

Genetic diversity in oocyte donors used in *in vitro* bovine embryo production programs in Brazil

Diversidad genética en donadoras de ovocitos utilizadas en programas de producción de embriones bovinos in vitro en Brasil

Diversidade genética em doadoras de oócitos utilizadas em programas de produção de embriões bovinos in vitro no Brasil

Wilder Hernando Ortiz Vega^{1*}, MVZ, PhD;

Celia Raquel Quirino¹, Agronomist Engineer, PhD;

Aylton Bartholazzi Junior¹, MV, PhD;

Clara Slade Oliveira², MV, PhD;

Raquel Varella Serapião³, MV, PhD;

Caroline Marçal Gomes David¹, MV, PhD;

Miguel Alejandro Silva Rua¹, MV, PhD;

Ana Carolina Barros de Freitas¹, MV, PhD.

✉ To cite this article: Ortiz Vega WH, Quirino CR, Bartholazzi Junior A, Oliveira CS, Varella Serapião R, Gomes David CM, Silva Rua MA, Barros de Freitas AC. Genetic diversity in oocyte donors used in *in vitro* bovine embryo production programs in Brazil. Rev Colomb Cienc Pecu 2019; 32(2):90-99.

* Corresponding author: Wilder Hernando Ortiz Vega. Laboratório de Reprodução e Melhoramento Genético Animal, Universidade Estadual do Norte Fluminense (LRMGA/UENF), Av. Alberto Lamego, 2000, Campos dos Goytacazes, Rio de Janeiro, CEP: 28013-602, Brazil. Tel.: 229991711888. E-mail: wilortvet@yahoo.es

¹*Laboratório de Reprodução e Melhoramento Genético Animal. Universidade Estadual do Norte Fluminense. Campos dos Goytacazes, Rio de Janeiro. CEP: 28013-602. Brasil.*

²*Embrapa Gado de Leite, Laboratório de Reprodução Animal, Campo Experimental Santa Mônica (CESM), Valença, Rio de Janeiro. Brasil.*

(Received: September 8, 2017; accepted: November 2, 2018)

Abstract

Background: Current reproductive management of bovine elite populations involves the use of assisted reproductive technologies (ARTs), aiming to obtain the greatest genetic gain. However, inadequate use of ARTs may lead to loss of genetic diversity in the offspring.

Objective: To assess the genetic diversity in elite female cattle populations used in commercial *in vitro* embryo production.

Methods: Using genetic and ecological approaches for the study of populations based on microsatellite markers, we assessed the genetic diversity between and within populations of cows used in commercial *in vitro* embryo production programs in Brazil.

Results: Endogamy within populations varied from zero to 9.1%, while heterozygosity between populations (F_{ST}) was <0.05 in the different population interactions. AMOVA showed 1% variation between populations, 8% between individuals and 91% within individuals. The dimensionality reduction method utilized indicated a lack of structure in the populations analyzed, identifying two main clusters in the three populations.

Conclusions: Low genetic diversity between cow populations associated with commercial programs of *in vitro* embryo production in Brazil was evidenced. Variable levels of endogamy within the populations were observed. Approaches of population genetics as well as ecological diversity can be implemented to more thoroughly estimate genetic diversity in livestock populations.

Keywords: *allele frequencies, heterozygosity, inbreeding, microsatellite markers, oocyte.*

Resumen

Antecedentes: El actual manejo reproductivo en poblaciones de bovinos de élite incluye la utilización de tecnologías de reproducción asistida (ARTs) con el fin de obtener mayor ganancia genética. Sin embargo, el uso inadecuado de las ART puede llevar a la pérdida de diversidad genética en los descendientes.

Objetivo: Evaluar la diversidad genética en poblaciones de vacas de élite utilizadas en la producción comercial de embriones bovinos *in vitro*.

Métodos: Utilizando abordajes de la genética y ecología de poblaciones basados en marcadores microsatélites, evaluamos la diversidad genética entre y dentro de poblaciones de vacas participantes de programas comerciales de producción de embriones *in vitro* en Brasil.

Resultados: La endogamia dentro de las poblaciones varió de cero a 9,1%, mientras que la heterocigosidad entre poblaciones (F_{ST}) fue $<0,05$ en las diferentes interacciones de la población. El AMOVA mostró variación del 1% entre poblaciones, 8% entre individuos y 91% dentro de individuos. El método de reducción de dimensionalidad utilizado indicó una falta de estructura en las poblaciones analizadas, identificando dos

grupos principales en las tres poblaciones.

Conclusiones: Se evidenció una baja diversidad genética entre las poblaciones de vacas asociadas a programas comerciales de producción de embriones *in vitro* en Brasil. Se evidenciaron niveles variables de endogamia entre las poblaciones. Abordajes de la genética poblacional, así como de diversidad ecológica pueden ser implementados para estimar de manera más amplia la diversidad genética en poblaciones animales de interés pecuario.

Palabras clave: *endogamia, frecuencia alélica, heterocigosidad, marcadores microsatélites, ovocito.*

Resumo

Antecedentes: O atual manejo reprodutivo das populações de elite em bovinos envolve o uso de tecnologias de reprodução assistida (ARTs), visando obter o maior ganho genético. No entanto, o uso inadequado de ARTs pode levar à perda de diversidade genética na prole.

Objetivo: Avaliar a diversidade genética em populações de vacas de elite utilizadas na produção comercial de embriões bovinos *in vitro*.

Métodos: Utilizando abordagens da genética e ecologia de populações baseadas em marcadores microsatélites, foi avaliada a diversidade genética entre e dentro das populações de vacas participantes de programas comerciais de produção *in vitro* de embriões.

Resultados: A endogamia dentro das populações variou de zero a 9,1%, enquanto a heterocigosidade entre populações (F_{ST}) foi $<0,05$ nas diferentes interações populacionais. AMOVA mostrou variação de 1% entre populações, 8% entre indivíduos e 91% dentro de indivíduos. O método de redução de dimensionalidade utilizado indicou uma falta de estrutura nas populações analisadas, identificando dois clusters principais nas três populações.

Conclusões: Baixa diversidade genética entre populações de vacas associadas a programas de produção *in vitro* de embriões foi evidenciada. Níveis de endogamia variáveis dentro das populações foram observados. Abordagens da genética populacional assim como de diversidade ecológica podem ser implementadas na tentativa de estimar de maneira mais abrangente a diversidade genética em populações animais de interesse pecuário.

Palavras-chave: *endogamia, frequência alélica, heterocigosidade, marcadores microsatélite, oócito.*

Introduction

Currently, a large part of the genetic transfer material in cattle occurs through the use of female-mediated reproductive biotechnologies. In this regard, the Brazilian market for *in vitro* production of cattle embryos accounts for approximately 70% of the global production (IETS, 2016). The main impact female-mediated assisted reproductive technologies (ARTs), such as multiple ovulation and embryo transfer (MOET) and ultrasound-guided follicular puncture, or ovum pick-up (OPU), followed by *in vitro*

fertilization and embryo transfer (IVF) is focused on the creation and dissemination of the genetic gain, as well as reduction of the interval between generations in the populations of the animals involved, through the production of superior individuals (Pryce *et al.*, 2010).

The greater participation of females in genetic improvement programs associated with ARTs can be attributed, among other aspects, to the ease of estimating the genetic merit for traits directly linked to sex (Dassonneville *et al.*, 2012).

Recent studies using stochastic and deterministic simulation models have shown that the genetic gain obtained with the use of ARTs in cattle populations can be significant. Increases from 23% (with MOET) to 98% (with OPU-IVF) have been obtained (Pedersen *et al.*, 2012; Pryce *et al.*, 2010). However, due to the small number of progenitors that contribute to the next generation, the additional genetic gain can be accompanied by an increase in endogamy within the population and among descendants, with the consequent reduction of genetic diversity (Pedersen *et al.*, 2012).

Considering that sufficient genetic variation is necessary in animal populations, both for adaptation and resistance, and for ongoing genetic improvement of traits with economic importance (Biscarini *et al.*, 2015), the objective of this work was to assess the genetic diversity in elite female cattle populations used in commercial *in vitro* embryo production.

Material and methods

Ethical considerations

This study was approved by the Ethics Committee on Animal Experimentation of Norte Fluminense State University (UENF Protocol no. 243, March 11, 2014).

Population studied

Fifty adult dairy cows of the Gir breed (*Bos taurus indicus*) were evaluated, used as oocyte donors in commercial programs existing for more than four years in the Brazilian industry for *in vitro* embryo production (Table 1).

The samples were collected in three populations on farms located in three municipalities in the state of Rio de Janeiro, Brazil: Campos dos Goytacazes (Population 1, n= 6; 21°45' 16" South and 41°19'28" West), Rio das Flores (Population 2, n= 27; 22° 9' 32" South and 43° 34' 52" West) and Valença (Population 3, n= 27; 22°14'44" South and 43°42' 01" West). To assure representativeness of the samples, all donor cows of each farm were included. The animals studied were born in each farm; they were self-sufficient herds in repositioning of females.

DNA extraction and genotyping

The genomic DNA was obtained from tail hair follicles. The extraction and purification steps were carried using the NucleoSpin tissue kit (Macherey- Nagel, Düren, Germany), following the manufacturer's instructions. The DNA concentrations were measured with a spectrophotometer (NanoDrop™ 2000-Thermo Science).

The animals were genotyped using 20 microsatellite markers recommended for routine identification of kinship

and paternity in cattle by the International Society for Animal Genetics (ISAG), 14 of them belonging to the main group of markers (BM1818, BM1824, BM2113, ETH10, ETH225, ETH30, INRA23, INRA63, SPS115, TGLA122, TGLA126, TGLA227, TGLA53 and TGLA57), and 6 to the additional group (CSSM66, ETH152, ILSTS06, INRA05, INRA172 and AE129). The genotyping was performed with a capillary sequencer (MegaBACE 1000 DNA Analysis System – GE Healthcare). The marker ILTS06 was excluded from the analysis due to its scarcity of amplification in the genotyped animals.

Population indicators of genetic diversity

The basic descriptive statistics of the population genetics were used to estimate the diversity within the breed divided into sub-populations, using the GenAlEx software, version 6.502 (Peakall and Smouse, 2012). The following parameters were estimated: number of different alleles (N_a), number of effective alleles (N_e), number of private alleles (N_p), percentage of polymorphic loci (PI), observed heterozygosity (H_o), expected heterozygosity (H_e), and fixation index (F). To evaluate the genetic diversity and estimate the variance components of the populations, analysis of molecular variance (AMOVA) was applied based on microsatellite loci.

Initially, the genetic diversity distribution between the populations was studied by analysis of the genetic distances of Nei (Nei, 1977), and the F -statistics of Wright (F_{IS} - f , F_{ST} - θ and F_{IT} - F) (Wright, 1921); the significance was tested using bootstrapping of loci after 1,000 permutations of alleles within a population. Fisher's exact test was applied to detect global and population deviations (per locus) from Hardy-Weinberg equilibrium (HWE).

The population allelic richness (A_R) and private allelic richness (pA_R) for the 19 markers tested were estimated by the rarefaction method using the HP- RARE 1.0 software. Analysis of variance (ANOVA) was applied to the allelic richness measures involving the 19 markers. Statistical differences between populations were evaluated with the SNK test using the MIXED procedure of the SAS® software, versión 9.2 (SAS Institute Inc., Cary, NC, USA) (2009).

Gene flow and population structure

The gene flow between populations was estimated by the indirect method based on the number of migrants per generation (Nm), using the value of F_{ST} in the formula $Nm = [(1/ F_{ST}) - 1] / 4$ (Gaggiotti *et al.*, 1999).

The genetic differentiation pattern between individuals of the three populations was established by analysis of genetic distances of Nei, calculated with the GenAlEx software, version 6.502, using a shared allele distance matrix, which in turn was used to construct the dendrogram by means of the unweighted pair group method with arithmetic mean (UPGMA). The result of the matrix was visualized in a heat map (1-RD) (heatmap.2 function gplots in the R package), which revealed genetic relations between individuals of the populations.

An alternative model to analyze the population structure based on Bayesian grouping was used. For this analysis, the STRUCTURE 2.3.4 program was used. The simulation was carried out using the ancestry by mixture model (admixture) with correlated frequencies and burn-in period of 200,000 rounds followed by

500,000 MCMC (Markov chain Monte Carlo) iterations. Independent executions of “K” testing from 1 to 10 clusters with 25 repetitions were performed to confirm the consistency of the results. The ideal “K” value was chosen after analysis of the results file by the Evanno method with the web software STRUCTURE Harvester, versión 0.6.94. After identification of the ideal number of sub-populations, the last analysis was conducted by selecting the optimal “K” to generate the plot illustrating the population structure.

Results

Main characteristics of the markers

Among the 19 microsatellite markers tested, 18 were polymorphic in all the populations, for a total of 129 alleles detected. Only INRA63, in Population 1, was monomorphic. The markers were highly informative (polymorphic information content – PIC >0.5). The mean PIC value for the three populations was 0.58. The percentage of polymorphic loci (PI) varied from 95 (Population 1) to 100% (Populations 2 and 3). Deviations from Hardy-Weinberg equilibrium (HWE) were significant ($p < 0.001$) in Populations 2 and 3 (markers CSSM66 and TGLA126), and in Population 3 (marker INRA23, $p < 0.05$).

Genetic differentiation of the populations

Basic population statistics. The descriptive statistics of the genetic diversity of the three populations of oocyte donor cows used in the commercial programs for in vitro production of embryos are reported in Table 1. Variations were observed from 1.0 to 9.0 in the number of effective alleles (N_e), from 1.0 to 5.45 in the number of different alleles (N_a), and from 0 to 6 in the number of private alleles (N_p).

The potential for long-term adaptability and persistence of the populations was estimated by evaluating the allelic richness (AR) (Greenbaum et al., 2014). The final estimates of AR were significantly lower ($p < 0.05$) in Population 1 (3.53) than in Population 3 (4.77). The average allelic richness value in the three populations was 4.2 alleles. The estimate of private allelic richness (pAR), showed variations from 0.21 (Population 1) to 1.06 alleles (Population 3), and was also significantly lower ($p < 0.005$) in Population 1 compared to Population 3. The total value of pAR including all the individuals and all the loci tested was 1.88 alleles (Table 1).

The estimates of observed heterozygosity (H_o) were larger than the expected heterozygosity (H_e) in Populations 1 and 2, while in Population 3 the observed genetic diversity was lower than expected (H_o of 0.57 and H_e of 0.62).

The fixation index (F) presented the same pattern described for heterozygosity. It was negative in Populations 1 and 2 (because $H_o > H_e$), and 0.08 in Population 3 (Table 1).

Distribution of genetic diversity

Fixation indices F-statistics of Wright. The estimates of the F-statistics of Wright allow measuring the deficiency or excess of heterozygotes in a population in various ways: within sub-populations (F_{IS}), between populations in relation to the expected diversity in the total population (F_{ST}), and of a single

individual in relation to the total population (F_{IT}) (Nei, 1977). Negative values of F_{IS} and F_{IT} were found in Populations 1 and 2 (indicating high heterozygosity), while Population 3 was endogamic, with values of 0.091 and 0.035, respectively. In the case of F_{ST} , the values were negative in Populations 2 and 3, indicating weak or no differentiation between these populations, but in Population 1 the value was 0.086 (F_{ST} values greater than 0.05 for microsatellite markers are considered indicative of differentiation between populations, on a scale from zero to one (Holsinger & Weir, 2009). Table 2 presents the averages per population and the global values (including all the markers and all the individuals as a single population) of the F-statistics as well as the number of migrants per generation (Nm).

Table 1. Summary of genetic diversity indices (average) across three oocyte donor cow populations in 19 loci.

Population	Na	Ne	Np	A _R	pA _R	PI	Ho	He	F
1	3.53	2.51	0.11	3.53	0.21	0.95	0.59	0.54	-0.10
2	4.95	2.93	0.53	4.30	0.61	1.00	0.64	0.60	-0.07
3	6.11	3.19	1.58	4.77	1.06	1.00	0.57	0.62	0.08
Overall	4.86	2.87	2.22	4.20	1.88	0.98	0.60	0.59	-0.03

Na= number of different alleles; Ne= number of effective alleles; Np = number of private alleles; A_R= allelic richness; pA_R = private allelic richness; PI= percentage of polymorphic loci; Ho= observed heterozygosity; He= expected heterozygosity, and F= fixation index.

Table 2. Global F-statistics and estimates of Nm across three donor cow populations in 19 loci.

Population	F _{IS}	F _{IT}	F _{ST}	Nm
1	-0.097	-0.002	0.086	2.65
2	-0.072	-0.098	-0.02*	**
3	0.091	0.035	-0.06*	**
Overall	-0.022	0.015	0.039	11.036

* Negative F_{ST} values indicate weak or no differentiation, and are interpreted as zero. ** The result of the number of migrants (Nm) (indirect method to estimate the genetic flow between populations) when the F_{ST} value is negative lacks biological significance. F_{IS} , the correlation between alleles within an individual relative to the sub-population to which that individual belongs; F_{IT} , the correlation between alleles within an individual relative to the entire population; and F_{ST} , the correlation between alleles chosen randomly from within the same subpopulation relative to the entire population.

Gene flow and population structure

Gene flow is estimated from the number of migrants per generation (Nm), so an increase in this parameter is inversely related to the degree of diversity between geographically separated populations. A low Nm value leads to divergence between populations via selection and drift, possibly leading to speciation (Slatkin, 1985). When populations are small and the number of microsatellites is < 20 , the value of F_{ST} is considered the best estimator or Nm in a population (Gaggiotti *et al.*, 1999). The average Nm estimated between the populations in this study was 11,036 individuals, indicating a high genetic exchange rate.

Analysis of molecular variance (AMOVA)

The distribution of the variability of genetic diversity between and within populations measured using the

genetic distance matrix, including all the pairs of genotypes found by the AMOVA, revealed that 8% of the total variation was due to differences between individuals within the populations, 91% was due to heterozygosity in the individuals, and that variation between the populations only represented 1% of the total variation (Table 3).

Table 4 presents results of the diversity measures comparing the populations. The values of scaled diversity, F_{ST} and Nm between the populations were significant ($p < 0.001$) only between Populations 2 and 3. In general, the Shannon information analysis revealed low differentiation of the three populations.

The relationship between Populations 2 and 3 can be considered the least narrow due to the low Nm value (17.36), versus 89.91 for Populations 1 and 2, and 34.32 between Populations 1 and 3. The highest “species” diversity was found between Populations 2 and 3 (1.244).

Genetic distances and relations between individuals

The dendrogram based on Nei’s genetic distances grouped 50 individuals of the three populations in three clusters, two of them grouping the great majority of individuals (Clusters 2 and 3); with Cluster 1 only containing 10% of the animals (Figure 1). The proximity between individuals can also be assessed in the heat map, which shows the low level of structuring of the populations, evidenced by the mixture among the components in the three clusters formed (Figure 1).

The alternative approach used in delineating the clusters of individuals based on their genotypes as revealed by the 19 microsatellite markers, using the Bayesian method (employed in the STRUCTURE software), resulted in similar grouping to that obtained by the principal component analysis in relation to the structure. In turn, the grouping identified $K = 2$ as the most probable number of clusters (or sub-populations) in the individuals studied (Figure 2). In the plot, it can be observed that most of the population presents similar distribution of the genetic groups (in green and red in Figure 2), with proportions varying from about 30 to 60%.

Table 3. Analysis of molecular variance in oocyte donor cow populations based on 19 microsatellite markers.

Source of variation	Degrees of freedom	Sum of squares	Mean of squares	Estimated variance	Variation (%)	p-value
Between Populations	2	16.997	8.498	0.068	1	<0.005
Between Individuals	47	307.283	6.538	0.484	8	<0.001
Within Individuals	50	278.500	5.570	5.570	91	<0.001
Total	99	602.780		6.122	100	

Table 4. Shannon information and diversity analysis for 19 loci within and between oocyte donor cow populations.

Population interactions	Shannon within populations	Shannon between populations	F_{ST} values*	Nm values**	(0-1) scaled diversity	p-value
1 vs. 2	1.123	0.055	0.003	89.91	0.122	—
1 vs. 3	1.235	0.040	0.007	34.32	0.103	—
2 vs. 3	1.244	0.052	0.014	17.36	0.103	<0.001

* F_{ST} : (genetic distance between populations) values based on allelic distance matrix from F-statistics analysis.

** Nm: (Number of migrants) values based on F_{ST} values.

Discussion

Selection schemes based on reproductive biotechnologies are commonly used in cattle breeding. These are based on the continuous use of elite animals to generate products to supply markets for beef and milk around the world (Biscarini et al., 2015).

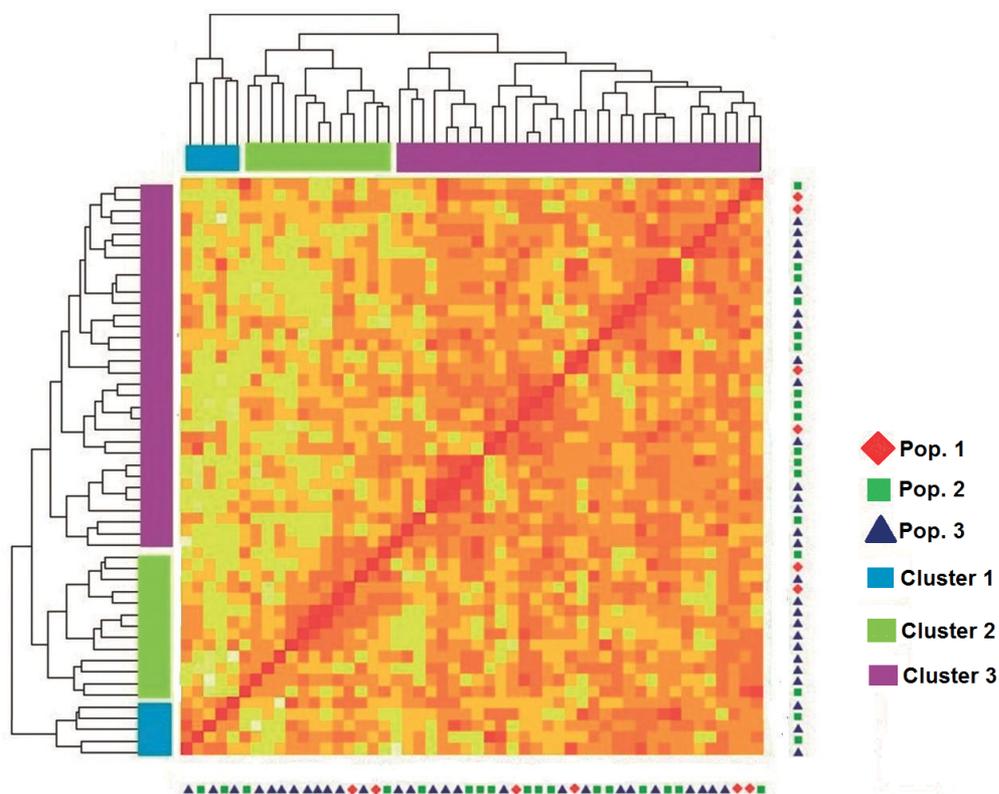


Figure 1. Heat map of genetic distances between three elite cow populations associated with embryo production programs. The heat map illustrates genetic relationship from lower (white/yellow) to higher (orange/red) among individuals of three elite cow populations.

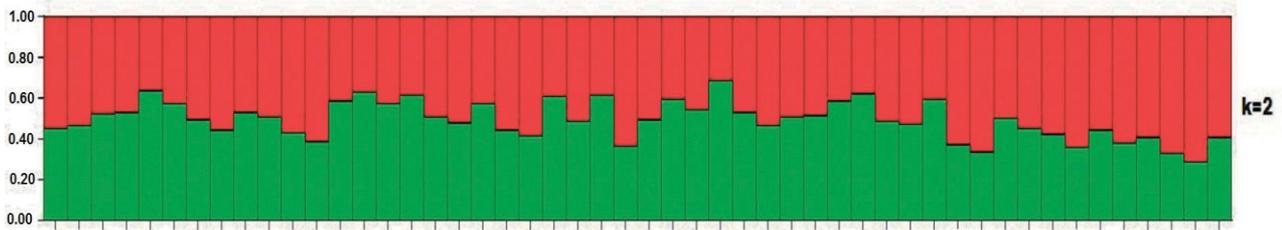


Figure 2. Clustering assignment based on the Bayesian model for 50 animals representing three oocyte donor cattle populations. Each individual animal is represented on the graph by a vertical bar divided into K colored segments (red and green) corresponding to K genetic clusters. The length of each colored segment is proportional to the individual's membership in the cluster of corresponding color. Numbers on the y-axis show coefficient of membership/assignment.

The results of this study (Table 1) suggest that the populations analyzed have variable levels of heterozygosity based on their different allele frequencies. Genetically, these differences in the frequencies of the loci evaluated indicate selection pressure that is also variable (natural or artificial) in each population. This can lead to fixation of some homozygous alleles (endogamy) (Pedersen *et al.*, 2012) as in Population 3. Since a fixation index near zero is expected in random coupling, and positive values indicate fixation of alleles in homozygosity (endogamy), due mainly to the population's genetic structure (Peakall & Smouse, 2012), it can be stated that Population 3 had the least diversity.

It should be mentioned that the type of selective pressure applied by the use of reproductive biotechnologies such as OPU-IVF can quickly modify the gene and allele frequencies in the populations due to the short intergenerational interval or the reduced number of bulls used for the mating between elite donors, among other reasons (Dassonneville *et al.*, 2012). The specific effect of the use of ARTs on genetic diversity is unknown, however, it is expected that this will be decreased after several inbreeding mating.

Other researchers analyzing microsatellite markers in Gir cattle reported observed heterozygosity values in all the loci, ranging from 0.60 to 0.62, in the case of Brazilian Gir cattle (Egito *et al.*, 2007; Villalobos-Cortés *et al.*, 2015), and 0.67 in a study carried out in India (Kale *et al.*, 2010). These results are similar to those presented in this study for Populations 1 (0.59) and 2 (0.64).

Considering the concept of allelic richness and the statistical differences between the populations (Populations 1 and 3), it can be concluded that the lower value found in the first population would reduce the chances of long-run adaptation and persistence, since the limits of selection are determined by the initial allelic composition rather than by the heterozygosity (Petit *et al.*, 2008) (a rare allele that is lost in a founder event probably will not greatly affect the heterozygosity, but the loss will reduce the allelic richness (Greenbaum *et al.*, 2014).

Measures of population allelic richness are not only considered important for preservation of breeds or species, but also in marker-assisted selection, since the existence of alleles, instead of their frequencies, determines a significant part of the potential to respond to selection (Petit *et al.*, 2008). The average value of A_R obtained in

this study (4.2 alleles) was similar to the average obtained in a study conducted in 2015 ($A_R=4.11$), also with Brazilian Gir cattle (Villalobos-Cortés *et al.*, 2015). Average values of A_R in seven indigenous Vietnamese breeds were higher than those found in this study (8.72 alleles), but the authors used a larger number of microsatellites (27 loci) (Pham *et al.*, 2013). The total value of pA_R involving the three populations (1.88), also considered a function of genetic diversity, was greater than that reported in other studies (0.002 to 0.006) (Gautier *et al.*, 2010; Porto-Neto *et al.*, 2013; Wang, 2015). However, these studies used single nucleotide polymorphism (SNP) markers, which are considered less polymorphic than microsatellites, and hence contain a smaller number of alleles distributed in the populations (Wang, 2015).

In relation to N_p and its association with pA_R , the latter estimate is considered more accurate, since the statistical technique of rarefaction considers the sample size, compensating the sampling disparity (Kalinowski, 2005).

When evaluating the level of endogamy within the populations (F_{IS}) and the individuals in relation to the total number of alleles of the populations (F_{IT}), it can be stated that for the first two populations studied, the result is within the range reported in other studies of the Gir breed, where F_{IS} values using genetic markers or pedigree analysis were negative (O'Brien *et al.*, 2015; Wang, 2015), smaller than 1% (Egito *et al.*, 2007; Porto-Neto *et al.*, 2013) or varied between 1 and 3% (Faria *et al.*, 2009; Santana *et al.*, 2014). Nevertheless, the endogamy within Population 3 (9.1%) due to the fixation of alleles in homozygosis lies out above the values reported for this breed recently (Wang, 2015), indicating a delicate situation regarding genetic management. The level and rate of endogamy are generally utilized as parameters for variation within breeds and are negatively correlated with the effective population size (Biscarini *et al.*, 2015). The greater allelic richness observed in this population should allow directed mating, seeking a balance between genetic merit and diversity. This can be attained by using tools such as optimum contribution selection.

In contrast to a recent study that compared seven Brazilian Gir breeds (Wang, 2015), the elite cow populations evaluated in this study did not present a clear pattern of differentiation between populations (values of $F_{ST} < 0.05$, AMOVA and Shannon diversity with variance between 1 and 7%, respectively). This low diversity might have been the result of common ancestors in the establishment of the three populations (identical by descent) (Slatkin, 1985).

It should be mentioned that the low values of F_{ST} could also represent purifying selective forces, which are applied simultaneously in the populations in the same direction, imposing strong similarity between the compared groups and resulting in low differentiation. In this case, deleterious genes also can be selected jointly, affecting traits of common interest (Porto-Neto *et al.*, 2013).

In the reproductive management of elite populations involving ARTs, it is routine to use the same reproducers in different farms due to traits such as high genetic value, reproductive efficiency and common selection objectives, thus facilitating the wide dissemination of the same alleles in different populations without the apparent influence of geographic isolation. This is confirmed by observing Populations 2 and 3, which although being nearest

geographically, presented a higher value of F_{ST} (0.014 and $p < 0.001$) compared to the more distant population pairs.

We therefore consider all the populations studied to be similar genetically due to the high gene flow between them (high migration rate) revealed in the analysis of the F-statistics (Table 2), and the Shannon diversity (Table 4) including all the markers, where the values of Nm fluctuated between 11.03 and 89.91 individuals per generation. The similarity of the populations is mainly due to migration rather than genetic drift.

It is common to evaluate the Nm in studies of diversity within breeds, but no consensus exists about the ideal value in this type of study. However, it is known that the higher the value of Nm is, the smaller the genetic separation between populations is.

Values of Nm reported among five Indian breeds varied from 5.40 to 32.80 individuals according to the degree of genetic proximity (Sharma *et al.*, 2013).

In turn, in Brazil values of Nm and F_{ST} published in previous studies involving Zebu Gir and Brahman breeds reported a low level of differentiation ($Nm=7.34$ and $F_{ST}=0.042$) (Villalobos-Cortés *et al.*, 2015). The reported value of Nm was considered high by those authors. The pattern of low genetic differentiation between Brazilian Zebu breeds was initially confirmed in 2007 (Egito *et al.*, 2007), and more recently in 2015 (O'Brien *et al.*, 2015) through the F_{ST} values (< 0.05) using different classes of markers.

From the standpoint of evolutionary genetics, the genetic structure of a population is characterized by the morphological and quantitative variability existing between individuals, the reproductive system, gene flow patterns and adaptive strategies to local environments (Holsinger & Weir, 2009). The lack of structure can be an indication of the influence of selective forces acting naturally or artificially on populations, which modify the allelic frequencies in favor of homozygosis (Slatkin, 1985).

The identification of the most likely number of clusters carried out in the STRUCTURE program included the allelic frequency model correlated between the populations in the previously established parameters. This frequency model has greater power to detect distinct populations that are particularly related, presenting the same results as the independent allele frequency model in the absence of high levels of correlation between the populations (Porrás-Hurtado *et al.*, 2013).

Besides the mentioned relevance of the stratification of the populations regarding the aspects of conservation of genetic resources and genetic diversity related to the long-term adaptation of populations (Slatkin, 1985), the differences in the allele frequencies between individuals in a population, derived from the systematic variation of the allele frequencies in their ancestors through sub-populations, are highly prized in searching for QTLs (quantitative trait loci), employing the genome-wide association study (GWAS) approach (Liu, *et al.*, 2013).

The oocyte donor cows included in the in vitro embryo production programs analyzed here presented low genetic differentiation between populations. Within the populations, they presented low to moderate diversity and inbreeding, indicating genetic management and non-standardized mating.

We suggest the use of mating guidance tools that use measures of genetic diversity and richness in elite populations dedicated to production of genetic material of interest to animal breeders. New studies assessing

the effect of the levels of endogamy on traits of interest in *in vitro* embryo production are necessary.

Acknowledgements

The authors thank the Office to Coordinate Improvement of University Personnel (CAPES) for the grant awarded to the first author and the National Council for Technological and Scientific Development (CNPq) for financial support.

Conflicts of interest

The authors declare they have no conflicts of interest with regard to the work presented in this report.

References

- Biscarini F, Nicolazzi E, Alessandra S, Boettcher P, Gandini G. Challenges and opportunities in genetic improvement of local livestock breeds. *Front Genet* 2015; 5:1-16.
- Dassonneville R, Baur A, Fritz S, Boichard D, Ducrocq V. Inclusion of cow records in genomic evaluations and impact on bias due to preferential treatment. *Genet Sel Evol* 2012;44(40):1-8.
- Egito AA, Paiva SR, Albuquerque MS., et al. Microsatellite based genetic diversity and relationships among ten Creole and commercial cattle breeds raised in Brazil. *BMC Genet* 2007; 8(8).
- Faria FJC, Filho AE V, Madalena FE, Josahkian LA. Pedigree analysis in the Brazilian Zebu breeds. *J Anim Breed Genet* 2009; 126(2):148-153.
- Gaggiotti OE, Lange O, Rassmann K, Gliddon C. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Mol Ecol* 1999; 8(9):1513-1520.
- Gautier M, Laloë D, Moazami-Goudarzi K, Brandstaetter A. Insights into the Genetic History of French Cattle from Dense SNP Data on 47 Worldwide Breeds. *PLoS One* 2010; 5(9):1-11.
- Greenbaum G, Templeton AR, Zarmi Y, Bar-David S. Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. *PLoS One* 2014; 9(12):1-23.
- Holsinger KE, Weir BS. Fundamental concepts in genetics: Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat Publ Gr* 2009; 10:639-650.
- IETS. 2016 Statistics of embryo collection and transfer in domestic farm animals. IETS Data Retrieval Report. 2016.
- Kale DS, Rank DN, Joshi CG, et al. Genetic diversity among Indian Gir, Deoni and Kankrej cattle breeds based on microsatellite markers. *Indian J Biotechnol* 2010; 9(2):126-130.
- Kalinowski ST. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 2005; 5(1):187-189.
- Liu L, Zhang D, Liu H, Arendt C. Robust methods for population stratification in genome wide association

studies. *BMC Bioinformatics* 2013; 14(1):132.

Nei M. F-statistics and analysis of gene diversity in subdivided populations. *Ann Hum Genet* 1977; 41(2):225-233.

O'Brien AMP, Höller D, Boison SA, et al. Low levels of taurine introgression in the current Brazilian Nelore and Gir indicine cattle populations. *Genet Sel Evol* 2015; 47(31):1-7.

Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. 2012; 28(19):2537-2539.

Pedersen LD, Kargo M, Berg P, Voergaard J, Buch LH, Sørensen AC. Genomic selection strategies in dairy cattle breeding programmes: Sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. *J Anim Breed Genet* 2012; 129(2):152-163.

Petit RJ, El Mousadik A, Pons O. Identifying Populations for Conservation on the Basis of Genetic Markers. *Conserv Biol* 2008; 12(4):844-855.

Pham LD, Do DN, Binh NT, et al. Assessment of genetic diversity and population structure of Vietnamese indigenous cattle populations by microsatellites. *Livest Sci* 2013; 155(1):17-22.

Porrás-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Àngel, Lareu M V. An overview of structure: Applications, parameter settings, and supporting software. *Front Genet* 2013; 4:1-13.

Porto-Neto LR, Sonstegard TS, Liu GE, et al. Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. *BMC Genomics* 2013; 14(876):1-12.

Pryce JE, Goddard ME, Raadsma HW, Hayes BJ. Deterministic models of breeding scheme designs that incorporate genomic selection. *J Dairy Sci* 2010; 93(11):5455-5466.

Santana ML, Pereira RJ, Bignardi AB, El Faro L, Tonhati H, Albuquerque LG. History, structure, and genetic diversity of Brazilian Gir cattle. *Livest Sci* 2014; 163(1):26-33.

SAS[®], Statistical Analysis System. SAS/STAT User's Guide. Version 9.2. Cary, NC, USA: SAS Institute Incorporation; 2009.

Sharma R, Maitra A, Singh PK, Tantia MS. Genetic diversity and relationship of cattle populations of East India: distinguishing lesser known cattle populations and established breeds based on STR markers. *Springerplus* 2013; 2(1):359.

Slatkin M. Gene flow in natural populations. *Ann Rev Eco/ Syst* 1985; 16:393-430.

Villalobos-Cortés A, Martínez A, Vega-Pla JL, et al. Genetic relationships among five zebu breeds naturalized in america accessed with molecular markers. *Ital J Anim Sci* 2015; 14(2):158-162.

Wang Y. Genetic and Geographic Diversity of Gyr (*Bos Indicus*) Cattle in Brazil. 2015.

Wright S. Systems of mating. I. The biometric relations between parent and offspring. *Genetics* 1921; 6:111-

