Population analysis of the *Retriever (Golden)* breed

Genetic analysis of the Kennel Club pedigree records of the UK ***Retriever (Golden)*** population has been carried out with the aim of estimating the rate of loss of genetic diversity within the breed and providing information to guide a future sustainable breeding strategy. The population statistics summarised provide a picture of trends in census size, the number of animals used for breeding, the rate of inbreeding and the estimated effective population size. The rate of inbreeding and estimated effective population size indicate the rate at which genetic diversity is being lost within the breed. The analysis also calculates the average relationship (kinship) among all individuals of the breed born per year and is used to determine the level of inbreeding that might be expected if matings were made among randomly selected dogs from the population (the expected rate of inbreeding).

Summary of results

The analysis utilises the complete computerised pedigree records for the current UK Kennel Club registered ***Retriever (Golden)*** population, and statistics were calculated for the period 1980-2014.

THE KENNEL CLUB

Making a difference for dogs

Figure 1: a plot of number of registrations by year of birth, indicative of any changing trend in popularity of the breed, followed by the yearly trend in number of animals registered (and 95% confidence interval).

Breed: *Retriever (Golden)*

Figure 1: Number of registrations by year of birth

Trend of registrations over year of birth (1980-2014) = -83.49 per year (with a 95% confidence interval of -193.46 to 26.48).

Table 1: census statistics by year, including sire use statistics.

Table 1: by year (1980-2014), the number of registered puppies born, by the number of unique dams and sires; maximum, median, mode, mean and standard deviation of number of puppies per sire; and the percentage of all puppies born to the most prolific 50%, 25%, 10% and 5% of sires.

| year	#born	#dams	#sires	puppies per sire					%puppies sired by most prolific sires			
				max	median	mode	mean	sd	50% sires	25% sires	10% sires	5% sires
1980	1986	1232	509	62	2	1	3.9	6.09	85.45	67.82	44.56	30.66
1981	6892	1897	677	270	5	1	10.18	17.91	88.31	70.92	48.56	34.46
1982	10025	2236	776	221	6	2	12.92	23.01	88.07	71.68	49.45	36.16
1983	10517	2356	811	271	6	1	12.97	22.75	88	71.33	49.21	35.29
1984	11088	2572	843	269	6	3	13.15	23.97	88.15	72.66	50.37	35.69
1985	11532	2658	941	284	6	3	12.26	20.63	87.11	69.66	46.4	32.61
1986	11503	2639	949	198	6	3	12.12	19.04	87.19	69.29	45.68	31.71
1987	11217	2515	912	229	6	2	12.3	19.39	87.03	69.24	45.84	32.42
1988	11170	2410	905	204	6	2	12.34	18.8	86.63	68.16	45.06	30.64
1989	16012	2489	887	308	9	8	18.05	24.5	84.75	66.67	42.41	27.39
1990	14871	2235	833	321	8	8	17.85	26.1	84.42	68.18	44.4	29.66
1991	15325	2259	820	269	9	7	18.69	25.64	84.69	66.62	42.17	28.48
1992	14110	2092	762	355	9	6	18.52	27.88	84.98	67.24	43.2	29.43
1993	14627	2139	788	427	9.5	8	18.56	27.06	84.02	64.59	40.9	27.34
1994	14676	2127	789	354	9	7	18.6	25.68	84.28	65.6	41.6	27.44
1995	15708	2287	806	269	10	7	19.49	26.51	84.07	64.67	40.82	27.3
1996	16619	2411	865	425	10	8	19.21	27.06	84.04	64.56	40.21	27.01
1997	15400	2249	814	406	10	7	18.92	26.08	83.82	64.48	39.9	26.69
1998	14220	2077	796	234	10	7	17.86	22.42	83.34	63.35	39.63	26.35
1999	12615	1818	738	238	9.5	8	17.09	20.78	82.75	62.83	38.08	24.49
2000	11837	1741	676	178	10	7	17.51	21.89	84.24	63.71	39.75	26.64
2001	10083	1528	644	138	9	8	15.66	16.89	81.65	61.33	35.72	22.47
2002	10529	1589	635	132	9	7	16.58	19.46	84.51	64.29	38.8	24.76
2003	10601	1574	640	160	9	8	16.56	19.07	83	62.48	38.3	24.08
2004	10531	1545	651	154	9	9	16.18	20.07	83.4	64.12	40.31	27.16
2005	10105	1483	626	159	9	9	16.14	19.76	83.45	64.04	40.22	26.12
2006	9282	1401	604	144	9	1	15.37	17.61	82.87	61.79	36.94	24.06
2007	9851	1454	613	199	9	7	16.07	19.31	83.51	63.44	38.81	24.77
2008	8714	1294	570	133	9	1	15.29	17.46	83.45	61.85	37.51	24.35
2009	8098	1230	556	152	9	8	14.56	16.9	82.5	61.43	36.64	23.62
2010	7829	1203	521	146	9	1	15.03	17.51	83.48	61.9	37.65	24.49
2011	8018	1211	552	139	9	1	14.53	17.55	83.77	63.03	38.7	25.8
2012	7141	1087	476	121	9	1	15	16.05	83.45	61.29	35.86	22.53
2013	7058	1034	466	210	9	7	15.15	18.83	83.04	63.73	39.22	24.88
2014	6786	971	397	213	9	9	17.09	21.52	83.08	63.45	39.82	26.05

Mean generation interval (years) = 4.19

Generation interval: the mean average age (in years) of parents at the birth of offspring which themselves go on to reproduce.

Figure 2: a plot of the annual mean observed inbreeding coefficient (showing loss of genetic diversity), and mean expected inbreeding coefficient (from 'random mating') over the period 1980-2014. 'Expected inbreeding' is staggered by the generation interval and, where >2000 animals are born in a single year, the 95% confidence interval is indicated.

Figure 2: Annual mean observed and expected inbreeding coefficients
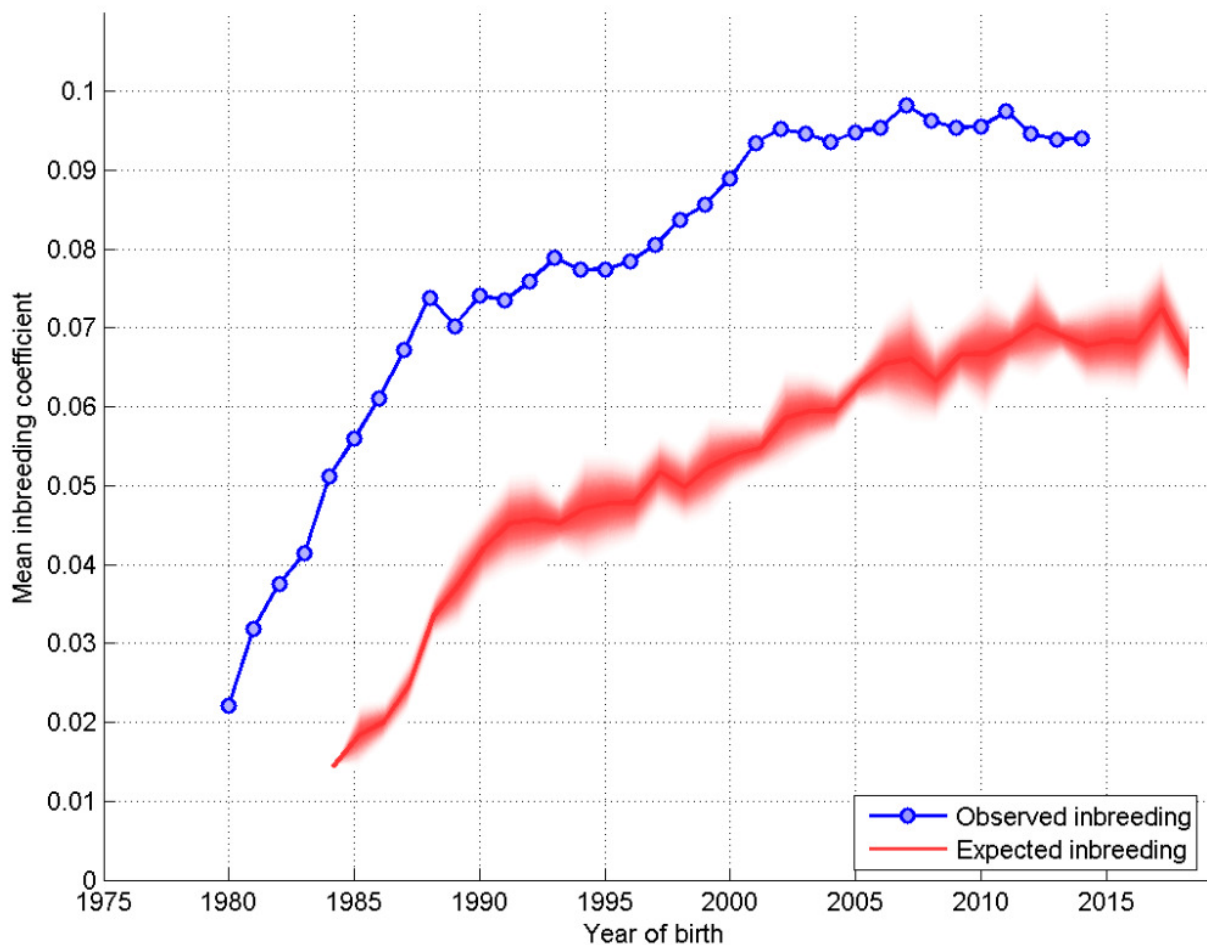
> **Estimated effective population size:** the rate of inbreeding (slope or steepness of the observed inbreeding in Figure 2) is used to estimate the effective population size of the breed. The effective population size is the number of breeding animals in an idealised, hypothetical population that would be expected to show the same rate of loss of genetic diversity (rate of inbreeding) as the breed in question. It may be thought of as the size of the 'gene pool' of the breed.
>
> Below an effective population size of 100 (inbreeding rate of 0.50% per generation) the rate of loss of genetic diversity in a breed/population increases dramatically (Food & Agriculture Organisation of the United Nations, "Monitoring animal genetic resources and criteria for prioritization of breeds", 1992). An effective population size of below 50 (inbreeding rate of 1.0% per generation) indicates the future of the breed many be considered to be at risk (Food & Agriculture Organisation of the United Nations, "Breeding strategies for sustainable management of animal genetic resources", 2010).
>
> Where the rate of inbreeding is negative (implying *increasing* genetic diversity in the breed), effective population size is denoted 'n/a'.

Estimated effective population size = 61.3
NB - this estimate is made using the rate of inbreeding over the whole period 1980-2014

Table 2: a breakdown of census statistics, sire and dam usage and indicators of the rate of loss of genetic diversity over 5 year periods (1980-4, 1985-9, 1990-4, 1995-9, 2000-4, 2005-9, 2010-14). Rate of inbreeding and estimated effective population size for each 5-year block can be compared with the trend in observed inbreeding in Figure 2.

Table 2: by 5-year blocks, the mean number of registrations; for sires the total number used, maximum, mean, median, mode, standard deviation and skewness (indicative of the size of the 'tail' on the distribution) of number of progeny per sire; for dams the total number used, maximum, mean, median, mode, standard deviation and skewness of number of progeny per dam; rate of inbreeding per generation (as a decimal, multiply by 100 to obtain as a percentage); mean generation interval; and estimated effective population size.

years	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014
mean #registrations	8101.6	12287	14722	14912	10716	9210	7366.4
Total #sires	2062	2498	2120	2060	1698	1514	1240
Max #progeny	844	722	1053	921	534	654	495
Mean #progeny	19.619	24.593	34.721	36.195	31.555	30.411	29.638
Median #progeny	6	8	10	13	11	11	11
Mode #progeny	1	3	7	7	1	1	1
SD #progeny	52.916	55.274	72.398	67.036	53.626	52.021	49.946
Skew #progeny	7.7444	6.091	5.9726	5.4377	3.9242	4.4372	4.0197
Total #dams	7331	8777	7327	7115	5480	4658	3881
Max #progeny	37	46	56	53	53	52	41
Mean #progeny	5.5183	6.9993	10.046	10.479	9.7774	9.8847	9.4695
Median #progeny	4	6	8	8	8	8	8
Mode #progeny	1	3	7	8	7	8	8
SD #progeny	4.5767	5.3973	6.939	7.1695	6.5074	6.8944	6.4329
Skew #progeny	1.8372	1.7113	1.6362	1.5754	1.32	1.4286	1.2012
Rate of inbreeding	0.026684	0.017504	0.0055983	0.010147	0.0051418	0.0009352	-0.0029858
Generation interval	3.7995	3.9625	4.3351	4.2764	4.5042	4.4275	4.0232
Effective pop size	18.738	28.565	89.313	49.278	97.241	534.63	n/a

THE KENNEL CLUB
Making a difference for dogs

Figure 3: a histogram ('tally' distribution) of number of progeny per sire and dam over each of the seven 5-year blocks above. A longer 'tail' on the distribution of progeny per sire is indicative of 'popular sires' (few sires with a very large number of offspring, known to be a major contributor to a high rate of inbreeding).

Figure 3: Distribution of progeny per sire (blue) and per dam (red) over 5-year blocks (1980-4 top, 2010-14 bottom). Vertical axis is a logarithmic scale.
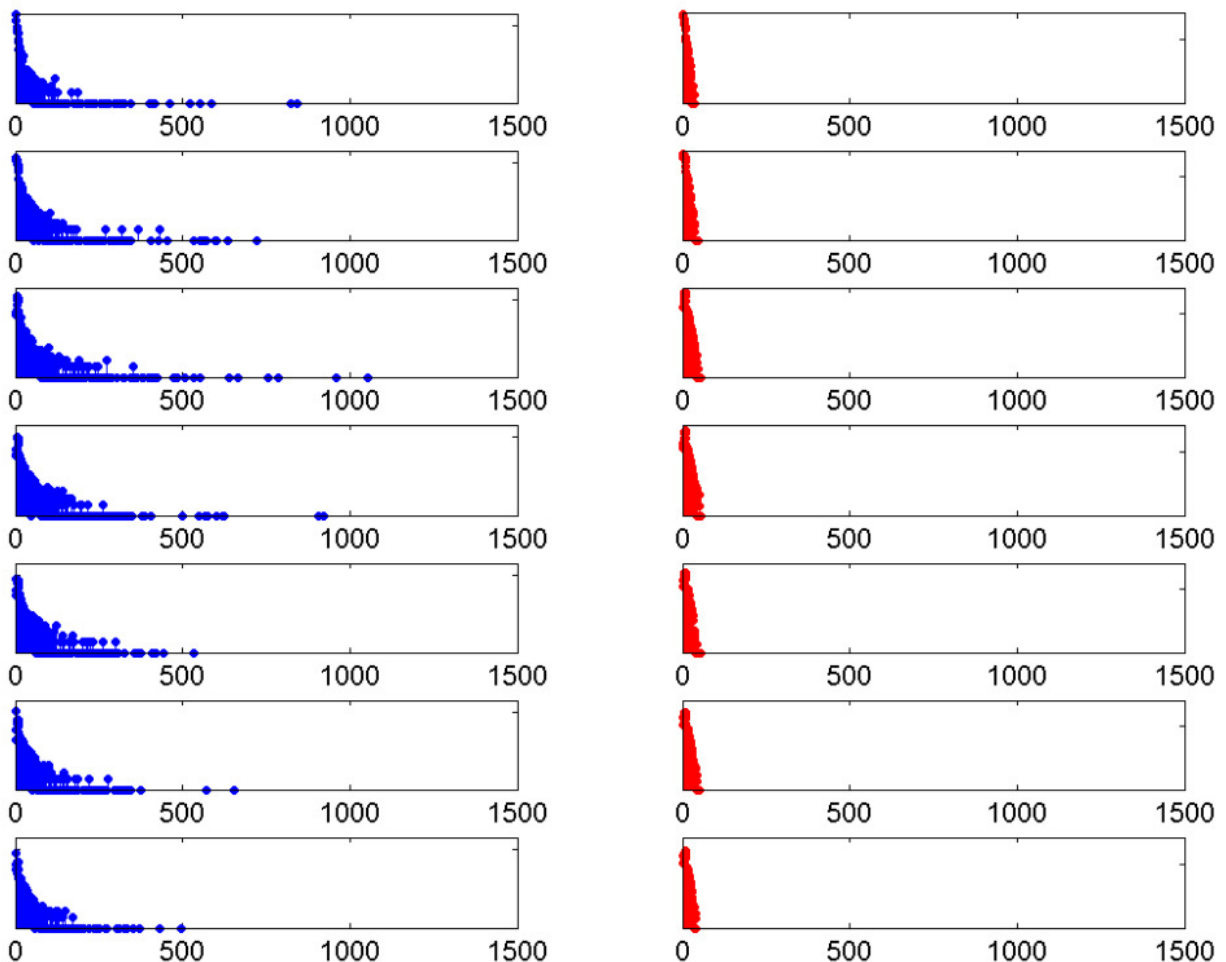
Comments

As with most breeds, the rate of inbreeding was at its highest in this breed in the 1980s and 1990s. This represents a 'genetic bottleneck', with genetic variation lost from the population. However, since the early 2000s, the rate of inbreeding has decreased implying a slowdown in the rate of loss of genetic diversity (possibly through the use of imported animals).

It appears that the extensive use of popular dogs as sires has eased a little (the 'tail' of the blue distribution shortening in figure 3).

It should be noted that, while animals imported from overseas may appear completely unrelated, this is not always the case. Often the pedigree available to the Kennel Club is limited in the number of generations, hampering the ability to detect true, albeit distant, relationships.