

## Non-hierarchical cluster analysis with the breeding values for reproductive and productive traits in beef cattle

C.C.P. Paz<sup>1,2§</sup>, H.L. Moreira<sup>2+</sup>, R. P. Savegnago<sup>3++</sup>, M. E. Buzanskas<sup>3+++</sup>, R. B. Lôbo<sup>2§</sup>, D. P. Munari<sup>3§</sup>

<sup>1</sup>Centro de Bovinos de Corte - Instituto de Zootecnia/APTA - SAA, Sertãozinho, SP, Brasil, <sup>2</sup>Departamento de Genética, FMRP/USP, Ribeirão Preto, SP, <sup>3</sup>Departamento de Ciências Exatas, FCAV/UNESP, Jaboticabal, SP, Brasil.

**ABSTRACT** The objective of this work was to explore the genetic patterns of the individuals using the non-hierarchical cluster analysis with the breeding values for productive and reproductive traits of Brazilian Nellore cattle to help in the selection process of the animals. The population was divided into three groups, each one with animals of different genetic patterns. The cluster 1 has grouped animals with EBV-BW210 and EBV-BW365 below the population mean (zero), EBV-AFC e EBV-GL around the population mean and EBV-SC365 and EBV-SC450 above the population mean. The cluster 2 has grouped animals with genetic patterns of EBV-SC365, EBV-SC450, EBV-BW210 and EBV-BW365 below the population mean and EBV-AFC and EBV-GL around the population mean. The cluster 3 has grouped animals with genetic patterns of EBV-AFC below the population mean, EBV-GL around the population mean, EBV-SC365, EBV-SC450, EBV-BW210 and EBV-BW365 above the population mean.

**Keywords:** breeding value, cluster analysis, gain genetic

### Introduction

Selection in beef cattle aims to improve the productive and reproductive indices of the herds, which represents an important source for development of Brazilian agribusiness. The reproductive performance of the herds needs to be evaluated according to traits associated directly with the productive efficiency of the herds and as a result economic gain by decreasing production cycles.

The age at first calving (AFC) is used to evaluate the Nellore cattle with respect to the sexual precocity in females. Consequences of selection on gestation length (GL) on Nellore cattle breeding program can be evaluated by genetic correlations that suggest reduction in the number of difficult calving and birth weight of calves, increasing reproductive efficiency of females.

The body weight at 210 (BW210) and at 365 (BW365) days of age have been used to evaluate the growth of the animals and the scrotal circumference at 365 (SC365) and at 450 (SC450) days of age have been used to evaluate sexual precocity in males.

The selection of animals by breeding values for traits of economic importance is used in the selection process to improve genetically the herds. The greatest difficulty about the selection process is to choose which traits should be selected. An alternative is the use of multivariate analysis techniques that can be employed to better understand the complex relationships among genetic values of the multi-traits data sets. (Savegnago et al., 2011; Buzanskas et al., 2013).

Cluster analysis can be used to explore data sets and define homogeneous groups of animals within a population, based on the grouping criteria (Johnson and Wichern, 2007).

The objective of this work was to explore the genetic patterns of the individuals using the non-hierarchical cluster analysis with the breeding values for productive and reproductive traits of Brazilian Nellore cattle to help in the selection process of the animals.

### Material and Methods

**Animals and data file.** The present study was conducted using data on animals of Nellore breed that participated in the Nellore Brazil Program, coordinated by the National Association of Breeders and Researchers (Associação Nacional de Criadores e Pesquisadores – ANCP).

**Statistical analysis.** Estimation of genetic parameters was performed using the Restricted Maximum Likelihood Method (REML) in a multi-trait animal model, using the WOMBAT software, described by Meyer (2006). Genealogy files containing identifications for each animal and its sires and dams, consisting of 139,210 animal records, were used in the kinship matrix. The mixed model used for body weight at 210 (BW210) and at 365 (BW365) days of age, scrotal circumference at 365 (SC365) and at 450 (SC450) days of age, gestation length calf traits (GL) and age at first calving (AFC), was  $y = Xb + Z_1a + Z_2m + e$ , in which:  $y$  is the vector of the dependent variable;  $X$  is the incidence matrix for fixed effects, thereby associating the elements of  $b$  and  $y$ ;  $b$  is the fixed-effects vector (contemporary group);  $Z_1$  and  $Z_2$  is the incidence matrix for direct and maternal random effects, thereby associating the elements of  $a$  and  $y$ ;  $a$  is the random-effects vector for direct additive genetic effects;  $m$  is the random-effects vector for maternal

§ CNPq Productivity Fellowship

+ CNPq Fellowship

++ FAPESP Fellowship 2013/20091-0

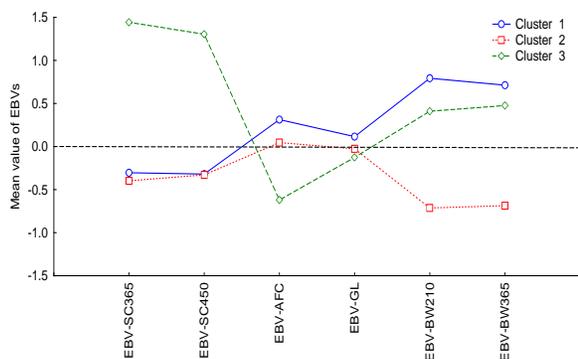
+++ FAPESP Fellowship 2013/19335-2

genetic effect and  $e$  is the residual-effects vector. For BW365, SC365, SC450, AFC, the maternal genetic effect was not included in the mixed model. For W210 and GL, a linear and quadratic effect of the covariate dam age was considered.

The Multivariate analysis by non-hierarchical clustering was performed based on breeding values of the traits W210 (EBV-BW210), W365 (EBV-BW365), SC365 (EBV-SC365), SC450 (EBV-SC450), AFC (EBV-AFC) and GL (EBV-GL), with the objective of to explore the genetic patterns of the group of animals within a population to assist in the selection process of the animals. This analysis uses the k-means algorithm. The cluster analyzes were performed using Statistica 8.0 software (StatSoft, Inc., 2008).

## Results and Discussion

The genetic patterns of each group using the breeding values of the animals are in Figure 1. There is no specific rule to determine the optimal number of groups in which the population should be divided. The interpretation of the results would be difficult if the population was divided into many groups. On the other hand, if the population was divided in few groups, it would have to uninformative conclusions about the groups. The division of this population was performed in three groups (cluster 1, cluster 2 e cluster 3).



**Figure 1.** Genetic patterns of the breeding values of the animals in each cluster. EBV-BW210 and EBV-BW365 = breeding values for body weight at 210 and 365 days of age; EBV-SC365 and EBV-SC450 = breeding values for scrotal circumference at 365 and 450 days of age; EBV-AFC = breeding values for age at first calving; EBV-GL = breeding values for gestation length.

The cluster 1 has grouped animals with EBV-BW210 and EBV-BW365 below the population mean (zero), EBV-AFC e EBV-GL around the population mean and EBV-SC365 and EBV-SC450 above the population mean. The cluster 2 has grouped animals with genetic patterns of EBV-SC365, EBV-SC450, EBV-BW210 and EBV-BW365 below the population mean and EBV-AFC and EBV-GL around the population mean. The cluster 3

has grouped animals with genetic patterns of EBV-AFC below the population mean (desired), EBV-GL around the population mean and EBV-SC365, EBV-SC450, EBV-BW210 and EBV-BW365 above the population mean.

Considering the genetic patterns of the animals in Figure 1, if the selection was performed based on the animals of the cluster 1, it would increase the body weight of the animals and would decrease the precocity of males and females. If the selection was performed considering the genetic patterns of the animals of the cluster 2, the body weight of animals as well as the precocity of males would decrease and if the selection was performed considering the genetic patterns of the animals of the cluster 3, the body weight of the animals as well as the precocity of males and females would increase.

The selection based on animals from cluster 1 and cluster 2 is not desirable because these two groups account animals of undesirable genetic patterns for productive and reproductive traits in beef cattle, resulting in unfavorable genetic gains

However, the selection based on the genetic patterns of the cluster 3 animals would be desirable, because the animals in this group presented genetic patterns that could result in desirable genetic gains to increase the body weight and the increase sexual precocity in males and females without change genetically the gestation period of females will be not affected.

## Conclusion

The analysis of non-hierarchical clustering can be an interesting alternative in the selection of beef cattle.

The selection for animals with the genetic patterns of the cluster 3 may result in favorable and balanced genetic gain for production and reproduction traits evaluated.

## Literature Cited

- Buzanskas, M.E., Savegnago, R.P., Grossi, D.A., et al. (2013). *Reprod. Fert. Dev.* 25:775-781.
- Johnson, R.A., Wichern, D.W. (2007). Pearson-Prentice Hall. 773p.
- Meyer, K. (2006). *J. Zhejiang Univ. Sci. B.* 8:815-821.
- SAS System for Windows®. Version 9.1.2008. SAS Inst., Cary, NC, USA.
- Savegnago, R.P., Caetano, S.L., Ramos, S.B., et al. (2011). *Poult. Sci.* 90:2174-2188.