

Genetic Variability of French Dairy Breeds Estimated From Pedigree Information

L. Maignel^{1,2}, D. Boichard¹, and E. Verrier^{1,2}

¹Institut National de la Recherche Agronomique, Station de Génétique Quantitative et Appliquée, 78 352 Jouy-en-Josas Cedex, France

²Institut National Agronomique Paris-Grignon, Département des Sciences Animales, 16 Rue ClaudeBernard, 75 231 Paris Cedex 05

SUMMARY

Pedigree information has been analyzed in the eight major French dairy cattle breeds (Holstein, Normande, Montbéliarde, Brown Swiss, Abondance, Tarentaise, Simmental, and Pie Rouge des Plaines) to estimate their genetic variability. Pedigree files included from 48,320 animals in Tarentaise to 11,180,346 animals in Holstein. Pedigree information was very limited before 1965, and reached 80-92% known parents for recent animals, according to the breed. The effective size (N_e) of the population was estimated from inbreeding change in the last generation. It was found to be very small in Holstein (46), Normande (47), Tarentaise (27) and Brown Swiss (56). Estimates were larger in the other breeds (106 to 2500) but probably meaningless because of either very incomplete pedigree or recent foreign gene infusion. The effective number of ancestors (N_a), derived from probabilities of gene origin, was found to be a more robust alternative approach. N_a was found to be very small in all breeds, from 17 to 64, and particularly in Holstein (43). As a conclusion, dairy breeds are genetically very small populations, because of the large impact of artificial insemination and short-term oriented selection procedures.

Introduction

Selection generally provides some genetic gain at the expense of the long term genetic variability. However, this problem is potentially more acute in dairy cattle, because of several reasons : (1) populations include a large number of reproductive cows and give the impression that their effective size should be high; (2) as the generation interval is long, the short-term genetic gain is expressed only after 8-10 years, and the long-term is too far to be efficiently integrated in the selection policy ; (3) artificial insemination is much more efficient in dairy cattle than in any other species, making it possible to concentrate the selection effort on very few key animals ; (4) as selection in dairy cattle is very competitive

across breeds, breeding organizations, and countries, breeders put emphasis on short-term results conditioning their survival ; (5) in the last 30 years, selection objective has remained quite simple and stable across herds, breeds, and even countries.

Genetic variability and its evolution over time may be estimated from pedigree information. The trend in inbreeding is the method the most frequently used to quantify the rate of genetic drift. This method relies on the relationship between the increase in inbreeding and decrease in heterozygosity for a given locus in a closed, unselected, and panmictic population of finite size. However, in cattle populations, the efficiency of this approach may be limited because of several reasons. First of all, the relationship between

the rate of inbreeding and the decrease in heterozygosity is only asymptotic. The cattle populations are managed with overlapping generations, and the breeding strategies have been strongly modified over the last 25 years. Therefore, it is very unlikely that these populations are currently under asymptotic conditions. Secondly, for a given generation, the value of the average coefficient of inbreeding may reflect not only the cumulated effects of genetic drift but also the effect of the mating system, which is rarely panmictic. Thirdly, the computation of the individual coefficient of inbreeding is very sensitive to the quality of the available pedigree information. In most practical situations, some information is missing, even for the most recent generations of ancestors, leading to large biases when estimating the rate of inbreeding.

An alternative approach is to analyze the probabilities of gene origin (James, 1972). In this method, the genetic contributions of the founders, i.e. the ancestors with unknown parents, to the current population are measured. Although the definition of a founder is also very dependent on the pedigree information, this method assesses how an original gene pool has been maintained across generations. As proposed by Lacy (1989), these founder contributions could be combined to derive the 'founder equivalents', i.e. the number of equally contributing founders that would be expected to produce the same level of genetic diversity as in the population under study. Boichard et al. (1995, 1996) adapted Lacy's approach to account for the bottlenecks in the pedigree and to derive an effective number of ancestors. MacCluer et al. (1986) and Lacy (1989) also proposed to estimate the 'founder genome equivalent', i.e. the number of equally contributing founders with no random loss of founder alleles in the offsprings, that would be expected to produce the same genetic diversity as in the population under study.

The purpose of this paper is to estimate the genetic variability of the French dairy cattle breeds, on the basis of their pedigree information, using both inbreeding and probabilities of gene origin approaches.

Material and Methods

Populations

Eight populations were analyzed, i.e. the eight main dairy breeds raised in France: Holstein, Normande, Montbéliarde, Brown Swiss, Abondance, Tarentaise, Simmental, and Pie Rouge des Plaines (PRP). Each data set included every female with a recorded lactation beginning between September 1975 and January 1995, and all their known ancestors. Data are described in table 1.

Pedigree information is described in table 2 and in figure 1. It was very limited for animals born before 1966, i.e. before the implementation of the present identification system, and reached 80-95% for females born in 1988-1991, according to the breed. The best criterion to characterize the pedigree information is probably the number of complete generation-equivalent, defined as the sum of the proportion of known ancestors over all generations traced. It varied from 3.8 to 5, except in Pie Rouge des Plaines and in Simmental. For these two breeds, which extensively use foreign genes, pedigree information was not automatically retrieved into the French data files.

Methods

Inbreeding was computed according to VanRaden (1992). This method is less efficient than that of Meuwissen & Luo (1992), but it is more flexible and makes it possible to consider some non zero relationship between unknown ancestors and then, to some extent, to account for missing pedigree. Inbreeding coefficients were computed twice. In the first run, unknown ancestors were assumed to be unrelated, whereas in the second run, they were assumed to be related as in Wiggans et al. (1995). The **effective size** (N_e) was derived from the trend in inbreeding (ΔF) during the last generation.

$$\Delta F / (1-F) = 1 / (2N_e)$$

Three other parameters were derived from

probabilities of gene origin. The first one (f_e) is the **effective number of founders** (Rochambeau et al., 1989) or the 'founder equivalent' (Lacy, 1989), i.e. the number of equally contributing founders that would be expected to produce the same genetic diversity as in the population under study. This criterion is computed from the contributions q_k of all f founders of the population under study.

$$f_e = 1 / \sum_{k=1}^f q_k^2$$

The more balanced the contributions of the founders, the higher the effective number of founders. The population considered here included all females born from 1988 to 1991.

An important limitation of the previous approach is that it ignores the bottlenecks in the pedigree. This overestimation is particularly strong in very intensive selection programmes, when the germplasm of a limited number of breeding animals is widely spread, as in dairy cattle. Boichard et al (1995, 1996) proposed a method to identify the major ancestors (potentially not founders) and to determine their marginal contributions to the gene pool of the reference population. Formula [1] could be applied to these marginal contributions (p_k) to determine the **effective number of ancestors** (f_a).

$$f_a = 1 / \sum_{k=1}^f p_k^2$$

However, this approach still underestimate the probability of gene loss by drift from the ancestors to the population under study. As a result, the effective number of ancestors may be overestimated. A third method is to analyze the probability that a given gene present in the founders, i.e. a "founder gene", is still present in the population under study. However, in a complex pedigree, an analytical derivation may be not feasible. MacCluer et al. (1986) proposed to use Monte Carlo simulation to estimate the probability of a founder gene remaining present in the population under study. At a given locus, $2f$

founder genes are generated. Then the segregation is simulated throughout the complete pedigree and the genotype of each progeny is generated by randomly sampling one allele from each parent. Gene frequencies f_k are determined by gene counting in the population under study. The effective number of founder genes N_a in the population under study is obtained as an effective number of alleles (Crow and Kimura, 1970):

$$N_a = 1 / \sum_{k=1}^{2f} f_k^2$$

As a founder carries two genes, the **effective number of founders genomes** (called 'founder genome equivalent' by Lacy, 1989) still present in the population under study (N_g) is simply half the effective number of founder genes.

$$N_g = \frac{N_a}{2} = 1/2 \sum_{k=1}^{2f} f_k^2$$

Results and discussion

Inbreeding trends and corresponding effective size are presented in Table 3. In the last generation, the increase in inbreeding varied from 0.40 to 1.83%, according to the breed, except in PRP and Simmental, two breeds with a limited pedigree knowledge. Inbreeding trend was highest in the Tarentaise breed, which is the smallest population. It was close to 1% per generation in Holstein, Normande, and Brown Swiss breeds, indicating an effective size close to 50. The lower trend observed in Abondance and Montbéliarde could be attributed to the recent and limited infusion of Red Holstein genes. More generally, crossbreeding and incomplete pedigree are two important factors responsible for a strong underestimation of inbreeding, limiting the meaning of the concept of effective size in such a population (Boichard et al., 1996).

Probabilities of gene origin show that very few ancestors contribute the major part of the gene pool of the population under study

(Table 4). Accordingly, the three criteria derived from the probabilities of gene origin provided very low results (Table 5).

Results were not very different across breeds and did not reflect at all the population size (Table 2). They were very low, indicating that all French dairy breeds could be considered as genetically small populations.

The effective number of founders (f_e) measures how the balance in founder contributions is maintained across generations. It accounts for selection rate (i.e. the probability of being a parent or not) and for the variation in family size, but it neglects the probability of gene loss from parent to progeny. The effective number of ancestors (f_a) accounts for bottlenecks in the pedigree, which is the major cause of gene loss in the dairy populations. Consequently f_a is always less than or equal to f_e . Finally, the effective number of founder genomes (N_g) measures how many founder genes have been maintained in the population for a given locus, and how balanced their frequency is. It accounts for all causes of gene loss during segregations and, consequently, provides a smaller number than f_a and f_e . For a detailed discussion, see Boichard et al. (1996).

Conclusion

Whereas effective size reflects very long term trends, parameters derived from probability of gene origin are very useful for describing a population structure after a small number of generations. They can detect recent changes in breeding strategy, before their consequences appear in term of inbreeding. From that point of view, they are very well suited to dairy cattle populations. Moreover, these parameters are less affected than effective size by incomplete pedigree information.

The main conclusion of this study is that, from a genetic point of view, dairy cattle are very small populations. Selection of very few bull sires and wide spreading of elite germplasm concentrate gene origins on a few families. Moreover, selection of bull dams

mostly on pedigree information and Animal Model-BLUP evaluation enhance this phenomenon. As a consequence, the maintenance of genetic variability in these intensively selected populations may be questioned. Several theoretical studies showed the interest for the middle run of some changes on the weight given to family information in selection criteria or on management policies, in comparison of current practices (see for example Verrier et al., 1993). Extensive studies on realistic situations are required to check some simple modifications of current selection programmes which could lead to both genetic gains still compatible with short-term objectives and maintenance of genetic variability on the middle run.

References

- Boichard, D., Maignel, L. and Verrier, E. 1995. Estimation of the effective number of founders of a population from pedigree information. 2nd European Workshop in Advanced Biometrical Methods in Animal Breeding, Salzburg, Austria, June 12-20 1995, 7p.
- Boichard, D., Maignel, L. and Verrier, E. 1996. Interest of probabilities of gene origin to measure the genetic variability in a population. *Genet. Sel. Evol.* (submitted).
- Crow, J.F. and Kimura, M. 1970. An introduction to population genetics theory. Harper & Row, New York, USA.
- James, J. 1972. Computation of genetic contributions from pedigrees. *Theor. Appl. Genet.* 42, 272-273.
- Lacy, R.C. 1989. Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo Biol.* 8, 111-123.
- MacCluer, J.W., Van de Berg, J.L., Read, B. and Ryder, O.A. 1986. Pedigree analysis by computer simulation. *Zoo Biol.* 5, 147-160.
- Meuwissen, T.H.E. and Luo, Z. 1992. Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* 24, 305-313.

Table 4. Contribution to the gene pool of the major ancestors

Breed	Most important ancestor (%)	5 most important ancestors (%)	Nr ancestors explaining 50% of the genes
Abondance	9.60	40.85	8
PRP	5.87	20.792	3
Brown Swiss	11.89	34.89	11
Tarentaise	15.92	44.60	7
Simmental	10.59	33.09	14
Montbéliarde	8.29	25.04	20
Normande	9.12	28.84	17
Holstein	9.81	28.84	29

Table 5. Criteria derived from probabilities of gene origin (see text for a detailed definition)

Breed	f_e	f_a	N_g (1)
Abondance	69	25	17.3
PRP	96	64	49.6
Brown Swiss	84	28	19.3
Tarentaise	50	17	15.2
Simmental	75	33	24.4
Montbéliarde	146	63	35.8
Normande	132	40	21.6
Holstein	140	43	30.4

(1) mean of 100 replicates, standard deviation = 1.3 to 3.9 according to breed

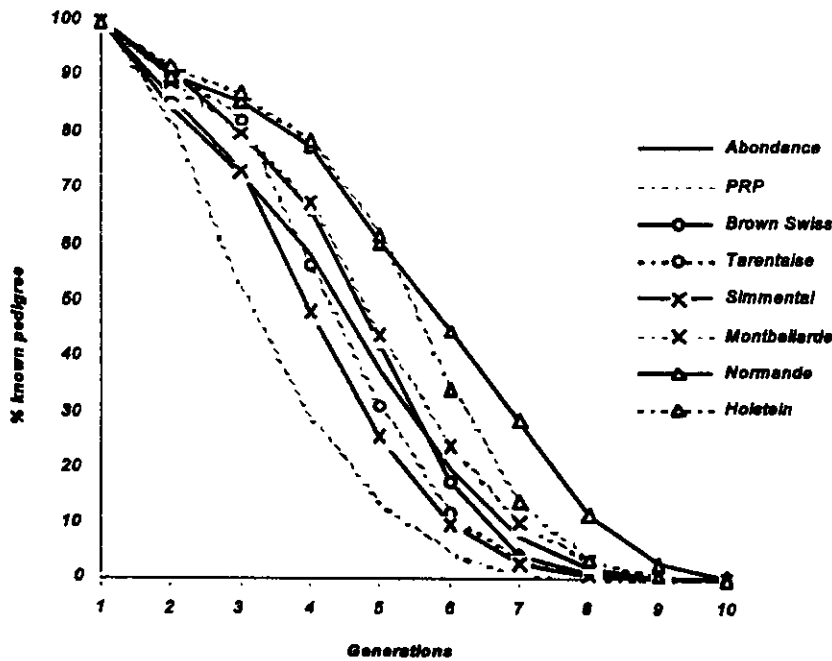


Figure 1. Proportion of known ancestors per generation and breed.

Rochambeau, H. de, La Fuente, L.F. de, Rouvier, R. and Ouhayoun, J. 1989. Sélection sur la vitesse de croissance post-sevrage chez le lapin. *Genet. Sel. Evol.* 21, 527-546.

VanRaden, P.M. 1992. Accounting for inbreeding and crossbreeding in genetic evaluation for large populations. *J. Dairy Sci.* 75, 3136-3144.

Verrier, E., Colleau, J.J. and Foulley, J.L. 1993. Long term effects of selection based on the Animal Model BLUP. *Theor. Appl. Genet.* 87, 446-454.

Wiggans, G.R., VanRaden, P.M. and Zuurbier, J. 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United States dairy cattle. *J. Dairy Sci.* 78, 1584-1590.

Table 1. Description of the data sets

Breed	Total Nr of animals	Total Nr of founders	Nr of Females born in 1988-1991
Abondance	106,520	6,109	12,345
PRP	102,270	5,442	12,502
Brown Swiss	91,579	6,291	13,856
Tarentaise	48,320	2,942	5,817
Simmental	121,239	7,604	13,947
Montbéliarde	1,664,506	122,981	331,600
Normande	2,338,305	138,291	358,900
Holstein	11,180,346	802,289	2,316,371

Table 2. Characterization of the pedigree information of females born in 1988-1991

Breed	Average Nr of ancestors	Maximum Nr of generations traced	Nr of complete generation-equivalent traced
Abondance	62.4	13	3.83
PRP	22.9	11	2.82
Brown Swiss	59.3	18	4.04
Tarentaise	45.8	12	3.78
Simmental	35.6	13	3.47
Montbéliarde	79.4	14	4.19
Normande	159.3	16	5.02
Holstein	98.5	17	4.75

Table 3. Inbreeding trends in the last generation and effective size of the 8 breeds

Breed	Inbreeding trend (%/generation)	Effective size (N_e)
Abondance	0.47	106
PRP	0.02(1)	2500
Brown Swiss	0.90	56
Tarentaise	1.83	27
Simmental	0.24	208
Montbéliarde	0.40	125
Normande	1.07	47
Holstein	1.09	46

(1)result probably meaningless because of very incomplete pedigree information.

95:1 and the test run 95:1* were essentially zero at 0.1, 0.01 and 0.00 for milk, fat and protein yield respectively. The standard deviations of the differences were very small at 8 kg milk, 0.32 kg fat and 0.22 kg protein.

A closer examination of individual bulls showed that 41 bulls had a significant change based on size of change and reliability in their evaluation as a result of the introduction of the management group structure.

A total of 15 bulls had a change in milk Predicted Transmitting Ability (PTA) of greater than 100 kg. Seven of these bulls were UK registered, which benefited due to reduced competition from imported contemporaries.

Table 1 shows the effect on several high profile foreign bulls. The largest reduction was BLACKSTAR at -137 kg milk, followed by MICHAEL at -118 kg milk. SUNNY BOY and SEXATION are examples of bulls which showed an increase in Predicted Transmitting Ability (PTA). The changes resulted in a re-ranking of a few bulls in the top list. BLACKSTAR's ranking for the UK PIN value changed from 6th (PIN=94) to 23rd (PIN=79). MICHAEL changed from 3rd (PIN=100) to 11th (PIN=85). The revised evaluations for these bulls were closer to their converted UK PTA equivalents. The converted and test PTAs for BLACKSTAR were as follows:

BLACK- STAR	REL %	Milk kg	Fat kg	Protein kg
Official (95:1)	95	1020	24.2	27.2
Converted	84	799	17.6	20.5
Test PTA	95	883	21.1	23.0

The bulls with large reductions in their PTAs in the test run showed a large reduction in average management and yield deviations for imported daughters compared with the official run (Table 2). This indicated that the mean performance of the imported daughters within herds for these bulls were higher than home bred daughters and the separate grouping removed any advantage exclusive to imports as a result of preferential treatment. However the bulls with increases in their PTAs, showed little reduction in the average

management and yield deviations for imported daughters in the test run compared with the official run. In addition there was an increase in management and yield deviations for home bred daughters of these bulls accounting for the increase in their proofs.

Comparison of January Official (95:1) and July Official (95:2) Evaluations

The average difference for matching bulls between the official runs of 95:1 and 95:2, which included the new management grouping procedure were small at -1 kg milk, -0.3 kg fat and -0.2 kg protein with standard deviations of 20, 0.9 and 0.6 respectively. This was similar to the difference reported for the test run.

Table 3 shows the average differences for bulls between the 95:2 PTAs and 95:1 PTAs, broken down by country of origin. The figures, in brackets, indicate the standard deviation of the difference. The analysis includes all bulls with reliability greater than 80%. The higher restriction was introduced to eliminate bulls with a normal and expected increase in daughters and therefore, potential large change in evaluation.

In practice, there was little overall change for UK registered Holstein Friesians. In general, for most foreign imports, most changes were within ± 100 kg milk. The main interest was the greater changes in bulls from the USA with five having a change in excess of 180 kg. At the same time, five other USA bulls had changes in the category 140-180 kg milk. Overall, the changes were very much in line with normal expectation except for a greater effect on the USA than any other country. This was not unexpected since the majority of imports have come from the USA.

In order to look at the effect on cows, a comparison was made of cows with a foreign dam on both the 95:2 and 95:1 files. A total of 8,767 cows matched when the two files were compared. The overall mean change was small at -7.5 kg milk, -0.8 kg fat, -0.03 kg protein, -0.01 fat % and 0.00 protein %. The standard deviation of the difference was considerably wider than normally found. For