GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

In-depth pedigree analysis in a large Brazilian Nellore herd

F.V. Brito[1], M. Sargolzaei[2,3], J. Braccini Neto[1], J.A. Cobuci[1], C.M. Pimentel[1], J. Barcellos[1] and F.S. Schenkel[2]

[1]Departamento de Zootecnia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
[2]Center for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, Canada
[3]L'Alliance Boviteq, Saint-Hyacinthe, QC, Canada

Corresponding author: F.V. Brito
E-mail: fernandabrito@gensys.com.br

Genet. Mol. Res. 12 (4): 5758-5765 (2013)
Received February 5, 2013
Accepted August 16, 2013
Published November 22, 2013
DOI http://dx.doi.org/10.4238/2013.November.22.2

ABSTRACT. A large herd of Nellore cattle was evaluated using in-depth pedigree analyses. Taking into account the incomplete pedigree due to the use of multiple young sires for mating, the average inbreeding coefficient was calculated as 1.73% for the last generation, which was higher than the regular inbreeding coefficient (0.25%). The effective population size was estimated to be 114, 245, and 101 for the time periods 1995-1999, 1999-2003, and 2003-2007, respectively. Parameters based on the probability of gene origin were used to describe the genetic diversity over time in the herd. The effective number of founders, ancestors, and founder genomes decreased over time, showing an overall loss of genetic diversity. In the last five-year period (2003-2007), based on available pedigree information, one prominent ancestor contributed 10.6% to the gene pool of the herd, and 30% of this pool was contributed by 31 ancestors. The analysis of inbreeding under random mating indicated that the mating strategies used in the herd are slowing down inbreeding rates. However, it is advisable to continue monitoring the inbreeding rates and genetic diversity in this herd in the future.

Key words: Effective population size; Population structure; *Bos indicus* cattle; Inbreeding; Probability of gene origin

INTRODUCTION

Due to advances in animal breeding methods and in techniques used by farmers in recent decades, an increasing number of Nellore cattle herds have been under intensive genetic improvement programs. These programs, which are based on mixed model methodology and the use of artificial insemination, facilitate more accurate genetic evaluations and decreased generation intervals, resulting in higher short-term genetic gains (Henderson, 1975; Fernando and Gianola, 1986; Kennedy et al., 1988). However, these programs also result in increased inbreeding rates and reduced genetic diversity, since related animals are more likely to be selected (Quinton et al., 1992; Verrier et al., 1993). Therefore, inbreeding and genetic diversity parameters should be monitored in animal breeding programs. One way to analyze the genetic diversity of a population is based on information provided by pedigree records. Besides the coefficient of inbreeding and effective population size, statistics based on the probability of gene origin (MacCluer et al., 1986; Lacy, 1989; Boichard et al., 1997) provide valuable information for studying a population undergoing several years of selection.

In Brazil, Nellore cattle form the majority of the beef cattle population. Artificial insemination (AI) is used as a tool for genetic improvement, but because large herds are managed in large paddocks where the use of single sire mating is not feasible, the use of multiple selected young sires (MS, groups of bulls mated to groups of cows) is common. Due to the average high genetic value of the young sires, their use results in genetic gain, but also in loss of pedigree information, leading to an underestimation of the inbreeding coefficient (Lutaaya et al., 1999); however, missing pedigree information can be taken into account while computing inbreeding coefficients (VanRaden, 1992; Colleau and Sargolzaei, 2011).

The present study was carried out to: 1) estimate the inbreeding coefficients and the effective population size of a large Nellore herd that is typical of Nellore herds under intensive genetic improvement programs in Brazil that utilize AI and MS matings, and 2) analyze the genetic variability of this herd using statistics derived from the probability of gene origin.

MATERIAL AND METHODS

Data

A large Nellore cattle herd, consisting of 146,291 animals born from 1984 to 2007, was investigated in this study. The herd belongs to Agropecuária Jacarezinho that participates in a large-scale breeding program in Brazil along with other breeders. In addition, the database of the Brazilian Association of Zebu Breeders was used to trace the pedigree of the herd back to a 1938 base, resulting in 162,244 genealogical records. The data set was divided into three five-year periods (1995-1999, 1999-2003, and 2003-2007), as a five-year period approximates the generation interval in the herd. All of the genetic diversity parameters were calculated for each period. The genetic evaluation in this population is based on BLUP estimates of breeding values, using an animal model. Every year, the best 20% two-year-old young bulls are used for MS mating, and the top 1% of bulls are used for AI. In addition, semen from bulls from other breeders participating in the breeding program and from insemination companies is used.

Methodology

Inbreeding coefficient

The regular inbreeding coefficients (Wright, 1922) were calculated for all animals in the population using the algorithm described by Sargolzaei et al. (2005), which is based on the indirect method proposed by Colleau (2002). The inbreeding coefficients were also calculated using the method proposed by VanRaden (1992), VRM, which accounts for missing pedigree information. This method assigns the average relationship coefficient of known parents of animals from the same generation to the unknown parent(s) of an animal. The inbreeding coefficients calculated under an assumption of random mating (F_{ST}), proposed by Wright (1969), were compared to regular inbreeding coefficients, resulting in the F_{IS} statistic, which describes deviations from random mating, allowing for the assessment of mating strategies used in the breeding program.

Effective population size (N_E)

The N_E was calculated as:

$$N_E = ½\, \Delta F_t\, L \qquad \text{(Equation 1)}$$

where ΔF_t is the annual rate of inbreeding at time t, and L is the average generation interval. F was calculated as proposed by VanRaden (1992).

Effective number of founders (E_f)

E_f is defined as the number of equally contributing founders that would produce the same amount of genetic variability as observed in the current population. This contribution was calculated as:

$$E_f = 1 / \sum_{j \in Fo} \left(\frac{\sum_{i \in G} t_{ij}}{n_g} \right)^2 \qquad \text{(Equation 2)}$$

where *Fo* is the set of founders, *G* is the set of individuals in the reference population (animals born in 2003-2007), n_g is the number of individuals in the reference population, and t_{ij} is an element of matrix T that represents the fraction of the genes that individual *i* inherited from founder *j* (Lacy, 1989).

Effective number of founder genomes (E_{fg})

E_{fg} is defined as the number of equally contributing founders with no loss of founder alleles that would result in the same amount of genetic diversity as in the current population. It was calculated as in Lacy (1989):

$$E_{fg} = \frac{1}{2\bar{f}_g} \qquad \text{(Equation 3)}$$

where $\bar{f}_g$ is the average co-ancestry in the reference population.

Effective number of non-founders (E_{nf})

E_{nf} accounts for the loss of genetic diversity due to drift accumulated over non-founder generations. It was calculated as in Caballero and Toro (2000):

$$E_{nf} = \frac{1}{\frac{1}{E_{fg}} - \frac{1}{E_f}} \qquad \text{(Equation 4)}$$

where E_{fg} and E_f are as previously defined.

Effective number of ancestors (E_a)

E_a was proposed by Boichard et al. (1997), and determines the minimum number of ancestors (founders or not) required to explain the complete genetic diversity in the current reference population. It was calculated as:

$$E_a = \left(\sum_{i=1}^{n_a} p_i^2\right)^{-1} \qquad \text{(Equation 5)}$$

where p_i is the marginal genetic contribution of ancestor i to the population, and n_a is the number of ancestors. In this study, the 1000 most contributing ancestors were used.

Software

The CFC software (Sargolzaei et al., 2006) was used to calculate pedigree statistics, the regular inbreeding coefficient, E_f, E_{fg}, and E_{nf}. Inbreeding expected under random mating was calculated using the EVA software (Berg et al., 2007). The PEDIG package (Boichard, 2002) was used to calculate generation intervals, E_a, and to identify ancestors with the highest genetic contributions.

RESULTS AND DISCUSSION

Of the total 162,244 animals, 9.2% were considered to be founders (both sire and dam were unknown), 34.4% had both parents known, and 56% had only the dam known. The regular average inbreeding value for animals born in 2007 was 0.25%, but was 1.73% when calculated by VRM, which was nearly seven times higher (Figure 1A). This sizeable difference can be attributed to the large number of unknown sires, and indicates that VRM recovered some of the inbreeding missed by the regular method. Lutaaya et al. (1999) tested the efficacy of this method by simulating the loss of dam information, ranging from 10 to 50%. The deletion of dam information was done randomly, and since dams could have come from any of the previous generations, the authors chose to assign the average inbreeding coefficient of the animals born in the four years preceding the birth of their progeny. They concluded that under these conditions, using VRM, the average inbreeding value decreased by only 30% when 20% of the dams were missing, and decreased by 78% when 50% of the dams were missing. VRM was able to recover most of the inbreeding with up to 20% of missing pedigree information, and was superior to the regular inbreeding calculation for all proportions of missing dams. Because in the current data set the bulls used in MS matings were all selected from the same generation that preceded the birth of their offspring, VRM seemed more efficient here than it was in Lutaaya et al. (1999), especially for the scenario with the highest levels of loss of pedigree records.

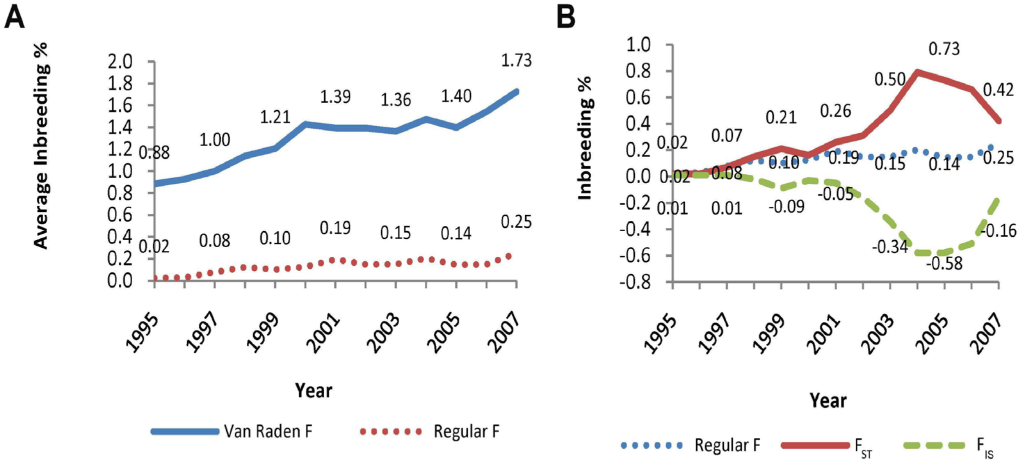

Figure 1. A. Average regular inbreeding coefficient and average inbreeding coefficient calculated by VanRaden's method. **B.** Average regular inbreeding coefficient, inbreeding expected under random mating (F_{ST}) and deviation from random mating (F_{IS}).

The average inbreeding increased across years, indicating a loss of diversity, which was possibly caused by selection and the consequent use of animals with some degree of relatedness. Average regular inbreeding levels in Nellore have been reported by Faria et al. (2009), and were slightly higher (2.13% in the 1994-1998 period) than those found using VRM in the current study (1.73% in 2007). For other Brazilian Zebu breeds in this same period, Faria et al. (2002) reported values of F ranging from 0.98% in the Polled Nellore to 10.13% in the Sindhi breed.

Although the rates of inbreeding fluctuated across the periods considered in the current study, they showed an overall increasing trend (Figure 1A). However, Figure 1B shows that deviations from random mating were mostly negative since 1995. This suggests that the average observed inbreeding of the herd was lower than the level expected under random mating, indicating the use of some form of minimum co-ancestry mating in the herd. The inbreeding under random mating was, on average, 1.3 times higher than the inbreeding coefficient calculated using actual matings, reaching levels up to 4 times higher in 2005 (Figure 1B). Since the criteria for using bulls from other herds or AI companies have remained the same throughout the whole period, this difference could be explained by an appropriate control of matings using a software package called PAD, which was adopted for this population in the late 1990s (Roso and Fries, 1998). Despite the fact that the inbreeding coefficients in Figure 1B were calculated by the regular method, the difference in inbreeding levels calculated under random mating to those based on actual matings is still relevant. The increase in inbreeding rates and the reduction of the difference between inbreeding under random and actual matings after 2005 (Figure 1A and B) should be monitored.

As for the rates of inbreeding, the N_E showed fluctuations since 1995 (Table 1). In the first period, with the intensification of the use of AI and the use of an important ancestor (Kulal AJ) that was born in 1993 (Table 2), the estimated N_E was only 114. N_E was increased in the subsequent period, which one could speculate to be a consequence of the more intense use of the PAD mating software. More recently, however, there was a decrease in N_E, which could be explained by an increasing use of descendants of Kulal AJ (Table 2) that contributed to 10.6% of the gene pool of the herd in the 2003-2007 period.

Table 1. Generation interval (Gen. Int.), inbreeding calculated by VanRaden's method (ΔF), population size (N), and effective numbers in the Nellore herd.

Period	Gen. Int.	ΔF	N	N_E	E_f	E_a	E_{fg}	E_{nf}
1995-1999	6.85	0.00064	33,775	114	607	239	152	204
1999-2003	6.47	0.00032	33,053	245	436	87	74	89
2003-2007	6.83	0.00073	39,209	101	479	77	67	78

N_E = effective population size; E_f, E_a, E_{fg}, E_{nf} = effective numbers of founders, ancestors, founder genomes, and non-founders, respectively.

Table 2. Eight ancestors with the highest genetic contributions to the gene pool of animals born between 2003 and 2007.

Ancestor	Sex	Birth (year)	No. of offspring	No. of descendants	Genetic contributions (%)		
					Total	Marginal	Cumulated
Kulal AJ	M	1993	7593	20,745	10.6	10.6	10.6
Gim de Garça	M	1976	333	29,154	2.1	2.1	12.7
Tiete	M	1987	1351	7,409	1.6	1.6	14.3
Cabral AJ	M	1997	892	1,722	1.5	1.5	15.7
Solimões AJ	M	1995	777	2,700	1.3	1.3	17.1
Karvardi	M	1955	89	50,127	1.8	1.2	18.3
DiCavalcanti AJ	M	1998	646	1,408	1.2	1.0	19.3
Rondon AJ	M	1997	533	956	0.9	0.9	20.1

In a genetic study of 19 breeds that differ geographically and biologically, Gibbs et al. (2009) observed that the recent N_E has decreased for all breeds, reflecting the occurrence of bottlenecks associated with domestication, breed formation, and, in some breeds, recent intense selection for production traits. Faria et al. (2009) estimated the N_E of the main Brazilian Zebu breeds and found a significant decrease due to the use of few sires. The N_E values for Nellore were 85, 98, 71, and 68 across the studied periods (79-83, 84-88, 89-93, and 94-98, respectively). According to Magnabosco et al. (1997), six bulls have significantly contributed to the formation of Nellore cattle in Brazil, with approximately 20% of genes of the current herds coming from these bulls.

For the period of 2003-2007 in the current study, 8 ancestors contributed to 20% of the gene pool (Table 2), while 31 ancestors contributed to 30% of the gene pool. These contributions might have been underestimated due to the high amount of missing pedigree information. However, studies involving other herds of Brazilian Nellore reported similar contributions from the first 30 main ancestors (Vozzi et al., 2006, 2007; Faria et al., 2009).

Despite the fact that a mating system is being used to control inbreeding rates in the herd, the statistics based on the probability of gene origin show a loss of genetic diversity over time (Table 1). The E_f values were high due to the high number of unknown parents. The E_{fg} decreased over time, and its value for animals born in recent years was around 70. This shows that, besides the unequal founder contributions, bottleneck effects and genetic drift were important causes of genetic diversity loss in this herd. E_{nf} was only slightly higher than E_{fg}, which indicates that most of the genetic drift that occurred in the herd accumulated over non-founder generations. The ratio between E_f and E_a, which assesses possible bottlenecks, increased with time, indicating that the herd did not go through a recent bottleneck.

The ratio between E_{fg} and E_f indicates the magnitude of genetic drift and the lower the value, the greater the impact of genetic drift on diversity loss. In this study, these ratios were 0.25, 0.17, and 0.14 for the three periods, respectively, indicating that the importance of genetic drift increased over time. Nevertheless, the recent values of E_f, E_a, and E_{fg} are larger than those reported by Faria et al. (2009) for the entire registered Nellore population in Brazil, which indicates a greater genetic diversity in this particular evaluated herd.

Even though the pedigree records contained a large number of unknown sires, the use of VanRaden's method for calculating inbreeding coefficients, which takes into account the missing pedigree records, along with statistics based on the probability of gene origin, seemed suitable for assessing the genetic diversity in the herd. Although the analyses showed an increase in inbreeding rates and loss of diversity, the mating strategies used in the herd seem to be lessening the rates of inbreeding and genetic diversity loss. It is, however, advisable to continue monitoring the inbreeding rates and diversity of this population in the future.

ACKNOWLEDGMENTS

The authors thank Agropecuaria Jacarezinho and the Brazilian Association of Zebu breeders (ABCZ) for providing the data for this study. F.V. Brito was the recipient of a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

Berg P, Sørensen MK and Nielsen J (2007). EVA Interface User Manual. Available at [http://eva.agrsci.dk/]. Accessed February, 2009.

Boichard D (2002). Pedig: A Fortran Package for Pedigree Analysis suited for Large Populations. In: Proceedings of the 7th WCGALP. CD-ROM, Communication, Montpellier, 28-13.

Boichard D, Maignel L and Verrier E (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.* 29: 5-23.

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

Colleau JJ (2002). An indirect approach to the extensive calculation of relationship coefficients. *Genet. Sel. Evol.* 34: 409-421.

Colleau JJ and Sargolzaei M (2011). MIM: an indirect method to assess inbreeding and coancestry in large incomplete pedigrees of selected dairy cattle. *J. Anim. Breed. Genet.* 128: 163-173.

Faria FJ, Filho AE, Madalena FE and Josahkian LA (2002). Pedigree Analysis in the Brazilian Zebu Breeds. In: Proceedings of 7th World Congress on Genetic Applied to Livestock Production. CD-ROM, Communication, Montpellier, 26-29.

Faria FJ, Filho AE, Madalena FE and Josahkian LA (2009). Pedigree analysis in the Brazilian Zebu breeds. *J. Anim. Breed. Genet.* 126: 148-153.

Fernando RL and Gianola D (1986). Optimal properties of the conditional mean as a selection criterion. *Theor. Appl. Genet.* 72: 822-825.

Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, et al. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324: 528-532.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31: 423-447.

Kennedy BW, Schaeffer LR and Sorensen DA (1988). Genetic properties of animal models. *J. Dairy Sci.* 71: 17-26.

Lacy RC (1989). Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zool. Biol.* 8: 111-123.

Lutaaya E, Misztal I, Bertrand JK and Mabry JW (1999). Inbreeding in populations with incomplete pedigrees. *J. Anim. Breed. Genet.* 116: 475-480.

MacCluer JW, VandeBerg JL, Read B and Ryder OA (1986). Pedigree analysis by computer simulation. *Zool. Biol.* 5: 147-160.

Magnabosco CU, Cordeiro CMT, Trovo JBF, Mariante AS, et al. (1997). Catálogo de Linhagens do Germoplasma Zebuino: Raça Nelore. Embrapa-Cenargen, Documentos, Brasília.

Quinton M, Smith C and Goddard ME (1992). Comparison of selection methods at the same level of inbreeding. *J. Anim. Sci.* 70: 1060-1067.

Roso VM and Fries LA (1998). PAD - Um Programa para Planejar Acasalamentos em Bovinos de Corte. In: Proceedings of the II Simpósio Nacional da SBMA, Uberaba, 359-360.

Sargolzaei M, Iwaisaki H and Colleau JJ (2005). A fast algorithm for computing inbreeding coefficients in large populations. *J. Anim. Breed. Genet.* 122: 325-331.

Sargolzaei M, Iwaisaki H and Colleau JJ (2006). CFC: A Tool for Monitoring Genetic Diversity. In: Proceedings of the 8th WCGALP. CD-ROM, Communication, Belo Horizonte, 27-28.

VanRaden PM (1992). Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. *J. Dairy Sci.* 75: 3136-3144.

Verrier E, Colleau JJ and Foulley JL (1993). Long-term effects of selection based on the animal model BLUP in a finite population. *Theor. Appl. Genet.* 87: 446-454.

Vozzi PA, Marcondes CR, Magnabosco CU, Bezerra LAF, et al. (2006). Structure and genetic variability in Nellore (*Bos indicus*) cattle by pedigree analysis. *Genet. Mol. Biol.* 29: 482-485.

Vozzi PA, Marcondes CR, Bezerra LAF and Lôbo RB (2007). Pedigree analysis in the breeding program for Nellore cattle. *Genet. Mol. Res.* 6: 1044-1050.

Wright S (1922). Coefficients of Inbreeding and relationship. *Am. Nat.* 56: 330-338.

Wright S (1969). Evolution and the Genetics of Populations. The University of Chicago Press, Chicago.