Genetic variability within french race and riding horse breeds from genealogical data and blood marker polymorphisms

Sophie Moureaux, Etienne Verrier, Anne Ricard, J-Claude J.-C. Meriaux

HAL Id: hal-02694593
https://hal.inrae.fr/hal-02694593

Submitted on 1 Jun 2020

Original article

Genetic variability within French race and riding horse breeds from genealogical data and blood marker polymorphisms

S Moureaux [1,2], É Verrier [1,2], A Ricard [1], JC Mériaux [3]

[1] *Station de génétique quantitative et appliquée, Institut national de la recherche agronomique, 78352 Jouy-en-Josas cedex;*
[2] *Département des sciences animales, Institut national agronomique Paris-Grignon, 16, rue Claude-Bernard, 75231 Paris cedex 05;*
[3] *Laboratoire d'analyses génétiques pour les espèces animales, 78352 Jouy-en-Josas cedex, France*

Summary – The genetic variability of five horse breeds raised in France was analysed: Thoroughbred, Trotteur Français, Arab, Anglo-Arab and Selle Français. Genealogical data and genotypes at seven blood group and nine protein loci were used. Paternal family sizes were found to be unbalanced, especially in Trotteur français, Selle Français and Thoroughbred. Average coefficients of inbreeding for offspring born from 1989 to 1992 were 1.02 (Thoroughbred), 1.86 (Trotteur Français), 3.08 (Arab), 1.17 (Anglo-Arab) and 0.70% (Selle Français). High individual coefficients ($> 6.25\%$) were found in substantial proportions only in Arab, where such high values represent one fifth of the total individual coefficients. Inbreeding was analysed according to the number of generations of ancestors considered. The results revealed the importance of close inbreeding in Arab and remote inbreeding in Thoroughbred. Arab was the only breed that showed evidence for a substantial amount of mating between close relatives. From 1974 to 1992, the rates of inbreeding, in percentage points per year, were $+ 0.026$ (Thoroughbred), $+ 0.052$ (Trotteur Français), $+ 0.071$ (Arab), $+ 0.029$ (Anglo-Arab) and $+ 0.024$ (Selle Français). The distribution of genetic contributions of founder animals was found to be unbalanced, especially in Trotteur Français where 25 founder animals only accounted for half the actual gene pool. No significant time-trend was found for blood markers allelic frequencies. The mean heterozygosity was highest in Trotteur Français and Selle Français and lowest in Thoroughbred and Anglo-Arab. The meaning of recent trends for genetic variability is discussed. The need for equalizing paternal family sizes in the future is outlined.

demography / inbreeding / probability of gene origin / heterozygosity / horse

Résumé – Analyse de la variabilité génétique de cinq races françaises de chevaux de course et de sport à partir des données généalogiques et du polymorphisme des marqueurs sanguins. *On a dressé un bilan de la variabilité génétique au sein des cinq principales races françaises de chevaux de course et de sport, le Pur-Sang, le Trotteur*

Français, l'Arabe, l'Anglo-Arabe et le Selle Français. On a utilisé les données généalogiques ainsi que les résultats de typage pour sept groupes sanguins et neuf protéines sanguines. La distribution du nombre de descendants, mâles ou femelles, par étalon est déséquilibrée, particulièrement pour le Trotteur Français, le Selle Français et le Pur-Sang. Le coefficient de consanguinité moyen des animaux nés entre 1989 et 1992 est de 1,02 % (Pur-Sang), 1,86 %, (Trotteur Français), 3,08 % (Arabe), 1,17 % (Anglo-Arabe) et 0,70 % (Selle Français). Des coefficients individuels élevés (> 6,25 %) n'ont été trouvés en proportion substantielle que chez l'Arabe, où ils représentent un cinquième des valeurs calculées. Une analyse des coefficients moyens en fonction du nombre de générations d'ancêtres considéré montre l'importance de la consanguinité éloignée chez le Pur-Sang et de la consanguinité proche chez l'Arabe, seule race où la pratique des accouplements entre proches apparentés semble être courante. De 1974 à 1992, le taux d'accroissement du coefficient de consanguinité moyen (en points de pourcentage) par année de naissance a été de +0,026 (Pur-Sang), +0,052 (Trotteur Français), +0,071 (Arabe), +0,029 (Anglo-Arabe) and +0,024 (Selle Français). La distribution des contributions des ancêtres fondateurs au patrimoine génétique actuel est déséquilibrée. La situation est particulièrement critique chez le Trotteur Français, où seulement 25 animaux fondateurs contribuent pour la moitié des gènes présents actuellement. Aucune tendance significative d'évolution des fréquences géniques des marqueurs sanguins n'a pu être mise en évidence. L'hétérozygotie moyenne est la plus élevée pour le Trotteur Français et le Selle Français et la plus faible pour le Pur-Sang et l'Anglo-Arabe. La signification et les causes de l'évolution récente de la variabilité génétique au sein de chaque race sont discutées. On insiste sur la nécessité de mieux équilibrer les tailles de familles paternelles afin de préserver la variabilité actuelle.

démographie / consanguinité / probabilité d'origine des gènes / hétérozygotie / cheval

INTRODUCTION

Race and riding horse breeding has expanded greatly in France during the last three decades. This activity involves five main breeds to different extents (table I) and with different origins and selection goals. Thoroughbred was imported from the British Isles during the last century, and is bred for galloping races. Trotteur Français is a native breed from Normandy, bred for trotting races; today, it is the most widespread horse breed in France. Arab was imported from the Near East during the last century; it is bred for several purposes, mainly leisure but also sporting activities and endurance racing. The last two breeds are composite breeds. Anglo-Arab was created at the end of the last century, mainly by crossing Thoroughbred and Arab; it is bred for several purposes, such as jumping, dressage, cross-country and galloping races reserved for this breed. Selle Français was more recently derived from the cross of local breeds (essentially the Normandy breed) and Thoroughbred. This is the most widespread riding horse breed in France, and is bred mainly for jumping, but also for dressage and cross-country. Management rules for these breeds are different. Thoroughbred and Arab are managed with closed studbooks, but at an international level. Trotteur Français is essentially managed with a closed studbook, at a national level. However, a few foreign Standardbred stallions may be used (since 1977). Anglo-Arab and Selle Français were managed

Table I. Basic information for the different breeds.

Breed	No of breeding animals in 1993	
	Stallions	Mares
Thoroughbred	468	8 354
Trotteur Français	846	21 588
Arab	469	1 344
Anglo-Arab	262	3 483
Selle Français	549	15 095

with open studbooks, but the Selle Français studbook has been partially closed (since 1994).

The analysis of the genetic structure of these five breeds and an investigation of current trends concerning their genetic variability are presented here. The purpose of this work is to provide an understanding of the background upon which selection is applied. Genealogical data and blood marker polymorphism will be used. In order to highlight the results, some demographical parameters will also be given.

MATERIAL AND METHODS

Genealogical data and their analysis

The data used came from the national horse register, *Système d'Identification Répertoriant les Equidés* (SIRE), as filed by the *Institut du Cheval*. At the time of this study, this file included all the animals born from 1974 to 1992 and their known ancestors (table II).

Table II. Numbers of animals in data files.

Breed	Genealogical data	Blood-typing data[a]	
	Total no of animals[b]	Born in 1992	Born from 1989 to 1992
Thoroughbred	149 482	3 770	14 200 − 15 810
Trotteur Français	203 486	5 100	15 940 − 17 480
Arab	12 557	860	3 030 − 3 040
Anglo-Arab	63 299	540	1 570 − 1 580
Selle Français	231 423	4 790	12 300 − 13 290

[a] For blood typing, numbers change from one marker to another: for the first sample, only the rounded mean value is given (no variation more than ± 3% of the mean); for the second sample, due to a larger range of variation, the rounded range of variation is given.
[b] Including foreign ancestors and ancestors from other breeds.

Demographical analysis was performed for each breed separately. For a given animal kept for breeding, only its 'useful' offspring were taken into account; an

offspring was considered as useful if it left at least one offspring. Generation lengths, in the four pathways, were computed as the average age of parents at the birth of their useful offspring. For this purpose, two cohorts of offspring were considered, born in 1974 and 1985. The distributions of numbers of useful offspring per sire and per dam were analysed considering all useful offspring born between 1974 and 1992.

Genetic structure was analysed on the basis of pedigree information. The pedigree completeness level was analysed by computing the average proportion of ancestors known per generation for a given cohort of offspring. Coefficients of inbreeding were computed for all the animals in the file, using the algorithm proposed by Quaas (1976). In order to distinguish close and remote inbreeding, these coefficients were computed for successive values of the number of generations of ancestors considered and for the total pedigree information available. The distribution of individual coefficients was analysed for offspring born between 1989 and 1992. The evolution of the average coefficient of inbreeding per birth year was observed from 1974 to 1992. The annual rate of change in inbreeding was estimated by linear regression over time. Ancestors with no parent known in the file were considered as founders and probabilities of origin of genes of offspring born in 1992 were computed in reference to these founders. The distribution of genetic contributions of founders was analysed and an effective number of founders (NF_e) was computed as:

$$NF_e = \frac{1}{\sum_i \theta_i^2}$$

where θ_i is the probability of a current gene coming from a given founder (i). If each founder had the same genetic contribution, the effective number would be equal to the actual number of founders. If not, it would be lower than the actual number.

Blood typing data and their analysis

In France, parentage control was made systematically in Thoroughbred and Arab breeds, for all animals born since 1985 and 1988, respectively, and for other breeds since 1988, when artificial insemination was used or when a mare was covered by more than one stallion. Before 1988, parentage control was requested by breeders for 20 to 40% of offspring in the different breeds, except in Arab, for which the rate of testing was around 60%. Blood typing data resulted from these tests, which were carried out by the *Laboratoire d'analyses génétiques pour les espèces animales* (Labogéna). Standard methods of starch gel electrophoresis and polyacrylamide gel electrophoresis were used to identify alleles of nine protein loci: albumin (Al), post-albumin (A1B), carboxylesterase (Es), Gc protein (Gc), glucosephosphate isomerase (GPI), 6-phosphogluconate dehydrogenase (PGD), phosphoglucomutase (PGM), protease inhibitor (Pi) and transferrin (Tf). Standard serological reactions were used to detect red cell alloantigens at seven blood group systems: A, C, D, K, P, Q, U. All loci considered are known to be polymorphic in the domestic horse.

For protein loci, allelic frequencies were estimated by direct counting from phenotypes. For blood group systems, an iterative procedure was used to assess the

conditional possible genotype(s) of each animal knowing its phenotype and parents' and its offspring's phenotypes, on the basis of the Mendel rules. The likelihood of the sample was then computed assuming Hardy–Winberg genotype frequencies and allelic frequencies were estimated by maximizing this likelihood. For each locus, the evolution of allelic frequencies across birth years was analysed from 1974 to 1992.

Two recent samples were analysed in detail (table II). The first sample included offspring born in 1992 and typed, amounting to 99, 41, 99, 26 and 46% of offspring born in the five breeds (listed in the same order as in table II). For this sample, Hardy–Weinberg heterozygosity (H) was computed for each locus according to the classical formula:

$$H = 1 - \sum_i p_i^2$$

and the effective number of alleles (n_e) was computed as follows:

$$n_e = \frac{1}{\sum_i p_i^2} = \frac{1}{(1 - H)}$$

In the formulae, p_i is the estimated frequency of allele i. Sampling variance of H was computed with the formula given by Nei and Roychoudhury (1974, cited by Hedrick, 1985, Eq 2.31, p 65). On a second sample, including all offspring born between 1989 and 1992 and typed (table II), Hardy–Weinberg proportions of genotypes for protein loci were checked by the χ^2 test. This second sample was larger than the first, in order to avoid problems due to too small expected numbers of animals for a given genotype. However, when such a case occurred, the rarest alleles were pooled and considered as a single allele. Two proteins were excluded from this analysis, Es and Pi, due to too large a number of necessary poolings, which systematically increased the 'observed' proportion of homozygotes.

RESULTS

Demographical results

Figure 1 shows the change in total number of animals born per year from 1974 to 1992, which provides a good view of the recent evolution in size of each breed. Trotteur Français and Selle Français showed similar trends, with the numbers of births a little more than doubling in 18 years. Anglo-Arab and Arab increased in numbers but remained at a moderate level. Only Thoroughbred has fluctuated in yearly number of births, from 3 000 to 4 500, with a decrease in the last four years.

Table III shows generation lengths between useful offspring born in 1985 and their parents. The paternal interval was significantly higher than the maternal one for female offspring of all breeds ($P < 0.01$ or $P < 0.001$). For male offspring, the difference between average intervals on paternal and maternal sides was only significant ($P < 0.01$) in Thoroughbred and Trotteur Français. Average generation length was smaller in Arab and Thoroughbred than in other breeds ($P < 0.05$ for the ten possible tests as a whole). No significant time-trend was observed for
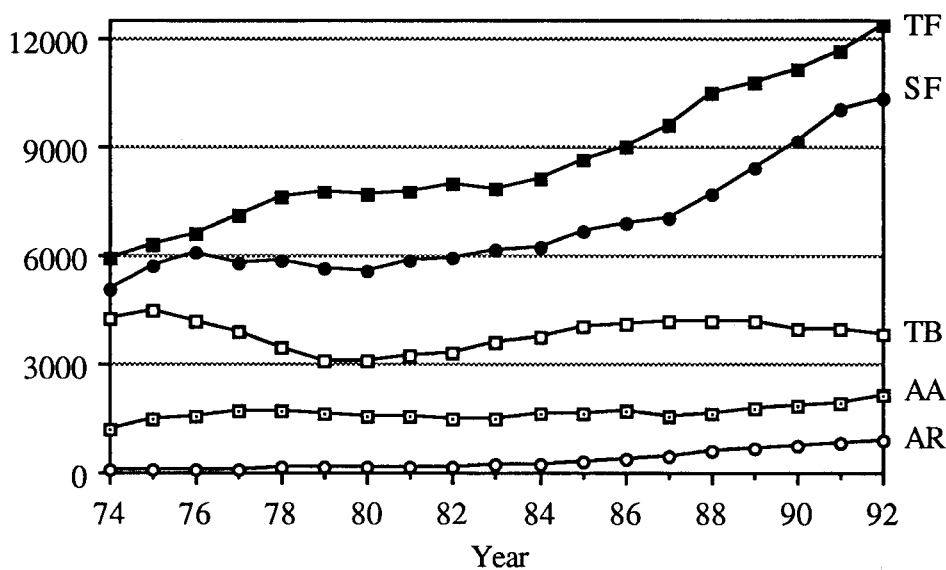
Fig 1. Evolution of the number of animals born per year from 1974 to 1992. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

Table III. Generation lengths (L, in years) between useful offspring born in 1985 and their parents.

Breed	L_{ss}	L_{sd}	L_{ds}	L_{dd}	$\bar{L}$
Thoroughbred	11.7 ± 0.6	10.8 ± 0.1	9.4 ± 0.5	10.3 ± 0.1	10.6 ± 0.2
Trotteur Français	13.1 ± 0.6	12.8 ± 0.1	10.4 ± 0.6	10.8 ± 0.1	11.8 ± 0.2
Arab	9.8 ± 0.7	10.3 ± 0.4	10.5 ± 0.7	8.2 ± 0.4	9.7 ± 0.3
Anglo-Arab	11.9 ± 1.0	12.0 ± 0.3	10.7 ± 0.9	11.3 ± 0.3	11.5 ± 0.4
Selle Français	11.9 ± 0.6	12.1 ± 0.1	11.7 ± 0.6	10.9 ± 0.1	11.7 ± 0.2

Mean ± standard error of the mean. Values for the four pathways (s = sires, d = dams) and mean over pathways ($\bar{L}$).

generation lengths from 1974 to 1985 except in Selle Français, where the average generation length increased by around six months.

The average number of useful male offspring per stallion with at least one useful male offspring ranged from 2.1 in Arab to 3.9 in Selle Français. The average number of useful female offspring per stallion with at least one useful female offspring ranged from 3.1 in Arab to 18.7 in Trotteur Français. Analysis of the distribution of numbers of useful offspring per mare showed a low variability between mares, both within and between breeds (results not shown here). On the other hand, the
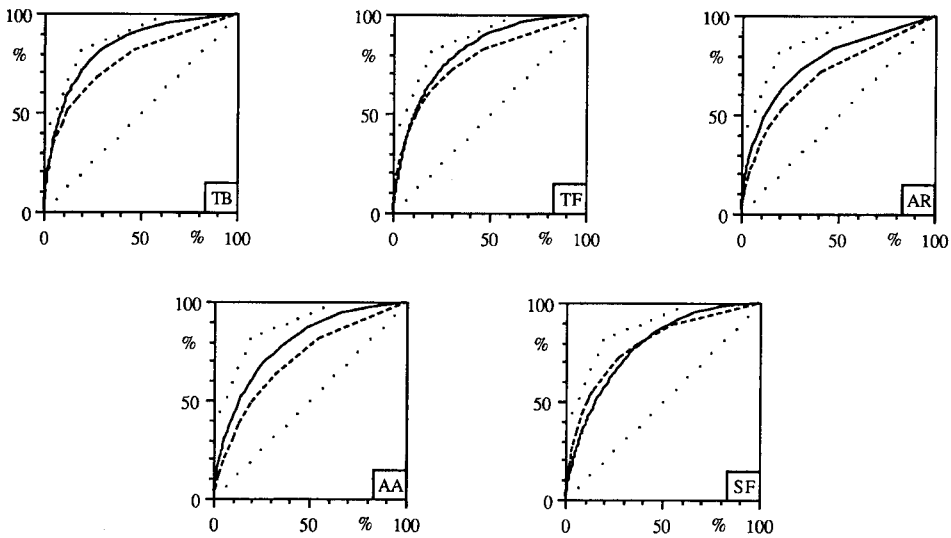
Fig 2. Cumulated proportion (%) of useful offspring of stallions versus cumulated proportion (%) of stallions. Stallions are ranked by decreasing number of offspring. Solid line = female offspring; dashed line = male offspring; dotted lines = reference marks, identical from one breed to another; the diagonal represents a situation of strict equilibrium; the upper left reference line was arbitrarily defined and the closer the curve to this line the more unbalanced the distribution of family sizes. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

distribution of family sizes for stallions was clearly unbalanced. Figure 2 shows the plot of the cumulated proportion of offspring against the cumulated proportion of sires. Except in Selle Français, the distribution was more balanced for male useful offspring than for female ones. For male offspring, the distributions were very similar in Arab and Anglo-Arab and were the most balanced. The distribution was more unbalanced in the other three breeds, especially in Selle Français and Trotteur Français. In these three breeds, a few stallions had more than 40 useful male offspring, ie, ten or more times the average number. For female offspring, the most unbalanced distributions were observed in Thoroughbred and Trotteur Français.

Pedigree completeness level and inbreeding

Considering the most recent cohorts of offspring, the pedigrees in each breed were found to be very complete up to the fifth generation of ancestors, as shown in figure 3 for offspring born in 1992. Up to the fourth generation, the pedigrees were very well known. From the sixth generation onward, the proportion of known ancestors was less than 80% and some differences appeared between breeds. The less complete pedigrees were observed in Trotteur Français and Arab, with proportions of known ancestors less than 6 and 12% respectively from the eighth generation onwards.

On the other hand, in Thoroughbred, the proportion of known ancestors remained higher than 50% up to the eighth generation. The pedigree completeness level in the other two breeds was intermediate. For older cohorts of offspring, pedigrees were less complete. Taking into account the above values of generation lengths (table III), the situation could be roughly described as shown by figure 3 with a shift of one generation per 10–12 years, according to the breed considered. This should be kept in mind when examining the computed coefficients of inbreeding.
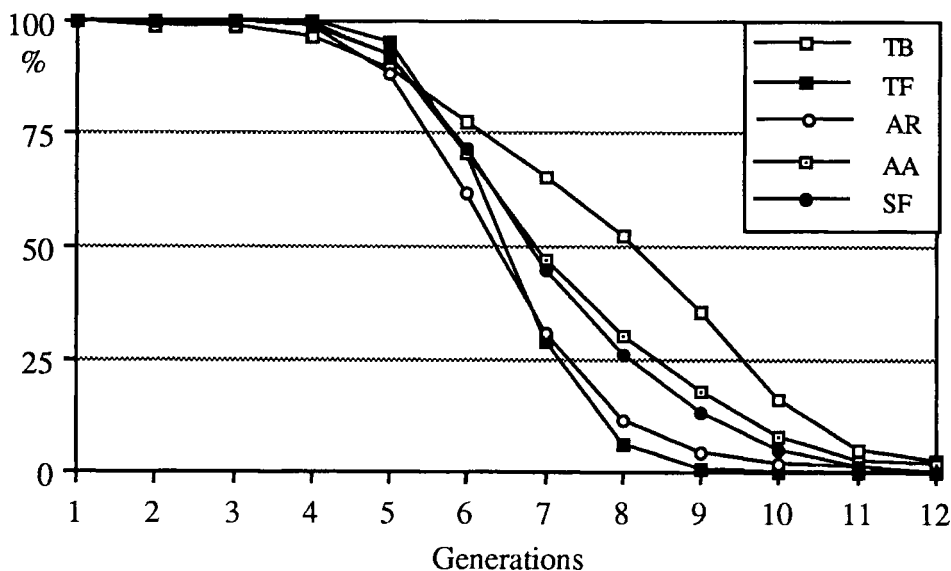

Fig 3. Percentage of ancestors known per generation (1 = parents, etc) for offspring born in 1992. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

Table IV shows the average coefficient of inbreeding in each breed for the four youngest cohorts of offspring taken as a whole. Considering all the animals, the lowest computed mean values were observed in Selle Français and Thoroughbred. The computed mean values were substantially higher in Trotteur Français and Arab. In Thoroughbred and Trotteur Français, very small differences were found between average coefficients computed for all animals and for inbred animals only. In other breeds, the difference between the two computed mean values was appreciable, especially in Arab. These phenomena are directly linked to the distribution of individual coefficients of inbreeding (fig 4). The proportion of animals with a zero computed value was the highest in Arab, and was found to be near to zero in Thoroughbred and Trotteur Français. The lowest variability of individual coefficients was observed in Thoroughbred, whereas the highest variability was observed in Arab. In Arab, coefficients higher than 6.25% (eg, mating between first cousins) and higher than 12.5% (eg, mating between half-sibs) were found to occur at as substantial proportions, around one fifth and one thirtieth, respectively. Such high values were rarely found in other breeds. Figure 5 shows plots of the

average coefficient of inbreeding expressed as a percentage of average coefficients computed from the whole pedigree against the number of generations of ancestors considered. In Arab, inbreeding coefficients computed at the grandparental and great-grandparental levels accounted for around a quarter and a half of the total inbreeding, respectively. On the other hand, in Thoroughbred, half the total inbreeding was reached at only the fifth generation. The situations of the other three breeds were close to Thoroughbred at the grandparental level and, next, were intermediate between Thoroughbred and Arab.

Table IV. Average coefficient of inbreeding (%) in offspring born from 1989 to 1992, computed for all animals (F) and for inbred animals only (F_i).

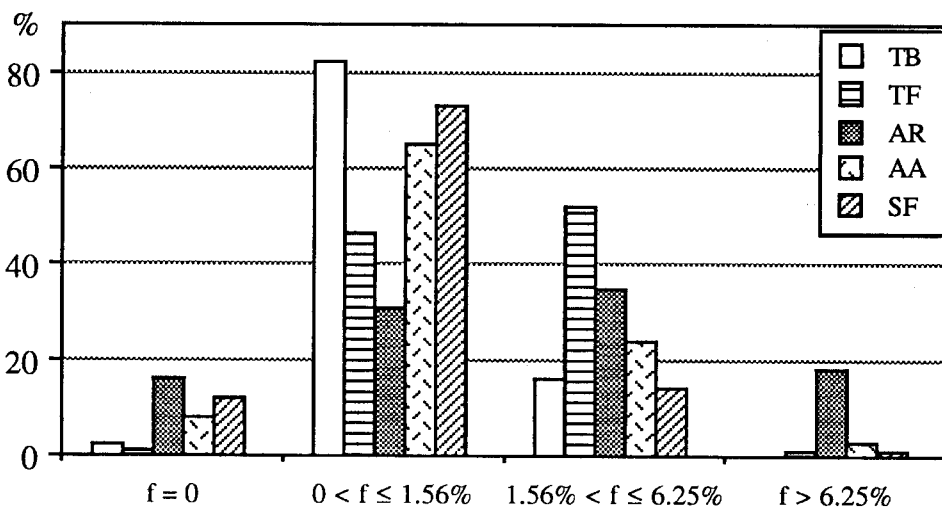|	Breed				
	Thoroughbred	Trotteur Français	Arab	Anglo-Arab	Selle Français
Total no of offspring	15 895	45 754	3 034	7 637	37 879
F	1.02	1.86	3.08	1.17	0.70
No of inbred offspring	15 642	45 255	2 543	7 060	33 508
F_i	1.04	1.88	3.67	1.27	0.80

Fig 4. Distribution of individual coefficients of inbreeding (f) in offspring born between 1989 and 1992. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

Figure 6 shows the evolution of the average coefficient of inbreeding per birth year, taking into account the whole pedigree or four generations of ancestors only. The evolution was less regular in the Arab breed than in the others. In each
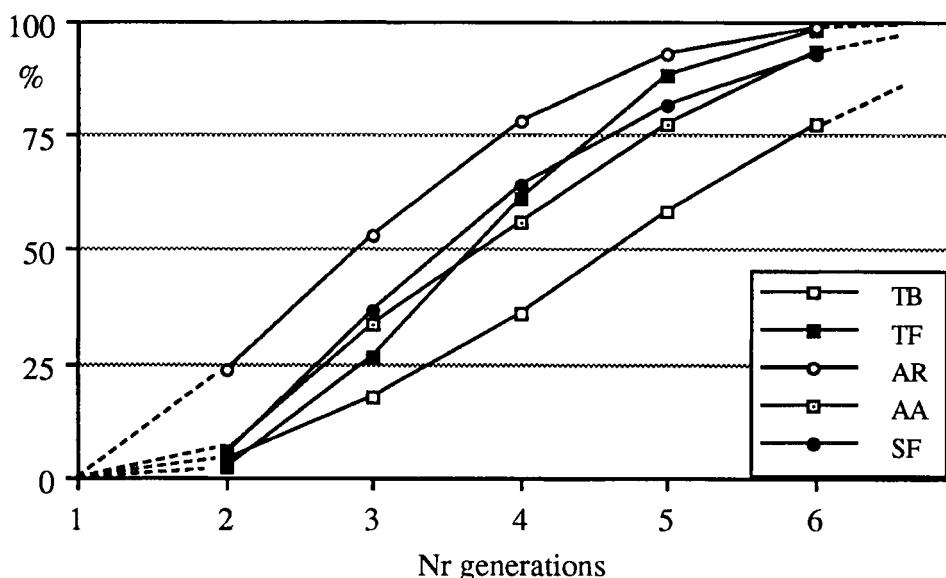
Fig 5. Average coefficient of inbreeding in offspring born between 1989 and 1992, expressed as a percentage of the value computed on the basis of the whole pedigree, according to the number of generations of ancestors considered. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

breed, when the whole pedigree was considered, the average coefficient of inbreeding increased with a statistically significant estimated coefficient of regression over time. In Arab and Trotteur Français, the annual change in inbreeding was approximately equal to respectively three and two times the annual change within other breeds. The variation between observed trends according to the amount of genealogical information considered was small in Arab, especially for the oldest cohorts of offspring. This differences was greater in Selle Français; it was still greater in Trotteur Français and Anglo-Arab, where the coefficients of regression over time estimated from the whole pedigree amounted in both cases to four times the coefficient estimated from only four generations of ancestors. The most extreme situation was observed in Thoroughbred, where the coefficient of regression over time estimated on the basis of four generations of ancestors was not found to differ significantly from zero.

Probabilities of gene origin

For offspring born in 1992 in each breed, the total number of founder animals and the effective number of founders are given in table V; figure 7 shows plots of the cumulated contribution to the gene pool against the cumulated proportion of founders. Clearly, the contributions to the current gene pool were the most balanced in Arab. The minimal proportion of founder animals for a cumulated genetic contribution of 80% was equal to 21.0% in this breed, whereas it ranged
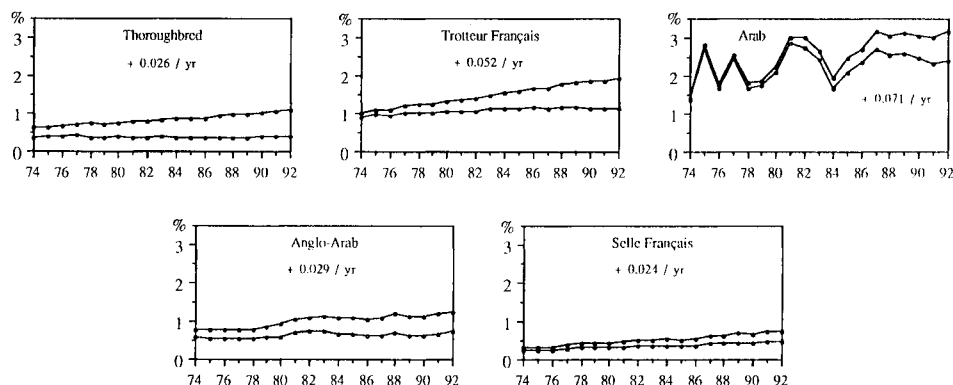
Fig 6. Evolution of the average coefficient of inbreeding (%) per birth year from 1974 to 1992. Bottom line = only four generations of ancestors considered; top line = the whole pedigree considered. Values given under breed names represent estimated regression coefficients of average coefficient of inbreeding (the whole pedigree considered) over time.

Table V. Total number of founders and effective number of founders for offspring born in 1992.

	Breed				
	Thoroughbred	Trotteur Français	Arab	Anglo-Arab	Selle Français
Total No of founders	3 548	2 513	962	3 669	8 933
Effective No of founders	236	70	135	129	333

from 5.6 to 10.5% in the other breeds. The genetic contributions were the most unbalanced in Trotteur Français, where only 25 animals (1% of founders) accounted for half the current gene pool, and next in Selle Français. At an intermediate level between these two breeds and Arab, the Thoroughbred and Anglo-Arab showed quite similar situations. However, the contributions of the most important founders were more balanced in Thoroughbred. Resulting from differences in the total number of founders and the balance of their contributions, large differences were observed between breeds for their effective number of founders (table V). The effective number was particularly small in Trotteur Français, due to very unbalanced contributions. On the other hand, the effective number was the highest in Selle Français, due to a high total number of founders and despite the lack of balance of their contributions. Due to less balanced contributions, the effective number of founders in Anglo-Arab was smaller than in Thoroughbred, with a similar total number of founders, and even smaller than in Arab, which clearly had a smaller total number of founders.
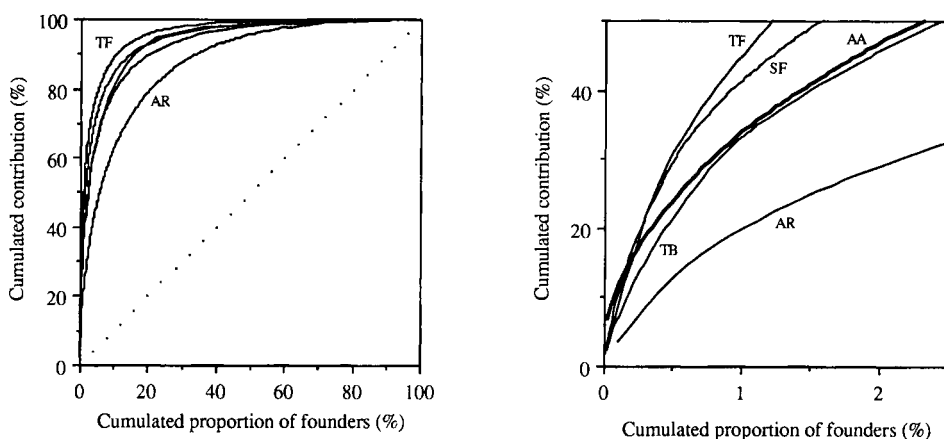
Fig 7. Cumulated contribution (%) to the gene pool of offspring born in 1992 versus cumulated proportion (%) of founder animals. Founder animals are ranked by decreasing contribution to the gene pool. Left: total figure; only the extreme two breeds are named because curves for other breeds overlap; the dotted diagonal represents a situation of strict equilibrium. Right: figure for founders contributing the most. TB = Thoroughbred; TF = Trotteur Français; AR = Arab; AA = Anglo-Arab; SF = Selle Français.

Genetic structure at blood markers

From one breed to another, a total of 60 alleles (Trotteur Français) to 76 alleles (Selle Français) were found across the 16 analysed loci. For the whole period considered (1974–1992), changes in allelic frequencies were generally small, lower than 0.01 in absolute value for most of alleles. Moreover, changes from one year to another were rather erratic and no significant time-trend was observed (plots not shown here). Table VI shows Hardy–Weinberg heterozygosity for the 16 loci analysed in offspring born in 1992 in each breed. Herterozygosity was generally higher for blood groups than for blood proteins. The Gc locus was barely polymorphic in all breeds. Some other loci were polymorphic in one or two breeds only, such as the A1B and PGM proteins or the K blood group. The mean heterozygosity was highest in Trotteur Français and Selle Français and lowest in Anglo-Arab and Thoroughbred. Correspondingly, the effective number of alleles was generally higher in Trotteur Français or Selle Français than in other breeds. For example, the average effective number of alleles across all protein loci was equal to 1.83, 1.81, 1.76, 1.74 and 1.70 in Selle Français, Trotteur Français, Arab, Thoroughbred, and Anglo-Arab, respectively. For blood group loci, the average effective number of alleles was equal to 2.33, 2.13, 2.04, 1.98 and 1.83 in Selle Français, Trotteur Français, Arab, Thoroughbred, and Anglo-Arab, respectively. As in the case of allelic frequencies, within-breed changes of heterozygosity from one year to another were small and rather erratic. For the whole period, in each breed, no significant time-trend was found for the mean heterozygosity at protein loci or at blood group loci.

Table VI. Within-breed heterozygosity at 16 blood marker loci in offspring born in 1992 and typed.

Marker	Breed				
	Thoroughbred	Trotteur Français	Arab	Anglo-Arab	Selle Français
Proteins					
Al	0.37	0.46	0.48	0.38	0.41
A1B	0.01	0.43	0.04	0.06	0.17
Es	0.18	0.19	0.09	0.09	0.28
Gc	0.04	0.08	0.04	0.03	0.03
GPI	0.00	0.16	0.03	0.00	0.01
PGD	0.48	0.19	0.45	0.42	0.34
PGM	0.00	0.06	0.18	0.00	0.10
Pi	0.68	0.73	0.70	0.69	0.70
Tf	0.75	0.68	0.65	0.75	0.76
Mean	0.28	0.33	0.30	0.27	0.31
(se)	(0.003)	(0.004)	(0.008)	(0.009)	(0.004)
Blood groups					
A	0.39	0.68	0.60	0.51	0.61
C	0.43	0.36	0.42	0.34	0.46
D	0.77	0.83	0.73	0.64	0.81
K	0.05	0.08	0.02	0.06	0.20
P	0.36	0.51	0.42	0.47	0.46
Q	0.57	0.59	0.59	0.57	0.59
U	0.19	0.26	0.20	0.16	0.22
Mean	0.39	0.47	0.42	0.39	0.48
(se)	(0.004)	(0.003)	(0.010)	(0.012)	(0.004)
Overall mean	0.33	0.39	0.35	0.32	0.38
(se)	(0.004)	(0.004)	(0.009)	(0.001)	(0.004)

Standard errors (se) are given only for mean values accross loci; for individual loci, standard errors ranged from 0.001 to 0.006 in Thoroughbred, Trotteur Français and Selle Français and from 0.004 to 0.014 in Arab and Anglo-Arab.

Table VII shows results of the check of Hardy–Weinberg proportions for seven protein loci in offspring born from 1989 to 1992. In 11 breed × locus combinations over 35, the test was not possible due to too small expected numbers for given genotypes, even after pooling the rarest alleles. This occurred mainly in Anglo-Arab, Arab and Thoroughbred, ie, the breeds with the smallest numbers of offspring born and typed during the period considered. From the results, it is not possible to reject the hypothesis of Hardy–Weinberg proportions in Anglo-Arab and Selle Français. In other breeds, significant departures from these proportions were observed, but the situation differs from one breed to another. In Arab, the different results were more consistent; low P values were obtained for each locus where the test was possible except PGD, and for each significant case the departure consisted of an appreciable

Table VII. Test of Hardy–Weinberg proportions for protein loci.

Breed	α	Al			A1B		Gc		GPI		PGD		PGM			Tf	
		S	P	Δ	S	P	S	P	S	P	S	P	S	P	Δ	S	P
Thoroughbred	0.013	21.91	$<10^{-3}$	−2.1			0.57	0.452			0.00	0.963				4.11	0.942
Trotteur Français	0.007	3.18	0.074		0.04	0.842	3.82	0.051	1.46	0.227	3.09	0.079	10.70	0.001	+0.2	13.18	0.214
Arab	0.013	7.12	0.008	+4.4							0.92	0.336	66.75	$<10^{-3}$	+3.9	15.95	0.014
Anglo-arab	0.017	0.02	0.878								0.00	0.999				4.45	0.616
Selles Français	0.009	2.17	0.141		0.25	0.618	0.43	0.512			1.26	0.261	3.64	0.057		5.25	0.874

When the test was not possible, no results are mentioned here. α = individual risk for a global risk of 0.05 on all within-breed tests. $S = \sum (N_{obs} - N_{ex})^2 / N_{ex}$. P = probability of a χ^2 exceeding the observed value of S ($df = 1$ for all loci, except for Tf where it was equal to 10 in Thoroughbred, Trotteur Français and Selle Français and to 6 in Arab and Anglo-arab). Δ = difference between observed and expected frequencies of homozygotes, as a percentage of the expected value (given only when $P < \alpha$).

excess of homozygotes. The situation was less clear in Trotteur Français: both low and high P values were obtained and for the only significant case the departure consisted of a small excess of homozygotes. The Thoroughbred is a special case in which high to very high P values were obtained for every locus, except one where the departure (statistically significant) consisted of an excess of heterozygotes.

DISCUSSION AND CONCLUSIONS

The results show some characteristics shared by the different breeds studied. However, on the basis of many criteria, important differences between these breeds appeared, concerning their demography, their management and their genetic variability.

The generation lengths are longer than in the other main domestic species. Our results are consistent with previous results on some of the breeds studied here, Thoroughbred (Langlois, 1976), Anglo-Arab (Langlois, 1979) and Trotteur Français (Langlois, 1985) and with results obtained for other breeds, eg, Norwegian Standardbred (Klemetsdal, 1992), Finnish Trotteur (Ojala, 1982) and the review by Ojala (1982). Such large values of generation length do not have solely biological causes, especially if we consider that intervals on the paternal side are generally larger than, or of the same magnitude as, intervals on the maternal side. These values are mainly due to the principal uses of horses in these breeds. In general, for racing and riding horses, there is a clear-cut division between sporting career and breeding life. For females, racing or jumping are incompatible with pregnancy. On the other hand, for males, developing artificial insemination could contribute to a shorter average generation length and thus to increased annual rate of genetic gain, as suggested by Langlois (1976, 1985) and Klemetsdal (1992).

Considering only useful offspring, ie, offspring kept for breeding, paternal family sizes were found to be unbalanced. This picture was shown in other studies that considered all offspring, eg, for Thoroughbred (Langlois, 1976), Norwegian Standardbred (Klemetsdal, 1992) and *Boulonnaise*, a draught-horse from the north of France (Tellier et al, 1993). Clearly, the selection of stallions is made through two steps. The first step is monitored by the French Ministry of Agriculture *(Service des Haras)* and consists of choosing stallions approved for breeding. The second step results from the individual choice of the breeders and consists of a variable use of these approved stallions. The phenomenon is particularly clear in breeds with a unique or almost unique selection goal, ie, race breeds (Thoroughbred and Trotteur Français) and a riding breed which is almost exclusively used for jumping *(Selle Français)*. Of course, if most of the stallions used are also the best, this second step leads to an increase in selection intensity. This procedure, however, is not optimal. By selecting fewer stallions and equalizing their family sizes, the same selection intensity could be achieved with a higher effective number of stallions, and this would be useful for maintaining genetic variability on the long term. Unfortunately, the small size of studs is not favourable to a balanced use of stallions since many breeders hold one or two mares only and most wish to use only the stallion they consider to be the best.

The pedigree completeness level was found to be very good in this study. It was higher than in other studies on populations from other domestic species,

eg, dairy cattle (Casanova et al, 1992; Miglior and Burnside, 1995), beef cattle (Giraudeau et al, 1991; Cañón et al, 1994), meat sheep (Djellali et al, 1994) and dairy sheep (Barillet et al, 1989). An almost as good or better knowledge of pedigrees was observed in other studies on horse breeds (references given below). This fact illustrates the antecedence in many countries of pedigree recording in horse breeding. As it is well illustrated for the horse by MacCluer et al (1983), computed values of inbreeding coefficients depend on pedigree completeness, a criterion which shows some differences between the five breeds studied here. Comparison of average coefficients with results from other studies should be done with care because pedigree completeness levels are sometimes different (when mentioned). Moreover, due to computing constraints, in several cases only a small sample of animals was considered, and approximate methods were often used, such as random sample of two-line pedigree. However, when recent studies are considered, in which pedigree completeness level was roughly of the same order as in our study, are considered two groups of breeds may be distinguished, according to their computed value of average coefficients of inbreeding. On the one hand, computed values of the same magnitude as in our study, ranging from 0.81 to 2.89%, were generally found in race breeds, which are often international, eg, Thoroughbred from the British Isles (Mahon and Cunningham, 1982), Italy (Galizzi Vechhiotti, 1977) or Poland (Kownacki and Jezierski, 1976), Swedish Standardbred (Ström, 1982), North-Swedish Trotter (Bohlin and Rönningen, 1975) and German Trotter (Fehlings et al, 1983). The only exception seems to be Norwegian Standardbred, which is managed with a few breeding animals and where an average coefficient of 5.8% was computed (Klemetsdal, 1992). On the other hand, higher values, ranging from 2.25 to 14.7%, were found in draught breeds or in other breeds with small population size, eg, Süddeutsches Kaltblut and German Haflinger (Fehlings et al, 1983), Italian Haflinger (Gandini et al, 1992), Italian Lipizzan (Galizzi Vecchiotti et al, 1989), Jutland (Johansen, 1985) and Boulonnaise (Tellier et al, 1993). However, in race breeds, when pedigrees were much more complete, higher average values were computed: 8.9% in American Standardbred (MacCluer et al, 1983) and 12.5% in a sample of 59 Thoroughbred mares by tracing four-line pedigrees up to the foundation of the breed (Mahon and Cunningham, 1982). Undoubtedly, if pedigrees were more complete, computed coefficients of inbreeding in our study would be higher.

On the basis of many of the results of this study, causes of inbreeding and change in genetic variability show different pictures in the five breeds. Here, pure breeds (Thoroughbred, Trotteur Français and Arab) and composite breeds (Anglo-Arab and Selle Français) should be considered separately. Within offspring born from 1989 to 1992, Thoroughbred and Trotteur Français showed the same almost null proportion of non-inbred animals, indicating the general relationship between the stallions and mares mated. However, from the detailed analysis of inbreeding (see figs 4 and 5), it seems that mating between close relatives is exceptional in these two breeds, a result observed in other Thoroughbred populations (Dusek, 1966; Gallazi Vecchiotti, 1977; Mahon and Cunningham, 1982). A check of Hardy–Weinberg proportions for blood protein loci in these two breeds (see table VII) provided results that were not clear enough to invalidate this conclusion. Inbreeding in Thoroughbred and *Trotteur Français* seems to originate from both a bottleneck

(at least at the foundation of the breeds) and unbalanced paternal family sizes. Remote inbreeding is particularly important in Thoroughbred, as shown by both analysis of the actual inbreeding according to the number of generations of ancestors considered (see fig 5) and the difference in the annual rate of inbreeding when considering either the whole pedigree or four generations of ancestors only (see fig 6). In comparison with Thoroughbred, the founder effect was particularly stringent in Trotteur Français, as shown by an analysis of gene origin. This fact explains the higher average coefficient of inbreeding, despite less complete pedigrees, and the higher annual rate of inbreeding, despite a considerable increase in the number of breeding animals. The third pure breed, Arab, shows a specific picture. The average coefficient of inbreeding was substantially higher in Arab than in all the other breeds, whereas its effective number of founders was equal to twice the effective number in Trotteur Français; paternal family sizes were found to be the most balanced and a high proportion of non-inbred animals was observed. The high proportion of offspring with a null computed inbreeding coefficient is due to the widespread use of stallions from many countries, assumed to be unrelated. On the other hand, the high proportion of coefficients higher than 6.25% and the large part of close inbreeding in the total inbreeding (see fig 5) show evidence for mating between close relatives at a substantial level in Arab. This conclusion is well supported by analysis of genotypic frequencies for blood protein loci (see table VII) clearly showing a significant excess of homozygotes in Arab. The two composite breeds, Anglo-Arab and Selle Français, were created from a large genetic basis. Their foundation involved animals from different, and probably barely related, breeds including mainly Thoroughbred, which represents at least 50% of their present gene pool (Moureaux et al, 1995). The basis was particularly large in Selle Français, and in the management rules of the breed, many crosses (such as Trotteur Français × Selle Français, Thoroughbred × Selle Français, etc) are allowed. Therefore, despite unbalanced paternal family sizes, the present average coefficient of inbreeding is low in Selle Français. In Anglo-Arab, crosses (such as Thoroughbred × Arab, Anglo-Arab × Arab, etc) are allowed too. However, Anglo-Arab was found to be more inbred than Selle Français, probably due to a smaller effective number of founders. It looks as if Anglo-Arab breeders were less interested in using breeding animals from origins other than their Thoroughbred or Arab foundation stocks.

Inbreeding depression has been studied in horses with various results. The effect of inbreeding on fertility in race breeds was shown to be significant (Klemetsdal and Johnson, 1989) or not (Mahon and Cunningham, 1982; Cothran et al, 1984). The effect on body measurements was established in Italian Haflinger by Gandini et al (1992); the effect on race winnings was studied only on very small samples in Thoroughbred and was found to be not significant (Dusek, 1966; Galizzi Vecchiotti, 1977). Further research is needed to analyse the effect of inbreeding on various performance traits in the five breeds studied here. From a practical point of view, this would be particularly useful in Arab, to assess consequences of the current practice of mating between close relatives. For mid-term purposes, breeders should pay more attention to the balance of paternal family sizes, in order to maintain genetic variability. This would be particularly necessary for Trotteur Français, the gene pool of which originated from a very small effective number of founders.

Results from blood marker data analysis are consistent with an earlier study by Guérin and Mériaux (1986) on three of the five breeds studied here, Thoroughbred, Trotteur Français and Selle Français. Using the seven blood group loci and four protein loci out of the nine studied here, these authors found values of heterozygosity very similar to ours in all breeds and at all loci. The average heterozygosity in Arab estimated in our study is of the same magnitude as the value computed for the Arab breed in Morocco on the basis of exactly the same markers (Ouragh et al, 1994). In comparison with the five breeds studied here, significantly higher values of average heterozygosity were computed in some American breeds on the basis of the seven blood group loci and seven protein loci only (Bowling and Clark, 1985). These differences could be partly explained by the fact that protein loci are less polymorphic than blood group loci, as shown in table VI. Including five more protein loci (total = 12), Bowling (1994) found values of 0.30 ± 0.06, 0.41 ± 0.05, 0.35 ± 0.06 and 0.39 ± 0.06 in American Thoroughbred, Standardbred, Arab and Saddle horse, respectively. These values are of the same magnitude as the values computed in this study. Notably, there is no significant difference for average heterozygosity between French and American Throughbred or between French and American Arab.

The links between results from blood marker data and results for inbreeding and gene origin are rather complex. The increase in inbreeding and the decrease in heterozygosity are generally assumed to be parallel. If we refer to the recent evolution of the breeds, it should be kept in mind that the period analysed (offspring born from 1974 to 1992) represents around 1.7 equine generations only. In each breed during this period, the annual rate of inbreeding was found to be statistically significant (see fig 6). In an ideal population and for a neutral locus (Wright, 1931), the total increase in inbreeding should correspond to an equivalent decrease in heterozygosity. Remembering that annual rates of inbreeding in figure 6 are expressed in percentage points per year, the corresponding total decrease in heterozygosity should be around 0.005 to 0.013 according to the breed considered, ie, values of the same magnitude as the actual standard error of the mean heterozygosity. Moreover, due to substantially smaller numbers of typed animals, in each breed from 1974 to the late eighties, the standard error of the mean heterozygosity was equal to approximately twice the value computed in 1992. Such a situation may largely explain why the observed changes in heterozygosity were not found to be statistically significant for the short period analysed. If we refer to the current status of the breeds and their history, the case of Selle Français, one of the two most polymorphic breeds and the least inbred, could be considered as consistent with the hypothesis of parallelism between increase in inbreeding and decrease in heterozygosity and partly due to the impact of crossing. On the other hand, Trotteur Français was found to be as polymorphic as Selle Français, whereas, among the five breeds, its effective number of founders was the smallest and its average coefficient of inbreeding was the second highest. This situation may be explained by the short history of Trotteur Français, the studbook of which was created in 1922 and closed in 1937, ie, respectively 6 and 4.5 equine generations ago (see table III). Assuming a broad genetic basis, polymorphism could remain at a substantial level, despite an appreciable decrease in gene origins. Thoroughbred shows an opposite pattern. Possibly due to an earlier foundation and studbook closure and an accumulation of remote inbreeding, this breed is now less polymorphic. In fact, inbreeding may be

used as an indicator of genetic variability but not in too rigorous a way, because it is defined for neutral loci and its computation does not take into account mutations and selection. Moreover, the computed values depend largely on the extent of known pedigree. If pedigree completeness is low, the true probability of gene identity is clearly underestimated. If pedigrees are traced back to distant generations of ancestors, hypotheses assumed for computing inbreeding coefficients are probably not valid. It would be very interesting to use more markers to get more comprehensive and more accurate results about links between biological polymorphism and indicators of genetic variability derived from genealogical data. Further research is needed on this topic, and molecular markers should be very useful in this respect. Due to high pedigree completeness level and high rate of parentage control in many domestic breeds, the horse should be a species of choice for this kind of study.

ACKNOWLEDGMENTS

REFERENCES

Barillet F, Vu Tien Khang J, Roussely M, Poivey JP, Chevalet C, Elsen JM, de Rochambeau H (1989) Variabilité génétique dans le noyau de sélection des ovins laitiers de race Lacaune: une étude rétrospective. In: *La gestion des ressources génétiques des espèces animales domestiques*, Lavoisier, Paris, 71-80

Bohlin O, Rönningen K (1975) Inbreeding and relationship within the north Swedish horse. *Acta Agric Scand* 25, 121-124

Bowling AT (1994) Population genetics of Great Basin feral horses. *Anim Genet* 25, 67-74

Bowling AT, Clark RS (1985) Blood groups and protein polymorphism gene frequencies for seven breeds of horse in the United States. *Anim Blood Groups & Biochem Genet* 16, 93-108

Cañón J, Guittiérez JP, Dunner S, Goayche F, Vallejo M (1994) Herdbook analyses of the Asturiana beef cattle breeds. *Genet Sel Evol* 26, 65-75

Casanova L, Hagger C, Kuenzi N, Schneeberger M (1992) Inbreeding in Swiss Braunvieh and its influence on breeding values predicted from a repeatability animal model. *J Dairy Sci* 75, 1119-1126

Cothran EG, MacCluer JW, Weitkamp LR, Pfenning DW, Boyce AJ (1984) Inbreeding and reproductive performance in Standardbred horses. *J Hered* 75, 220-224

Djellali A, Vu Tien Khang J, de Rochambeau H, Verrier E (1994) Bilan génétique des programmes de conservation des races ovines Solognote et Mérinos Précoce. *Genet Sel Evol* 26, Suppl 1, 255-265

Dusek J (1966) On the problem of inbreeding in the breeding of full-blooded horses with regard to the performance obtained. *Zivocisna Vyroba* 34, 121-130 (in Czech)

Fehlings R, Grundler C, Wauer A, Pirchner F (1983) Inzucht-und Vermdtschafts-verhältnisse in Bayerischen Pferde-rassen (Haflinger, Süddeutsches Kaltblut, Traber). *Z Tierz Zuchtungsbiol* 100, 81-86

Galizzi Vecchiotti G (1977) Consanguinité et performances chez les Pur-Sang : étude sur les étalons employés en Italie en 1971 et 1972. *Rivista Zootec Veter* 5, 560-566

Galizzi Vecchiotti G, Silvestrelli M, Pieramati C, Pelligra S (1989) Inbreeding and fertility in Lipizzan breed raised in Italy. In: *Societa Italiana di Ippologia, Atti IX Congresso Nationale, 217-221 (in Italian)*

Gandini GC, Bagnato A, Miglior F, Pagnacco G (1992) Inbreeding in the Italian Haflinger horse. *J Anim Breed Genet* 109, 433-443

Giraudeau L, Verrier E, de Rochambeau H, Ménissier F, Laloë D, Casañe D, Poupinot JP, Bougler J, Vu Tien Khang J (1991) Survey and management of the genetic variability of a cattle breed of small population size: a successful program in the Parthenaise breed. In: *42nd Annual Meeting EAAP, Berlin, September 9-12*, 11 p

Guérin G, Mériaux JC (1986) La distribution des marqueurs sanguins dans les races équines: analyse sur un échantillon de Pur-Sang, Trotteur Français et Selle Français. In: *CEREOPA, 12ᵉ journée d'étude, Paris, 12 March 1986*, 2-13

Hedrick PW (1985) *Genetics of Populations.* Jones and Bartlett Publishers, Boston

Johansen EO (1985) Inbreeding and relationship within the national Danish draught-horse: the Jutland breed. In: *35th Annual Meeting EAAP, The Hague, The Netherlands*, Comm H 3b-4, 5 p

Klemetsdal G (1992) Demographic parameters and inbreeding in the Norwegian Trotter. *Acta Agric Scand* 43, 1-8

Klemetsdal G, Johnson M (1989) Effect of inbreeding on fertility in Norwegian Trotter. *Livest Prod Sci* 21, 263-272

Kownacki M, Jezierski T (1976) Inzuchtverhältnisse in der Zucht des englischen Voll-blutpferdes der Vr Polen und Beziehungen zur Leistung. In: *International Wissenchaft-liches Symposium. Leipzig, Deutschland, 13-14 July 1976*, 177-180

Langlois B (1976) Estimation de quelques paramètres démographiques du Pur-Sang anglais en France. *Ann Génét Sél* Anim 8, 315-329

Langlois B (1979) Zuchtmethodische Grundsätze bei der Züchtung des Angloarabers in Frankreich. In: *International Wissenchaftliches Symposium, Leipzig, Deutschland, 22-23 May 1979*, 153-159

Langlois B (1985) Effets de la sélection sur les données de performances chez le Trotteur Français. In: *CEREOPA, 11ᵉ journée d'étude, Paris, 6 March 1985*, 38-51

MacCluer JW, Boyce AJ, Dyke B, Weitkamp LR, Pfenning DW, Parsons CJ (1983) Inbreeding and pedigree structure in Standardbred horses. *J Hered* 74, 394-399

Mahon GAT, Cunningham EP (1982) Inbreeding and the inheritance of fertility in Thoroughbred mare. *Livest Prod Sci* 9, 743-754

Miglior F, Burnside EB (1995) Inbreeding of Canadian Hostein cattle. *J Dairy Sci* 78, 1163-1167

Moureaux S, Ricard A, Mériaux JC, Verrier E (1995) Caractérisation génétique des races françaises de sport et de course et analyse de leur variabilité génétique. In: *Institut du cheval, 21ᵉ journée d'étude, Paris, 1ᵉʳ mars 1995*, 100-105

Ojala MJ (1982) Some parameters estimated from a restricted set of race records in Trotters. *Acta Agric Scand* 32, 215-224

Ouragh L, Mériaux JC, Braun JP (1994) Genetic blood markers in Arabian, Barb and Arab-Barb horses in Morocco. *Anim Genet* 25, 45-47

Quaas RL (1976) Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32, 949-953

Ström H (1982) Changes in inbreeding and relationship within the Swedish Standardbred Trotter. *Z Tierz Zuchtungsbiol* 99, 55-58

Tellier X, Lefaudeux F, Tavernier A, Verrier E, Stievenard R, Pourchet B (1993) Analyse de la structure génétique de la race chevaline Boulonnaise et programme mis en place pour préserver sa variabilité. In: *CEREOPA, 19ᵉ journée d'étude, Paris, 3 mars 1993*, 130-139

Wright S (1931) Evolution in Mendelian populations. *Genetics* 16, 97-159