2001; Calboli *et al.* 2008).

As an alternative method, the probabilities of gene origin were investigated. Through the effective and actual number of founders and the ratio, unequal contributions of breeding animals, like the use of popular sire and champion stud dogs, can be assessed (Lacy 1989). It has to be noted that some breeds, such as the Bouvier des Ardennes and Bichon frisé have a low number of total founders, which was also reflected by their negative scores for PCA1, which does not allow for unbalanced contributions (Leroy *et al.* 2006). Comparison of the effective number of founders and ancestors (f_a/f_e) reveals genetic bottlenecks (Boichard *et al.* 1997), as can be seen in Australian shepherd, Malinois shepherd, Papillon and Phalène breed. Moreover, these breeds as well as the Tervueren shepherd and the Bouvier des Flandres have a lower N_g/f_e ratio, indicating that a strong genetic drift is present in these breeds (Lacy 1989).

The values of mean inbreeding coefficients (F) differ strongly between breeds, likely due to the varying number of founder animals and founder genomes. Breeds with low inbreeding coefficients (<0.05) had higher numbers of founders and founder genomes compared to the other breeds. This was also reflected in the opposite directions of both variables along PCA1. Also, breeds with high inbreeding coefficients in general had a lower effective population size N_e. In the Bouvier des Ardennes the limited number of dogs used to establish this breed at its recent (re-)creation
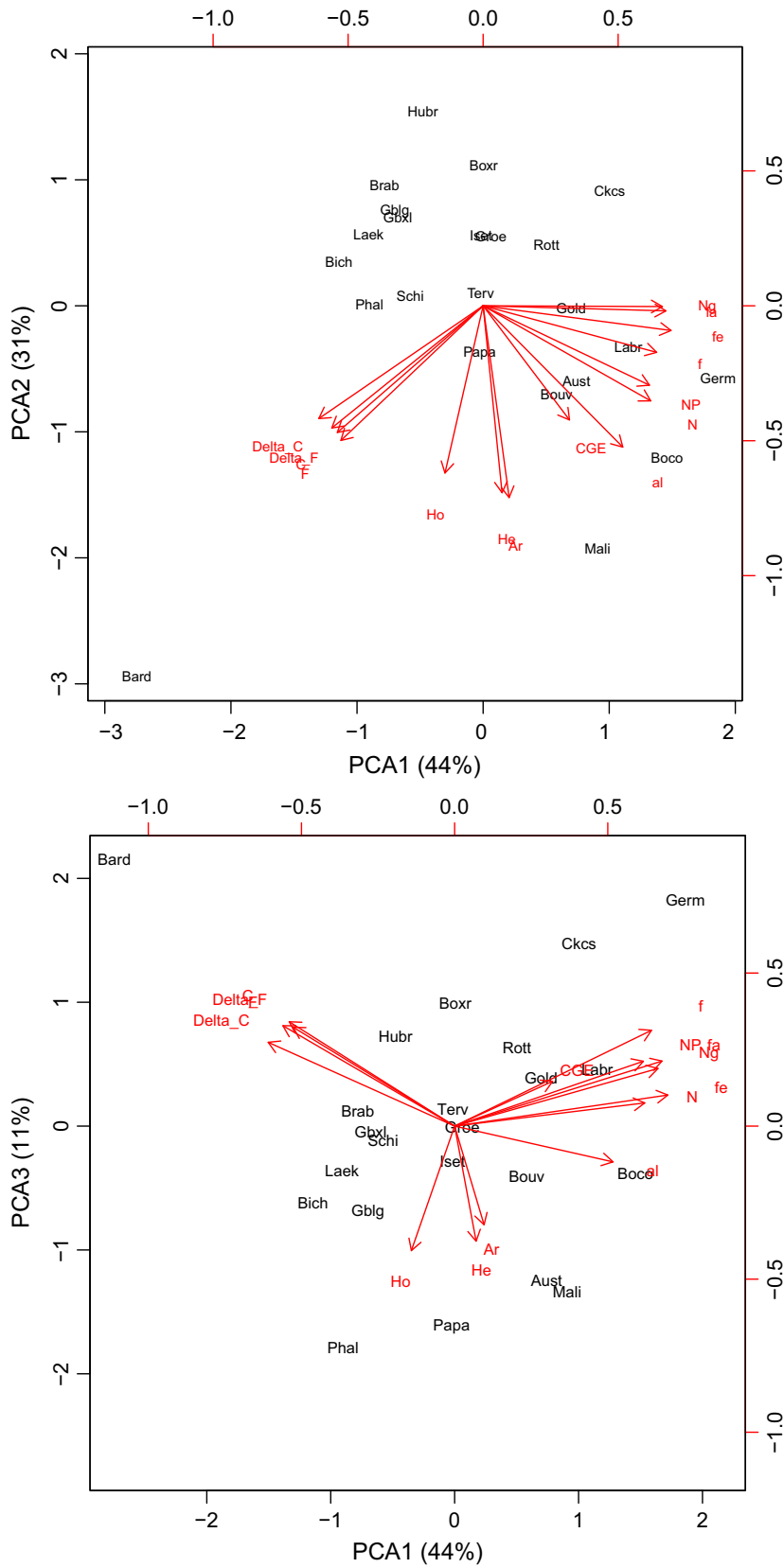
Figure 1 Principal Component Analysis (PCA) presenting the distribution of the breeds according to the correlation matrix between the different diversity parameters using the first three axes. Abbreviations for the breeds can be found in Table 1. Furthermore, N_g = number of founder genomes, f_a = effective number of ancestors, f_e = effective number of founders, f = number of founders, N_ped = number of dogs in the pedigree file, NP = mean number of pups per year (period 2003–2012), NP = the mean number of puppies per year (2003–2012), CGE = complete generation equivalent, al = number of alleles, A_r = allelic richness, H_e = expected heterozygosity, H_o = observed heterozygosity, Delta_F = increase in inbreeding, Delta_C = increase in coancestry, C = coefficient of kinship and F = average coefficient of inbreeding.

resulted in a high level of inbreeding. For the Australian shepherd, the low CGE likely led to an underestimation of the inbreeding, since import dogs from abroad without pedigree information are treated as founders in this study.

In general, the native Belgian breeds have higher inbreeding and coancestry and lower N_e values compared to the popular international breeds, indicating a lower genetic diversity, which was also reflected by their lower scores for PCA1. Our results are in line with other genealogical studies in various (national) dog populations. For example, for the Golden retriever breed, the inbreeding coefficient in our study was 1.6% ($N_e = 178$), while it reached 1.27% ($N_e = 39$) in a Dutch population (Windig & Oldenbroek 2015), 1.3% ($N_e = 219$) in a French population (Leroy *et al.* 2009b), 5.1% ($N_e = 1090$) in an Australian population (Shariflou *et al.* 2011), and 3.5% ($N_e = 67$) in a UK population (Calboli *et al.* 2008). Another example is the Boxer breed, with an inbreeding coefficient of 2.5% ($N_e = 119.8$) in our study, 2.4% ($N_e = 231$) in a French population (Leroy *et al.* 2009b), 4.3% ($N_e = 113$) in an Australian population (Shariflou *et al.* 2011) and 4.8% ($N_e = 45$) in a UK population (Calboli *et al.* 2008).

In terms of the molecular indicators clear differences between breeds were found. In general, within-breed variation was higher in breeds with a larger number of individuals. Furthermore, breeds are found to be more differentiated in allelic richness than in heterozygosity (Leroy *et al.* 2009b) as the allelic richness is more sensitive to bottlenecks than heterozygosity measurements (Foulley & Ollivier 2006). An example is the Bouvier des Ardennes, a breed that has only recently been re-established (in 2008). This breed probably has not suffered from bottlenecks yet, which explains its relatively high A_r value compared to the other breeds.

Negative F_{is} values were found in the Bouvier des Ardennes, Bichon frisé, and Phalène breed, indicating an excess of heterozygotes compared to HWE; however, caution is needed because of the low number of sampled animals. Positive F_{is} values show a considerable shortage of heterozygotes compared to HWE, indicating preferential mating of relatives (inbreeding or linebreeding) or presence of population structure (Wahlund effect) (Björnerfeldt *et al.* 2008). In general, heterozygosity values and inbreeding coefficients (F_{is}) in the Belgian populations were similar to other molecular studies. For example, for the Border Collie, observed heterozygosity values were 0.67 in an American population (Irion *et al.* 2003), 0.60 in a French population (Leroy *et al.* 2009b) and 0.65 in a UK population (Mellanby *et al.* 2013), compared to 0.64 in our study, whereas for the Boxer, observed heterozygosity values were 0.47 in an American population (Irion *et al.* 2003), 0.46 in a French population (Leroy *et al.* 2009b) and 0.51 in a UK population (Mellanby *et al.* 2013), compared to 0.48 in our study.

Comparison of the genealogical and molecular analysis show differences within the investigated breeds. Genealogical analysis revealed that native Belgian breeds in general showed lower diversity measures compared to popular international breeds; however, this was not confirmed in the molecular analysis. These differences are likely due to the different characteristics of the two approaches (Leroy *et al.* 2009b). In theory, breeds with small population size, high inbreeding and low genealogical diversity parameters should have low heterozygosity values and allelic richness. However, factors related to either the initial conditions or to the breed management may influence this correlation (Leroy *et al.* 2009b). By chance, the breeds with the highest current average inbreeding could be, the breeds with the highest initial heterozygosity, consequently their current value of H_e remains higher.

To facilitate the advice for the Belgian Kennel Club, a PCA was performed on the diversity parameters in order to identify groups of breeds in which a similar breeding strategy could be applied. No breed groups could be clearly identified, confirming the necessity to assess the genetic diversity for each breed separately. However, the PCA showed that the number of parameters that should be analysed per breed can be reduced, by keeping only one of the highly correlated measures. We would propose the following subset of criteria: ΔF, H_o, A_r, N_g, f_e and N.

There are several strategies to maintain or increase the genetic diversity going from the introduction of more animals in breeding, the promotion of seldom used lines, a 5% limit on the number of puppies in a 5-year period to counteract the popular sire breeding practice, to equalizing the contributions of the reproducing animals (optimal contribution selection) (Mäki 2010; Leroy 2011; Leroy & Baumung 2011). In practice however, optimal contributions are very difficult to realize, as the choice of the animal is made by individual breeders, and also influenced by the success and availability of matings (Windig & Oldenbroek 2015). Although several of these methods constrain the choice of breeders, they can be recommended for the native Belgian breeds in this study. For breeds with small and/or decreasing population size, such as the Phalène, Bichon frisé, Griffon Belge and Petit Brabançon among others, matings should be made

minimising the relatedness of the dogs. Also, the promotion of crosses between closely related breeds or breed varieties, such as the Griffon populations, the breed varieties of the Continental Toy Spaniel (Papillon and Phalène), and Belgian shepherd. This could be especially of interest in these first two breeds since they are at risk. To conclude, the results of this study allowed us to identify and asses the Belgian breeds at risk for which special efforts should be made to preserve genetic diversity, and emphasised the need to investigate each breed separately. These results are of great interest for the KMSH, the Belgian breeders and breed clubs.

Acknowledgements

References

Björnerfeldt S., Hailer F., Nord M., Vilà C. (2008) Assortative mating and fragmentation within dog breeds. *BMC Evol. Biol.*, **8**, 28.

Boichard D. (2002) PEDIG?: a fortran package for pedigree analysis suited for large populations. In: 7th World Congress on Genetics Applied to Livestock Production, Montpellier, 19–23 August 2002, pp. 28–13.

Boichard D., Maignel L., Verrier E. (1997) The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.*, **29**, 5–23.

Calboli F.C.F., Sampson J., Fretwell N., Balding D.J. (2008) Population structure and inbreeding from pedigree analysis of purebred dogs. *Genetics*, **179**, 593–601.

Cervantes I., Goyache F., Molina A., Valera M., Gutiérrez J.P. (2011) Estimation of effective population size from the rate of coancestry in pedigreed populations. *J. Anim. Breed. Genet.*, **128**, 56–63.

El Mousadik A., Petit R.J. (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor. Appl. Genet.*, **92**, 832–839.

Excoffier L., Laval G., Schneider S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online*, **1**, 47–50.

FAO (1998) Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans: Management of Small Populations At Risk, FAO. Italy, Rome.

Foulley J.-L., Ollivier L. (2006) Estimating allelic richness and its diversity. *Livest. Sci.*, **101**, 150–158.

Goudet J. (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *J. Hered.*, **86**, 485–486.

Gutiérrez J.P., Cervantes I., Goyache F. (2009) Improving the estimation of realized effective population sizes in farm animals. *J. Anim. Breed. Genet.*, **126**, 327–332.

Irion D.N., Schaffer A.L., Famula T.R., Eggleston M.L., Hughes S.S., Pedersen N.C. (2003) Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J. Hered.*, **94**, 81–87.

Lacy R.C. (1989) Analysis of Founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo Biol.*, **8**, 111–123.

Leroy G. (2011) Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. *Vet. J.*, **189**, 177–182.

Leroy G., Baumung R. (2011) Mating practices and the dissemination of genetic disorders in domestic animals, based on the example of dog breeding. *Anim. Genet.*, **42**, 66–74.

Leroy G., Rognon X., Varlet A., Joffrin C., Verrier E. (2006) Genetic variability in French dog breeds assessed by pedigree data. *J. Anim. Breed. Genet.*, **123**, 1–9.

Leroy G., Verrier E., Meriaux J.C., Rognon X. (2009a) Genetic diversity of dog breeds: between-breed diversity, breed assignation and conservation approaches. *Anim. Genet.*, **40**, 333–343.

Leroy G., Verrier E., Meriaux J.C., Rognon X. (2009b) Genetic diversity of dog breeds: within-breed diversity comparing genealogical and molecular data. *Anim. Genet.*, **40**, 323–332.

Leroy G., Danchin-Burge C., Palhiere I., Baumung R., Fritz S., Mériaux J.C., Gautier M. (2012) An ABC estimate of pedigree error rate: application in dog, sheep and cattle breeds. *Anim. Genet.*, **43**, 309–314.

Leroy G., Mary-Huard T., Verrier E., Danvy S., Charvolin E., Danchin-Burge C. (2013) Methods to estimate effective population size using pedigree data: examples in dog, sheep, cattle and horse. *Genet. Sel. Evol.*, **45**, 1.

Lüpke L., Distl O. (2005) Microsatellite marker analysis of the genetic variability in Hanoverian Hounds. *J. Anim. Breed. Genet.*, **122**, 131–139.

Mäki K. (2010) Population structure and genetic diversity of worldwide Nova Scotia Duck Tolling Retriever and Lancashire Heeler dog populations. *J. Anim. Breed. Genet.*, **127**, 318–326.

Mäki K., Groen A.F., Liinamo A., Ojala M. (2001) Population structure, inbreeding trend and their association with hip and elbow dysplasia in dogs. *Anim. Sci.*, **73**, 217–228.

Mellanby R.J., Ogden R., Clements D.N., French A.T., Gow A.G., Powell R., Corcoran B., Schoeman J.P.,

Summers K.M. (2013) Population structure and genetic heterogeneity in popular dog breeds in the UK. *Vet. J.*, **196**, 92–97.

Raymond M., Rousset F. (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.*, **86**, 248–249.

Shariflou M.R., James J.W., Nicholas F.W., Wade C.M. (2011) A genealogical survey of Australian registered dog breeds. *Vet. J.*, **189**, 203–210.

Weir B.S., Cockerham C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Windig J.J., Oldenbroek K. (2015) Genetic management of Dutch golden retriever dogs with a simulation tool. *J. Anim. Breed. Genet.*, **132**, 428–440.