ABSTRACT: The aim of this study was to explore the genetic curve pattern of EBVs for test-day milk yield of Guzerá cattle using cluster analysis. Test-day milk yields of 5,274 first-lactation Guzerá cows were recorded in a progeny test from 1987 to 2012. A total of 34,193 monthly records were analyzed with a random regression animal model using Legendre polynomials to fit additive genetic and permanent environmental random effects and the mean trend. Residual variances were modeled by a step function. Hierarchical and non-hierarchical cluster analyses were performed based on the EBVs for monthly test-day milk yield, peak yield, lactation persistency, and partial cumulative and 305-day yields. The heritability estimates for test-day milk yields ranged from 0.24 to 0.52. Cluster analysis identified animals in the population that belong to different groups according to milk production level and lactation persistency. Keywords: lactation curve; persistency; random regression; Zebu cows

Introduction

The Guzerá Zebu breed has been selected in Brazil for milk and meat production and is traditionally considered a dual purpose breed. In addition, this breed is known for its rusticity, rendering it adequate for the different and adverse tropical environments that predominate in Brazil. Guzerá cattle have been widely used in crossings with taurine breeds specialized in the production of dairy crossbreds. However, the frequency of short lactations is high in this breed and, since it is not a specifically dairy breed, lactation curves are less persistent.

Persistency is a trait that is directly related to economic aspects of dairy farming (Jakobsen et al., 2002; Muir et al., 2004). Since selection of Guzerá cattle for milk production is recent, simultaneous selection for lactation persistency using measures of persistency that are poorly correlated with 305-day milk yield would be important (Pereira et al., 2010).

Since the selection goals for Guzerá animals are designed to improve meat and/or milk production, major differences may exist in the shapes of the lactation curves of these animals depending on a higher or lower selection pressure for one or the other trait. In this respect, the grouping of genetically similar animals by multivariate clustering analysis seems to be an interesting approach. The division of animals into groups would permit to identify animals based on the genetic curve pattern of milk yield. Thus, the objective of the present study was to explore the genetic curve patterns of milk yield in Guzerá cattle based using cluster analysis with the estimated breeding values (EBVs) for components of the lactation curve.

Material and Methods

Data. A database of the National Dairy Guzerá Breeding Program (PNMGuL), which is coordinated by Embrapa Gado de Leite in cooperation with the Brazilian Center for Guzerá Improvement (CBMG) and the Brazilian Association of Zebu Breeders (ABCZ), was used for analysis. The data comprised 34,193 test-day milk yield records of 5,274 purebred first-lactation Guzerá cows, daughters of 628 sires and 3,302 dams, which had calved between 1987 and 2012.

Statistical analysis. Test-day milk yields of first-lactation Guzerá cows were analyzed using a single-trait random regression animal model. The model included direct additive genetic, permanent environmental, and residual effects as random effects. The fixed effects consisted of the contemporary group and the linear and quadratic effects of age of cow at calving as covariate. The contemporary group was defined by herd, year and season of test-day record (dry and rainy season). The average population trajectory was fitted using a Legendre polynomial with 4 coefficients. Test-day milk yields were divided into 10 monthly classes, ranging from day 6 to day 305 of lactation. Random additive genetic and permanent environmental effects were modeled by orthogonal Legendre polynomials with 4 and 5 coefficients, respectively. The structure of residual variances was considered to be heterogenous and contained seven variance classes. Variance components were estimated by the restricted maximum likelihood method (REML) using the WOMBAT statistical package (Meyer, 2007).

Cluster analysis, divided into two phases, was used to study the genetic profile of the lactation curves. Hierarchical cluster analysis was used to choose the number of clusters in which the population could be divided. After definition of the number of groups, non-hierarchical cluster analysis was performed using the k-means algorithm. The cluster analyzes were performed using SAS software (2008). The measures used for cluster analysis were the EBVs for test-day milk yield (M1 to M10); partial cumula-
tive milk yields in the first (M100), second (M200) and third (M300) trimester of lactation; 305-day cumulative milk yield (RR305); peak milk yield (PEAK), and lactation persistency (PS). The last parameter was adapted from Pereira et al. (2002) and was calculated by the formula: \( PS = \sum_{t=30}^{270} (EBV_t - EBV_{PEAK}) \). The lactation peak was determined based on the average EBVs between days 5 and 30 of lactation.

Results and Discussion

Genetic parameters. The peak yield was observed in the first month of lactation, with a mean yield of 9.04 ± 4.62 kg milk. Milk production remained constant in the second month of lactation (8.98 ± 4.14 kg), with subsequent declines after the third month until the end of lactation (5.44 ± 2.53 kg). The heritability estimates for test-day milk yield were 0.52 ± 0.04, 0.38 ± 0.03, 0.33 ± 0.03, 0.33 ± 0.03, 0.29 ± 0.03, 0.26 ± 0.03, 0.24 ± 0.03, 0.24 ± 0.03, 0.27 ± 0.04 and 0.38 ± 0.05 for M1 to M10, respectively. These estimates were higher than those observed by Santos et al. (2013) for the same breed. These authors estimated higher heritability in the second month of lactation (0.35) and a lower value in the eighth month (0.18). However, the trajectory and magnitude of the estimates were similar in the two studies after the second test-day. The genetic correlations between monthly milk yields ranged from -0.03 to 0.95. The lowest or negative correlations between yields were observed at the beginning and end of lactation. Similar results have been reported by Santos et al. (2013) for Guzerá cattle. The phenotypic correlations were all positive and ranged from 0.18 to 0.81. The same was observed for permanent environmental correlations which ranged from 0.48 to 0.98, with lower estimates between more distant test-days. The simple correlations between EBVs for RR305 and peak yield, M100, M200 and M300 were positive and high (0.76, 0.90, 0.95 and 0.76, respectively). The higher correlations for milk yields in the first and second trimester of lactation (M100 and M200) indicate a strong association with the level of milk production during lactation. The correlation between EBVs for peak yield and lactation persistency was -0.96. Since persistency was expressed as the sum of deviations in relation to peak yield, a strong and antagonistic association between these values was expected. In contrast, the correlations of persistency with RR305, M100, M200 and M300 were -0.42, -0.74, -0.15 and 0.10, respectively. Since peak yield and M100 are measures that are close to each other and correspond to milk production at the beginning of lactation, an antagonistic and strong association between EBVs was expected. The association with RR305 was negative, but of low magnitude, as desired (Pereira et al., 2002). However, the antagonistic association between the two traits should serve as an alert to breeders who have exclusively used the increase in cumulative milk yield during lactation as a selection criterion for dairy Zebu breeds.

Cluster analysis. Hierarchical cluster analysis was used to divide the population in groups based on EBVs for different lactation measures (Figure 1). The population was divided into two clusters (Figure 2). Cluster 1 contained animals with negative EBVs for partial lactation periods and peak yield, but with EBVs above the average for persistency, indicating that these animals are more persistent despite the lower level of production. Cluster 2 showed EBVs above the average for partial lactation periods and for peak yield, but below the average for lactation persistency. With respect to partial measures (M100, M200, M300, and M1 to M10), the EBVs obtained for cluster 2 were higher at the beginning of lactation and decreased during lactation. Lower breeding values were observed at the end of lactation (Figure 2). For deeper analysis with the candidates to selection, i.e. animals from cluster 2 which presented higher breeding values for almost all the traits, was divided into two subgroups (Figure 3).

The two subgroups presented negative EBVs for lactation persistency and positive values for partial yields, RR305 and peak yield (Figure 3). However, positive EBVs close to the average were observed for animals of subgroup 1, whereas animals of subgroup 2 exhibited a genetic curve pattern with higher EBVs. Lactation persistency was lower for animals of subgroup 2 compared to those of subgroup 1. Since animals of subgroup 1 presented lower, but still positive, production levels when compared to the population average, they may be used in simpler production systems in which the animals remain on pasture throughout the year,
and probably belong to productions systems that explore dual purpose breeds. Breeding programs of dairy Zebu breeds, whose selection for increased milk yield is still recent when compared to taurine breeds, should emphasize the level of production, as observed for animals of group 2, subgroup 2. However, animals of this group exhibited lower breeding values for lactation persistency. Thus, persistency should be part of a selection index in an attempt to combine production level (RR305) and lactation persistency.

From an economic point of view, selection of more persistent animals in breeding programs would be more indicated since cows with high lactation persistency require less feed and health and reproductive costs are lower. As a consequence, these animals are more profitable for breeders (Dekkers et al., 1998). However, these proposals have been made for taurine breeds, which have long been selected for high milk production. For dairy Zebu breeds, the level of milk production is still low, especially in the case of Guzerá cattle which consist of dual purpose lineages (beef and milk). The level of production should therefore continue to be the main selection criterion for this breed, giving less emphasis on lactation persistency in order to permit the long-term correction of problems related to the lower lactation persistency and shorter lactation duration in animals of this breed.

**Conclusion**

Cluster analysis identified animals in the population that belong to different groups according to production level and lactation persistency, and may be used to assist the selection process.

**Acknowledgements**

National Council for Scientific and Technological Development (CNPq), the State Funding Agency of Minas Gerais (FAPEMIG) for financial support.

**Literature Cited**