N.L. Ribeiro
https://orcid.org/0000-0002-6410-244X
G.R. Medeiros
https://orcid.org/0000-0001-6544-1518
G.V. Nascimento
https://orcid.org/0000-0003-3548-3615
J.K.G. Arandas
https://orcid.org/0000-0002-0582-0978
M.N. Ribeiro
https://orcid.org/0000-0003-3373-3246

Analysis of the population structure of a cattle conservation nucleus Curraleiro Pé Duro

[*Análise da estrutura populacional de um núcleo de conservação de bovino Curraleiro Pé Duro*]

N.L. *Ribeiro* [1], G.R. *Medeiros*[1], G.V. *Nascimento*[1], J.K.G. *Arandas*[2], M.N. *Ribeiro*[2]

[1]Pesquisador do Instituto Nacional do Semiárido - Campina Grande, PB
[2]Universidade Federal Rural de Pernambuco - Recife, PE

ABSTRACT

The objective of this research was to study the population structure of the Cattle Conservation Nucleos Curraleiro Pé Duro of the Instituto Nacional do Semiárido (NCP_INSA) based on pedigree data. Genealogical information from 338 animals registered in the period from 1991 to 2019 was used. The number of founding animals (N_f), the effective number of founders (f_e), effective number of ancestors (f_a), inbreeding coefficient (F), and average relatedness coefficient (AR), in addition to F_{is}, F_{it} and F_{st} were estimated. It was possible to identify ancestors up to the third generation, with an increase in information over the generations. Of the total pedigree information evaluated, 90.53% had the identification of the father and mother. The effective size of the population was smaller than those proposed by FAO, suggesting the need to redefine the herd management and genetic management plan strategies, promoting gene flow and breed expansion.

Keywords: founders, inbreeding, pedigree

RESUMO

O objetivo com essa pesquisa foi estudar a estrutura populacional do Núcleo de Conservação de Bovinos Curraleiro Pé-Duro (NCP) do Instituto Nacional do Semiárido (INSA), por meio de dados de pedigree. *Utilizaram-se informações genealógicas de 338 animais registrados no período de 1991 a 2019. Foi estimado o número de animais fundadores (N_f), o número efetivo de fundadores (f_e), o número efetivo de ancestrais (f_a), o coeficiente de endogamia (F) e o coeficiente de parentesco médio (AR), além do F_{is}, F_{it} e F_{st}. Foi possível identificar ancestrais até a terceira geração, com aumento crescente das informações ao longo das gerações. Do total de informações avaliadas, 90,53% possuíam identificação do pai e da mãe. O tamanho efetivo da população foi inferior ao mínimo proposto pela FAO, o que sugere a necessidade de redefinir as estratégias do plano de gestão e de manejo genético do rebanho, de modo a promover fluxo gênico e expansão da raça.*

Palavras-chave: endogamia, fundadores, pedigree

INTRODUCTION

The several locally adapted breeds in Brazil were developed from specimens that were introduced by Portuguese settlers. The animals brought by the colonizers were subjected to natural and artificial selection in different environments, thus developing specific adaptation traits to different conditions, such as rusticity, prolificacy, and resistance to endo and ectoparasites (Egito *et al.*,

2002). Local breeds constitute a unique genetic heritage with substantial genetic variability for the maintenance of breeds of different domestic species. The Curraleiro Pé-Duro is a local and adapted breed, low production cost, and has tasty meat with low-fat content. Demographic studies developed in all northeastern region reveled around 6500 animals (Fioravanti *et al.,* 2011; Salles, 2013).

Despite great regional importance, the breed goes through a process of genetic dilution and a decrease in its effective size, which represents a significant threat to its survival. Conservation and preservation actions have been carried out with the breed as maintenance of genetic material in germplasm banks, genetic studies (Egito *et al.*, 2011; Montesinos *et al.*, 2011) and dissemination of genetic material by assigning and lending animals to breeders in the region. Efforts need to be implemented to increase this database and associate it with productive and genealogical data. Besides, it is necessary to monitor the breed status, mainly the population size, which is inversely proportional to inbreeding and is a manner to monitor the variability. Additionally, it allows the identification of founding, ancestral individuals, that is, those animals with relevant genetic contributions to the population (Tino *et al.*, 2020).

There are many studies of population structure and genetic variability, using pedigree data in cattle in Brazil (Malhado *et al.*, 2010; Rezende *et al.*, 2017). However, for local breeds of cattle raised in the Brazilian semiarid, studies are limited. The objective of this research was to study demographic and population structure of the Curraleiro Pé Duro cattle nucleus, from the National Institute of the Semi-Arid - INSA through pedigree data, available from the moment of implantation of the nucleus to the present day.

MATERIAL AND METHODS

Genealogical information from the Curraleiro Pé Duro Cattle Conservation nucleus, belonging to the National Semi-Arid Institute (INSA) located in Campina Grande, Paraíba State, Northeast Brazil, were used. 338 animals records were taken, where 125 were identified as breeding animals (13 fathers and 112 as mothers) and 99 young animals. These animals were submitted to the natural breeding season (1 male to 12 females), which occurs from July to September in order for calving to occur from May to July (rainy season) of the following year. The feeding was based on the native pasture of the region, but in the dry season, the animals were supplemented with buffel hay (*Cenchrus ciliaris*) and forage palm Mexican elephant ear (*Opuntia stricta Haw*).

Genealogical information between 1991 and 2006 was made available by the Brazilian Agricultural Research Corporation, the institution responsible for the founding herd of the current INSA conservation nucleus. The database consisted of the animal number, identification of father and mother, sex of the animal, and birth date of each animal. The data were submitted to the analysis of the effective number of founding (f_e) animals, obtained by the expression: $fe = 1/\Sigma(pi^2)$ (equation 1), where: p_i = proportion of genes of the living descendant population, with contribution from the founder i. The effective number of ancestors was estimated using the equation: $fa = \frac{1}{\Sigma_{j=1}^{a} q^2 j}$ (*equation* 2), where: q_j = marginal contribution of an ancestor j, which is the genetic contribution made by an ancestor that is not explained by other ancestors chosen before.

The inbreeding coefficient (F) is defined as the probability that an individual has two identical alleles by descent, and was estimated according to Meuwissen and Luo (1992), in which the average F of a given generation t (F_t) was determined by using equation 2 (Falconer and Mackay, 1996): $F_t = 1 - (1 - \Delta F)t$. The increase in inbreeding (ΔF) is calculated for each generation using the classical formula $\Delta F = (F_t - F_{t-1})/(1 - F_{t-1})$ (equation 3), where F_t and F_{t-1} represent the average inbreeding estimates for the current and the previous generation, respectively (Falconer and Mackay, 1996). The obtained values for ΔF were used to calculate the effective number (N_e) (Gutiérrez *et al.*, 2003), using equation 4: $Ne = 1/2\Delta F$.

The average relatedness coefficient (AR) of each individual is defined as the probability that an allele randomly was chosen from the whole population in the pedigree belongs to a given animal (Gutiérrez and Goyache, 2005). This parameter was calculated using an algorithm, described in Goyache *et al.* (2003). Fixation indexes or Wright's F statistics were calculated according to Wright (1978) using the formulas: $F_{IS} = \frac{\tilde{F} - \bar{f}}{1 - \bar{f}}$ (equation 4), $F_{ST} = \frac{\tilde{f} - \bar{f}}{1 - \bar{f}} = \frac{\bar{D}}{1 - \bar{f}}$ (equation 5) and $F_{IT} = \frac{\tilde{F} - \bar{f}}{1 - \bar{f}}$ (equation 6). Where: $\tilde{F}$ is the consanguinity coefficient and kinship of the subpopulations and $\tilde{f}$ the average kinship for the metapopulation, so that: $(1 - F_{IT})$

$= (1 - F_{IS})(1 - F_{ST})$ (equation 7). The data were analyzed using ENDOG v4.6 software.

RESULTS AND DISCUSSION

It was possible to identify ancestors up to the third generation, with a high percentage of fathers and mothers identified in the current population and an increase in information over the generations (Figure 1). Of the total pedigree information evaluated for the 338 animals, 90.53% had the identification of the father and mother.

As the ancestry increases, the percentage of known father and mother decreases, characterizing the low structure of the pedigree, which can compromise the assessment of the current situation of breed. Therefore, the higher the number of known generations more accurate the estimated parameters. The information of an individual's ancestors (parents, grandparents, and great-grandparents) allows better mating monitoring, avoiding increased inbreeding. The depth of pedigrees is directly reflected in the quality estimates of important demographic and population parameters, such as inbreeding and the generation interval, for example (Oliveira, 2012).
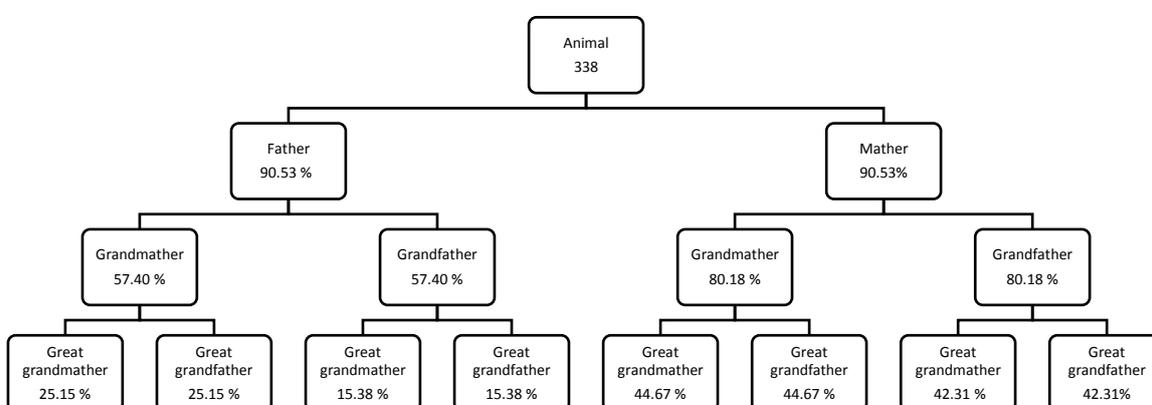

Figure 1. Pedigree structure of Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

The individuals responsible for the breed genetic variability were determined from the base population; that is, the number of founding and ancestral animals (Table 1). The f_a/f_e ratio was equal to 1, indicating that ancestors and founders contributed equally to the formation of the studied population. However, the values are low concerning the total number of individuals in the population, indicating a loss of origin genes.

Table 1. Population parameters of the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

Parameters	
Total number of animals	338
Base population (with at least one unknown relationship) - Nf	32
Number of animals in the reference population	306
Number of ancestors that gave rise to the reference population	30
Effective number of founding animals ($\widetilde{f e}$)	7
Effective number of ancestors ($\widetilde{f a}$)	7
Number of ancestors that explain 50%	3
F (%)	4.77
AR (%)	16.17

F = inbreeding coefficient; AR = average relatedness coefficient

The f_e/f_a ratio equal to 1 is desirable because it indicates an equal use of founders and ancestors. However, f_a and f_e values were low, characterizing a founding effect. Rezende *et al.* (2017) evaluated the population structure of Nellore cattle raised on Northeastern Semi-Arid and observed that the *fe/fa* ratio of 1.99, suggesting unequal use of founders and ancestors. The average inbreeding values (F) obtained in this study were less than 10%, even with the excessive use of some males. Higher values represent a high risk of inbreeding depression (Santana *et al.,* 2010). However, the almost inexistence of individual inbreeding is not only since there are no inbreeding animals within the herd, but also due to the difficulty of estimating this parameter in the first generation. This fact was also verified by Tino *et al.* (2020) in Nellore cattle in the semiarid.

The significant difference between the number of animals (N_f) and the actual number of founders (f_e) (Table 2) indicates that few founders contributed to the genetic diversity of the studied population. These results may be due to the loss of genetic diversity since the founders contributed differently, probably due to the use of a few paternal strains, as mentioned by Texeira Neto *et al.* (2013). The number of ancestors that explain 50% of the genetic variability of the population studied is concentrated in only three animals. This result is an indication that the current population has developed from a narrow genetic base, which may have contributed to the loss of genetic material over the generations.

The relationship between the inbreeding rate and the average relatedness (AR) coefficient is shown in Figure 2. The null values for F in the early years are common once founder animals have an unknown genealogy, so it is impossible to estimate the inbreeding coefficient. It was observed high F from 2006 to 2011 (12.5%), reflecting on the increase in the AR coefficient. After the year 2012, the F coefficient increased due to the increase in mating between relatives. In the period from 2013 to 2017, a severe drought occurred, generating the need to reduce the number of animals and enable the maintenance of the nucleus herd. This measure was responsible for an increase in the number of inbreeding animals (F >25%) and an increase in the F/AR ratio in the period.
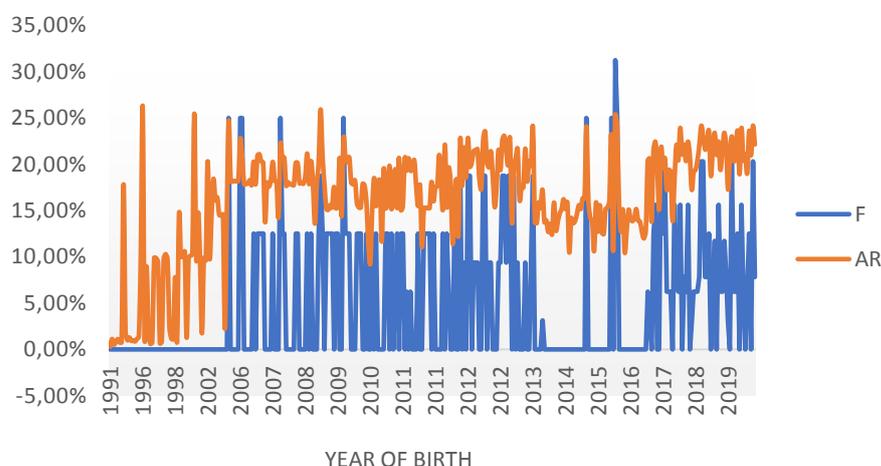

Figure 2. Relationship between inbreeding (F) and average relatedness coefficient (AR) per year of birth in the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

The AR values from 1991 to 2002 increased, indicating that animals born in that period had a strong genetic contribution to the population, a fact confirmed in Figure 3. From 2006, there was a clear tendency of increase the inbreeding coefficient. This increase was expected since the individual inbreeding coefficient depends on pedigree information. Thus, the more the genealogy of the individual's ancestors is known, the more reliable the estimated inbreeding coefficient (Muniz *et al.,* 2012). Table 2 shows the founding ancestors that most contributed to the variability of the study population. Among the 30 founders, much of the 50% variability was concentrated in just three individuals.
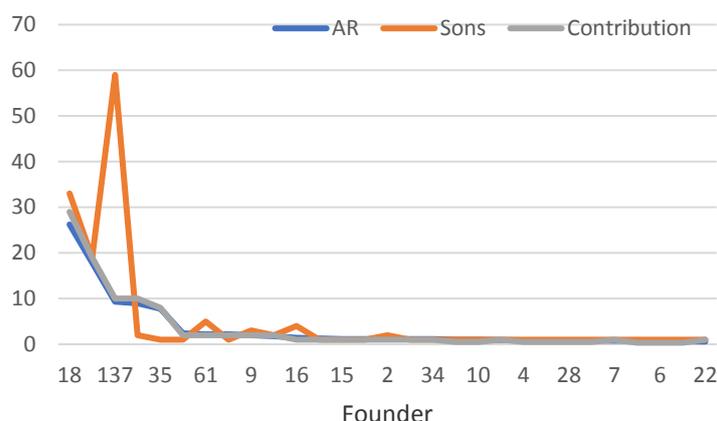
Figure 3. Relationship between the founders of the herd, the kinship coefficient, the number of sons, and the genetic contribution of the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

Table 2. Description of ancestors and founders who most contributed to the genetic variability of the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

Classification	Founder	Year of birth	Sex	AR (%)	Number of sons	Contribution (%)
1	18	1996	Male	26.24	33	29
2	8	1993	Male	17.83	19	19
3	137	2010	Male	9.32	59	10
	Ancestral					
1	18	1996	Male	26.24	33	29
2	8	1993	Male	17.83	19	19
3	137	2010	Male	9.32	59	10

AR = Average relatedness coefficient

There were no females in the group of founding animals, indicating an imbalance in the male: female ratio in the base population. This fact suggests an unfavorable situation to the breed conservation, with little capacity to respond to selection. According to Faria *et al.* (2006), the preservation of genetic variability in a given population is essential for long-term conservation and selection. Moreover Teixeira Neto *et al.* (2013) suggest greater control of mating as a strategy to control inbreeding when the number of founders is reduced. Animals number 18 and 137

were the ones that most contributed genetically to the population, with 33 and 59 offspring, respectively. The estimated AR for animal 18 was 26.24%, which is a very high value. The null value for F in the first generation occurred because they are founding animals, without pedigree information to make it possible to calculate. However, the AR values were increasing (3 to 21%), indicating that animals born at that time had a high genetic contribution to the population (Table 3).

Table 3. Values of inbreeding (F) and the average coefficient of kinship (AR), percentage of inbreeding individuals (End), an average of F for inbreeding individuals (FEnd) and Effective number (N_e) of the Curraleiro Pé Duro cattle breed conservation center by the number of known complete generations

Generation	Number	F (%)	AR (%)	End (%)	FEnd(%)	Ne
0	32	0	3	9	20	-
1	116	2	15	54	13	28
2	143	7	19	85	10	9
3	47	8	21	9	20	32

The average relatedness (AR) has been of practical use in conservation programs and has been used for population management, replacing inbreeding as proposed by Fenandez *et al.* (2009). In these cases, animals with lower AR values should be preferred in the mating, allowing greater participation of animals that are being underused. This measure contributes to reducing inbreeding, minimizing possible losses of origin genetic material. In small herds, this trend is common, since with the advancement of generations, the chances of inbreeding mating increase, demanding more rigorous mating control. However, in herds with well-planned genetic management, it is possible to keep lows levels of AR (Rodrigues *et al.*, 2009). It was observed that the average inbreeding coefficient

(F) suffered a gradual increase until the fourth generation (Table 4); the same trend occurred with the percentage of inbreeding individuals. The average inbreeding coefficient (F) ranged from 6% in the second generation to 3% in the fifth, reaching a peak in the fourth generation (8%).

In the present study, the AR varied from 3 to 21%, with its peak in the fifth generation. Rezende *et al.* (2017) analyzing the population structure of a Nelore herd raised in the semiarid region of Northeast Brazil, observed AR values between 0.80% and 0.35% from the first to the fourth generation, decreasing after the fifth generation. The small number of known ancestors is corroborated by the low average number of generations (Table 5).

Table 4. Values of inbreeding (F), relatedness coefficient (AR), percentage of inbreeding individuals (End), the average of F for inbreeding individuals (FEnd), and Effective number (N_e) of the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido, by the number of known maximum generations

Generation	Number	F (%)	AR (%)	End (%)	FEnd(%)	Ne
0	32	0	3			-
1	31	0	12			-
2	109	6	18	45	14	8
3	111	4	18	42	10	-
4	54	8	20	56	15	22
5	1	3	21	1	3	-

Table 5. The average number of traced generations (TG), increase in inbreeding (F), and effective size (N_e) by the traced generation of the Curraleiro Pé Duro cattle breed nucleus of the Instituto Nacional do Semiárido

Generation types	TG	F(%)	N_e
Complete[1]	1,61	3,47	14,39
Maximum[2]	2,38	1,76	28,48
Equivalents[3]	1,96	3,34	14,99

[1] Farthest generation in which all ancestors are known

[2] Number of generations that separate the individual from his most remote ancestor

[3] Sum of terms (1/2) n of all known ancestors, where n is the number of generations that separate the individual from each known ancestor

Therefore, when the pedigree is shallow, inbreeding and N_e are underestimated (Rezende *et al.*, 2017). The values found for the different types of generations are lower than the minimum values proposed by FAO (Guidelines…, 2012) (N_e minimum = 50). The N_e affects the dynamics of the population, which in turn can influence the rate and size of the allele frequency of a given genotype. When below the minimum recommended by FAO, it directly affects variability with adverse effects on the conservation and genetic improvement (Oliveira, 2012).

The obtained values for F_{is}, $F_{st,}$ and F_{it} were -0.156844; 0.104356 and -0.036120, respectively. The fixation index between subpopulations (F_{is}) considers values close to 1 with a higher degree of inbreeding within subpopulations. Despite the minimal number of animals since the foundation of the nucleus, a high frequency of heterozygotes was observed, as shown by the negative F_{is} and F_{it} values (Rodrigues *et al.*, 2009). To increase the breed variability, it is necessary to introduce and use genetic material from other herds, to promote gene flow, an important strategy to avoid isolation of the herd and maintain genetic diversity. The

exchange of genetic material favors the increase in the effective number and consequently decreases the rate of inbreeding in the herd. Despite the small numbers, there is still diversity, which can be safeguarded for use in future conservation and breeding programs.

CONCLUSION

The effective population size is smaller than that proposed by FAO to maintain a minimum loss of genetic diversity per generation. The balanced use of available males and females and the exchange of males between existing nuclei in Brazil may be the first step to recover the effective number and minimize genetic losses in the short term.

REFERENCES

EGITO, A.A.; FIORAVANTI, M.C.S.; GRATTAPAGLIA, D. *et al*. Origem e diversidade genética materna de populações de bovinos da raça Curraleira de diferentes regiões do Brasil. *Actas Iberoam. Cons. Anim*., v.1, p.110-113, 2011

EGITO, A.A.; MARIANTE, A.S.; ALBUQUERQUE, M.S.M. Programa brasileiro de conservação de recursos genéticos animais. *Arch. Zootec.*, v.51, p.39-52, 2002.

FALCONER, D.S; MACKAY, T.F.C. *Introduction to quantitative genetics.* Harlow: Longman, 1996

FARIA, F.J.C.; VERCESI FILHO, A.E.; MADALENA, F.E. *et al*. Genetic structure of the registered polled Gir breed. *Bol. Ind. Anim*., v.63, p.135-141, 2006.

FERNANDEZ. C.; MARTINÉZ, B.; GÓMEZ, E. *et al*. Qualitative analysis of official milk control in Valencia Community (Spain) by self-organizing maps. *Trop. Subtrop. Agroecosyst.,* v.11, p.91-94, 2009.

FIORAVANTI, M.C.S.; JULIANO, R.S.; COSTA, G.L. *et al*. Conservación del bovino Curraleiro: cuantificación del censo y caracterización de los criadores. *Anim. Genet. Res.*, v.48, p.109-116, 2011.

GOYACHE, F.; GUTIÉRREZ, J.P.; FERNÁNDEZ, I. *et al.* Using pedigree information to monitor genetic variability of the endangered population: the Xalda sheep breed of Asturias as an example. *J. Anim. Breed. Genet.*, v.120, p.95-105, 2003.

GUIDELINES for the in vivo conservation of animal genetic resources (Draft). Rome: FAO, 2012. 160p.

GUTIÉRREZ, J.P.; ALTARRIBA, J.; DÍAZ, C. *et al*. Pedigree analysis of eight Spanish beef cattle breeds. *Genet. Select. Evol.*, v.35, p.43-63, 2003.

GUTIÉRREZ, J.P.; GOYACHE, F.A. Note on ENDOG, a computer program for analyzing pedigree information. *J. Anim. Breed. Genet.,* v.122, p.172-176, 2005.

MALHADO, C.H.M.; CARNEIRO, P.L.S.; MALHADO, A.C.M. *et al*. History of registered Gyr breed in Brazilian Northeast: population structure and genetic improvement of growth traits. *Cienc. Rural*, v.40, p.1385-1391, 2010.

MEUWISSEN, T.I.; LUO Z. Computing inbreeding coefficients in large populations, *Genet. Select. Evol.,* v.24, p.305-313,1992.

MONTESINOS, I.S.; NETO, J.O.C.; DODE, M.A.N. *et al*. Avaliação espermática de sêmen criopreservado de touros da raça Curraleira em banco de germoplasma: dados preliminares. In: CONGRESSO DE PESQUISA, ENSINO E EXTENSÃO, 8., 2011, Goiânia. *Anais...* Goiânia: SBPC, 2011. Disponível em: http://www.sbpcnet.org.br/livro/63ra/resumos.htm. Acessado em: 17 maio 2020

MUNIZ, L.M.S.; SOUZA, L.A.; BARBOSA, A.C.B. *et al.* A raça Gir mocha na região nordeste do Brasil: estrutura genética populacional via análise de *pedigree. Arq. Bras. Med. Vet. Zootec.,* v.64, p.1656-1664, 2012.

OLIVEIRA, R.R. *Demografia e estrutura populacional da raça caprina Murciano-Granadina na Espanha com base em análise de pedigree.* 2012. 86f. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Recife, PE.

REZENDE, M.P.G.; CONDE, E.A.S.L.; BORGES, A.C. *et al*. Estrutura populacional do rebanho nelore criado no Semiárido do Nordeste brasileiro. *Ciênc. Anim. Bras.,* v.18, p.e-38048, 2017.

RODRIGUES, D.S.; RIBEIRO, M.N.; OLIVEIRA, S.M.P. *et al*. Estrutura populacional de um rebanho da raça Morada Nova como contribuição para a conservação. *Ciênc. Anim.*, v.19, p.103-110, 2009.

SALLES, P.A. *Estado atual de conservação da raça bovina Curraleiro Pé Duro na região Nordestina brasileira*. Campina Grande: Instituto Nacional do Semiárido – INSA, 2013. 15p. (Documentos Técnicos, n.3).

SANTANA, M.L.; OLIVEIRA JR., P.S.; ELER, J.P. *et al*. Pedigree analysis and inbreeding depression on growth traits in Brazilian Marchigiana and Bonsmara breeds. *J. Anim. Sci.*, v.90, p.99-108, 2010.

TEIXEIRA NETO, M.R.; CRUZ, J.F.; CARNEIRO, P.L.S. *et al*. Parâmetros populacionais da raça ovina Santa Inês no Brasil. *Pesqui. Agropecu. Bras.,* v.48, p.1589-1595, 2013.

TINO, C.S.R.; CAVANI, L.; FONSECA, R. *et al*. Análise de estrutura populacional de ovinos deslanados do núcleo de conservação. *Arq. Bras. Med. Vet. Zootec.*, v.72, p.560-564, 2020.

WRIGHT, S. *Evolution and the genetics of populations. Variability within and among natural populations.* Chicago: University of Chicago Press, 1978. v.4.