

# Population structure and inbreeding from pedigree analysis of purebred dogs

Federico C. F. Calboli<sup>1</sup>, Jeff Sampson<sup>2</sup>, Neale Fretwell<sup>3</sup> and David J. Balding<sup>1</sup>

<sup>1</sup>Dept Epidemiology & Public Health, Imperial College, St Mary's Campus, Norfolk Place, London W2 1PG, UK; <sup>2</sup>The Kennel Club, Clarges St. London W1J 8AB, UK; <sup>3</sup>WALTHAM Centre for Pet Nutrition, Freaby Lane, Waltham-on-the-Wolds, LE14 4RS, UK

## Introduction

The UK Kennel Club registration database provides one of the world's most extensive resources for canine population-genetics studies.

We chose ten representative breeds and analysed their pedigrees since electronic records were established around 1970 (~8 generations before present).

Our goal is to inform:

- the design of canine genetic association studies;
- debate about the implications of breeding practices for canine welfare.

## Dogs Included in the Study

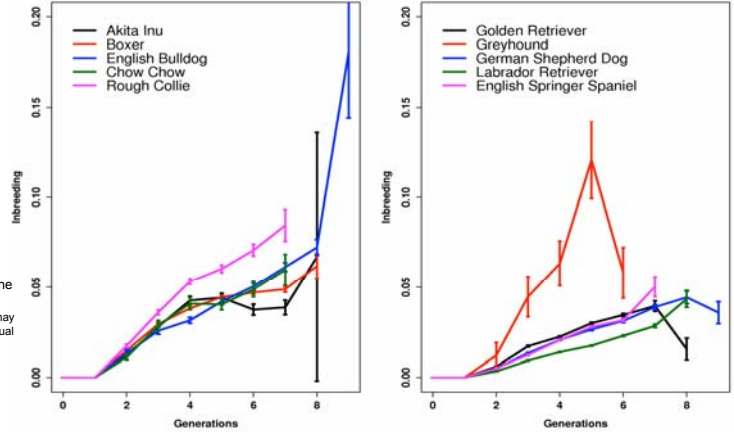
Breed	Group	Total # dogs	# Founders
Chow Chow	Oriental	18K	1.0K
Akita Inu	Oriental	21K	223
Boxer	Mastiff	5K	4.0K
English Bulldog	Mastiff	46K	882
Greyhound	Shepherd	1K	81
Rough Collie	Shepherd	84K	5.3K
German Shepherd Dog	Shepherd	474K	16K
English Springer Spaniel	Hunting	276K	10K
Golden Retriever	Hunting	318K	6.9K
Labrador Retriever	Hunting	704K	15K

## Popular Sires

Breed	Sires as % of male dogs	# Offspring max (median)	# Sires with >100 offspring
Chow Chow	11	212 (8)	22
Akita Inu	13	306 (9)	16
Boxer	9	1101 (8)	372
English Bulldog	11	430 (9)	77
Greyhound	21	45 (7)	0
Rough Collie	11	775 (7)	128
German Shepherd Dog	9	1479 (8)	851
English Springer Spaniel	13	2538 (7)	271
Golden Retriever	5	1386 (10)	792
Labrador Retriever	8	1911 (9)	1338

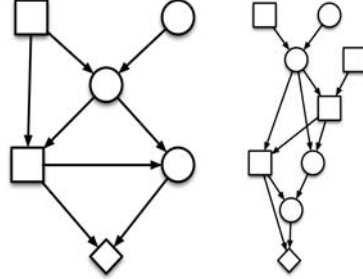
## Inbreeding, kinship and diversity

To calculate the coefficients of inbreeding  $f$  and kinship  $\phi$ , we used the "Meuw" in the Pedig package (Boichard 2002), which implements the algorithm of Meuwissen (1992).



Right: Coefficient of inbreeding averaged over the dogs in each generation. NB the final 1 or 2 generations may be atypical because most eventual members are not yet recorded.

Below: the pedigrees of the two most inbred dogs, both boxers, and both with  $f=0.5$ .



Below: Inbreeding ( $f$ ) and kinship ( $\phi$ ) coefficients for dogs in generations 6 & 7.

Breed	N	max f (%)	% dogs f > 0.1	average f (%)	average phi (%)
Chow Chow	1.7K	38	19	5.1	2.8
Akita Inu	2.9K	32	9.0	3.8	2.3
Boxer	45K	50	16	4.8	1.7
English Bulldog	12K	41	18	5.7	3.8
Greyhound	16	8	0	5.8	7.2
Rough Collie	4.7K	38	29	7.3	2.0
German Shepherd Dog	43K	47	12	3.3	1.4
English Springer Spaniel	24K	38	6.0	3.3	1.7
Golden Retriever	31K	39	8.2	3.5	1.3
Labrador Retriever	98K	39	5.2	2.4	1.2

NB average  $f$  minus average  $\phi$  measures consanguinity beyond that due to population size

## Within-breed population structure

We applied several multivariate statistical techniques to try to identify within-breed population structure from the pedigree only:

1. K-means clustering ( $K=2$  results shown as red and blue in plots on right)
2. Principal components analysis (PCA) based the ancestor/descendant incidence matrix (the descendants are treated as individuals and the ancestors as binary variates)
3. Multidimensional scaling (MDS, where distance between 2 individuals is  $1-\phi$  ( $\phi$  = kinship).

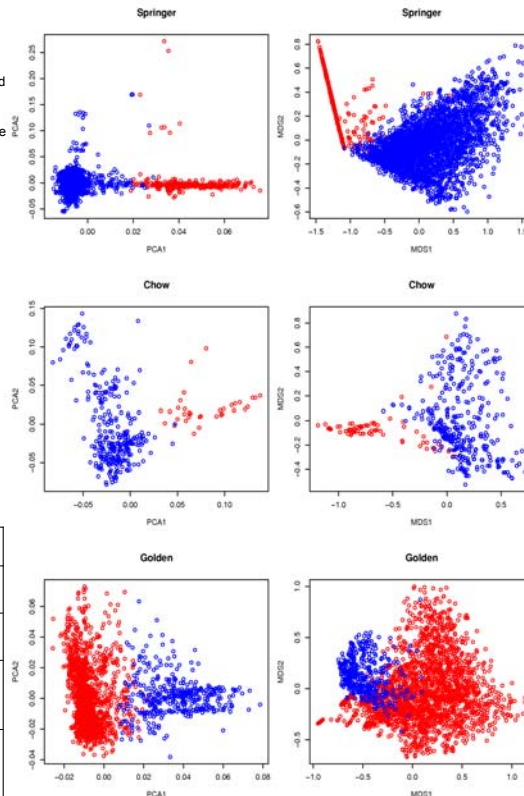
To quantify the level of clustering from 1. above we assumed that the count of descendants of ancestor  $j$  in cluster 1 has the beta-binomial distribution with parameters  $N_j$ ,  $\lambda p$ , and  $\lambda(1-p)$ , where

- $N_j$  = total number of descendants of  $j$ ,
- $p$  = Pr(any descendant is in cluster 1).

We estimate  $\lambda$  via maximum likelihood, treating the cluster memberships of descendants of different ancestors as independent (not strictly valid – but better if male and female ancestors are analysed separately). If we define

$$\Psi = 1/(1+\lambda)$$

then  $0 < \Psi < 1$  and we call  $\Psi$  the "Pedigree Structure Index". It is analogous to  $F_{ST}$ , the classical measure of population differentiation.



Breed	$\Psi$ male ancestors	$\Psi$ female ancestors
Chow Chow	0.43	0.45
Akita Inu	0.10	0.13
Boxer	0.36	0.40
English Bulldog	0.15	0.14
Greyhound	0.40	0.43
Rough Collie	0.34	0.38
German Shepherd Dog	0.41	0.35
English Springer Spaniel	0.55	0.28
Golden Retriever	0.37	0.43
Labrador Retriever	0.22	0.33

## Census & effective population sizes; genetic diversity

**Inbreeding  $N_e$ :** size of each generation in a random-mating population that would generate the observed average per-generation increase in  $f$ .

**Variance  $N_e$ :** as above except "variance in allele frequencies" replaces "increase in  $f$ ".

**Census  $N$ :** geometric mean of per-generation population sizes.

Breed	Census N	Inbreeding $N_e$	Variance $N_e$	% unique variants retained over 6 generations
Chow Chow	2.5K	50	44	6.2
Akita Inu	1.8K	45	59	30
Boxer	20K	45	92	11
English Bulldog	3.6K	48	65	17
Greyhound	154	17	20	6.5
Rough Collie	12K	33	71	2.9
Germ. Shepherd Dog	60K	76	257	5.6
Eng. Springer Spaniel	35K	72	83	8.0
Golden Retriever	38K	67	115	6.3
Labrador Retriever	77K	114	197	9.2

## Conclusions

In purebred dogs, we find

- high inbreeding/low  $N_e$ 
  - many highly-inbred dogs in every breed except Greyhound
  - very high average  $f$  in Collies – not attributable to demography
- high loss of genetic diversity – e.g. less than 3% of unique variants survive 6 generations in Collies
- evidence of population structure in every breed
  - Springers show distinct subpopulation with systematic pattern of sire selection

Implications for association mapping in dogs:

- within-breed structure should be accounted for (it may still limit effectiveness of naive methods of analysis)
- inbreeding can be advantageous through generating extra rare-allele homozygotes

Is intense inbreeding of dogs unacceptable given public concern about animal welfare?

- dramatic loss of genetic diversity over generations, little is replaced by mutation and no immigration
- effective population sizes much smaller than census sizes, and may be regarded as dangerously low

Possible ameliorative measures:

- limits on the use of popular sires
- encouragement of matings across national borders
- allow controlled outcrossing.

**Publication:** Further details of calculations and interpretation of results can be found in: Calboli FC, Sampson J, Fretwell N, Balding DJ. Population structure and inbreeding from pedigree analysis of purebred dogs. *Genetics* 179(1), May 2008.

**Acknowledgment:** This work has been funded by the UK Biotechnology and Biological Science Research Council under the Link Applied Genomics programme.