J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

Application of individual increase in inbreeding to estimate realized effective sizes from real pedigrees

I. Cervantes[1,3], F. Goyache[2], A. Molina[3], M. Valera[4] & J. P. Gutiérrez[1]

1 Departamento de Producción Animal, Facultad de Veterinaria, Madrid, Spain
2 SERIDA-Somió, Gijón (Asturias), Spain
3 Departamento de Genética, Universidad de Córdoba, Córdoba, Spain
4 Departamento de Ciencias Agro-Forestales, EUITA, Universidad de Sevilla, Sevilla, Spain

Keywords
effective size; increased in inbreeding; overlapped generation.

Correspondence
J. P. Gutiérrez, Departamento de Producción Animal, Facultad de Veterinaria, Avda. Puerta de Hierro s/n, E-28040-Madrid, Spain.
Tel. and Fax: +34913943767;
E-mail: gutgar@vet.ucm.es

Received: 7 February 2008;
accepted: 30 May 2008

Summary

The objective of this study was to test the performance of a recently proposed methodology for the estimation of realized effective size (N_e) based on individual increase in inbreeding (ΔF_i) on several real pedigrees: (a) an experimental mice population; (b) a closed pedigree of fighting bulls; (c) the Spanish Purebred (SPB, Andalusian) horse pedigree; (d) the Carthusian strain of SPB pedigree; (e) the Spanish Arab horse pedigree; and (f) the Spanish Anglo-Arab horse pedigree. Several reference subpopulations were defined on the basis of generation length in order to consider only animals in the last generation, to assess the influence of the pedigree content on the estimates of N_e. The estimates of realized N_e computed from ΔF_i ($\overline{N_e}$) tended to be higher than those obtained from regression on equivalent generations. The new parameter $\overline{N_e}$ remained approximately stable when pedigree depth achieved about five equivalent generations. Estimates of $\overline{N_e}$ take into account the genetic history of the populations, the size of their founder population, and the mating policy or bottlenecks caused by poor use of reproducing individuals. The usefulness of the realized N_e computed from individual increase in inbreeding in real pedigrees is also discussed.

Introduction

The size of an idealized population, which would give rise to the rate of inbreeding, or the rate of change in variance of gene frequencies observed in the population under consideration defines the effective population size (N_e) (Wright 1969). This concept is a key parameter widely used as a criterion for characterizing the risk status of livestock breeds (FAO 1998; Duchev *et al.* 2006). This is basically due to the direct relationship between N_e and the rate of inbreeding, fitness and the amount of genetic variation lost because of random genetic drift (Falconer & Mackay 1996; Caballero & Toro 2000), which makes it useful to infer the future of genetic stock.

However, the theoretical assumptions on which N_e is based only fit with idealized populations (Caballero 1994), which is extremely important in the development of methods enabling the computation of effective size from pedigrees of actual populations under generation overlapping, selection and non-random mating. A quick overview of recent literature shows that the generally accepted methods to compute N_e in real pedigrees are based on regressing the inbreeding coefficients on birth dates or generations (Pérez-Enciso 1995; Gutiérrez *et al.* 2003), even though they are always dependent on the definition of a reference population within pedigree and further adjustment by the generations interval (L). Moreover, the methods have not been proven to ascertain a reliable value when there are variations in the breeding policy affecting mean inbreeding in the reference population (Gutiérrez *et al.* 2008). In such a scenario where planning matings to minimize co-ancestry after a period in which mating between close relatives was preferred, the regression-based methods to compute effective size can give unrealistic negative N_e values (Gutiérrez *et al.* 2008).

Many attempts have been made to deal with the different real-world scenarios in order to obtain reliable estimates of the effective population size (Caballero 1994; Caballero & Toro 2000). However, there is no standard method for general application to obtain the effective population size. Gutiérrez *et al.* (2008) have recently described a method for the estimation of effective size from pedigrees based on the definition of the individual increase in inbreeding coefficient (ΔF_i; González-Recio *et al.* 2007) using simulated standard pedigrees and in a unique real pedigree.

The objective of this study was to test the methodology for estimating N_e based on ΔF_i in several real pedigrees of varying depth in pedigree depth, overlapping generations, subdivision, migration, unbalanced representation of founders, or other unknown effects affecting the breeding policy, in order to further establish the usefulness of this parameter in real livestock breeding scenarios. To achieve these goals, data from an experimental mice population, the Domecq fighting bull herd pedigree, the Spanish Purebred horse (SPB) pedigree, the Carthusian strain of SPB pedigree, the Spanish Arab horse pedigree and the Spanish Anglo-Arab horse pedigree were used to test the methodology developed by Gutiérrez *et al.* (2008). A comparison of the obtained estimates with other well-known methods for computing N_e will be performed.

Materials and methods

Individual increase in inbreeding

The original idea of the methodology is to assume for each individual that all the animals in the referred population are under the same situation regarding inbreeding and genealogical information; this is measured over *t*, the discrete equivalent generation (Maignel *et al.* 1996; Boichard *et al.* 1997). The discrete generation equivalent is computed for the pedigree of each individual as the sum over all known ancestors of the term of $(½)^n$, where *n* is the number of generations separating the individual from each known ancestor.The steady increase in inbreeding that would have led to the inbreeding coefficient of animal *i* (F_i) in *t* generations should be:

$$\Delta F_i = 1 - \sqrt[t]{1 - F_i} \qquad (1)$$

Following Gutiérrez *et al.* (2008), an overall mean of this parameter ($\overline{\Delta F}$) can be easily computed by averaging the ΔF_i values of the individuals included in a defined reference subpopulation. From this, as a 'realized' effective population size, $\overline{N_e}$ can be straightforwardly computed as $\overline{N_e} = 1/2\overline{\Delta F}$. The realized effective size should be understood as the global effective size over time which has led to the present inbreeding level in the reference population from the founder population.

Moreover, given that an individual increase in inbreeding exists for each animal in the reference subpopulation, a standard error of the realized $\overline{N_e}$ can be further computed from the standard deviation of the individual increase in inbreeding and the square root of the reference subpopulation size as:

$$\sigma_{\overline{N_e}} = 2\overline{N}_e^2 \sigma_{\Delta F} \frac{1}{\sqrt{\overline{N}_e}}.$$

This is a modification of the expression reported by Gutiérrez *et al.* (2008):

$$\sigma_{\bar{N_e}} = 2\overline{N}_e^2 \sigma_{\Delta F} \frac{1}{\sqrt{n}}$$

where *n* has been changed to $\overline{N_e}$, given that this value would be an equivalent number of independent individuals instead of *n*. Note also that this standard error gives an approximate idea of certainty but has a limited informative value given that the distribution of ΔF_i is skewed to the right.

Other methods to estimate N_e using pedigree information

Several approaches have been used to check estimates of N_e obtained from individual increase in inbreeding:

1 First, N_e was estimated from the variances of family size as (Hill 1979):

$$\frac{1}{N_e} = \frac{1}{16ML}\left[2 + \sigma^2_{mm} + 2\left(\frac{M}{F}\right)cov(mm, mf) + \left(\frac{M}{F}\right)^2 \sigma^2_{mf}\right] + \frac{1}{16FL}\left[2 + \left(\frac{F}{M}\right)^2 \sigma^2_{fm} + 2\left(\frac{F}{M}\right)cov(fm, ff) + \sigma^2_{ff}\right]$$

where *M* and *F* are the number of male and female individuals born or sampled for breeding each time period, L the average generation interval, σ^2_{mm} and σ^2_{mf} are the variances of the male and female offspring of a male, σ^2_{fm} and σ^2_{ff} are the variances of the male and female offspring of a female, cov(mm, mf) and cov(fm, ff) the respective covariances. Note that the family size of a parent (male or female) consists of sons and daughters kept for reproduction (James 1977). This should be a good method when all the other conditions are in compliance with the idealized population but generations are overlapping (Falconer & Mackay 1996).

2a Moreover, N_e was also estimated from ΔF following Gutiérrez *et al.* (2003). The increase in inbreeding between two generations ($F_t - F_{t-1}$) was obtained from the regression coefficient (b) of the inbreeding coefficient over the date of birth obtained in the reference subpopulation, and considering the average generation interval (L) as $F_t - F_{t-1} = L \times b$ with F_{t-1} computed from the mean inbreeding in the reference subpopulation (F_t) as: $F_{t-1} = F_t - L \times b$.

2b N_e was also computed in a very similar way as in (2a) but using t directly instead of approaching the generations through generation intervals. Using this approach, N_e was computed from the regression coefficient (b) of the individual inbreeding values over the individual equivalent complete generations t. In this case,

$$\Delta F = \frac{F_t - F_{t-1}}{1 - F_{t-1}} \approx \frac{b}{1 - (F_t - b)},$$

where F_t is the average F of the reference subpopulation.

3 Estimation of N_e could also be approached by using $1 - F_t = (1 - (1/2N_e))^t$ to derive its value from a log regression of $(1 - F)$ over generation number (Pérez-Enciso 1995). Here we computed the log regression of $(1 - F)$ over birth date and corrected further by the length of the generation interval (Pérez-Enciso 1995).

Examples

To assess this methodology it was applied to several pedigrees in common scenarios in real populations:

(a) An experimental mice population (Moreno 1998; Ibáñez-Escriche *et al.* 2008) consisting of 18 discrete generations with about 72 males and 144 females per generation structured in nine replicates of selection. Mating policy was different in three replicates of each of the selection method: at random, weighted selection and minimum coancestry.

(b) The pedigree information available from the beginning of the 20th century in the Fighting Bull herd of Juan Pedro Domecq, including a total of 33 330 individuals.

(c) The pedigree of Spanish Purebred horse (SPB, Andalusian horse) from their foundation in the XIXth century until December 2004 (Ministry of Defence: http://www.librogenealogico.com). The total available records were 140 629.

(d) The pedigree of the Carthusian strain of the Spanish Purebred horse, which is a subset of the pedigree listed before having a particular management, including those founder individuals classified as Carthusian because of their particular historical origin and those individuals resulting from matings between Carthusian parents (Valera *et al.* 2005). This is the example population already used for this purpose by Gutiérrez *et al.* (2008), but has also been included here for a quick comparison among the different methodologies and pedigrees.

(e) The pedigree of the Spanish Arab horse recently analysed by Cervantes *et al.* (2008) including a total of 18 880 individuals.

(f) The pedigree of the Spanish Anglo-Arab horse, including a total of 10 568 individuals. The data file was obtained from the Ministry of Defence (website http://www.librogenealogico.com).

The main parameters describing the real pedigrees analysed are given in Table 1. The examples listed above are expected to characterize a classical example of population with discrete inbred generations (mice), closed population with overlapping generations and matings between relatives (SPB, Carthusian strain), closed population with overlapping generations and matings planned to avoid inbreeding (Fighting Bull) and, finally, open populations with different pedigree depths varying from 'sufficiently' deep (Spanish Arab horse) to very shallow (Spanish Anglo-Arab horse) as a consequence of different rates of migration. The mice data set also included a strong population subdivision in nine independent lines from the foundation of the population (Moreno 1998; Ibáñez-Escriche *et al.* 2008). The Carthusian population was reported to be managed using minimum co-ancestry mating policies during recent generations (Valera *et al.* 2005) and the Fighting Bull population is a case of a population of constant size with a mating policy avoiding common ancestors in the last three generations of each individual.

Several reference subpopulations were defined trying to take into account the last generation interval, and also trying to assess the influence of the pedigree content on the estimates of N_e. For the mice pedigree the 15–16th and the 18th generations were used as two reference subpopulations (RS). For the

Table 1 Number of individuals (N), average number of equivalent generations and standard deviation ($t \pm$ SD), maximum number of equivalent generations (Max. t), average inbreeding (F; in percentage), average individual inbreeding rate (ΔF_i; in percentage) and number of male and female reproductive individuals and average family size for males (A_m) and females (A_f) for the whole (WP) and the reference (RS) populations used as examples in the present analyses

Population	N	$t \pm$ SD	Max. t	F (%)	ΔF (%)	Males	A_m	Females	A_f
Mice									
WP	9152	9.2 ± 5.25	18.3	12.6	1.27	1395[a]	2.9	2243[b]	1.8
RS_{18}	283	17.9 ± 1.46	18.3	26.2	1.67				
$\text{RS}_{15–16}$	985	15.7 ± 0.71	16.3	24.0	1.74	160[a]	2.7	239[b]	1.8
Fighting Bull									
WP	33 330	5.0 ± 3.20	10.9	1.3	1.29	348[a]	15.3	3072[b]	1.9
RS_{10}	5663	9.0 ± 0.91	10.9	1.8	1.82	24[a]	8.6	100[b]	1.3
$\text{RS}_{92–01}$	6641	8.3 ± 0.81	10.3	1.8	1.77	51[a]	11.3	392[b]	1.5
SPB									
WP	140 629	9.0 ± 1.82	11.8	8.2	0.95	5475[a]	7.5	14 689[b]	2.8
RS_{10}	89 324	9.9 ± 0.58	11.8	8.4	0.90	210[a]	2.2	738[b]	1.1
$\text{RS}_{83–93}$	55 737	9.3 ± 0.67	11.3	8.8	1.00	1755[a]	4.4	5631[b]	1.7
Carthusian									
WP	6318	6.6 ± 2.75	10.9	13.0	1.89	424[a]	4.9	933[b]	2.3
RS_{10}	1721	9.1 ± 0.68	10.9	18.6	2.25	1[a]	1.0	5[b]	1.2
$\text{RS}_{79–89}$	1464	8.2 ± 0.64	9.8	17.5	2.34	97[a]	4.1	167[b]	2.1
Arab									
WP	18 880	5.7 ± 2.78	10.3	7.0	1.07	1560[a]	4.2	3260[b]	2.0
RS_{10}	6240	7.9 ± 1.28	10.3	9.8	1.28	22[a]	1.8	36[b]	1.1
$\text{RS}_{86–97}$	5557	7.1 ± 1.60	9.7	9.7	1.37	307[a]	2.7	663[b]	1.5
Anglo-Arab									
WP	10 568	1.7 ± 1.51	6.2	0.8	0.24	757[a]	2.8	1325[b]	1.5
RS_{10}	2862	2.4 ± 1.59	6.2	0.9	0.24	2[a]	1.0	5[b]	1.2
$\text{RS}_{84–96}$	2920	1.7 ± 1.64	5.6	0.9	0.28	129[a]	1.8	225[b]	1.3

RS_{18}: animals in the 18th generation; $\text{RS}_{15–16}$: animals in the 15th and 16th generation; RS_{10}: animals born in the last decade. Other $\text{RS}_{\text{xx-yy}}$ animals born between the years 19xx and 19yy.

[a]Individuals born in the defined period that acted sequentially as male parents.

[b]Individuals born in the defined period that acted sequentially as female parents.

other pedigrees that were analysed various RS were defined according to three different strategies: (a) pedigree depth of the individuals; (b) the individuals born in the last 10 years of the available records; and (c) individuals born in a given period of years allowing their use for reproduction. RS included individuals with four equivalent generations or more, and eight equivalent generations or more except for the Spanish Anglo-Arab horse for which only the first RS was defined, given their extremely shallow pedigree.

Program used

The analyses were performed using the ENDOG program (current version v4.5) (Gutiérrez & Goyache 2005), which can be freely downloaded from the World Wide Web at http://www.ucm.es/info/prodanim/html/JP_Web.htm.

Results

Table 1 gives the main parameters describing the pedigrees that were analysed. Whole pedigreed set sizes varied from 6318 (Carthusian) to 140 629 (SPB), while the sizes for the fitted RS varied from 283 individuals for RS_{18} in mice to 89 324 for RS_{10} in SPB. The mean equivalent generations ranged between 1.7 ± 1.51 for WP in Spanish Anglo-Arab horse to 17.9 ± 1.46 for RS_{18} in mice. The maximum values for equivalent generations were observed in the mice population (from 16.3 to 18.3), it was over 10–11 for the different RS fitted for Domecq's Fighting Bull, SPB, Carthusian and Spanish Arab horses and reached lower values (from 5.6 to 6.2) for the Spanish Anglo-Arab horse. The highest average inbreeding values at the WP level were found in Carthusians (13.0%) and the lowest (0.8%) in the population with the shallowest pedigree (Spanish Anglo-Arab horse). The lowest increases in inbreeding were found in the Spanish Anglo-Arab horse (0.2%) and in the SPB (lower than 1%), while the highest was found in the Carthusian subpopulation (around 2%).

Figure 1 shows the plot of the values of individual increase in inbreeding per number of equivalent generations in all the pedigrees analysed. Notice that the trend of the values of ΔF_i tended to stabilize

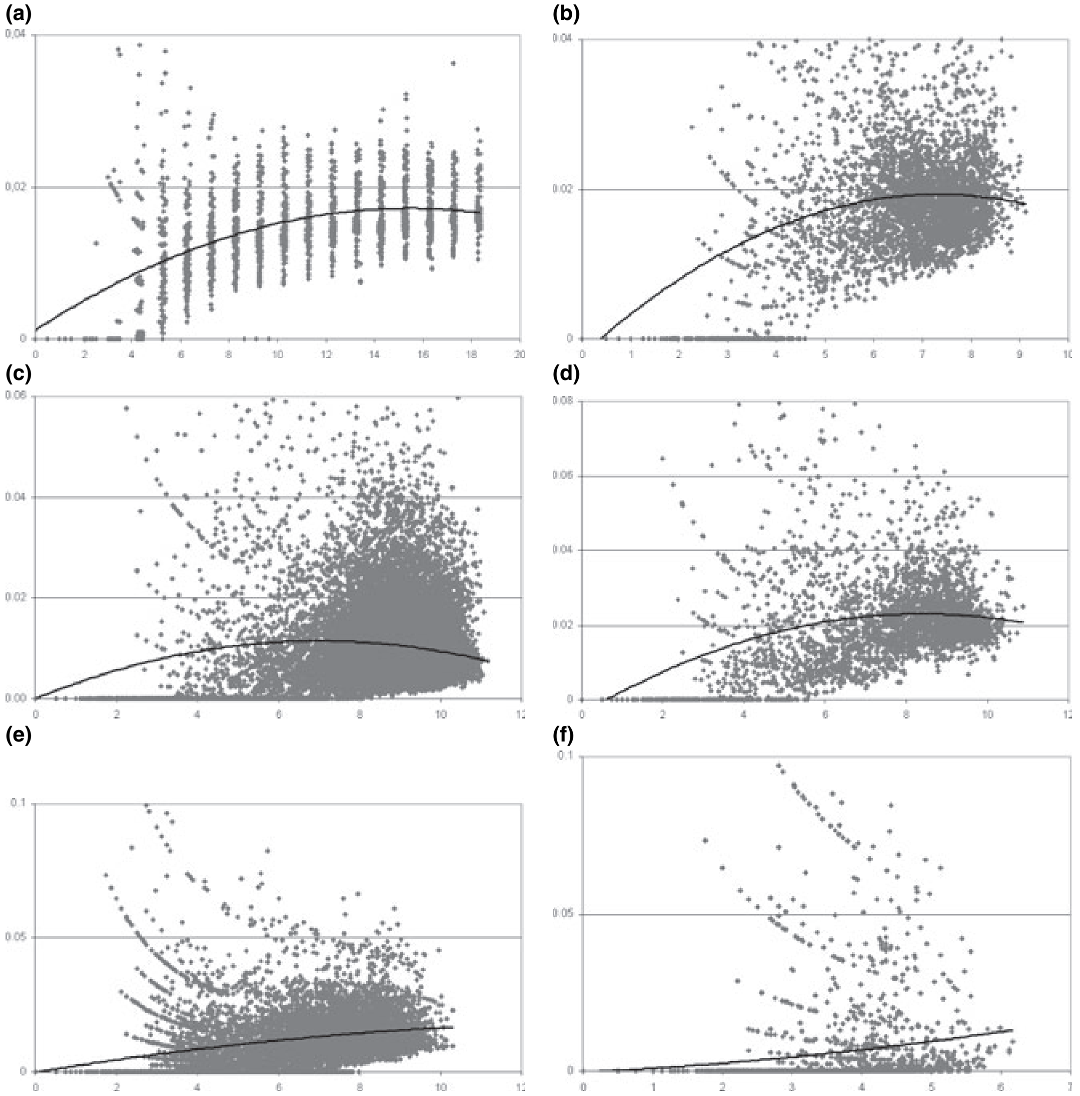

Figure 1 Plots summarizing the dispersal and the trend line of the individual increase in inbreeding (ΔF_i; on the *y*-axis) per individual number of equivalent generations (on the *x*-axis) in the whole analysed pedigrees: (a) mice population; (b) Fighting Bull population; (c) Spanish Purebred horse; (d) Carthusian strain of the Spanish Purebred horse; (e) Spanish Arab horse; (f) Spanish Anglo-Arab horse.

with the increase in t (about 5) because of the correction for pedigree depth of the individuals.

Figure 2 shows the evolution of mean inbreeding, mean individual increase in inbreeding, mean equivalent generations, and realized effective size ($\overline{N_e}$) across years of birth (generations in the mice population) for all the pedigrees. It is worth pointing out that the mean individual increase in inbreeding became approximately stable in the last period of time, while mean inbreeding increased and mean equivalent generations linearly increased during this last period. Furthermore, $\overline{N_e}$ remains approximately stable when known pedigree reaches about five equivalent generations. This effect is not so clear in the Spanish Anglo-Arab population with lower pedigree knowledge. A decrease in mean individual increase in inbreeding (and a related increase in $\overline{N_e}$) was observed in the SPB population.

Table 2 includes the estimates of N_e obtained using regression of the individual coefficients of inbreeding on equivalent generations, variance of family size and individual increase in inbreeding (ΔF_i) in the different RS analysed here. Furthermore, restrictions in pedigree depth of the individuals included in the corresponding RS are included in this table. Regardless of the RS being considered, the estimates

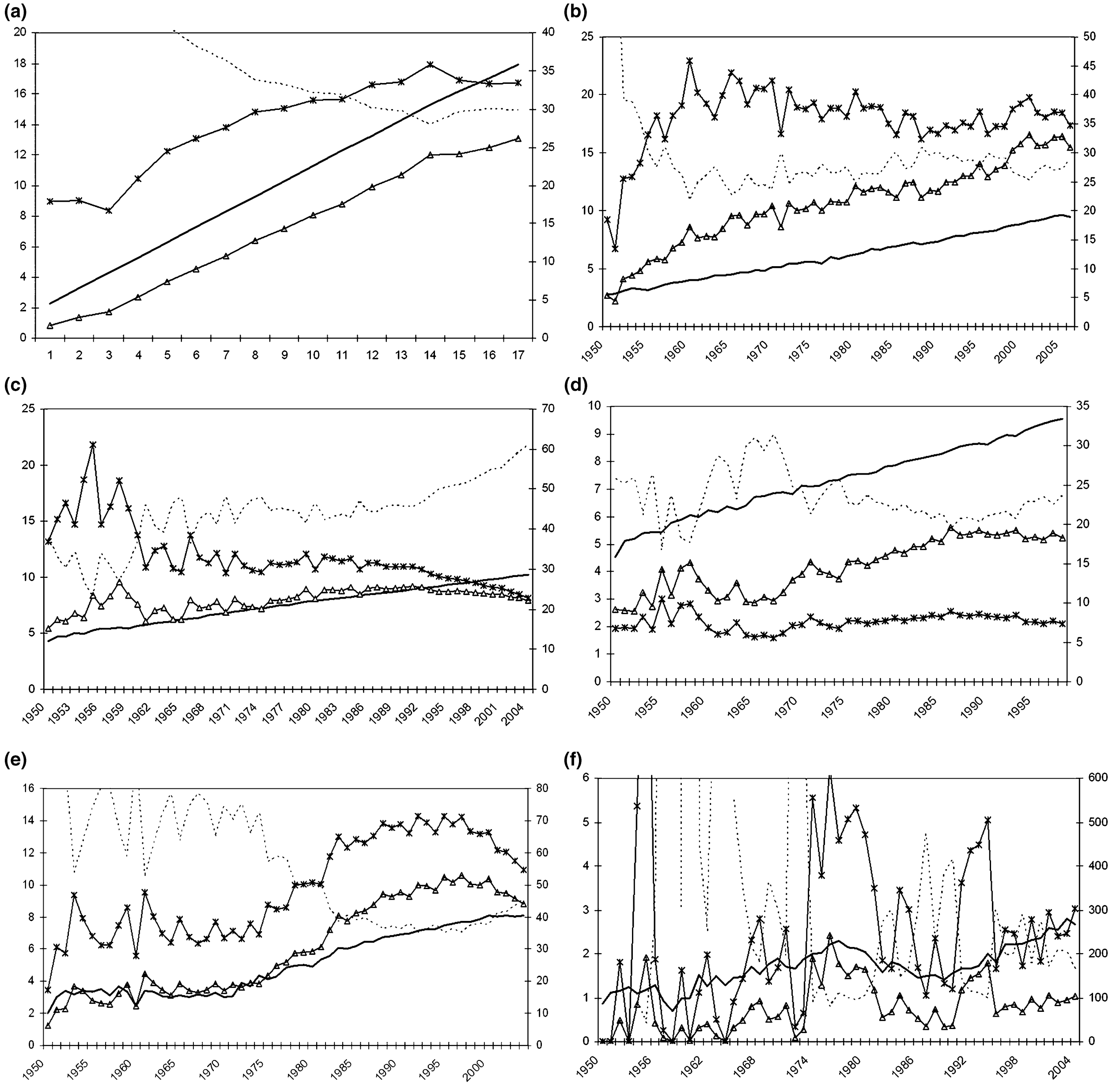

Figure 2 Evolution of the mean inbreeding (-△-, right axis in mice and Carthusian examples and left axis for the others), mean individual increase in inbreeding (×10, except Carthusian, -*-, left axis), mean equivalent generations (–, left axis), and $\overline{N_e}$(---, right axis) across year of birth (across generations in mice), in the analysed pedigrees: (a) mice population; (b) Fighting Bull population; (c) Spanish Purebred horse; (d) Carthusian strain of the Spanish Purebred horse; (e) Spanish Arab horse; (f) Spanish Anglo-Arab horse.

obtained using variances in family size are larger than those obtained using both regression on equivalent generations and using ΔF_i except for the population with shallower pedigree (Spanish Anglo-Arab horse). In this latter population, the regression coefficients are affected by the existence of a few highly inbred individuals while most individuals in the populations have $F = 0$, as about 30% of Spanish Anglo-Arab horse are founders. In this respect, the larger estimates for the obtained $\overline{N_e}$ using ΔF_i, reflect more appropriately the genetic situation of such populations with shallow pedigrees than those based on regression of individual values of inbreeding on equivalent generations. A possible explanation for this is that the regression of individual values assumes a quadratic loss function (applying the

Table 2 Estimates of N_e using the whole pedigrees (WP) and reference subpopulations fitted using the individuals born in the last 10 years (RS_{10}) or in a given period of years ($RS_{xx–yy}$) obtained using regression of the individual coefficients of inbreeding on equivalent generations [$b(t)$], variance of family size (V) and individual increase in inbreeding ($\overline{N_e}$)

	No pedigree depth restriction			$t \geq 4$			$t \geq 8$		
Population	N_e [$b(t)$]	N_e (V)	$\overline{N_e}$ ± SD	N_e [$b(t)$]	$\overline{N_e}$ ± SD	N	N_e [$b(t)$]	$\overline{N_e}$ ± SD	N
Mice									
WP	27.1	198.4	39.4 ± 5.66	24.8	34.2 ± 2.42	7307	24.2	31.3 ± 1.52	5289
RS_{18}	12.3		29.9 ± 1.33	12.3	29.9 ± 1.33	283	12.3	29.9 ± 1.33	283
$RS_{15–16}$	19.7	223.1	28.8 ± 1.21	19.7	28.8 ± 1.21	985	19.7	28.8 ± 1.21	985
Fighting Bull									
WP	25.0	22.2	38.8 ± 6.16	29.0	26.9 ± 2.23	20 482	26.1	27.2 ± 1.30	7654
RS_{10}	19.0		27.5 ± 1.46	17.8	27.5 ± 1.43	5648	33.4	26.9 ± 1.20	5300
$RS_{92–01}$	17.7	101.1	28.2 ± 1.70	16.5	28.1 ± 1.67	6624	24.1	27.3 ± 1.38	4776
SPB									
WP	64.0	257.9	52.4 ± 6.01	147.4	51.2 ± 5.31	135 813	−533.1	52.3 ± 4.92	121 371
RS_{10}	146.4		55.4 ± 4.96	150.9	55.4 ± 4.96	89 318	172.3	55.4 ± 4.94	89 080
$RS_{83–93}$	62.9	2570.6	50.1 ± 4.79	67.9	50.0 ± 4.78	55 681	85.0	50.0 ± 4.72	54 722
Carthusian									
WP	19.5	32.5	26.4 ± 3.96	20.4	22.4 ± 2.64	5172	273.5	21.4 ± 1.54	2518
RS_{10}	33.4		22.2 ± 1.60	40.0	22.2 ± 1.59	1719	−142.2	22.0 ± 1.53	1671
$RS_{77–89}$	12.2	88.1	21.4 ± 1.71	11.6	21.3 ± 1.71	1463	21.9	20.3 ± 1.49	924
Arab									
WP	30.6	121.7	46.6 ± 7.04	22.0	37.7 ± 3.93	13 624	21.6	33.1 ± 2.67	4138
RS_{10}	18.1		39.1 ± 4.19	15.3	38.4 ± 4.00	6093	19.2	33.8 ± 2.80	3495
$RS_{86–97}$	22.2	569.9	36.5 ± 3.85	16.2	34.5 ± 3.13	5095	11.4	30.8 ± 2.27	1364
Anglo-Arab									
WP	71.1	142.7	210.0 ± 72.30	870.9	61.6 ± 13.81	1022			
RS_{10}	80.5		209.8 ± 68.20	91.2	84.7 ± 19.37	612			
$RS_{84–96}$	74.1	372.8	175.8 ± 59.70	25.5	96.8 ± 20.34	255			

Analyses were carried out without pedigree depth restrictions and using only individuals with four equivalent generations or more ($t \geq 4$) and eight equivalent generations or more ($t \geq 8$).

RS_{18}: animals in the 18th generation; $RS_{15–16}$: animals in the 15th and 16th generation; RS_{10}: animals born in the last decade. Other RS_{xx-yy} animals born between the years 19xx and 19yy; N: number of individuals.

largest weight on extreme values), while taking the mean of individual values assumes a linear loss function (applying equal weight on all values).

The estimates of realized N_e computed from ΔF_i ($\overline{N_e}$) tend to be higher than those obtained from regression on equivalent generations. Moreover, when RS is analysed, the values of $\overline{N_e}$ are quite stable regardless of any restriction of the pedigree depth of the individuals included in the corresponding RS, while the N_e obtained from regression of F on equivalent generations show an undesired behaviour. In the Carthusian population it is not only possible to find a noticeable increase of the N_e value from 20.4 to 40.0 when RS_{10} includes individuals with four or more equivalent generations in the pedigree but also negative values of N_e (−142.2) when RS includes individuals with eight or more equivalent generations in the pedigree. This was similar to that assessed in the SPB, where N_e ranged between 147.4 and −533.1 for WP. Note that a negative estimate of N_e can be obtained when the increase in inbreeding is attained by regression of the inbreeding coefficient on the year of birth and younger individuals are less inbred than the older individuals. Thus, the N_e obtained depends partially, on the effect of the changes in the mating policy. The criterion of minimal coancestry has recently been used to define the mating policy (Valera *et al.* 2005) in the Carthusian and SPB populations. However, $\overline{N_e}$ shows a more stable value of about 22 for the Carthusian population and 52 for the SPB population.

Discussion

The estimation of the realized N_e from individual increase in inbreeding has been shown to perform reasonably well in simulated examples (Gutiérrez *et al.* 2008) from the comparison between expected and estimated values. Its performance in real populations is rather complex because the real value of the parameter is unknown and no point of comparison exists. To assess this performance, N_e has been computed from individual increase in inbreeding in several real populations analysing the influence of the definition of different RS relating to more or less recent periods and pedigree knowledge. Some properties of the realized N_e computed from individual increase in inbreeding were tested in these real pedigrees.

This methodology addresses the realized N_e directly from ΔF that theoretically becomes constant in an ideal population instead of addressing it from F that increases from one generation to another in the idealized population (Falconer & Mackay 1996). By definition, ΔF is thus completely independent of pedigree depth in the idealized population (Falconer & Mackay 1996). Therefore, the use of ΔF_i contributes to overcoming the problem of using the F_i coefficients because the latter are nonlinearly dependent on the pedigree depth of each individual. Of course, ΔF_i are still dependent on the completeness of the analysed pedigree and need a few generations to become constant at the population level (Figures 1 and 2). In this respect, the simple plot of ΔF_i over generations serves as a guide for correctly defining RS to obtain a reliable estimate of the realized N_e (Figure 1). The definition of RS is still critical to obtain useful values of N_e, but this approach makes it possible to assess the confidence level of $\overline{N_e}$ from the standard deviation of $\overline{\Delta F}$ (Gutiérrez *et al.* 2008).

Regression of inbreeding values based on time or generation to compute N_e can be useful when the aim is to know the influence of temporal mating policies in some reference subpopulations, because this method provides estimates that are highly dependent on those policies. This last concern is particularly noticeable when, after mating of relatives, a new mating strategy, based on low co-ancestry, is implemented, which leads to a decrease in the mean inbreeding, resulting in negative N_e values. On the contrary, when we compare $\overline{N_e}$ obtained from $\overline{\Delta F}$ with that obtained using other methods, $\overline{N_e}$ appears more stable regardless of the particular situation and depth of the analysed pedigrees (Table 2). In fact, $\overline{N_e}$ becomes a final measure of the accumulated drift, originating from the foundation of the population until the present, because it is unable to accommodate the consequences of a recent change in the mating policy, especially if the pedigree depth is high. If the aim for estimating the effective population size is to ascertain the effect of changes in the mating policy, it would be recommendable to truncate the pedigree, thus considering as founders those animals which were used to establish the new mating policy. However, always make sure that the depth of the remaining pedigree is enough to achieve reliable estimates.

The estimates of N_e computed from variances of family size are not useful to characterize the 'real' effective size of the population being studied (Gutiérrez *et al.* 2008). This is clearly reflected by the results obtained from the mice population analysed here. The values obtained using this method are approximately the real number of reproductive individuals in each mice generation (Table 2), because during the experiment the number of individuals that were selected to be parents of the following generation tended to be balanced across families. However, the value of N_e obtained using this method did not account for the structure of the population that has a permanent subdivision in nine different replicates per generation. However, the analysed fighting bull population is managed under a minimum co-ancestry mating policy that tends to avoid subdivision. As a consequence, the value of N_e obtained using variances of family size is fairly reasonable for this population. Known subdivision could be taken into account, but this is an uncommon situation in non-idealized situations. Thus, this method only uses differences in family size to estimate N_e, ignoring the subdivision of the population which, in fact, causes the reduction of N_e regarding the number of reproductive animals and therefore, tends to overestimate the parameter. A second important reason explaining why this approach fails in predicting N_e is that it ignores the part of the selection process that causes correlations across generations. The Hill (1979) formula ignores the fact that animals with the highest number of selected sibs are also the animals most likely to have offspring selected.

Besides the number and unbalanced contributions of founders or ancestors, there are several causes that influence the effective size, e.g. population structure, mating policy and changes in population size. This is shown by some of the databases analysed here. The effective number of ancestors computed for the Fighting Bull population (data not shown) is the lowest but its N_e is not so; this is a case of a population of constant size where there is a mating policy, which avoids common ancestors in the last three generations of each animal. A similar N_e is obtained for the mice population with much higher effective number of ancestors and moreover, a balanced contribution of parents to the next generation and, in some cases, with a minimum coancestry mating policy. However, the population is structured in nine lines and N_e is practically identical to the real number of reproductive animals within lines. $\overline{N_e}$ explains the number of individuals that would give rise to the rate of inbreeding (Falconer & Mackay 1996) when addressed from inbreeding coefficients, but when subdivision is present, the mean co-ancestry is lower than mean inbreeding and the Hill method should be preferred when a global idea of inbreeding is desired. Thus, the Hill method guesses better the real number of breeding individuals in the mice example. These results show that estimates of $\overline{N_e}$ take into account the genetic history of the populations, in particular, the size of their founder population, their mating policy, bottlenecks caused by abuse of reproductive individuals. But the estimates are less affected by, for example, changes in the mating policy or the accumulation of pedigree. On the other hand, realized effective population size has to be understood in the context of pedigree knowledge. If for instance, a bottleneck has occurred in the middle of a pedigree, it will affect the genetic stock only from this point in the past while the method also accounts from this point onwards. Then, if a breeder wishes to know the effective size in a population in a specific period, he should define the reference population at the end of the desired period, and truncate the pedigree in the beginning of the period, always making sure that there are enough generations considered within the period to obtain a reliable estimate.

Comparison of $\overline{N_e}$ with other estimates of N_e is problematic as the real value is unknown. Looking at Figures 1 and 2 it can be noted that the individual increases in inbreeding become stable after some generations of pedigree depth. Thus, we have compared the $\overline{N_e}$ values with others based on regression or log regression over generations or dates of birth when only individuals with more than four equivalent generations known are considered (Table 3). Estimates of N_e based on regressions using the whole pedigreed set are also shown in Table 3.

In general, the values of $\overline{N_e}$ obtained in the reference subpopulation are in close agreement with those of N_e obtained via regression (Table 3). In fact, values of $\overline{N_e}$ seem more reliable than the others when partial disagreement exists if a visual inspection of the evolution of N_e is used (Figure 2), but some of these values are worth commenting on. For instance, in the mice population, given that minimum co-ancestry is carried out for three of the replicates while random mating is in the other six replicates, stabilization of $\overline{N_e}$ is achieved in the latter generations and values lower than 30 are rare. $\overline{N_e}$ reaches a value of 34.2 while all the N_e values obtained by regression remain lower than 30 (Table 3). The case of SPB is also worth discussing. The change in the breeding policy brings a reduction in the average inbreeding of the last time points and generations with a consequent increase in N_e values. Looking at the evolution of N_e throughout the time points these values never exceeded more than 50 (Figure 2) and a $\overline{N_e}$ value of 51.2, whereas the lowest value based on regression is 147.4 (Table 3). This seems to be because the regression is more sensitive to the decline in average inbreeding in recent years. A mention is also needed in the case of the two populations subject to migration in a low rate (Spanish Arab horse) or a high rate (Spanish Anglo-Arab horse), in both of them the increase in pedigree knowledge is not linear showing changes in the policy of migration. Again, the evolution of N_e is unstable especially for the Spanish Anglo-Arab breed with very low pedigree knowledge. Thus, if considering N_e based on regression over time for the whole pedigreed set the values are higher than 300, but if only individuals with more than four equivalent generations are considered, the same methodology

Table 3 Estimates of N_e obtained using regression of the individual coefficients of inbreeding on equivalent generations [$b(t)$], regression on birth date, log regression on birth date and individual increase in inbreeding ($\overline{N_e}$)

	Whole pedigreed set				Reference subpopulations ($t \geq 4$)				
	N	$b(t)$	b(date)	log b(date)	N	$b(t)$	b(date)	log b(date)	$\overline{N_e} \pm$ SD ($\overline{N_e}$)
Mice	9152	27.1	28.1	27.3	7307	24.8	26.1	25.4	34.2 ± 2.42
Fighting Bull	33 330	25.0	28.1	27.7	20 482	29.0	30.6	30.6	26.9 ± 2.23
SPB	140 629	64.0	72.8	75.2	135 813	147.4	299.3	396.2	51.2 ± 5.31
Carthusian	6318	19.5	19.3	19.4	5172	20.3	23.0	23.6	22.4 ± 2.64
Arab	18 880	30.6	37.3	36.6	13 624	22.0	43.4	42.5	37.7 ± 3.93
Anglo-Arab	10 568	71.1	343.6	316.3	354	25.5	23.6	22.1	61.6 ± 13.81

Analyses were carried out using the whole pedigrees (WP) and reference subpopulations (RS), including only the individuals with four equivalent generations (t).

provides a value close to 20. Estimation of $\overline{N_e}$ results in 61.6, a coherent value but with a standard error much higher than those obtained for other populations. To conclude, we tested here the new approach to estimate the realized N_e in real populations thus enabling us to obtain stable estimates of N_e.

Acknowledgements

The authors wish to thank the 'Subdirección General de Medios de Producción Ganaderos' of the Spanish Ministry of Agriculture for their support. The authors are indebted to the group 'Modelos de Selección y Evaluación de Reproductores' of the U.C.M. and to J.P. Domecq and to the staff of the Fighting Bull Breeder Association (UCTL) for providing access to their pedigree data. We also thank the 'Fondo de Explotación de los Servicios de Cría Caballar y Remonta' of the Spanish Ministry of Defence and the Horse Breeders' Association for their collaboration. This work was partially funded by a grant from the Junta de Andalucia given to I. Cervantes (BOJA, n°. 120, 21/06/2004). The authors also wish to express their sincere thanks and acknowledgement to one of the referees for their significant contributions to improve this paper.

References

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.

Caballero A. (1994) Developments in the prediction of effective population size. *Heredity*, **73**, 657–679.

Caballero A., Toro M.A. (2000) Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res.*, **75**, 331–343.

Cervantes I., Molina A., Goyache F., Gutiérrez J.P., Valera M. (2008) Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. *Livest. Sci.*, **113**, 24–33.

Duchev Z., Distl O., Groeneveld E. (2006) Early warning system for loss of diversity in European livestock breeds. *Archiv. Anim. Breed.*, **49**, 521–531.

Falconer D.S., Mackay F.C. (1996) Introduction to Quantitative Genetics, 4th edn. Longman Group Ltd., Edinburgh.

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

González-Recio O., López de Maturana E., Gutiérrez J.P. (2007) Inbreeding depression on female fertility and calving ease in Spanish dairy cattle. *J. Dairy Sci.*, **90**, 5744–5752.

Gutiérrez J.P., Goyache F. (2005) A note on ENDOG: a computer program for analysing pedigree information. *J. Anim. Breed. Genet.*, **122**, 357–360.

Gutiérrez J.P., Altarriba J., Díaz C., Quintanilla R., Cañón J., Piedrafita J. (2003) Pedigree analysis of eight Spanish beef cattle breeds. *Genet. Sel. Evol.*, **35**, 43–64.

Gutiérrez J.P., Cervantes I., Molina A., Valera M., Goyache F. (2008) Individual increase in inbreeding allows estimating effective sizes from pedigrees. *Genet. Sel. Evol.*, **40**, 359–378.

Hill W.G. (1979) A note on effective population size with overlapping generations. *Genetics*, **92**, 317–322.

Ibáñez-Escriche N., Moreno A., Nieto B., Piqueras P., Salgado C., Gutiérrez J.P. (2008) Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response. *Genet. Sel. Evol.*, **40**, 279–293.

James J.W. (1977) A note on selection differentials and generation length when generations overlap. *Anim. Prod.*, **24**, 109–112.

Maignel L., Boichard D., Verrier E. (1996) Genetic variability of French dairy breeds estimated from pedigree information. *Interbull Bulletin*, **14**, 49–54.

Moreno A. (1998) Optimización de la respuesta a la selección en ''*Mus musculus*'' con consaguinidad restringida. PhD Thesis, Colección Tesis Doctorales, T22287, Univ. Complutense, Madrid.

Pérez-Enciso M. (1995) Use of the uncertain relationship matrix to compute effective population size. *J. Anim. Breed. Genet.*, **112**, 327–332.

Valera M., Molina A., Gutiérrez J.P., Gómez J., Goyache F. (2005) Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. *Livest. Prod. Sci.*, **95**, 57–66.

Wright S. (1969) Evolution and the Genetics of Populations: The Theory of Gene Frequencies, Vol. 2. University of Chicago Press, Chicago, IL.