

ESTIMATIVA DAS DISTÂNCIAS GENÉTICAS E COMPONENTES PRINCIPAIS EM BOVINOS DE CORTE NO BRASIL

GENETIC DISTANCES AND PRINCIPAL COMPONENT ESTIMATIONS IN BRAZILIAN BEEF CATTLE

Souza, J.C.^{1,4*}, Perotto, D.², Abrahão, J.J.², Freitas, J.A.¹, Ferraz Filho, P.B.³, Weaber, R.L.⁴ e Lamberson, W.R.⁴

¹Universidade Federal do Paraná. Campus Palotina. UFPR. Palotina, PR. Brasil. *jcs@ufpr.br

²Instituto Agronômico do Paraná. IAPAR. Curitiba. Brasil.

³Departamento de Ciências Biológicas. UFMS. Três Lagoas, MS. Brasil.

⁴Department of Animal Science. University of Missouri/Columbia. USA.

PALAVRAS CHAVE ADICIONAIS

Análise multivariada. Peso.

ADDITIONAL KEYWORDS

Multivariate analysis. Weight.

RESUMO

O objetivo desse trabalho foi estimar os componentes principais e as correlações entre o peso ao nascer (PN), aos 205 (P205), 365 (P365), 550 (P550) e aos 730 (P730) dias de idade para 2237 animais, provenientes do IAPAR, utilizando análise multivariada. A raça dos touros incluídos no programa foram: Nelore, Guzerá, Red Angus, Marchigiana e Limousin e das vacas: Nelore, ½ Guzerá x ½ Nelore, ½ Red Angus x ½ Nelore, ½ Marchigiana x ½ Nelore, ½ Simmental x ½ Nelore, ¾ Nelore x ¼ Red Angus, and ½ Red Angus x ¼ Guzerá x ¼ Nelore. O menor valor encontrado para as correlações foi entre PN e P730 dias, 0,22; e o maior, entre P365 e P550, 0,73. O primeiro componente principal (PC1) justificou 56%; o primeiro e o segundo (PC2), 72%; os três primeiros PC justificaram 87% e, os quatro primeiros PC 95%. O procedimento MANOVA permitiu separar os componentes de variância de cada par. O teste de Wilk's Lambda indicou uma diferença significativa na distância entre as raças das vacas baseando-se no desempenho dos bezerros. Matrizes pertencentes aos grupos ½ Marchigiana x ½ Nelore e ½ Simmental x ½ Nelore tiveram comportamento materno mais próximos, (0,058) ao passo que matrizes Nelore e ½ Simmental x ½ Nelore foram as que apresentaram maior distância (2,738). Utilizando o procedimento MANOVA é possível separar a variância para cada tratamento dos cinco avaliados permitindo a separação e estimativa das co-variâncias entre eles. O uso de

análise multivariada na seleção de gado de corte através dos componentes principais permite a geração de índices precisos, ponderando as variáveis econômicas a serem selecionadas. O uso da *arvore* permitiu uma melhor visualização das respostas, no caso, do desenvolvimento das progênes de acordo com o grupo racial das matrizes.

SUMMARY

The objective of this study was to estimate principal components for, and correlations between, weights at birth (BWT), 205d (WT205), 365d (WT365), 550d (WT550) and 730d (WT730) using multivariate analysis of data from 2,237 animals from IAPAR, PR. Calves were produced from Nelore, Guzerá, Red Angus, Marchigiana, and Limousin bulls mated to Nelore, ½ Guzerá x ½ Nelore, ½ Red Angus x ½ Nelore, ½ Marchigiana x ½ Nelore, ½ Simmental x ½ Nelore, ¾ Nelore x ¼ Red Angus, and ½ Red Angus x ¼ Guzerá x ¼ Nelore cows. Correlations ranged from 0.22 (BWT with WT730) to 0.73 (WT365 with WT550). The proportions of variation justified by the principal components were: first, 56%; first and second, 72%; first through third, 87%; and first four, 95%. The MANOVA tests model a separate variance for each trait (in this case, a separate variance for each of the five traits, and a separate covariance for each of the pairs of traits). The

Recibido: 6-2-08. Aceptado: 17-9-08.

Arch. Zootec. 59 (228): 479-485. 2010.

Wilks' Lambda indicates a significant difference in genetic distance between breeds of dam based on the performance of calves. Dams composed of $\frac{1}{2}$ Marchigiana x $\frac{1}{2}$ Nellore versus $\frac{1}{2}$ Simmental x $\frac{1}{2}$ Nellore had the least genetic distance (0.058), while Nellore versus $\frac{1}{2}$ Simmental x $\frac{1}{2}$ Nellore were most distant (2.738). Multivariate analysis can be used to estimate principal components for developing economic weights and accurate indices important in beef selection. Trees generated from cluster analysis of progeny weights can be used to visualize genetic grouping of breed of dam.

INTRODUÇÃO

O uso de análise multivariada para avaliação de dados pode ser uma ferramenta adicional para auxiliar no melhoramento genético. Estimativas de correlações, distâncias genéticas e de índices utilizando resultados a partir desse tipo de análise podem promover bons resultados em programas de melhoramento. Através da análise de componentes principais pode-se obter a matriz de variância e co-variância ou a matriz de correlação. Para Meyer (2006) os componentes principais de um grupo de k efeitos correlacionados são simplesmente um grupo de k variáveis com as funções lineares dos efeitos não correlacionados entre si e que conseguem explicar sucessivamente o máximo de variação entre os k efeitos. O principal componente representa o maior eixo de variação entre as observações no espaço multidimensional; o segundo principal componente mostra a menor variação entre as observações (Baker *et al.*, 1988; Timm, 2002). Principal componentes é uma combinação linear de um conjunto de covariáveis que permitem identificar as variáveis que mais contribuem na avaliação.

Para Chase *et al.* (2002) principal componente tem sido usado para obter o posto das estimativas da matriz de co-variância e obter também a decomposição dos seus vetores. Alternativamente principal componente tem sido uma opção para se estimar os parâmetros fenotípicos e genéticos quando se estima novas variáveis. A melhor

aproximação pode ser obtida estimando-os diretamente, e, ao mesmo tempo, restringindo-se aos mais importantes (Kirkpatrick e Meyer, 2004).

Ferraz Filho *et al.* (2006) reportaram que o primeiro principal componente explicou uma mudança de 30,33%, o segundo 24,09% e o terceiro 16,75%. Souza *et al.* (2007a) avaliando o desenvolvimento de búfalos criados em diferentes regiões brasileiras utilizando análise multivariada verificaram valores de eigenvalues dos primeiros quatro principais componentes iguais a 5,29; 2,54; 1,66 e 1,01; estes justificaram 95%. Mascioli *et al.* (2000) avaliando parâmetros genéticos de bovinos de corte constataram que o primeiro principal componente justificou 74% da variação total. Souza *et al.* (2007b) avaliando animais Hereford, utilizando análise multivariada encontraram valores iguais a 5,33; 3,75 e 2,03, para os eigenvalues e os 3 principais componentes juntos justificaram 86%. O objetivo deste trabalho foi estimar os componentes principais e as correlações entre os pesos ao nascer (PN), aos 205 dias (P205), aos 365 dias (P365), aos 550 dias (P550) e aos 730 dias (P730) dias de idade e as distâncias entre os grupos genéticos de vacas.

MATERIAL E MÉTODOS

Dados de pesos ao nascer (PN), aos 205 (P205), 365 (P365), 550 (P550) e aos 730 (P730) dias de idade, de 2237 animais foram avaliados utilizando análise multivariada. Esses animais são filhos de touros das raças Nelore, Guzerá, Red Angus, Marchigiana, and Limousin acasalados com vacas da raça Nelore, $\frac{1}{2}$ Guzerá x $\frac{1}{2}$ Nelore, $\frac{1}{2}$ Red Angus x $\frac{1}{2}$ Nelore, $\frac{1}{2}$ Marchigiana x $\frac{1}{2}$ Nelore, $\frac{1}{2}$ Simmental x $\frac{1}{2}$ Nelore, $\frac{3}{4}$ Nelore x $\frac{1}{4}$ Red Angus, and $\frac{1}{2}$ Red Angus x $\frac{1}{4}$ Guzerá x $\frac{1}{4}$ Nelore, provenientes da Estação Experimental de Paranavai, IAPAR. A estimativa das correlações entre os pesos, estimativas de componentes principais e as distâncias genéticas foram obtidos utilizando-se os

DISTÂNCIAS GENÉTICAS E COMPONENTES PRINCIPAIS COM ANIMAIS CRUZADOS

procedimentos PRINCOMP, CANDISC e TREE do programa SAS (2002). Para confeccionar o dendograma (árvore) e estimar as distâncias entre os grupos genéticos de vaca (GGV), estimou-se as médias ajustadas pelos quadrados médios, utilizando o procedimento GLM, com um modelo

$$Y_{ijklm} = \mu_1 + GGV_j + GGB(GGV)_{jk} + \text{Sexo}_i + b1(X_{ijklm} - \bar{X}) + b2(X_{ijklm} - \bar{X})^2 + e_{ijklm}$$

onde:

Y_{ijklm} : peso nas diferentes idades;

μ_1 : média geral;

GGV_j: jésima raça da vaca;

GGB(GGV)_{jk}: jésima raça da vaca dentro do késima raça do bezerro;

Sexo_i: lésimo sexo do bezerro;

IV: idade da vaca (efeito linear);

IV2: idade da vaca (efeito quadrático);

e_{ijklm} : erro independentemente distribuído ($\sigma^2 = 0$ e $\sigma = 1$).

As estimativas das distâncias, duas a duas, entre grupos foram obtidos com a equação:

$$D(i | j) = (\bar{X}_i - \bar{X}_j)' \text{cov}^{-1} (\bar{X}_i - \bar{X}_j)$$

onde:

$\bar{X}_i - \bar{X}_j$: são as médias das características i e j , respectivamente.

Para alguns grupos que possuíam valores perdidos, estimou-se a média através do procedimento NEXTED, do programa SAS, (2002).

RESULTADOS E DISCUSSÃO

A média e desvio padrão obtidos pelo procedimento PRINCOMP foram $31,9 \pm 5,1$ kg; $152,0 \pm 28,4$ kg; $202,3 \pm 35,2$ kg; $257,8 \pm 41,5$ kg; $383,7 \pm 60,82$ kg para os pesos PN, P205, P365, P550, P730, respectivamente. Os valores das correlações estimadas foram PN e P205 (0,32), PN e P365 (0,31), PN e P550 (0,27) e, PN e P730 (0,22); P205 e P365 (0,61),

P205 e P550 (0,67) e, P205 e P730 (0,31); P365 e P550 (0,72) e P365 e P730 (0,37); P550 e P730 (0,48). Os valores dos *eigenvalues* foram 2,79; 0,83; 0,74; 0,38 e 0,25. O primeiro componente principal (PC1) justificou 56%; o primeiro e o segundo (PC2), 72%; os três primeiros PC justificaram 87% e, os quatro primeiros PC 95%. Tais valores encontram-se próximos aos reportados por Mascioli *et al.* (2000), Ferraz Filho *et al.* (2006), Souza *et al.* (2007a,b). As equações de cada um dos quatro principais componentes são:

$$\text{PC1: } 0,300 \cdot \text{PN} + 0,487 \cdot \text{P205} + 0,507 \cdot \text{P365} + 0,530 \cdot \text{P550} + 0,368 \cdot \text{P730};$$

$$\text{PC2: } 0,938 \cdot \text{PN} - 0,032 \cdot \text{P205} - 0,118 \cdot \text{P365} - 0,228 \cdot \text{P550} - 0,232 \cdot \text{P730};$$

$$\text{PC3: } 0,149 \cdot \text{PN} - 0,393 \cdot \text{P205} - 0,245 \cdot \text{P365} - 0,093 \cdot \text{P550} + 0,869 \cdot \text{P730};$$

$$\text{PC4: } -0,049 \cdot \text{PN} + 0,728 \cdot \text{P205} - 0,661 \cdot \text{P365} - 0,105 \cdot \text{P550} + 0,140 \cdot \text{P730}.$$

Considerando os índices de maior valor dentro de cada componente principal, para o PC1 o peso aos 550 dias foi o componente mais importante (0,5295), seguido pelo peso aos 365 dias (0,5070). Para o segundo principal componente o peso ao nascer mostrou valor bastante elevado (0,9375), no entanto, todos os demais apresentaram valores negativos. Para o terceiro principal componente o peso aos 730 dias foi que apresentou maior valor (0,8690) e os pesos aos 205, 365 e 550 apresentaram valores negativos. Por último, o quarto principal componente, o peso aos 205 dias foi a primeira característica (0,7277). Na **figura 1** pode se comparar o comportamento das estimativas dos principais componentes plotados juntos. O maior valor da correlação foi entre os pesos aos 365 e 550 (0,73) e peso aos 205 e 550 (0,67) evidenciando a importância da fase de cria para os pesos em idades futuras. Também deve se ponderar que o maior crescimento dos animais ocorre ainda na fase de cria, Souza *et al.* (2002) avaliando animais da raça Guzará reportaram ganho líquido dos animais, na fase de cria, igual à 26,7% de uma UA (450 kg). Os mesmos autores afirmam que um dos fatores que

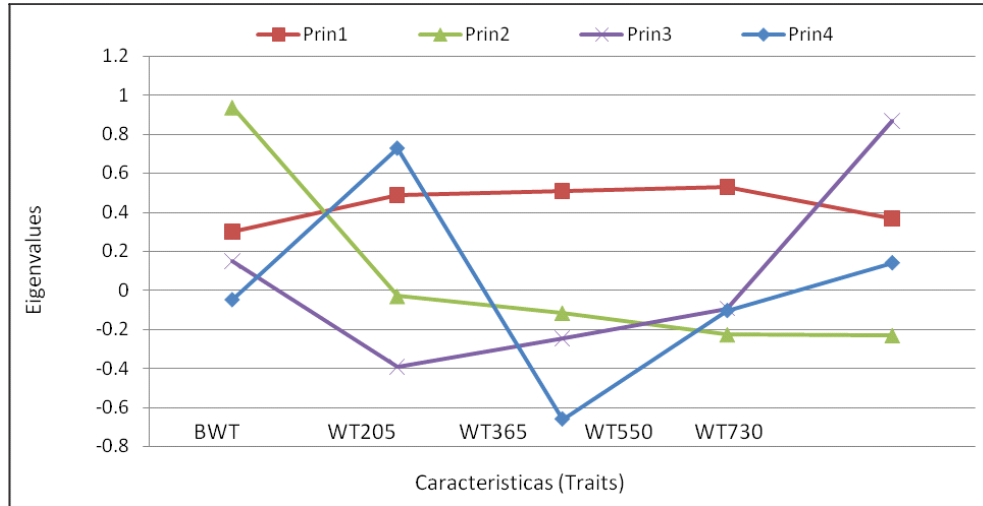


Figura 1. Valores estimados dos eigenvalues na matriz de correlação para animais cruzados do IAPAR. (Eigenvalues of the correlation matrix of crossbreed animals from IAPAR).

contribui para um maior ganho de peso, nesta faixa etária, é a característica fisiológica de maior crescimento neste período, além da ingestão do leite, um alimento completo disponível para o bezerro de corte nessa fase de criação, quando comparado com outras fases de crescimento do animal. No presente trabalho, estimando para todos os animais em conjunto, observou-se um ganho de 26,7% de uma UA.

Considerando que o desempenho dos bezerros na fase de cria tem uma correlação considerável com as idades posteriores e que o desempenho a essa idade encontra-se intimamente ligado à habilidade materna da vaca. As distâncias entre os grupamentos genéticos de vaca variaram de 0,057 a 2,217. Vacas pertencentes ao GGV4 (½ Marchigiana x ½ Nelore) e GGV5 (½ Simmental x ½ Nelore) foram as que apresentaram compor-

Tabela I. Estimativas das distâncias entre grupo genético de vaca (número de progênie por dam). (Estimates of distances between genetic groups of dam (number of progeny by breed of dam)).

GGV	1 (1189)	2 (220)	3 (428)	4 (113)	5 (178)	13 (23)	32 (65)	33 (21)
1		0,298	1,415	2,414	2,717	1,497	0,245	1,884
2			1,254	2,484	2,629	1,623	0,379	2,641
3				0,375	0,338	0,231	0,599	0,975
4					0,057	0,305	1,492	0,652
5						0,342	1,601	0,823
13							0,704	0,437
32								1,210

1: Nelore; 2: ½ Guzerá x ½ Nelore; 3: ½ Red Angus x ½ Nelore; 4: ½ Marchigiana x ½ Nelore; 5: ½ Simmental x ½ Nelore; 13: ¾ Nelore x ½ Red Angus; 32: ½ Red Angus x ¼ Guzerá x ¼ Nelore and, 33: ¼ Nelore x ¾ Red Angus.

DISTÂNCIAS GENÉTICAS E COMPONENTES PRINCIPAIS COM ANIMAIS CRUZADOS

Tabela II. Estimativas das distâncias entre os diferentes grupos genéticos com base na média de pesos de cada grupo materno (entre parênteses número de bezeros por GGV). (Distances estimates by CANDISC Procedure, by breed of dam (Number of calves by GGV) for crossbred animals).

GGV	1 (1189)	2 (220)	3 (435)	4 (113)	5 (178)	13 (25)	32 (22)	33 (21)
1		0,292	1,403	2,438	2,738	1,485	0,856	1,895
2			1,264	2,509	2,654	1,660	1,183	2,638
3				0,381	0,346	0,205	0,248	0,933
4					0,058	0,337	0,585	0,638
5						0,355	0,687	0,805
13							0,152	0,378
32								0,387

tamento materno mais próximos (0,057), indicando efeito materno semelhante para o desenvolvimento das progênies. Por outro lado, vacas do GGV1 (Nelore) e GGV5 (½ Simental x ½ Nelore) foram as que apresentaram valor mais distante igual a 2,717 (tabela I).

Estimativas das distâncias entre os diferentes grupos genéticos com base na média

de pesos de cada grupo materno revelou que a maior distância ocorreu entre os grupos 1 e 5 (Nelore x Simental) e as menores, foi entre os grupos 4 e 5 (Marchigiana x Simental) com um valor de 0,058, conforme

tabela II.

A **figura 2** mostra a distância entre os grupamentos, verifica-se proximidade entre GGV1 e GGV2, e ambos, próximos de GGV33;

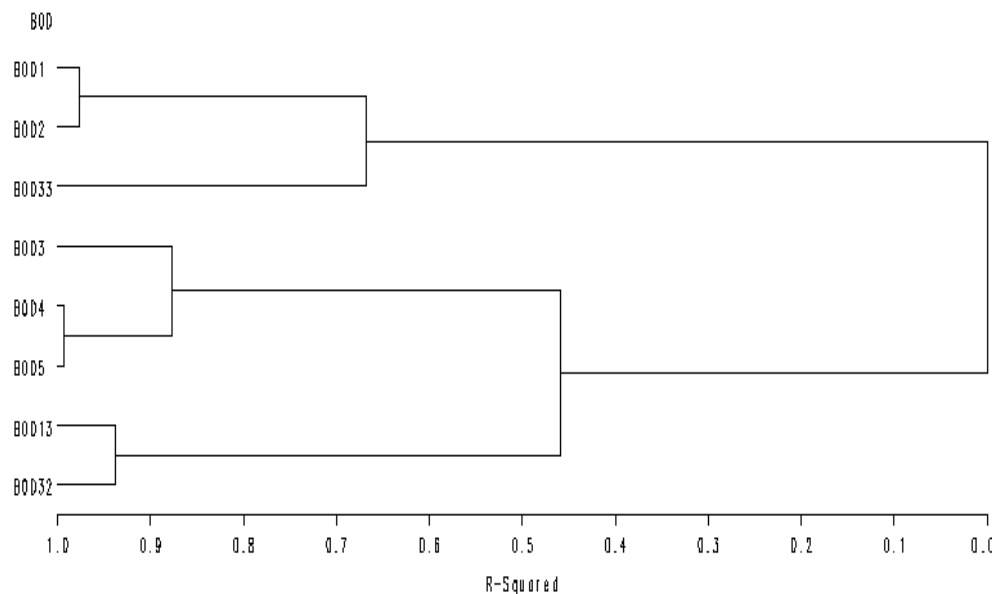


Figura 2. Agrupamento dos diferentes grupamentos racias de vacas (BOD) de acordo com o desenvolvimento da progeny. (Grouping by different group of breed of dam (BOD) according progeny development).

Tabela III. Testes estatísticos utilizados para avaliar a distância entre raças da vaca (GGV) com base na performance dos bezerros, utilizando análise multivariada. (Statistics test to evaluate the distance between breed of cow with basis in the performance of calves, using multivariate traits).

Teste estatístico	Valor	Valor	Num_GL	Den_GL
Wilks' Lambda	0,6774	25,56	35	9219,1***
Pillai's Trace	0,3360	22,59	35	10 975***
Hotelling-Lawley Trace	0,4568	28,58	35	6223,3***
Roy's Greatest Root	0,4106	128,74	7	2195***

Num_GL: Graus liberdade do numerador; Den_GL: Graus liberdade do denominador; ***p<0,0001.

porém distante de GGV13 and GGV32.

O uso de vacas cruzadas proporcionou progênes mais pesadas, no entanto, para se explorar corretamente essa opção deve ser ponderado o efeito da interação genótipo x ambiente, proporcionando condições às matrizes cruzadas e a suas progênes condições para expressar em sua total plenitude o seu potencial genético e o efeito da heterose. A não observação disso, pode levar a resultados de desempenho bastante insatisfatórios.

Um teste utilizando estatística multivariada e aproximação de F para avaliar as diferenças dos diferentes grupamentos genéticos de vacas mostrou-se significativo (p<0,001). Uma vez feito as estimativas das distâncias genéticas, realizaram-se os testes de significância, utilizando análise multivariada, entre as diferentes raças da mãe (GGV) e todos os teste foram significativos p<0,0001.

Para esse modelo em particular foi utilizado cinco dimensões canônicas e todas

apresentaram resultados significativos. Utilizando o procedimento MANOVA foi possível separar a variância para cada tratamento dos cinco avaliados permitindo a separação e estimativas das co-variâncias entre eles (**tabela III**).

O teste de Roy's foi o que apresentou a mais alta significância, embora, o teste de Wilk's Lambda seja o mais exato. Baseando se no teste de Wilks' Lambda, verifica-se diferença entre as raças das vacas no desempenho dos bezerros.

CONCLUSÕES

O uso de análise multivariada na seleção de gado de corte através de principais componentes permite a geração de índices precisos, ponderando as variáveis econômicas a serem selecionadas. O uso da *árvore* permitiu uma melhor visualização das respostas, no caso, do desenvolvimento das progênes de acordo com o grupo racial das matrizes.

BIBLIOGRAFIA

- Baker, J.F., Stewart, T.S., Long, C.R. and Cartwright, T.C. 1988. Multiple regression and principal components analysis of puberty and growth in cattle. *J. Anim. Sci.*, 9: 2147-2158.
- Chase, K., Carrier, D.R., Adler, F.R., Jarvik, T., Ostrander, E.A., Lorentzen, T.D. and Lark, K.G. 2002. Genetic basis for systems of skeletal quantitative traits: principal component analysis of the canid skeleton. *Proc. Natl. Acad. Sci. USA*, 99: 9930-9935.
- Ferraz Filho, P.B., Silva, L.O.C., Souza, J.C. e Malhado, C.H.M. 2006. Divergência genética de touros Nelores com sêmen. Disponível em Centrais de Inseminação no Brasil. In: 43 Reunião

DISTÂNCIAS GENÉTICAS E COMPONENTES PRINCIPAIS COM ANIMAIS CRUZADOS

- Anual da Sociedade Brasileira de Zootecnia. Anais... SBZ. Joao Pessoa:SBZ. Viçosa. v. 1.
- Kirkpatrick, M. and Meyer, K. 2004. Direct estimation of genetic principal components: Simplified analysis of complex phenotypes. *Genetics*, 168: 2295-2306.
- Mascioli, A.S., El Faro, L., Alencar, M.M., Fries, L.A. and Barbosa, P.F. 2000. Genetic and phenotypic parameters and principal components analyses for body weight gains in Canchim cattle. *Braz. J. Anim. Sci.*, 29: 1654-1660.
- Meyer, K. 2006. To have your steak and eat it: genetic principal component analysis for beef cattle data. In: 8th World Congress on Genetics Applied to Livestock Production. Anais... SBMA. Belo Horizonte. Brazil. pp. 13-18. (CD-Rom).
- SAS. 2002. SAS user's guide for Windows environment. V. 9.1. SAS Inst. Inc. Cary, NC. USA.
- Souza, J.C., Silva, L.O.C., Pinheiro, V.G., Grando, R.T., Malhado, C.H.M., Ferraz Filho, P.B. e Freitas, J.A. 2002. Estimativas de parâmetros genéticos dos pesos aos 205, 365 e 550 dias de idade em bovinos da raça Guzerá na micro-região de Araraquara. Em: IV Simpósio Nacional de Melhoramento Animal. Anais... SBMA. Campo Grande, MS. v. 1. pp. 188-191.
- Souza, J.C., Ramos, A.A., Malhado, C.H.M., Jorge, A.M., Ferraz Filho, P.B., DeFreitas, J.A., Flournoy, N. and Lamberson, W.R. 2007a. Evaluation of development traits of beef buffalo in Brazil using multivariate analysis. *Ital. J. Anim. Sci.*, Supp. 2 - Part 2. pp. 321-324.
- Souza, J.C., Campos, L.T., Freitas, J.A., Weaver, R. and Lamberson, W.R. 2007b. Evaluation of growth traits of Brazilian Herefords using multivariate analysis. In: Annual Meeting of the American Society of Animal Science, 2007. San Antonio, TX. Anais... American Society of Animal Science. JAS. Champagnat.
- Timm, N.H. 2002. Applied multivariate analysis. Springer-Verlag. New York.