

Genetic Diversity of Four Cattle Breeds Using Microsatellite Markers

Marco Antonio Machado^{1*}, Ivan Schuster², Mário Luiz Martinez¹, Ana Lúcia Campos¹

ABSTRACT - Four cattle breeds (Gyr, Nelore, Guzerat and Holstein) were analyzed by amplification of genomic DNA using microsatellite loci to evaluate the genetic diversity within and among them. DNA samples of 18 animals from each breed were collected to access the genetic content of them. Allele frequencies were calculated and used to generate a Nei's genetic distance matrix what was used to build a dendrogram following UPGMA clustering. As expected, Holstein breed was the most distinct from the other breeds: 1.15 in relation to Gyr, 1.12 in relation to Nelore and 0.94 in relation to Guzerat. The closest genetic distance was 0.25 between Guzerat and Nelore. A total of 64 alleles in all four breeds were detected using nine microsatellite primers. Each breed showed 53% of the total number of alleles. The average number of alleles per locus was 7.11 ± 3.21 . The most informative locus was BMS1237 with 53% of observed heterozygosity and the least informative locus was BMS3004 with 12% only. The average heterozygosity detected for the nine loci were 35% and the expected value for Hardy-Weinberg equilibrium was 53%. This low heterozygosity suggests a high endogamy level among the animals sampled within each breed.

Key Words: bovine, DNA, genetic diversity, marker, microsatellite

Diversidade Genética de Quatro Raças Bovinas Utilizando Marcadores Microssatélites

RESUMO - As raças bovinas Gir, Nelore e Guzerá (*Bos indicus*) e Holandesa (*Bos taurus*) foram analisadas pela amplificação do DNA genômico utilizando nove locos de microssatélites, visando avaliar a diversidade genética existente dentro e entre as raças. Foram utilizados 18 indivíduos puros para cada raça visando uma amostragem da constituição genética. As frequências alélicas foram calculadas e utilizadas para gerar uma matriz de distâncias genéticas de Nei, que foi utilizada para a construção de um dendrograma com agrupamento pelo método UPGMA (média das distâncias). A raça Holandesa foi a mais distante geneticamente das demais: 1,15 em relação à Gir; 1,12 em relação à Nelore e 0,94 em relação à Guzerá. A menor distância genética obtida (0,25) foi entre as raças Guzerá e Nelore. Os nove primers de microssatélites utilizados geraram um total de 64 alelos para as quatro raças, sendo que cada raça apresentou em média 53% do total dos alelos. A média do número de alelos por loco foi de $7,11 \pm 3,21$. O loco mais informativo foi BMS1237 com uma heterozigidade observada de 53%, enquanto que o loco menos informativo foi BMS3004 com 12%. A heterozigidade média para os nove locos foi de 35% sendo que o esperado para o equilíbrio de Hardy-Weinberg é 53%. Esta baixa heterozigidade sugere que, dentro de cada raça, os animais avaliados apresentam uma elevada endogamia.

Palavras-chave: bovino, DNA, diversidade genética, marcador, microssatélite

Introduction

Dairy and beef industry requires the development of very standardized cattle herds to fulfill their commercial needs that reflects on selection practices in breeding programs. Throughout generations, breeders have accomplished the industry needs and have developed cattle breeds suited to intensive production systems. The extensive selection and multiplication of superior animals cause a significant decrease on the genetic basis of the germplasm that is the sole of the genetic variation needed to the improvement of economic traits and breeds.

Evaluations of traditional breeding traits such as yield, type and morphology have little power to detect

subtle changes in the genetic variation of populations. DNA-based molecular markers have a very high level of polymorphism and have been successfully used to evaluate genetic variation of populations in breeding programs and to access the level of genetic conservation schemes (Martín-Burriel et al., 1999; Almeida et al., 2000; Regitano et al., 1999).

In tropical regions, Zebu (*Bos indicus*) breeds are very important due to their heat and parasite tolerance. Nelore, Gyr and Guzerat are Zebu breeds very common in Brazil and they are being used to generate *Bos taurus* X *Bos indicus* crosses in order to combine good production with parasite and heat tolerance in synthetic breeds. In order to maintain breed diversity it is necessary to know how far apart they are.

¹ Embrapa Gado de Leite - Rua Eugênio do Nascimento, 610 - Dom Bosco - 36038-330 - Juiz de Fora, MG, Brazil.

² BIOAGRO - Universidade Federal de Viçosa - 36571-000 - Viçosa, MG, Brazil.

* Corresponding author. E.mail: machado@cnppl.embrapa.br

In Brazil, the first Nellore, Gyr and Guzarat animals were imported from India in the 1930's and introduced in herds that aimed meat production (Santiago, 1985). Later in the 60's some Gyr breeders started a selection for milk production seeking morphological traits not included in the former objectives of meat production. At the same time, few Guzarat breeders also started a selection line for milk production making this breed suitable for milk and meat production. From the 70's until now, Nellore breed is exclusively explored for meat production, Guzarat for meat and milk purposes and Gyr breed has selected lines for both milk and meat production. Although the number of animals imported from India was relatively small (~6,000), these animals adapted very well in the Brazilian environment and multiplication developed rapidly (Santiago, 1985). The intensive selection practices and the small number of ancestral animals may have caused a decrease in the genetic diversity of Zebu breeds in Brazil.

In this work we investigated the allele frequency of nine microsatellite markers in *Bos indicus* and *Bos taurus* breeds in order to evaluate the genetic diversity within and among breeds. This will be important to explore heterosis in future complementary crosses within Zebu breeds and also for crosses between Zebu and European breeds.

Materials and Methods

Animals

Eighteen sires from three Zebu breeds (Nellore, Gyr and Guzarat) and eighteen sires from one *Bos taurus* breed (Holstein) were analyzed in this study. The most popular sires were picked based on the number of semen samples distributed from the breeders associations to farmers what represents most of the genetic component of these breeds throughout the country. Selected sires were considered unrelated because no common ascendance up to third generation was found among them. Four semen pallets were collected from each animal as source of DNA.

Microsatellite markers

A total of 9 microsatellite markers were studied in the four different breeds. Markers were chosen based on the size of the product (100 - 250 bp), level of Polymorphic Information Content (PIC) and easiness of scoring. Both di and tri nucleotides microsatellites were picked. ARO23, ARO42, (Band

& Ron, 1996) BMS2684, BMS518, BMS3004, BMS1120, BMS1237, BMS1126 and BL4 were picked from USDA map (Kappes et al., 1997).

Microsatellite typing

DNA samples were extracted from semen using a salt/proteinase K protocol and proteins were removed with phenol-chloroform treatment. Polymerase Chain Reaction were performed using 50 ng of template DNA in 15 µL reaction volume using 1 unit of Taq polymerase with reaction buffer containing 500 mM KCl; 100 mM Tris-HCl, pH 8.0; 20 mM MgCL₂; 1% Triton X-100; 0.2 mM dNTPs. A 3 min denaturation step at 94°C was followed by 10 cycles of 30 s denaturation step at 94°C; 30 s annealing (-1°C/cycle) until reaching primer optimal temperature - "touchdown cycle profile" (Don et al., 1991); and 60 s extension step at 72°C. Final amplification consisted of 25 cycles of 30 s denaturation step at 94°C; 30 s annealing at primer optimal temperature; 60 s extension at 72°C followed by 7 min of final extension step at 72°C. A total of 8 mL of the amplification reaction of each sample were loaded onto a native 13% acrylamide gel 0.75 mm thick. In each gel it was also included a 25 bp ladder to allow determination of allele lengths. All 72 samples plus 5 standards were loaded in the same gel to facilitate scoring. Gels were fixed in one glass plate, stained with ethidium bromide for 30 min and photographed under UV light using a digital camera apparatus (EagleEye II, Stratagene Co.).

Data analysis

Analysis of the bands were done using the software RFLP scan (Scanalytics Co.) and the results were exported to an excel data sheet. Allele frequencies, direct count heterozygosity and Hardy-Weinberg expected heterozygosity were estimated for each marker and breed on the basis of frequency data.

Direct count heterozygosity was obtained as:

$$H(\text{direct}) = \sum_i \sum_{j \neq i} \frac{N_{ij}}{N}$$

where N_{ij} is the number of heterozygous individuals in the l locus; N is the number of individuals analyzed.

The Hardy-Weinberg expected heterozygosity, also defined as Gene Diversity (Weir, 1996) or Polymorphism Index Content (PIC) (Botstein et al., 1980), was obtained from observed allele frequencies:

$$H(HW) = 1 - \sum_{i=1}^n p_{li}^2$$

where p_{li} is the frequency of the i allele at the l locus; n is the number of alleles at the l locus.

The Hardy-Weinberg equilibrium was evaluated by the likelihood ratio test (Lynch and Walsh, 1997), using G statistics:

$$G = -2 \sum_{i=1}^n \sum_{j \geq i}^n N_{ij} \ln \left(\frac{\hat{N}_{ij}}{N_{ij}} \right)$$

where N_{ij} and \hat{N}_{ij} are the observed and expected numbers of the genotype $A_i A_j$ in the sample; n is the number of alleles at the locus. G statistics has a χ^2 distribution with $n(n-1)/2$ degrees of freedom (Weir, 1996).

Standard genetic distance of Nei and Li (1979) were calculated from the allele frequencies:

$$D = -\ln \frac{J_{PQ}}{\sqrt{(J_P J_Q)}}$$

where

$$J_{PQ} = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^{m_i} p_{ij} q_{ij}; \quad J_P = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^{m_i} p_{ij}^2; \quad J_Q = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^{m_i} q_{ij}^2;$$

p_{ij} and q_{ij} are the frequencies of the i allele at the j locus in the p and q populations; m is the number of loci.

Genetic dissimilarities among all individuals were obtained by the simple match coefficient complement using GQMOL software (Cruz & Schuster, 2001).

Distance matrix data were used to construct a dendrogram using the unweighted pair-group with arithmetic mean (UPGMA).

Results and Discussion

The touchdown cycle profile minimized most of the artifacts from the PCR reaction improving the product quality and the easiness of gel scoring. Electrophoresis on native acrylamide gel showed enough resolution to visualize 2 bp differences enabling allele scoring. For that, a 40 cm tall x 50 cm wide plate was used in an electrophoresis period of 7 hours. The results obtained with native acrylamide gel and ethidium bromide staining showed no stutter bands

that are very common in di and tri-nucleotides microsatellites loaded on denaturing gels and complicate much the scoring.

A total of 64 alleles in all 4 breeds were obtained with 9 primers used and each breed showed an average of 53% of the total number of alleles. The average number of alleles per locus was 7.11 ± 3.2 . The most informative locus was BMS1237 (53% of direct count heterozygosity) and the least informative one was BMS 3004 (12%) (Table 1). Two loci were found to be very little informative. BMS3004 showed very little heterozygosity (5%) for Gyr, Guzerat and Holstein breeds and BL4 showed 5% heterozygosity in Gyr, 0% in Nellore and 11% in Guzerat. BMS1237, on the other hand, showed very high heterozygosity for Gyr (33%), Nellore (53%), Guzerat (41%) and Holstein (82%) (Table 1). The most informative primers BMS1237, BMS1126 and BMS518 are good candidates to be used for paternity testing due to the high direct count heterozygosity, high number of alleles, and allele frequencies well distributed (data not shown).

The average direct count heterozygosity for the nine loci was 35% and the expected Hardy-Weinberg heterozygosity was 53%. For most loci the direct count and Hardy-Weinberg heterozygosity showed similar values although few loci showed less heterozygosity than expected in each breed.

Hardy-Weinberg Equilibrium (HWE) was tested in each locus using G statistics (Table 1). Four loci in Nellore breed (ARO23, BMS1120, BMS1126 and BL4), two loci in Guzerat breed (BMS2684 and BL4), and three loci in Gyr breed (BMS2684, BMS1237 and BL4) have shown deviation from HWE. For the Holstein breed, five loci showed deviation from HWE (BMS2684, BMS518, BMS1120, BMS1126 and BL4). For Nellore and Holstein breeds, most loci showed very low HWE probabilities. These results suggest that within these breeds the selected animals have shown endogamy that could be a result of the intensive selection process in order to fulfill dairy and beef industry needs. Guzerat and Gyr breeds showed deviation from HWE mainly for loci that had a general consistent deviation for most of the breeds (BL4 and BMS2684) indicating that these breeds are less endogamic than Nellore and Holstein breeds. Gyr breed also showed deviation from HWE at loci BMS1237 what might be due to sampling problems, once most loci showed equilibrium. One loci (BL4) showed very high deviation from HWE indicating that it might be associated with traits under selection common to all four breeds.

Table 1 - Hardy-Weinberg heterozygosity H(HW); direct count heterozygosity H(DC); G statistics; and Hardy-Weinberg equilibrium probability (P) for each locus combined to four cattle breeds Gyr, Nellore, Guzerat and Holstein

Locus	Nellore			Guzerat			Gyr			Holstein			Average	
	H(HW)	H(DC)	G (df)†	P (%)	H(HW)	H(DC)	G (df)	P (%)	H(HW)	H(DC)	G (df)	P (%)	H(HW)	H(DC)
ARO23	0.1172	0.1250	20.03 (3)	1.67*	0.2778	0.3333	0.13 (1)	71.84	0.5859	0.3750	1.21 (1)	27.13	0.5527	0.5000
ARO41	0.4753	0.4444	3.38 (1)	6.60	0.4444	0.3333	0.08 (1)	77.73	0.4861	0.2778	1.10 (1)	29.43	0.4985	0.2778
BMS2684	0.6270	0.3125	23.41 (15)	7.58	0.6938	0.4118	15.82 (6)	1.47*	0.6578	0.3333	21.03 (6)	0.18*	0.7457	0.2353
BMS518	0.6867	0.5000	12.90 (15)	61.00	0.7253	0.5556	13.33 (10)	20.58	0.6605	0.6111	29.26 (36)	77.94	0.6620	0.2222
BMS3004	0.4377	0.2941	0.03 (1)	86.71	0.0571	0.0588	1.78 (1)	18.21	0.0540	0.0556	0.03 (1)	86.25	0.0540	0.0556
BMS1120	0.7076	0.4444	20.63 (10)	2.38*	0.7515	0.4444	10.62 (6)	10.08	0.7266	0.4375	18.06 (10)	5.39	0.2022	0.0556
BMS1237	0.6384	0.5294	17.74 (10)	5.95	0.6869	0.4118	4.26 (6)	64.15	0.7356	0.3333	15.93 (6)	1.41*	0.6384	0.8235
BMS1126	0.4892	0.6667	9.56 (3)	2.27*	0.7128	0.8235	6.28 (6)	39.26	0.5867	0.2667	11.27 (10)	33.69	0.4753	0.3333
BL4	0.3704	0.0000	20.64 (3)	1.01*	0.6559	0.1111	24.61 (3)	1.86*	0.4550	0.0588	30.98 (6)	2.56*	0.7901	0.5556
Average	0.5055	0.3685			0.5562	0.3871			0.5498	0.3055			0.5132	0.3399

† G statistics for Hardy-Weinberg equilibrium. Degrees of freedom in parenthesis.

* Significant at 5% level.

The Nei's genetic distance matrix (Table 2) shows that Holstein breed was the most divergent from the others: 1.15 in relation to Gyr, 1.12 in relation to Nellore, 0.94 in relation to Guzerat. This result was expected since Holstein is a *Bos taurus* breed while Gyr, Nellore and Guzerat are *Bos indicus* breeds. The closest related breeds were Guzerat and Nellore (0.25) that agrees with the close geographical origins of these two breeds in India. Cluster analysis (Figure 1) showed that *Bos indicus* breeds cluster independently from the *Bos taurus* Holstein breed at 0.33 of genetic distance. Kemenes et al. (1999) used allele frequencies of three candidate genes (k-casein, b-lactoglobulin and growth hormone) and showed a very distinct clustering between *Bos taurus* and *Bos indicus* populations. Within *Bos indicus* breeds the allele frequencies were not sufficient to show a good discrimination among them. This was probably due to the fact that their results were based on three loci only and also because candidate genes are generally less polymorphic than microsatellites markers that were used in our study.

These results may be confirmed by a more detailed dendrogram (Figure 2). Using simple match coefficient

Table 2 - Nei's genetic distance matrix obtained from allele frequencies of 9 loci in four cattle breeds

	Gyr	Nellore	Guzerat
Nellore	0.35		
Guzerat	0.25	0.21	
Holstein	1.15	1.13	0.94

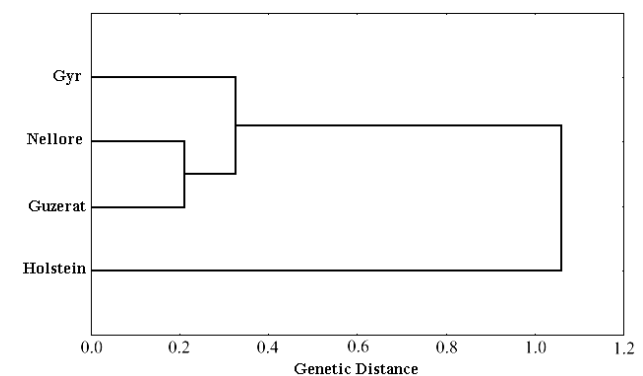


Figure 1 - UPGMA dendrogram generated from Nei's genetic distances of four cattle breeds.

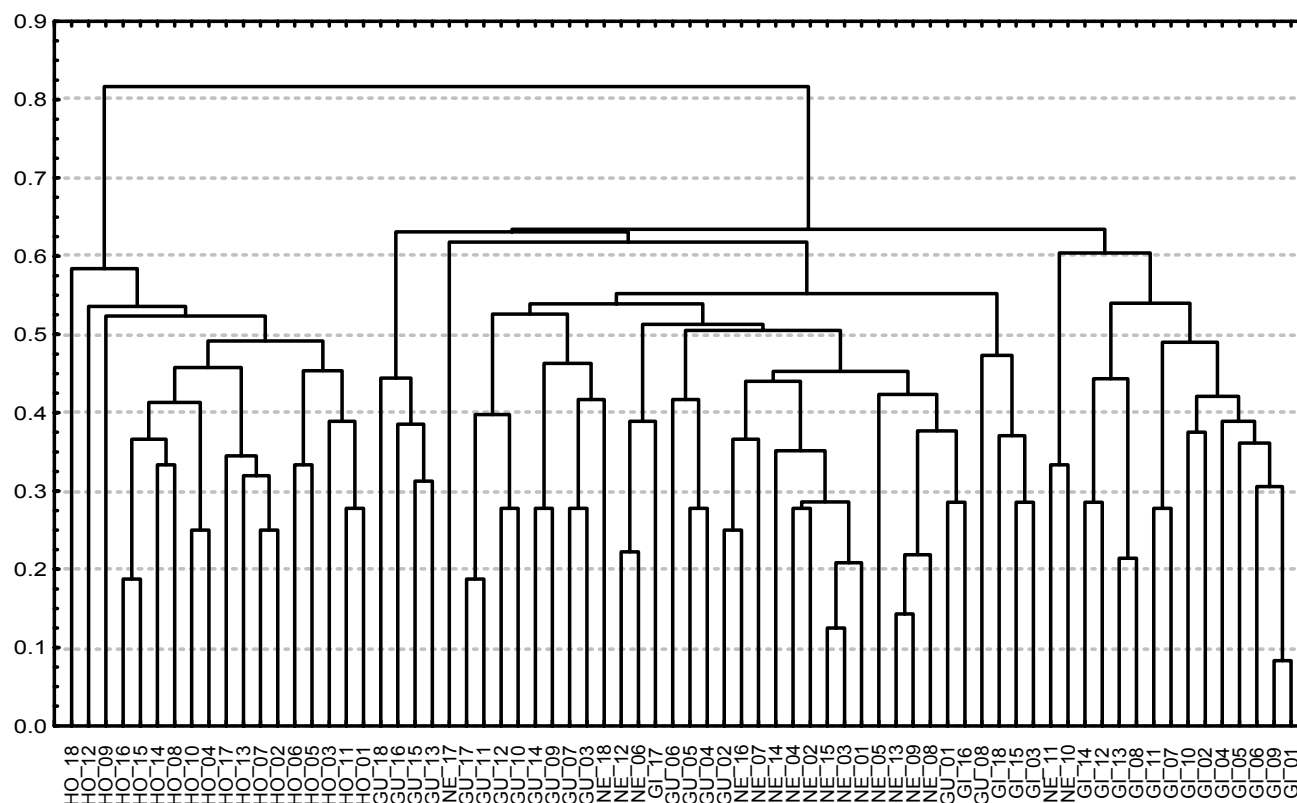


Figure 2 - UPGMA dendrogram using simple match coefficient complement among all individuals of four cattle breeds: Gyr (Gi), Nellore (Ne), Guzerat (Gu) and Holstein (Ho).

complement it was estimated the genetic divergence among individuals within each breed. By the results shown in this dendrogram it might be seen that all Holstein individuals constitute a single group. For the other three breeds some overlapping are shown making individuals from different breeds appear in the same group. These results indicate a smaller variation among the Zebu breeds and that the Holstein breed has enough divergence from the Zebu breeds, once no overlapping was found between them.

Conclusions

Nellore and Holstein breeds showed endogamy due to the deviation from HWE in some loci while Gyr and Guzerat breeds were less affected. Nellore and Holstein are the most explored breeds for meat and dairy industry respectively and this could explain their higher endogamy. Gyr and Guzerat on the other hand are less explored breeds and showed less deviation from HWE. The BL4 loci might be associated with traits undergoing selection because it showed

consistent deviation from HWE in all breeds.

The *Bos taurus* Holstein breed grouped very distantly from the *Bos indicus* breeds. Even though *Bos indicus* breeds showed some relatedness they were grouped individually in the cluster analysis.

Acknowledgements

This work was supported by Projeto de Apoio ao Desenvolvimento de Tecnologias Agropecuárias para o Brasil (PRODETAB), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Embrapa Gado de Leite.

Cited Literature

- ALMEIDA, S.E.M.; MACHADO, M.S.N.; STEIGLEDER, C.S. et al. Genetic diversity in a Brazilian bovine herd based on four microsatellite markers. **Genetics and Molecular Biology**, v.23, n.2, p.347-350, 2000.
- BAND, M.; RON, M. Creation of a SINE enriched library for the isolation of polymorphic (AGC)_n microsatellite markers in the bovine genome. **Animal Genetics**, v.27, p.243-248, 1996.

- BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M. et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v.32, p.314-331, 1980.
- CRUZ, C.D.; SCHUSTER, I. QMOL: genética quantitativa e molecular. <http://www.ufv.br/dbg/home3.html> (2001).
- DON, R.H.; COX, P.T.; WAINWRIGTH, B.J. et al. "Touchdown" PCR to circumvent spurious priming during gene amplification. **Nucleic Acids Research**, v.19, p.4008, 1991.
- KAPPES, S.M.; KEELE, J.W.; STONE, R.T. et al. A second-generation linkage map of the bovine genome. **Genome Research**, v.7, p.235, 1997.
- KEMENES, P.A.; REGITANO, L.C.A.; ROSA, A.J.M. et al. k-casein, b-lactoglobulin and growth hormone allele frequencies and genetic distances in Nellore, Gyr, Guzerat, Caracu, Charolais, Canchim and Santa Gertrudis cattle. **Genetics and Molecular Biology**, v.22, n.4, p.539-541, 1999.
- LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland, MA: Sinauer Association, 1997. 980p.
- MARTIN-BURRIEL, I.; GARCÍA-MURO, E.; ZARAGOZA, P. Genetic diversity analysis of six Spanish native cattle breeds using microsatellites. **Animal Genetics**, v.30, p.177-182, 1999.
- NEI, M.; LI, W. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proceedings of National Academy of Science USA**, v.76, p.5269-5273, 1979.
- REGITANO, L.C.A.; AZEVEDO, J.L.; VENCOVSKY, R. et al. Selection for breed-specific growth hormone and IGF-1 alleles in a synthetic beef cattle cross, Canchim. **Genetics and Molecular Biology**, v.22, n.4, p.531-537, 1999.
- SANTIAGO, A.A. **O Zebu na Índia, no Brasil e no mundo**. Campinas: Instituto Campineiro de Ensino Agrícola, 1985. 744p.
- WEIR, B.S. **Genetic data analysis**. Sunderland, MA: Sinauer Association, 1996. 445p.

Recebido em: 19/11/01

Aceito em: 01/10/02