Estimating the Effective Size of Farm Animals Populations from Pedigree or Molecular Data: a Case Study on two French Draught Horse Breeds

E. Verrier[*†], G. Leroy[*†], C. Blouin[†], J.C. Mériaux[‡],
X. Rognon[*†], and F. Hospital[†]

Introduction

The effective size (Ne) is a parameter of paramount importance to manage populations and understand their evolution. A large number of methods for estimating the effective size of real populations have been proposed in the literature. Unfortunately, these methods are rarely compared to each other, and not in a systematic way. The aim of the present study was to compare a few methods based on different concepts and using different information. The case of two French draught horse breeds, for which both pedigree and molecular data were available, was used to make the comparisons.

Material and methods

Populations studied. Two French draught horse breeds showing several differences in population size and breeders' management practices, namely the Comtois and the Boulonnais breeds, have been considered for this study. With about 7,000 mares owned by 3,700 breeders, the Comtois breed is the French draught horse breed with the highest actual population size. The Boulonnais breed is considered as endangered, due to both a small actual population size, with about 600 mares owned by 250 breeders, and an unbalanced use of stallions leading to a high rate of inbreeding (Verrier et al., 2005).

Information available and animals sampled for estimating *Ne*. Pedigree data came from the national horse database, as filed by the "*Institut français du Cheval et de l'Equitation*". This database included a total of 81,160 and 8,282 animals in the Comtois and Boulonnais breeds, respectively, born from 1900 to 2004. Molecular data resulted from the stallion's parentage controls which were carried out by the "Labogena" laboratory and concerned a total of 1,194 and 563 animals in the Comtois and Boulonnais breeds, respectively, born from 1988 to 2004. A set of 11 microsatellite markers were used, including the 10 markers recommended by ISAG for parentage controls. The group of animals born from 2002 to 2004 was called "reference population" and was considered as a "starting point" for the estimation of *Ne* (see next).

[*] AgroParisTech, UMR1313 Génét. Anim. et Biol. Intégrative, 16 rue Claude Bernard, 75231 Paris 05, France

[†] INRA, UMR1313 Génét. Anim. et Biol. Intégrative, 78350 Jouy-en-Josas, France

[‡] Labogéna, Domaine de Vilvert, 78350 Jouy-en-Josas, France

Methods for estimating the effective size. Four different methods were used to estimate *Ne* (Table 1). The first two methods, referenced here as "A" and "B", are known to provide an estimation of the "Inbreeding" effective size (Ne_I). They were based on the evolution observed during one generation of (A) the average coefficient of inbreeding computed from pedigree or (B) the expected heterozygosity averaged across the 11 markers. The other two methods, referenced here as "C" and "D", are known to provide an estimation of the "Variance" effective size (Ne_V). They were based on (C) demographic parameters computed between two consecutive generations of breeding animals or (D) the temporal variance of allele frequencies at the 11 markers from one generation to the other.

Table 1: Summary of the methods used to estimate *Ne*

Code of the method		A	B	C	D
Observed phenomenon		Rate of inbreeding	Decrease of expected heterozygosity	No of breeding animals and (co)variances of their progeny sizes	Temporal variance of allele frequencies
Kind of *Ne*	Ne_I	x	x		
	Ne_V			x	x
Information used	Pedigree	x		x	
	Markers		x		x
Set of animals		Reference population and animals born one generation before	Same as A, restricted to genotyped animals	Parents and grand-parents of the reference population	Same as A, restricted to genotyped animals
Reference		From Wright (1931)	From Wright (1931)	Hill (1972)	Waples (1989) + Nei & Tajima (1981)

Results, discussion and conclusions

The average generation length was estimated to 7.0 and 9.2 years in the Comtois and Boulonnais breeds, respectively. Therefore, for methods A, B and D, the group of animals born one generation before the reference population, used to compute *Ne* (see Table 1), was defined as the group of animals born in 1995-1997 in the Comtois breed and born in 1993-1995 in the Boulonnais breed. In both breeds, due to well known biological reasons, the variance of progeny size was very small for female parents. For male parents, the Boulonnais breed showed larger variances of both male and female progeny size than the Comtois breed. Finally, the effective size estimated from demographic parameters (Method "C", Hill, 1972), was equal to 916 and 159 in the Comtois and Boulonnais breeds, respectively.

Figure 1.a shows the evolution of the average coefficient of inbreeding over time. The annual rate of inbreeding in the Boulonnais breed was found to be almost twice the one in the Comtois breed (+0.15 *vs.* +0.09 per year). In the last generation (method A), the effective size was estimated to 79 and 34 in the Comtois and Boulonnais breeds, respectively.
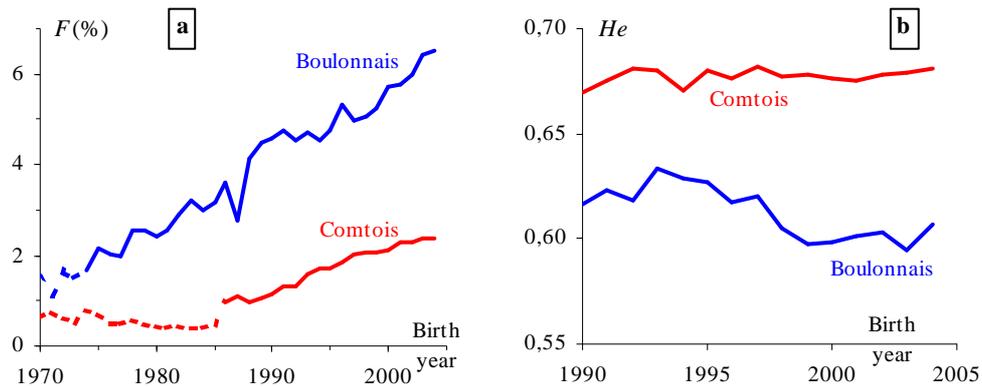
Figure 1: Evolution of the average coefficient of inbreeding (*F*, in %) and the expected heterozygozyty at 11 markers (*He*) according to birth year in the Boulonnais and Comtois horse breeds. Left: dashed lines correspond to the period when the pedigree completeness level was considered as low (less than the equivalent of 3 full generations of ancestors known). Right: to smooth out fluctuations due to small annual sample size, each point represents the mean of three consecutive years.

Depending on the marker locus, a number of alleles ranging from 6 to 11 in the Comtois breed and 3 to 7 in the Boulonnais breed was observed. The average number of alleles per locus was 7.5 and 5.7 in the Comtois and Boulonnais breeds, respectively. These differences between breeds may be partly explained by the larger size of the sample in the Comtois breed. The expected heterozygosity was always higher in the Comtois breed than in the Boulonnais breed (Figure 1.b). Almost no evolution was observed over time in the Comtois breed, whereas *He* decreased in the Boulonnais breed by about 4.5% in one generation. From this evolution of *He* over time, *Ne* was estimated to 427 in the Comtois breed and only 11 in the Boulonnais breed. From the temporal variance of allele frequencies, *Ne* was estimated to 5,712 and 161 in the Comtois and Boulonnais breeds, respectively.

The different estimated values of *Ne* are summarized on Figure 2. Whatever the method, the estimated effective size was always larger in the Comtois breed than in the Boulonnais breed, which is consistent with the known history and the current demographic parameters of these breeds (Verrier et al., 2005). For a given breed, a very large range of variation was observed between estimated values: the ratio of the highest one over the lowest one was 14.6 in the Boulonnais breed and even 77.4 in the Comtois breed. The estimated values of the "Variance" effective size were always higher than those of the "Inbreeding" effective size. For a given kind of *Ne* ("Variance" or "Inbreeding"), in the Comtois breed, the estimated value was higher when using molecular data than pedigree data. In the Boulonnais breed, the rank of estimated values according to the information used depended on the kind of *Ne*.
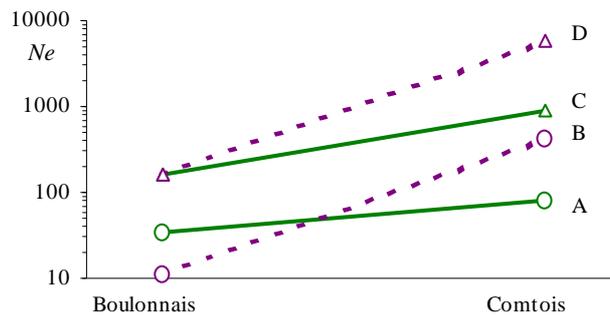
Figure 2: Estimated values of the effective size of the Boulonnais and Comtois horse breeds, using four different methods (A to D, see Tab.1). Circle = Inbreeding effective size; triangle = variance effective size. Straight (green) line: estimation from pedigree data; dashed (purple) line: estimation from molecular data (11 markers). The scale for the y axis is logarithmic.

Several reasons may explain the large differences in estimated values of the effective size of a given breed. First, while the concept of effective size has an asymptotic meaning (Wright, 1931), we made estimates over a very short time period (a single generation) and in breeds which are unlikely to be at an asymptotic state. Second, some hypothesis underlying the estimation methods, e.g. random mating, could be not met in the populations under study. Third, while the "Variance" and "Inbreeding" effective sizes are generally either identical or quite similar, under some conditions they may differ substantially (Kimura and Crow, 1969). Fourth, the sample sizes were not so large, especially for molecular data and in the Boulonnais breed, allowing large estimation errors. More detailed investigations are required to check the relative impact of these different reasons in the case of these two populations. More generally, further research is necessary to compare on a systematic way a number of methods under various circumstances. It would be very useful to determine the method(s) to be recommended according to the populations studied.

References

Hill, W.G. (1972) *Theor. Pop. Biol.* 3, 278-289.

Kimura M. and Crow J.F. (1969) *Evolution* 17, 279-288.

Nei, M. and Tajima, F. (1981) *Genetics* 98, 625-640.

Verrier, E., Loywyck, V., Donvez, J. *et al.* (2005) *Journée Rech. Equine* 31, 161-171.

Waples, R.S. (1989) *Genetics* 121, 379-391.

Wright, S. (1931) *Genetics* 16, 97-159.