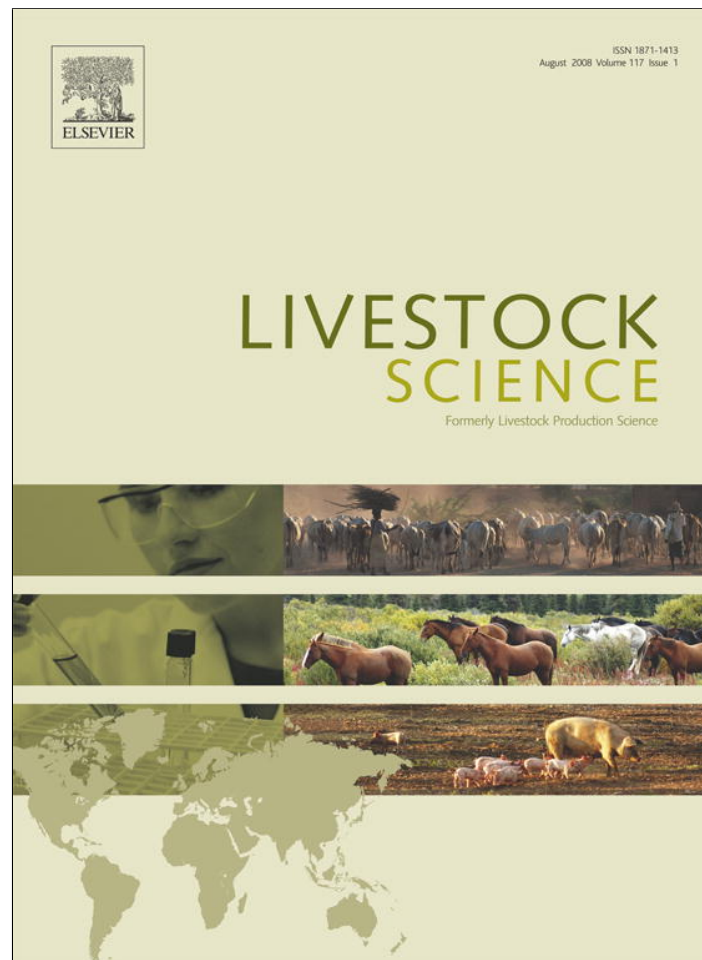


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Estimation of effective population size using bivariate discrete distributions for modeling family size in beef cattle

L. Ron Garrido^a, A.N. Birchmeier^b, S. Munilla^b, R.J.C. Cantet^{b,c,*}

^a *Facultad de Ingeniería en Biotecnología, Escuela Politécnica del Ejército, Ecuador*

^b *Departamento de Producción animal, Facultad de Agronomía, Universidad de Buenos Aires, Argentina*

^c *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina*

Received 3 January 2007; received in revised form 5 November 2007; accepted 8 November 2007

Abstract

Pedigree records of 72,808 animals (45,668 females and 27,140 males) from the genetic evaluation program of the Argentine Brangus Association were used to estimate effective number of founders (N_f), effective number of ancestors (N_a), and effective population size under random mating (N_e) or selection (N_{eS}), in order to assess genetic variability. The average level of completeness of the pedigree was low (0.17) and the average level of inbreeding (F) calculated from the pedigree was equal to 0.24%. Animals in the reference population were 21,662 calves born from 2001 to 2005. The estimated measures of variability were $N_f=765.7$ and $N_a=387.5$. The numbers of ancestors responsible for 100%, 50%, or 20% of the genes in the reference group, were equal to 12,471, 273, and 22, respectively. Direct estimates of N_e and N_{eS} were calculated using the variances and covariances of family sizes, i.e. male and female progeny numbers for bulls and cows. Estimates of the dispersion parameters were from the Bivariate Poisson model for the cows, and from the Generalized Bivariate Negative Binomial (GBIVARNB) distribution for the bulls. The latter probability mass function accounted for overdispersion, a characteristic present in the sampling distribution of family size of bulls. The estimated variances of male and female progeny and the covariance between them for the bulls were 5.70, 271.28, and 30.15, respectively, and 1.15, 2.10, and 1.06 for the cows. Generation intervals (in years) were: sires of bulls=5.0, sires of cows=5.7, dams of bulls=4.4, and dams of cows=5.2. The estimated N_e was 274, which corresponds to a rate of inbreeding (F) of 0.18%, whereas $N_{eS}=125$ and $F=0.40\%$. As a check of the proposed methodology, all analyses were also performed using the pedigree records of 10,483 Angus animals from a herd with an average level of completeness of 0.68. Using the GBIVARNB model for both bulls and cows the estimated $N_e=95.4$, thus $F=0.5\%$ in perfect agreement with the calculated average inbreeding from pedigree records. Under selection, $N_{eS}=79.3$ and $F=0.6\%$. The larger difference between estimated N_e and N_{eS} in the Brangus was related to the smaller bull to cow ratio in the breed. Therefore, it seems desirable to continue monitoring the effective size of the Argentine Brangus to prevent problems of inbreeding and lack of variability in the future.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Brangus; Effective population size; Family size; Bivariate discrete distributions; Overdispersion; selection

1. Introduction

In order for a composite breed not to dissipate the initial advantage of increased heterozygosity by becoming inbred, it is essential that heterozygosity (heterosis)

* Corresponding author. Av. San Martín 4453, C1417DSE, Buenos Aires, Argentina. Tel.: +54 11 4524 8000x8184; fax: +54 11 4514 8737.

E-mail address: rcantet@agro.uba.ar (R.J.C. Cantet).

be retained by maintaining an effective population size (N_e) sufficiently large (Gregory et al., 1993). Also, Hill (2000) observed that for a trait with heritability 1/3, (for example weaning weight in beef cattle), a value of $N_e=250$ is required to maintain the additive variance at its initial value. The Brangus breed is the largest composite population of beef cattle in the subtropics of Argentina. Since its creation in 1978, the Argentinean Brangus Breeders Association (AAB) has kept an open registry policy to maintain high levels of variability and to retain a high level of heterosis. The initial motivation for the research presented here was to quantify what is the current level of genetic variability, and to evaluate the pedigree structure of Brangus in Argentina.

Estimation of the N_e when pedigree records are available, can be accomplished indirectly by calculating the change in inbreeding (ΔF) since the breed formation, and then solving the expression $\Delta F=1/(2N_e)$ (Wright, 1931). However, this estimate is affected by the level of completeness in the pedigree, which causes underestimation of inbreeding (Miglior and Burnside, 1995; Lutaaya et al., 1999; Cassell et al., 2003). There is an extensive literature on direct calculation of N_e (see the review by Caballero, 1994) and several formulae are available depending on the assumptions of the data at hand. Hill (1979) obtained an expression for N_e that takes into account overlapping generations and the structure of the mating system through the variances and covariances of progeny numbers, or family sizes. In case of selection, the expression by Hill (1979) does not account for inherited selective advantage, i.e. the process by which the progeny from selected parents tend to have larger family size than those offspring from parents with smaller family size. Based on Santiago and Caballero (1995), Nomura (1996) derived an expression for N_e in selected populations with overlapping generations.

It is frequently assumed that the random variable family size in different animal species follows a Poisson distribution (Harris and Allenford, 1989; Goddard and Smith, 1990; Caballero, 1994; Joshi et al., 1999) in which mean and variance are equal. This assumption is unlikely to be fulfilled for most farm animal species where a small group of sires have a large contribution to the progeny pool. This, in turn, may induce overdispersion from a Poisson probability mass function. In addition, the family sizes of sires and dams in the formulae of Hill (1979) and Nomura (1996) require a bivariate specification, so that the covariance of male and female progeny numbers for both bulls and cows can be calculated. Possible discrete bivariate distributions for family sizes are the Bivariate Poisson (BP) and the Generalized Bivariate Negative Binomial (GBI-

VARNB, Gurmu and Elder, 2000). Both distributions account for correlations between male and female progeny numbers, but only the GBIVARNB takes overdispersion into account. The goals of this research are threefold: 1) to evaluate the pedigree structure through the effective number of founders (N_f), and the effective number of ancestors (N_a); 2) to estimate the (co) variances of family sizes of bulls and cows using either the BP or the GBIVARNB distribution to account for overdispersion, and 3) to estimate N_e using the (co) variances from the previous step. In doing so we used the data from purebreds and grades from the genetic evaluation program of the Brangus breed in Argentina. As a control population for the proposed methodology, we analyzed an Angus herd that has a more complete pedigree information than the Brangus breed.

2. Methods

2.1. Data

Data used for the study consisted of the pedigree records from 72,808 animals (45,668 females and 27,140 males) supplying records to ERBra, the genetic evaluation program of AAB. The animals were born from 1959 to 2005, most of them in Argentina but with some individuals originated in USA, Brazil or Bolivia. A large number of Brangus animals included in the ERBra are grades, which usually lack either sire or dam identification. The number of participating herds in 2005 was 56, from which 17 produce both purebred and grade animals whereas the rest raise grade cattle only. It is estimated that more than 75% of all purebred and grade Brangus animals participate of the ERBra. Most herds register both purebreds and grades, with a ratio of about 1:5. The AAB keeps the registry for the grades, whereas the national association of cattlemen keeps the herdbook for the purebreds of all breeds including the Brangus. Since its beginning in 1978, the AAB has kept an open policy of registering grade animals in order to maintain a high level of variability, and to retain the maximum possible levels of heterosis between Angus and Zebu cattle. Selection policies in the breed have resulted in the most popular bull sires being born in Argentina. Two herds are involved in an active embryo transfer program from a US based cattle company and register animals born locally but out of US parents. The total fraction of calves that are born to US bulls, either by embryo transfer or by artificial insemination, is about 3%.

The control purebred Angus herd consisted on records of 10,483 animals (4700 males and 5783 females) born between 1938 and 2005. The herd is located in Pasteur, western Buenos Aires province. The number of cows at any given year was about 250 up to 1990, moment at which was reduced to its actual size of about 100 females. A large proportion of all matings were to popular US bulls through artificial insemination. Since 1990, embryo transfer has become a common management practice with the 10% superior cows.

2.2. Pedigree structure and inbreeding

Most analyses to assess pedigree structure and to calculate F were performed using the program ENDOG (Gutiérrez and Goyache, 2005). The level of completeness of the Brangus and Angus pedigrees are seen in Fig. 1. In Brangus, almost 40% of the animals have either parent unknown due to the open policy of registration and genetic evaluation. Going further back, the amount of information on ancestors dramatically decreases. A totally different picture is observed in the Angus herd, with a reasonably informative pedigree. A global measure of pedigree completeness is the index proposed by MacCluer et al. (1983). The coefficient is defined in the 0–1 range, and is interpreted as the ability of the pedigree to measure the inbreeding of the animal. For any individual, the measure is computed as $\frac{2C_{\text{sire}}C_{\text{dam}}}{C_{\text{sire}}+C_{\text{dam}}}$, and $C = \frac{1}{d} \sum_{i=1}^d a_i$. The value a_i is the proportion of ancestors in generation i that are known, whereas d is the number of generations traced backward (Sørensen et al., 2005). We set $d=5$ generations, and averaged the coefficients of all individuals. Compared with the values reported for dairy cattle

breeds (0.85–0.93 in UK Holsteins, Kearney et al., 2004; 0.93–0.95 in Danish Holstein, Danish Jersey and Danish Red, Sørensen et al., 2005; 0.79 to 0.99 for Holsteins of 12 countries associated to Interbull, VanRaden, 2005), the level of completeness of the Argentinean Brangus was low: 0.17. In comparison, the Angus herd presented a completeness level equal to 0.68.

The effective number of founders (N_f), and the effective number of ancestors (N_a), were calculated to get further insight into the pedigree structure under such a loss of parent identification. A founder is an animal without known parents. As founders usually have unbalanced contributions, Lacy (1989) defined N_f to be the theoretical number of founders with balanced contributions that would be expected to produce the same genetic diversity as in the population under study. In case all founders have equal contributions, the actual number and the N_f are equal. The N_f is calculated as $N_f = \left(\sum_{i=1}^f p_i^2\right)^{-1}$ where p_i is the fractional contribution of the genes of founder i to a reference population, and f is the total number of founders. However, the measure does not take into account events that limit the genetic

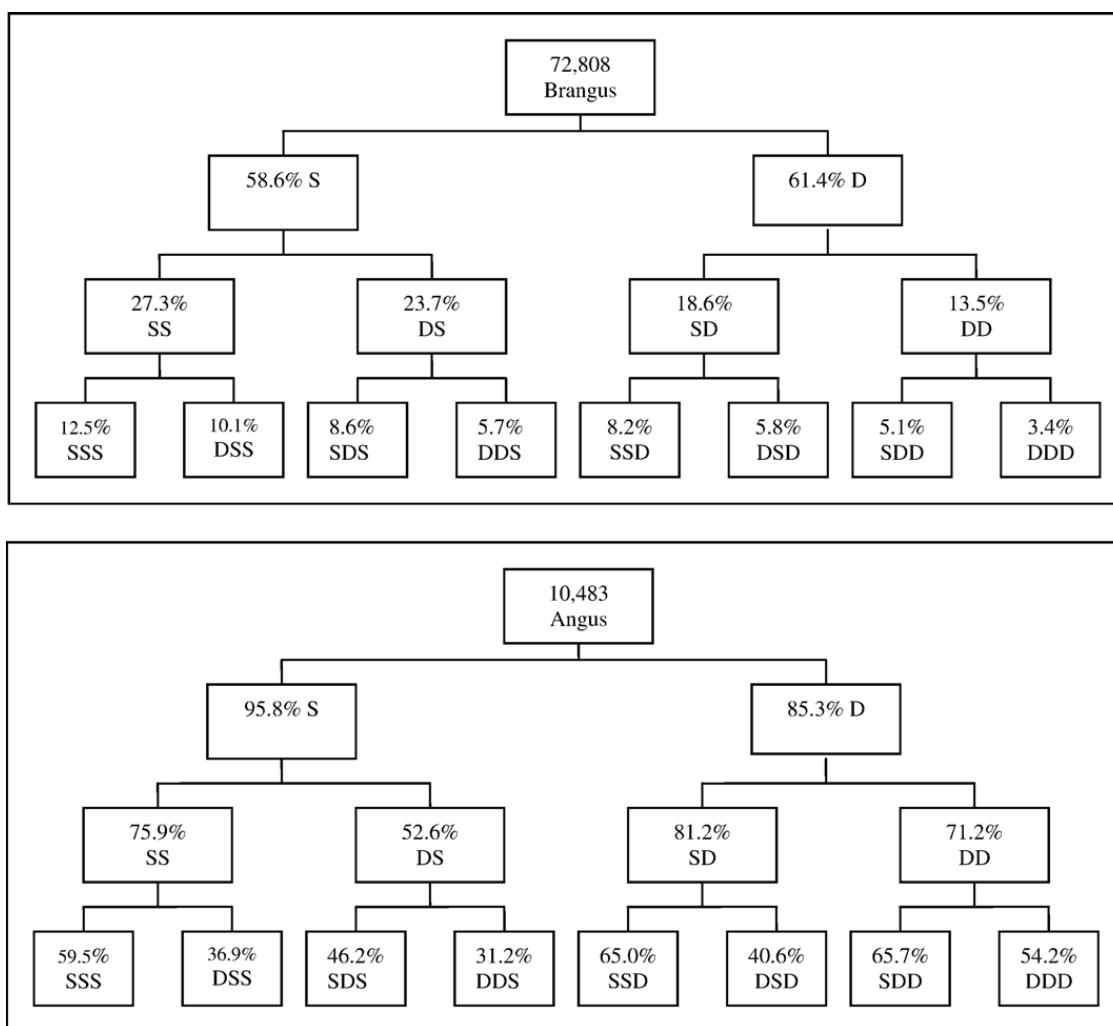


Fig. 1. Level of completeness of the Brangus and Angus pedigrees. S stands for Sire and D for Dam, so that SS means Sire of Sire; DSD dam of the Sire's dam, etc.

Table 1
Age distribution of Brangus cows from 2000 to 2004

Age (in years)	2000	2001	2002	2003	2004
2	0.028	0.031	0.030	0.029	0.031
3	0.245	0.225	0.215	0.213	0.216
4	0.145	0.148	0.206	0.187	0.185
5 to 8	0.478	0.471	0.435	0.458	0.452
9 or more	0.103	0.123	0.114	0.113	0.115

variation in the population such as drift and bottlenecks (for example, due to breed formation, or differential use of sire and dams). Boichard et al. (1997) suggested calculating the *effective number of ancestors* (N_a): “the minimum number of ancestors (founders or not) necessary to explain the genetic diversity under study” (Caballero and Toro, 2000, page 339). Similar to N_f the calculus of N_a is $N_a = \left(\sum_{i=1}^a q_i^2\right)^{-1}$ with q_i being now the marginal (or non-redundant) contribution of i to the pool of a ancestors. However, this measure does not account for additional losses of genes due to drift (Boichard et al., 1997; Caballero and Toro, 2000). Both N_f and N_a were calculated with respect to the contributions to the calves born in the reference population, which consisted of the calves born from 2001 to 2005: 21,662 in Brangus and 670 in the Angus herd.

The *effective number of herds* supplying fathers, grandfathers and great-grandfathers (Robertson, 1953) was computed for the Brangus using the inverse of the probability that two animals taken at random in the population have their sires (or grandsires, or great-grandfathers) in the same herd for each path (Gutiérrez and Goyache, 2005).

2.3. Effective population size

Boichard et al. (1997) observed that is difficult to assess drift using inbreeding measures when pedigree information is highly incomplete. Therefore, we attempted an alternative approach: the direct estimation of *effective population size* (N_e) using demographic parameters. In its original definition, Wright (1931) indicated that N_e is “the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration”. As in any other breed of cattle, the Brangus and Angus in Argentina have overlapping generations and the estimation of N_e as proposed by Hill (1972, 1979) was selected for the endeavor. Harris and Allenford (1989) compared several estimators of N_e and noted that the expression of Hill (1979) provided one of the most accurate estimates. Estimation by the approach of Hill (1979) results from solving for N_e in

$$\frac{1}{N_e} = \frac{1}{16ML} \left[2 + V_{mm} + 2 \left(\frac{M}{F}\right) C_{mm,mf} + \left(\frac{M}{F}\right)^2 V_{mf} \right] + \frac{1}{16FL} \left[2 + V_{ff} + 2 \left(\frac{F}{M}\right) C_{fm,ff} + \left(\frac{F}{M}\right)^2 V_{fm} \right]. \quad (1)$$

In expression (1), M and F are the number of bulls and cows that produce progeny during any given breeding season, and L

is the generation interval, or mean age of the parents when their reproductive progeny are born. Dispersion parameters of *family size* (or number of *reproductive progeny* from each sex) are the variance of male progeny among bulls (V_{mm}), the covariance between male and female progeny among bulls ($C_{mm,mf}$), the variance of female progeny among bulls (V_{mf}), the variance of male progeny among cows (V_{fm}), the covariance between male and female progeny among cows ($C_{fm,ff}$), and the variance of female progeny among cows (V_{ff}). All these demographic statistics carry information on the number of breeding individuals and their reproductive contributions, and allow assessing the genetic variability (de Rochambeau et al., 2000). In his derivation of (1), Hill (1979) assumed constant population size and a stable age distribution, so that the values of M and F are constant. A check of the age distribution of Brangus cows at different years suggests that the first assumption is correct for the females in these data (Table 1), but not for the bulls (Table 2). Besides, the requirement of a constant population size is not fulfilled in Argentine Brangus, as the number of evaluated animals has increased from 2001 to 2005.

To account for the possibility of selection, it was also computed the expression of N_e under selection (N_{eS}) given by Nomura (1996), which is equal to:

$$N_{eS} = \frac{4 N_M N_F}{N_M + N_F} \quad (2)$$

where

$$N_M = \frac{4 M L^2}{\left[\left(\frac{1}{\mu_{mm}} + \frac{1}{\mu_{mf}}\right) (1 - \alpha_m) + \left(\frac{V_{mm}}{\mu_{mm}^2} + \frac{V_{mf}}{\mu_{mf}^2} + \frac{2 C_{mm,mf}}{\mu_{mm}\mu_{mf}} + 4 Q^2 C_M^2\right) (1 + \alpha_m) \right]}$$

and

$$N_F = \frac{4 F L^2}{\left[\left(\frac{1}{\mu_{fm}} + \frac{1}{\mu_{ff}}\right) (1 - \alpha_f) + \left(\frac{V_{fm}}{\mu_{fm}^2} + \frac{V_{ff}}{\mu_{ff}^2} + \frac{2 C_{fm,ff}}{\mu_{fm}\mu_{ff}} + 4 Q^2 C_F^2\right) (1 + \alpha_f) \right]}$$

Formulae for N_M and N_F were different from Nomura’s (1996) in the sense that they include the term for the covariance between male and female reproductive progeny numbers divided but their respective means. The scalars μ_{mm} , μ_{mf} , μ_{fm} , and μ_{ff} , are the expected family sizes of sires of males, sires of females, dams of males and dams of females, respectively. Besides, α_m and α_f are the respective deviations from Hardy–Weinberg proportions in male and female parents. The term Q^2 accounts for the cumulative effects of selection on an inherited trait, and C_m^2 and C_f^2 are the variances of relative selective

Table 2
Age distribution of Brangus bulls from 2000 to 2004

Age (in years)	2000	2001	2002	2003	2004
2	0.06	0.08	0.08	0.04	0.04
3	0.13	0.12	0.14	0.23	0.14
4	0.20	0.20	0.11	0.17	0.29
5 to 8	0.45	0.42	0.51	0.34	0.30
9 or more	0.16	0.18	0.16	0.17	0.22

advantage of among families of males and females, respectively (Santiago and Caballero, 1995). The parameters Q^2 , C_m^2 and C_f^2 were estimated as suggested by Nomura (1996).

The generation interval L was calculated as the average age of parents when their replacing progeny was born (James, 1977) from the four selection pathways (sires of bulls mm, sires of cows mf, dams of bulls fm, and dams of cows ff): $L = \frac{1}{4}(L_{mm} + L_{mf} + L_{fm} + L_{ff})$.

2.4. Variances and covariances of family size

There seems to be no experimental estimates of the variance of family size in beef cattle. In this research, mean, variance and covariances of family sizes were calculated for sires of bulls (bull sires), dams of bulls (bull dams), sires of cows, and dams of cows. In theoretical models, family size is assumed to follow a Poisson distribution (Caballero, 1994), in which the expected value and the variance are equal. Estimates of the empirical sampling variances of family size and the covariance between female and male progenies were calculated using data on 80 bulls and 40 cows in Brangus and 178 bulls and 60 cows in Angus, as a large number of sires and dams do not produce both reproductive male and reproductive female progeny. The mean–variance equality assumption of the Poisson model was empirically challenged in this sample, and the results are displayed in Table 3.

With the exception of the statistics for the male progeny of cows, all variances were greater than their corresponding means, the largest difference being observed among the statistics of female progeny from bulls. These results suggest that family size of Brangus and Angus bulls displays overdispersion from a Poisson distribution. In cows, the results are less clear. There were also positive correlations of 0.45 and 0.16 between male and female progenies of Brangus bulls and cows, respectively. Corresponding sampling correlations in Angus were 0.27 and 0.03. To further collect evidence of overdispersion and of covariance in the numbers of male and

female progenies, we fitted two bivariate discrete distributions to the family sizes. The first one is the Bivariate Poisson (BP) model (Cameron and Trivedi, 1998, page 256). Let y_{1i} and y_{2i} be the number of progeny of bull or cow i for sex 1 (male) and 2 (female), respectively. Then, the BP model is obtained as

$$y_{1i} = x_{1i} + x_{3i} \quad y_{2i} = x_{2i} + x_{3i}$$

where x_{1i} , x_{2i} and x_{3i} are independent Poisson random variables with parameter λ_1 , λ_2 , and λ_3 , respectively. First and second moments of this joint density are equal to:

$$E(y_{1i}) = \lambda_1 + \lambda_3 = \text{Var}(y_{1i}); \quad E(y_{2i}) = \lambda_2 + \lambda_3 = \text{Var}(y_{2i}); \\ \text{Cov}(y_{1i}, y_{2i}) = \lambda_3.$$

The parameters of the joint distribution were estimated by maximum likelihood with the EM algorithm by means of the program *bivpois* (Karlis and Ntzoufras, 2005), which is written in the software *R*. Two Bivariate Poisson models were fit: the BP₁ allowing λ_3 to be estimated from the data, i.e. no covariance between male and female progeny numbers is assumed, and the BP₂ with $\lambda_3 \neq 0$.

Another model fitted to the distribution of male and female progeny numbers was the Generalized Bivariate Negative Binomial (GBIVARNB; Gurmu and Elder, 2000), which unlike the BP model allows overdispersion. Parameters of the GBIVARNB are θ_1 , θ_2 , α , and ρ , and the variances and covariance among the variables are equal to

$$\text{Var}(y_{1i}) = V\theta_{1i}^2 - \theta_{1i}^2 - \theta_{1i}; \quad \text{Var}(y_{2i}) = V\theta_{2i}^2 - \theta_{2i}^2 - \theta_{2i}; \\ \text{Cov}(y_{1i}, y_{2i}) = \theta_{1i}\theta_{2i}(V - 1)(2).$$

Gurmu and Elder (2000) indicated that V is equal to

$$V = \frac{\alpha + 1}{[\lambda^2(1 + \rho^2)]} [\alpha - 4\rho\sqrt{\alpha} + \rho^2(\alpha + 6)]$$

where $\lambda = [\alpha - 2\rho\alpha^{1/2} + \rho^2(\alpha + 2)] / (1 + \rho^2)$. The parameters of the GBIVARNB model were estimated using the program *gbcount.prg* written in GAUSS by Gurmu and Elder (2000). The following form of the Akaike information criterion (AIC, Burnham and Anderson, 1998, page 46) was used to compare the fit from the different models

$$\text{AIC} = -2l + 2p$$

where l denotes the maximum of the log-likelihood and p is the number of parameters. The model with the minimum AIC is to be selected.

3. Results

There were 29,127.5 equivalent founders (23,096 founder plus 12,063 half-founders or animals with one unidentified parent) in the Brangus breed. The reference population consisted of 21,662 calves born from 2001 to 2005. The number of ancestors responsible for 100% of

Table 3
Sampling statistics of family size of bull sires and bull dams

			Male progeny	Female progeny
<i>Brangus</i>				
Bulls	N=84	\bar{m}	2.12	17.81
		V_m	3.96	428.35
		$C_{mm,mf}$	18.49 (0.45) ^a	
Dams	N=40	\bar{f}	1.14	2.10
		V_f	0.23	3.27
		$C_{ff,fm}$	0.14 (0.16) ^a	
<i>Angus</i>				
Bulls	N=178	\bar{m}	2.06	8.02
		V_m	4.43	203.40
		$C_{mm,mf}$	8.16 (0.27) ^a	
Dams	N=60	\bar{f}	1.13	2.20
		V_f	0.22	5.86
		$C_{ff,fm}$	0.04 (0.03) ^a	

^a Correlation between male and female family sizes.

the genes in the reference group was 12,471. However, 50% of that variation was explained by 273 animals, and 20% by 22 ancestors. The animal with the largest individual contribution to the 2001–2005 calves was a bull responsible for 2.4% of the genetic variability. The effective number of founders was $N_f=765.7$, whereas the effective number of ancestors was $N_a=387.5$. The average value of F calculated from the entire pedigree was equal to 0.24%.

There were 992 equivalent founders (285 founders plus 1414 half-founders) in the Angus herd. For the 670 calves born in the period 2001–2005, the effective numbers of founders ($N_f=46.3$) and ancestors ($N_a=46$) were similar. The numbers of ancestors responsible for 100%, 50%, or 20% of the genes were respectively equal to 144, 17, and 4 animals. The individual with the largest individual contribution to the 2001–2005 Angus calves was a bull responsible for 4.9% of the genetic variability. The average inbreeding calculated with all pedigree information was 0.5%.

With respect to the differential contributions of bulls from different herds, Table 4 shows the actual and effective (Robertson, 1953) number of herds contributing sires, grandsires and great-grandsires to the Brangus population. The largest reduction from the actual to the effective number is for the sires (52 to 9.04). Interestingly enough, the 9 elite herds contributing most sires to the breed provided 50% of the animals with records in the ERBra. From these herds, those six that contribute most grandsires and those four that contribute most great-grandsires provided 44% and 40%, respectively, of the animals with records in the ERBra. Thus, although the nucleus is constituted by 16% (9 out of 56) of the herds, they contributed 50% of the animals to the breed.

The numbers of bulls (M) and cows (F) in any given year were estimated by the average number of reproductive males and females respectively that were parents of the calves born in the period 2001–2005 for both Brangus and Angus. The estimates in the Brangus breed were $M=210$ sires and $F=5415$ dams. Corresponding estimates in the Angus herd were $M=23$

Table 4

Actual and effective number of Brangus herds contributing sires, grandsires and great-grandsires

Ancestral generation	Actual number of herds	Effective number of herds
Sires	52	9.04
Grandsires	29	6.44
Great-grandsires	15	4.23

Table 5

Parameter estimates of the Bivariate Poisson (BP) models for family size

Parameters	BP ₁		BP ₂	
	Bulls	Cows	Bulls	Cows
<i>Brangus</i>				
λ_1	2.12	1.14	1.29	0.09
λ_2	17.81	2.10	16.98	1.04
λ_3	0.00	0.00	0.83	1.06
Log-likelihood	-1080.08	-115.21	-1.066.49	-112.13
AIC	2164.16	234.43	2138.98	230.27
<i>Angus</i>				
λ_1	8.02	2.20	7.66	2.17
λ_2	2.06	1.13	1.70	1.11
λ_3	0.00	0.00	0.34	0.03
Log-likelihood	-1906.45	-187.97	-1893.64	-187.96
AIC	3816.90	379.95	3793.28	381.93

and $F=106$. Generation intervals (in years) were $L_{mm}=5.0$, $L_{mf}=5.7$, $L_{fm}=4.4$, and $L_{ff}=5.2$ for the Brangus, and $L_{mm}=5.2$, $L_{mf}=5.9$, $L_{fm}=3.1$, and $L_{ff}=3.9$ for the Angus. The smaller value of L_{fm} in both breeds may be due to the high level of embryo transfer used to produce bulls in the nucleus herds with selected young heifers. Average generation intervals were 5.1 years in Brangus and 4.8 years in Angus.

Estimates of the parameters for family sizes using the BP models are presented in Table 5. The smaller values of AIC for the BP₂ model in Brangus suggest that family numbers of male and female progeny of both bulls and cows are not independent random variables. The same was observed for the Angus bulls, but the BP₁ model had a slightly smaller (i.e. better) AIC than the BP₂ model for the cows. This later result was consistent with the findings from the GBIVARNB model where the estimated covariance of male and female family size of the cows was zero.

The estimates of the parameters for the distributions of family sizes of bulls and cows under the GBIVARNB

Table 6

Parameter estimates of the Generalized Bivariate Negative Binomial (GBIVARNB) model for family size

Parameters	Brangus		Angus	
	Bulls	Cows	Bulls	Cows
θ_1	2.12	1.14	1.58	1.00
θ_2	17.81	2.10	4.56	1.38
α	1.25	19.18	0.86	453613.14
ρ	0.00	0.98	0.00	2.46
Log-likelihood	-498.22	-114.42	-833.01	-145.38
AIC	1004.45	236.84	1674.03	298.75

model are shown in Table 6. In Brangus, the AIC values for comparing the fit of different distributions favored the GBIVARNB model for the bulls. On the contrary, the BP₂ model provided a slightly better fit for the cows. Thus, the distribution of reproductive male and female progeny numbers for the bulls displayed overdispersion with respect to the Bivariate Poisson model in which mean and variance are equal (see expression (2)), and there was evidence for positive covariances among male and female progenies of bulls and cows. On the other hand, the GBIVARNB model produced the best fit for bulls and cows in the Angus.

Variances and covariances of family sizes were then calculated assuming the GBIVARNB model for the Brangus bulls, the BP₂ model for the Brangus cows, and the GBIVARNB model for the Angus bulls and cows. The results are shown in Table 7. In both breeds variances of male and female progeny numbers, as well as the covariance between family sizes, were larger for bulls than for cows, with the largest difference being in the variance between female progenies. Comparison of the estimated parameters under the GBIVARNB model and those displayed in Table 3, in which the implicit sampling distribution is the Bivariate Normal, shows that means are similar but variances and covariances are different.

When calculating the formula of Hill (1979) in Brangus, by using the estimates in Table 7, the number M of bulls and F of cows, and the generation interval as calculated above, resulted in $N_e=273.9$, which then produces $\Delta F=1/(2 N_e)=0.0018$, or 0.18%. This latter value is smaller than the observed average F which was

equal to 0.24%. When attempting to account for selection, the estimate of N_{eS} using the expression of Nomura (1996), resulted in a value of 124.8 so that $\Delta F=0.40\%$. For Angus, with a more complete pedigree, $N_e=95.4$, and $\Delta F=0.5\%$ in perfect agreement with the observed average inbreeding of all 10,483 animals. The effect of selection was less pronounced in the Angus herd than in Brangus breed as $N_{eS}=79.3$ and $\Delta F=0.63\%$.

4. Discussion

The indexes calculated using the Brangus calves born in the period 2001–2005 as the reference population were $N_f=765.7$ and $N_a=387.5$. In Austrian breeds, Sölkner et al. (1998) reported $N_f=221$ and $N_a=114$ for Simmental, $N_f=97$ and $N_a=52$ for Braunvieh, and $N_f=113$ and $N_a=39$ for Pinzgauer. Gutiérrez et al. (2003) obtained values of N_f ranging from 48 to 265 and N_a ranging from 25 to 163, in eight Spanish breeds. For the Italian breeds Chianina, Marchigiana and Romagnola, Bozzi et al. (2006) estimated N_f to be 152.1, 70.9 and 89.8; corresponding values of N_a were 73.6, 48.0 and 59.5. In the Mexican composite breed Tropicarne, Ruíz-Flores et al. (2006) found $N_f=48$ and $N_a=20$. In Brazil, Vozzi et al. (2006) reported $N_f=87.2$ and $N_a=59.8$ for Nellore, and $N_f=107.9$ and $N_a=61.5$ for Polled Nellore. In Ireland, Mc Parland et al. (2006) observed values of N_f varying from 55 (Simmental) to 357 (Charolais), whereas N_a ranged from 35 (Simmental and Hereford) to 82 (Limousin). Hence, the values of N_f and N_a in Brangus are higher than most reported estimates for beef cattle. Difference in population size, production systems for beef cattle, and above all in the policy of registration, may explain the reasons for the larger figures in Brangus. Sørensen et al. (2005) observed that the total number of founders contains limited information on the genetic basis for the population as founders are assumed to be unrelated, because their parentage is unknown. The largest source of founders in Argentinean Brangus is from grade dams. These are cows bringing grade calves with records to ERBra. Most of these are born following natural mating, so that groups of contemporary cows may be paternal half-sibs, and this may have induced a slight overestimation of N_f in the Brangus population.

The amount of inbreeding in populations with incomplete parent identification is underestimated (Miglior and Burnside, 1995; Lutaaya et al., 1999; Cassell et al., 2003). Rather than trying to amend the lack of information by using the average F of animals born in the same year as proposed by VanRaden (1992), we attempted an indirect method by calculating the N_e under the assumption of

Table 7
Estimates of the variances and covariances for family sizes assuming different distributions for bulls and cows

Family size	Model	Parameters	Estimate
<i>Brangus</i>			
Bulls	GBIVARNB	V_{mm}	5.70
		V_{mf}	271.28
		$Cov_{mm,mf}$	30.15
Cows	BP ₂	V_{fm}	1.15
		V_{ff}	2.10
		$Cov_{fm,ff}$	1.06
<i>Angus</i>			
Bulls	GBIVARNB	V_{mm}	4.48
		V_{mf}	28.67
		$Cov_{mm,mf}$	8.36
Cows	GBIVARNB	V_{fm}	1.00
		V_{ff}	1.38
		$Cov_{fm,ff}$	0.00

overlapping generations (Hill, 1979) and selection (Nomura, 1996). These methodologies are based on the assumptions of equal age distribution for bulls and cows and constant population size. The first assumption was verified in cows (see Table 1) but not in bulls (Table 2), and the changing herd size with time in both Brangus and Angus do not comply with the second assumption. It is unlikely to find beef cattle populations of constant size over extended periods of time. The consequences of those violations on the estimates of N_e are unknown. Key elements of the formulae by Hill (1979) and Nomura (1996) are the variances and covariances of family sizes (number of reproductive progeny) of bulls and cows, parameters that convey information on the mating system. A weakness of the approach is that a lot of information is discarded in the process of estimating the (co)variances of family size in beef cattle, as only data from bulls and cows that have reproductive progeny of both sexes (i.e. bull sires and bull dams) are included, and these constitute a relatively small fraction of all breeding individuals. This may induce an overestimation of N_e , as the information from a portion of sires of cows and most dams of cows is not used in the estimation process. Family size of dairy bulls has been considered to follow a Poisson distribution (Goddard and Smith, 1990), and there was experimental evidence for a Poisson male family size in *Drosophila melanogaster* (Joshi et al., 1999). However, our analysis shows that this was not the case for the family sizes of beef bulls as there was considerable overdispersion, i.e. the variance was higher than the mean. In addition, covariances of male and female family sizes of bulls and cows are needed in Hill's (1979) formula. Therefore, a proper specification of these discrete random variables necessarily involves a bivariate distribution such as the BP or GBIVARNB. The latter probability mass function also accounts for overdispersion (Gurmu and Elder, 2000). The methodology used here (i.e. estimating inbreeding by means of the formula of Hill (1979) for N_e , using variances and covariances of family sizes under a GBIVARNB model) seems to be an able procedure when pedigree information is weak, and gives similar results to the calculated inbreeding when pedigree information is more complete as in the Angus data.

The direct estimate of N_e under the assumption of no selection is above the value of $N_e=250$ mentioned by Hill (2000) to keep the additive variance at its initial value for a trait with $h^2=1/3$. However, the estimate from a selection model (N_{eS}) is strikingly different (less than half) than the value of N_e . A much smaller difference between estimated N_e and N_{eS} was obtained in Angus where N_{eS} was about 83% of N_e . When comparing all elements in formula (2), the ratio of bulls

to cows ($M:F$) was markedly different in both populations: 0.038 in Brangus vs. 0.216 in Angus. Had it been $M:F$ in the Brangus equal to the one observed in the Angus herd while keeping all remaining parameters the same, would have resulted in an increase in N_{eS} to 798. Therefore, the heavier bull usage in the Brangus resulted in a decrease in effective population size. There seems to be no reported direct estimates of N_e (rather than equating it to the change in F) for composite cattle breeds. However, the estimated N_{eS} is in the range of 50 to 150 in which most estimates of the parameter are reported for beef cattle breeds (Boichard et al., 1997; Nomura et al., 2001; Gutiérrez et al., 2003; Bozzi et al., 2006; Mc Parland et al., 2006; Vozzi et al., 2006). Notwithstanding this, and as it would have been expected a larger effective population size in a composite breed with an open policy registration, it seems desirable to continue monitoring the effective size of the Brangus to prevent problems of inbreeding and lack of variability in the future.

Acknowledgments

This research was made possible by grants UBACyT G018-Program 2004–2007 of Universidad de Buenos Aires, and CONICET PIP 5338 2005–2006 from Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina. The authors would like to thank the associated editor and to an unknown reviewer for their insightful comments which largely helped to improve an earlier version of the present 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.
- Bozzi, R., Franci, O., Forabosco, F., Pugliese, C., Crovetto, A., Filippini, F., 2006. Genetic variability in three Italian beef cattle breeds derived from pedigree information. *Ital. J. Anim. Sci.* 5, 129–137.
- Burnham, K.P., Anderson, D.R., 1998. Model selection and inference: A practical information–theoretic approach. Springer-Verlag, New York.
- 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 tool for the management of conserved populations. *Genet. Res.* 75, 331–343.
- Cameron, A., Trivedi, P., 1998. Regression Analysis of Count Data. Cambridge University Press, New York.
- Cassell, B.G., Adamec, V., Pearson, R.E., 2003. Effect of incomplete pedigrees on estimates of inbreeding and inbreeding depression for days to first service and summit milk yield in Holsteins and Jerseys. *J. Dairy Sci.* 86, 2967–2976.

- de Rochambeau, H., Fournet-Hanocq, F., Vu Tien Khang, J., 2000. Measuring and managing genetic variability in small populations. *Ann. Zootech.* 49, 77–93.
- Goddard, M., Smith, C., 1990. Optimum number of bull sires in dairy cattle breeding. *J. Dairy Sci.* 73, 113–1122.
- Gregory, K.E., Cundiff, L.V., Koch, R.M., 1993. What does the research tell us? Proceedings of the range beef cow symposium XIII. Available from <http://beef.unl.edu/beefreports/symp-1993-05-XIII.shtml>.
- Gurmu, S., Elder, J., 2000. Generalized bivariate count data regression models. *Econ. Lett.* 68, 31–36.
- Gutiérrez, J., Goyache, F., 2005. A note on ENDOG: a computer program for analyzing pedigree information. *J. Anim. Breed. Genet.* 122, 172–176.
- Gutiérrez, J., Altarriba, J., Díaz, C., Quitanilla, R., Cañón, J., Piedrafita, J., 2003. Pedigree analysis of eight Spanish beef cattle breeds. *Genet. Sel. Evol.* 35, 43–63.
- Harris, R., Allenford, F., 1989. Genetically effective population size of large mammals: an assessment of estimators. *Cons. Biol.* 3, 181–191.
- Hill, W.G., 1972. Effective size of populations with overlapping generations. *Theor. Popul. Biol.* 3, 278–289.
- Hill, W.G., 1979. A note on effective population size with overlapping generations. *Genetics* 92, 317–322.
- Hill, W.G., 2000. Maintenance of quantitative genetic variation in animal breeding programmes. *Livest. Prod. Sci.* 63, 99–109.
- James, J.W., 1977. A note on selection differentials and generation length when generations overlap. *Anim. Prod.* 24, 109–112.
- Joshi, A., Do, M., Mueller, L., 1999. Poisson distribution of male mating success in laboratory populations of *Drosophila melanogaster*. *Genet. Res.* 73, 239–249.
- Karlis, D., Ntzoufras, I., 2005. Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R. *J. Statistical Software*. 14, Issue 10, paper 1495. Available from <http://www.jstatsoft.org/>.
- Kearney, J.F., Wall, E., Villanueva, B., Coffey, M.P., 2004. Inbreeding trends and application of optimized selection in the UK Holstein population. *J. Dairy Sci.* 87, 3503–3509.
- Lacy, R.C., 1989. Analysis of founder representations in pedigrees: founder equivalents and founder genome equivalents. *Zoo Biol.* 8, 111–123.
- Lutaaya, B.E., Misztal, I., Bertrand, J.K., Mabry, J.W., 1999. Inbreeding in populations with incomplete pedigree. *J. Anim. Breed. Genet.* 116, 475–480.
- MacCluer, J.W., Boyce, A.J., Dyke, B., Weitkamp, L.R., Pfennig, D.W., Parsons, C.J., 1983. Inbreeding and pedigree structure in Standard bred horses. *J. Heredity* 74, 394–399.
- McParland, S., Kearney, J.F., Rath, M., Berry, D.P., 2006. Inbreeding trends and pedigree analysis of Irish dairy and beef cattle populations. *J. Anim. Sci.* 367 v1.
- Miglior, F., Burnside, E.B., 1995. Inbreeding of Canadian Holstein cattle. *J. Dairy Sci.* 78, 1163–1167.
- Nomura, T., 1996. Effective size of selected populations with overlapping generations. *J. Anim. Breed. Genet.* 113, 1–16.
- Nomura, T., Honda, T., Mukai, F., 2001. Inbreeding and effective population size of Japanese black cattle. *J. Anim. Sci.* 366–370.
- Robertson, A., 1953. A numerical description of breed structure. *J. Agric. Sci.* 43, 334–336.
- Ruiz-Flores, A., Núñez-Domínguez, R., Ramírez-Valverde, R., Domínguez-Viveros, J., Mendoza-Domínguez, M., Martínez-Cuevas, E., 2006. Levels and effects of inbreeding on growth and reproductive traits in Tropicarne and Brown Swiss cattle. *Agrociencia* 40, 289–301.
- Santiago, E., Caballero, A., 1995. Effective size of populations under selection. *Genetics* 139, 1013–1030.
- Sölkner, J., Filipčić, L., Hampshire, N., 1998. Genetic variability of populations and similarity of subpopulations in Austrian cattle breeds determined by analysis of pedigrees. *Anim. Sci.* 67, 249–256.
- Sørensen, A.C., Sørensen, M.K., Berg, P., 2005. Inbreeding in Danish dairy cattle breeds. *J. Dairy Sci.* 88, 1865–1872.
- VanRaden, P., 1992. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. *J. Dairy Sci.* 75, 1163–1167.
- VanRaden, P., 2005. Inbreeding adjustments and effect on genetic trend estimates. *Interbull Bulletin*. 33, 81–84.
- Vozzi, P.A., Marcondes, C.R., Ulhôa Magnabosco, C., Framartino Bezerra, L.A., Barbosa Lôbo, R., 2006. Structure and genetic variability in Nellore (*Bos indicus*) cattle by pedigree analysis. *Genet. Mol. Biol.* 29, 482–485.
- Wright, S., 1931. Evolution in Mendelian populations. *Genetics* 16, 98–160.