Biometric variability of goat populations revealed by means of principal component analysis

Luanna Chácara Pires¹, Théa M. Medeiros Machado², Adriana Mello Araújo¹, Timothy A. Olson⁴, João Batista Lopes da Silva³, Robledo Almeida Torres² and Márcio da Silva Costa⁶

¹Departamento de Agronomia, Universidade Federal de Piauí, Bom Jesus, PI, Brazil.
²Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, MG, Brazil.
³Empresa Brasileira de Pesquisa Agropecuária, Embrapa Meio Norte, Teresina, PI, Brazil.
⁴Department of Animal Science, University of Florida, Gainesville, FL, USA.
⁵Departamento de Engenharia de Agricultura, Universidade Federal de Viçosa, Viçosa, MG, Brazil.
⁶Departamento de Ciência Animal, Universidade Federal de Piauí, Teresina, PI, Brazil.

Abstract

The aim was to analyze variation in 12 Brazilian and Moroccan goat populations, and, through principal component analysis (PCA), check the importance of body measures and their indices as a means of distinguishing among individuals and populations. The biometric measurements were wither height (WH), brisket height (BH) and ear length (EL). Thorax depth (WH-BH) and the three indices, TD/WH, EL/TD and EL/WH, were also calculated. Of the seven components extracted, the first three principal components were sufficient to explain 99.5% of the total variance of the data. Graphical dispersion by genetic groups revealed that European dairy breeds clustered together. The Moroccan breeds were separated into two groups, one comprising the Drâa and the other the Zagora and Rhâali breeds. Whereas, on the one side, the Anglo-Nubian and undefined breeds were the closest to one another the goats of the Azul were observed to have the highest variation of all the breeds. The Anglo-Nubian and Boer breeds were similar to each other. The Nambi-type goats remained distinct from all the other populations. In general, the use of graphical representation of PCA values allowed to distinguish genetic groups.

Keywords: biometrics, genetic resources, goat breeds, morphometrics, population characterization.

Received: March 14, 2012; Accepted: May 29, 2012.

Introduction

Geographic isolation over time, along with small population size and the consequences of natural and artificial selection in domesticated animals, have contributed to the appearance of a large variety of genetic types within species, sometimes called breeds. These same factors have also contributed to variation within the breeds themselves (Hall and Bradley, 1995; Bruford et al., 2003).

The standard procedure used in Brazil for defining a goat breed is first there cognition of a standard population phenotype by the Ministry of Agriculture, followed by its registration with the Goat Genealogical Registration Service (Brasil, 1977). This definition of breed, a group of animals with standardized qualitative characteristics, was developed in England during the eighteenth century (Audiot, 1985; Sierra Alfranca, 2001).

Farm animals of colonial Brazil originally came from Portugal and the Cape Verde islands, and later from Spain. Sahelian goats from the Western Sahara and their crosses entered Brazil by way of the Cape Verde Islands, where they had been previously introduced by the Portuguese. The Portuguese had acquired Arabian, Persian (Iran) and Maghrebian (Morocco, Algeria and Tunisia) horses, before Brazil was discovered and later were thought to have introduced their descendants into Brazil. An ancient North African influence is also generally assumed regarding the origin of asinine, equine and ovine animals of the Iberian Peninsula (Machado TMM, 1995, PhD thesis, University of Paris XI, Paris). Andrade (1982) stated that the introduction of goats and other livestock species to Brazil coincided with the arrival of the first African slaves.

At the beginning of the nineteenth century, goat breeds selected for improving production began to arrive in Brazil. The European breeds, Alpine, Saanen and Toggenburg, were selected for increasing milk production and the Anglo-Nubian for both milk and meat. The Anglo-Nubian breed was formed in England by crossing African long-drooping-eared goats from the Nile and common English goats (Mason, 1988). The Boer breed, developed by South African farmers of the eastern Cape, came about...
by crossing indigenous Hottentot and Bantu animals with imported Nubian and Indian goats (Skinner, 1972; Erasmus, 2000; Malan, 2000).

Most Brazilian goats do not have a standardized phenotype, and thus, are classified as an undefined breed (UDB). Not with standing, some goats that were introduced into northwestern Brazil in the distant past have acquired peculiar phenotypic characteristics, e.g. the Marota goat is a white-coated ecotype. The Azul ecotype designates roan-coated goats with a fawn pigment pattern, thereby assuming a gray or “Azul” appearance (Machado TMM, 1995, PhD thesis, University of Paris XI, Paris). The Nambi goat’s distinguishing characteristic is its tiny ears. Among the Moroccan goat populations, only the Drâa have well-defined phenotypic characteristics and production (milk and meat) information (Hossaini-Hilari and Benlamlilh, 1995; Hossaini-Hilari and Mouslish, 2002). On comparing French and Moroccan goats using INRA microsatellite markers and casein polymorphism, it was noted that the Drâa-Zagora sample clustered with the Rhâali, separate from the French goats; among which, goats from the Pyrenees formed a separate stem from the Saanen, Alpine and Poitevine (Ouali et al., 2002).

The physical measurements were taken based in the knowledge that considerable variation exists among the various groups, from long-legged to midget goats and from goats with long, droopy ears to those with short, erect ears (Epstein, 1971; Mason, 1988).

Over the past two decades, improvements in phenotype marker methodology have facilitated the simultaneous measurement of many characteristics (Lauvergne, 1988; Rodero et al., 1992; Jordana et al., 1993; Herrera et al., 1996; Jordana and Parés, 1999; Bedotti et al., 2004; Dossa et al., 2007; Rocha et al., 2007; Chàcon et al., 2011), the establishment of indices between two body measurements (Bourrat et al., 1993; Bouchet et al., 1997; Lauvergne et al., 2000), and the integration of genetics statistics and informatics solver the last two decades. Subsequently, phenotypic and molecular markers began to be analyzed jointly using multivariate analyses. Breeds have acquired a quantitative measure for uniformity among their individuals (reduced intra-population variability) and become different from individuals of other breeds (variability amongst populations). The biological importance of biometric measures submitted to multivariate analyses is the observation of clustering of some individuals, and the distance between them and those of other populations. Those individuals clustered together (i.e. similar for phenotypic measures) could be recognized as an official breed.

Phenetic distances establish relations of similarity among populations with the purpose of classification, studied on the genetic and evolutionary distances (Meyer, 1996).

Principal component analysis (PCA) consists of transforming a set of variables $Z_1, Z_2, \ldots, Z_p$ into a new set of uncorrelated variables $Y_1 (CP_1), Y_2 (CP_2), \ldots, Y_p (CP_p)$ that are arranged in descending order of variance (Liberato et al., 1999). Each principal component is a linear combination of all the original variables, independent of each other, and estimated, with the aim of retaining the maximum of information and total variation contained in the original data. The main i.e. of this procedure, however, is that a few of the first principal components contain most of the variability of the original data. PCA can also unveil relationships not previously identified, contributing for a better interpretation of the collected data (Baker et al., 1988).

The aim of this study was to verify the ability of body measurements and their indices using PCA to discriminate among individuals and populations.

Material and Methods

Obtained data were of 796 Brazilian and Moroccan female goats, all over two-years-old. Additional data, collected between 1995 and 2008, included information on phenotype and ownership of each animal sampled.

Exotic breeds were sampled (34 Toggenburg, 86 Saanen, 28 Anglo-Nubian, 78 Alpine and 26 Boer) in Brazil (Minas Gerais State and in the Federal District). In addition, in the State of Piauí, 29 heads of the Azul type, 32 of Marota type, 35 of Nambi type and 123 goats lacking defined breed characteristics (UDB) were measured. The exotic breed goats sampled in Brazil, the Alpine, Saanen, Toggenburg, Anglo-Nubian and Boer, are registered in their respective Brazilian herdbooks. The local Brazilian goats, Azul, Marota, Nambi and UDB lack herdbook information.

In Morocco, the sample comprised 102 local Drâa goats from the Center for Goat Research in the village of Sidi Flah, in Skoura. A further 34 Zagora, considered to be Drâa crossbreds were studied in Demnate and Ouarzazate in Tahaanouat the Center for Goat Research in the Marrakech area. Finally, 189 Rhâali goats from the Zagora region were also included. There is no herdbook information for any goat breed in Morocco.

The characteristics evaluated were wither and brisket height, thorax depth and ear length (all measured in cm). Body morphometric measurements were taken using a tape measure on animals that were standing on the ground. Wither height (WH) was considered as the distance between the highest point of the wither to the front-leg distal extremity, brisket height (BH), as the distance from the brisket to the ground, ear length (EL), as the distance from ear base to end. Thorax depth (TD), was calculated from the difference between wither (WH) and brisket (BH) heights, as direct measurement of TD is not easily accomplished in the field. Therefore, TD was calculated from the difference between two measures (WH-BH). Indices/relationships were also calculated between these measurements, such as TD/WH, EL/TD and EL/WH, for each individual. The absence of valid records for birth-dates required estimating
age through dental chronology according to the methodology described by Quittet (1978).

Due to the use of variables measured in different units (linear measures as well as ratios between them), standardization of the variables \( Z_{ij} \) was necessary. This standardized measurement: \( Z_{ij} = (X_{ij} - \bar{X}_j)/S_j \), where \( X_{ij} \) is the value of \( i \)-th observation of the variable \( X_j \), \( \bar{X}_j \) is the mean of variable \( X_j \), and \( S_j \) is the standard deviation of variable \( X_j \).

Before the collected physical measurements were subjected to analysis using multivariate methods, they were submitted to preliminary analyses to remove invalid data, particularly to discard information in compatible with stated age and to eliminate animals with missing information. Biometric data were first analyzed using simple descriptive statistics (mean, standard deviation and coefficient of variation). The coefficient of variation (CV) was calculated via the expression \( CV = (S_j/\bar{X}_j)\times 100 \), where \( S_j \) and \( \bar{X}_j \) were defined previously. Variables were then submitted to principal component analysis (PCA) to so reduce data dimensionality and enable discrimination of groups by individuals and populations.

The criterion for discarding variables from the PCA followed the recommendations of Jolliffe (1973), based on actual and simulated data from the correlation matrix. It was defined that the number of discarded variables should be the same as the components whose variance (eigenvalue) is not greater than 0.7.

Multivariate analysis was with SAS version 8.0 software, under license to the Universidade Federal de Viçosa (SAS Institute Inc., Cary, NC, USA, 1999) and GENES version 6.0 software (Cruz, 2008).

Results and Discussion

The coefficients of variation of characteristics and indices showed these to be precisely estimated. The body morphometric measurements (WH, BH, EL and TD) presented coefficients of variation no greater than 12.8% (Table 1). The highest variability was in EL. In one group of goats, the Nambi-type, ears were substantially shorter, whereas other groups characteristic ally had average-sized to long ears. The coefficient of variation values of WH, EL and TD were similar to those found by Dossa et al. (2007).

Based on the PCA results, the respective eigenvalues, and the percentages of explained variance (Table 2) from the seven principal components, four of the components (57.14%) yielded variances no greater than 0.7 (eigenvalues no greater than 0.7). The first three principal components were selected and explained 99.5% of the total variation.

As presented in Table 3, the four variables that presented the highest coefficients, in absolute value, from the last principal component can be discarded. Hence, the most appropriate characteristics to discard, in order of the least important for explaining total variation were TD, EL/WH TD/WH and EL/TD. Based on our results, wither height, brisket height and ear length are recommended for use in future studies. The indices and TD accounted for only 0.5% of the total variability. In this study, they were of little importance in the evaluation of interbreed differences. Notably, those characteristics recommended for exclusion were highly correlated with the selected variables. Based on the results, it is not recommended that thoracic depth be calcul-

### Table 1 - Means, standard deviation and coefficient of variation of body measures and indices in Brazilian and Moroccan goats.

<table>
<thead>
<tr>
<th>Characters</th>
<th>Mean (cm)</th>
<th>Standard deviation (cm)</th>
<th>Coefficient of variation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WH</td>
<td>68.0</td>
<td>7.5</td>
<td>6.5</td>
</tr>
<tr>
<td>BH</td>
<td>36.6</td>
<td>4.8</td>
<td>9.1</td>
</tr>
<tr>
<td>EL</td>
<td>16.3</td>
<td>3.9</td>
<td>12.8</td>
</tr>
<tr>
<td>TD</td>
<td>31.4</td>
<td>5.3</td>
<td>10.3</td>
</tr>
<tr>
<td>TD/WH</td>
<td>0.5</td>
<td>0.1</td>
<td>8.2</td>
</tr>
<tr>
<td>EL/TD</td>
<td>0.5</td>
<td>0.2</td>
<td>18.2</td>
</tr>
<tr>
<td>EL/WH</td>
<td>0.2</td>
<td>0.1</td>
<td>13.0</td>
</tr>
</tbody>
</table>

WH - wither height, pH - brisket height, EL - ear length, PT - thoracic depth.

### Table 2 - Principal Components (PC), eigenvalues (\( \lambda_i \)) and variance percentage explained by components (simple variance and accumulated variance) of measured characteristics in Brazilian and Moroccan goats.

<table>
<thead>
<tr>
<th>Number</th>
<th>( \lambda_i )</th>
<th>Simple variance (%)</th>
<th>Accumulated variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>3.479</td>
<td>49.69</td>
<td>49.69</td>
</tr>
<tr>
<td>PC2</td>
<td>1.971</td>
<td>28.15</td>
<td>77.84</td>
</tr>
<tr>
<td>PC3</td>
<td>1.517</td>
<td>21.67</td>
<td>99.51</td>
</tr>
<tr>
<td>PC4</td>
<td>0.025</td>
<td>0.36</td>
<td>99.87</td>
</tr>
<tr>
<td>PC5</td>
<td>0.005</td>
<td>0.07</td>
<td>99.94</td>
</tr>
<tr>
<td>PC6</td>
<td>0.003</td>
<td>0.05</td>
<td>99.99</td>
</tr>
<tr>
<td>PC7</td>
<td>0.001</td>
<td>0.01</td>
<td>100.00</td>
</tr>
</tbody>
</table>

### Table 3 - Coefficients of weight of the variables with the four less important principal components to explain total variation.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
</tr>
</thead>
<tbody>
<tr>
<td>WH</td>
<td>0.1410</td>
<td>0.0299</td>
<td>-0.0929</td>
<td>-0.6264</td>
</tr>
<tr>
<td>PH</td>
<td>-0.1504</td>
<td>0.4627</td>
<td>0.1377</td>
<td>0.4606</td>
</tr>
<tr>
<td>EL</td>
<td>-0.5216</td>
<td>-0.3762</td>
<td>0.3704</td>
<td>0.0012</td>
</tr>
<tr>
<td>TD</td>
<td>0.3366</td>
<td>-0.3752</td>
<td>-0.2566</td>
<td>0.7101</td>
</tr>
<tr>
<td>TD/WH</td>
<td>-0.0159</td>
<td><strong>0.5930</strong></td>
<td>0.3966</td>
<td>-0.0010</td>
</tr>
<tr>
<td>EL/TD</td>
<td><strong>0.7538</strong></td>
<td>0.0371</td>
<td>0.3693</td>
<td>0.0018</td>
</tr>
<tr>
<td>EL/WH</td>
<td>-0.0598</td>
<td>0.3868</td>
<td><strong>-0.6897</strong></td>
<td>-0.0017</td>
</tr>
</tbody>
</table>

WH - wither height, BH - brisket height, EL - ear length, TD - thoracic depth.

*The bold-face values are of the variable with the highest scores in the last four principal components.
lated in future studies, but, instead, be substituted by brisket height. A population data set with additional animals and populations, and possibly much more diverse than that applied here, might have given rise to a different set of characteristics.

As can be seen in the graphical representation of the individual distribution according to principal components 1 and 2 (Figure 1), individuals from the European dairy breeds, Alpine, Saanen and Toggenburg, predominantly occupied the upper left quadrant. The relationships between these three breeds were corroborated by the grouping of the Saanen and Alpine, and the further grouping of the latter with the Toggenburg, as observed by electrophoresis of serum and erythrocyte proteins, and by microsatellite markers (Igarashi et al., 2000; Oliveira et al., 2007). The fact that our principal components data, based on physical measurements of these purebred animals of known pedigree origin, grouped these breeds together in the same way as did the protein and microsatellite markers, corroborates the use of biometric data and PCA for discerning differences among groups of individuals within a variety of populations.

The Moroccan goats were separated into two groups (Figure 1). Observations from the Drâa breed were located in the upper quadrants, mainly to the right, whereas those of the Zagora and Rhâali were located in the center of the lower quadrants. Some Drâa individuals were situated among the European dairy breeds. These results are partially in accordance with Machado et al. (2000), who, by comparing allelic frequencies of morphological characters, found that, where as the Zagora and Rhâali breeds were the closest grouped among the Moroccan types, the Drâa grouped together with Mediterranean-European goats. On the other hand, Ouali et al. (2002) report that in their phylologic tree the Drâa-Zagora and Rhâali goats were very much separated from the Alpine and Saanen goats. Considering the data obtained from this experiment and those obtained by Machado et al. (2000), arbitrary pre-grouping of Drâa and Zagora samples can be considered inadvisable.

Data from the Boer breed occupied the central area of the upper quadrants, especially that on the right, together with the Zagora, Drâa and Anglo-Nubian groups. The African component of the Boer and Anglo-Nubian breeds could possibly clarify this result. Thus, in this study, relationships among populations through PCA relate well with their respective breed descriptions and the geographical origins of these populations.

Although as a whole, the Anglo-Nubian and undefined breed goats from Piauí were predominantly in the upper right quadrant, some Anglo-Nubian individuals were placed among the European dairy breeds. On the other hand, although some undefined individuals from Piauí were grouped in the right lower quadrant together with the Zagora and Rhâali, they occupied more extreme positions than was the case with other populations. In contrast, Nambi-type goats occupied positions in the far-left portion of the lower left quadrant.

The Brazilian goat types, Marota and Azul, were also positioned in the lower left quadrant, in an intermediate position between the Namí and Moroccan populations. However, in the case of the Azul goats, some individuals overlapped into the right-hand quadrant, thus denoting greater variability than the rest. In this case, the distribution of individuals through principal components 1 and 2, which

Figure 1 - Graphical dispersion of scores of 12 genetic groups of goats in Brazil and Morocco, related to principal components (PC) 1 (abscissa) and 2 (ordinate). Genetic groups: T = Toggenburg, S = Saanen, G = Anglo-Nubian, A = Alpine, B = Boer, D = Drâa, Z = Zagora, R = Rhâali, U = Azul, M = Marota, N = Namí, P = Undefined Breed Goats from Piauí.
retain greater variance, was useful for providing a visual-
ization of intra-population and inter-population variability.

By relating principal components to known breed variabil-
ity (Figures 2 and 3), the Brazilian Nambi-type appears to be removed from all the other groups, due to its short ears. The undefined breed goats from Piauí were also isolated from the other groups and placed in opposition opposite to the Nambi type goats. The closest groups to the undefined breed goats from Piauí were the Moroccan and the Brazilian Azul and Marota goats. The only difference between the groupings in Figures 2 and 3 is the positioning of the Drâa breed, which was included with the other Moroccan breeds and some Brazilian goats in the first analysis and then with the European breeds in the second analysis. This difference is easily explained by the variability among

Drâa individuals, as shown in Figure 1. In Figure 3, the
group constituted by the Moroccan Rhâali plus Zagora and
the Brazilian Azul and Marota goats occupied an interme-
diate position between the European dairy breeds together
with the Drâa, and the Anglo-Nubian together with the
Boer. These results are in accordance with the literature
(Igarashi et al., 2000; Ouali et al., 2000; Oliveira et al.,
2007), and occasionally also with the origins of the respective
populations.

The graphical representation of PCA observations
permits distinguishing among genetic groupings/breeds of
goats, from the most similar to the most dissimilar. It also
facilitates preliminary visualization of the uniformity or
lack of uniformity among individuals within each popu-
lation. Based on results from the Brazilian and Moroccan
goat populations, the biometric indices, wither height, brisket height and ear length are recommended to be continued
to be collected in future evaluations of breed and popula-
diversity. In contrast, including thoracic depth is not
recommended in further studies of this kind. Due to its
greater variability, the Azul goat, an unselected type, dem-
onstrated the least standardization of all the breeds. Among
the Brazilian populations, the Nambi-type goats demon-
strated the highest degree of uniformity, whence the trend
to becoming a designated breed.

Acknowledgments

The authors would like to thank Mohamed Chakir for
his contribution to data collection in Morocco during his post
doctorate at INRA, Jouy-in-Josas, France.

References

Andrade MC (1982) Área de Domínio da Pecuária Extensiva e
Semi-Intensiva na Bahia e Norte de Minas. SUDENE,
Recife, 468 pp.

Audiot A (1985) La variant ‘oreilles raccourcies’ de la chèvre

Baker JF, Stewart TS, Long CR and Cartwright TC (1988) Multi-
ple regression and principal components analysis of puberty

Bedotti D, Gómez Castro AG, Sánchez Rodriguez M and Martos
Peinado J (2004) Morphologic and phaneroptic charac-

Brasil (1977) Leis, Decretos, etc. Portaria n° 11 de 1° Dezembro
de 1977. Padrões Raciais dos Caprinos. Diário Oficial, Bra-
silia, pp 17187-17189.

morpho-biométrique de la chèvre du rove. Hauteur au garrot
(HT), profondeur du thorax (PT), vide sous-sternal (VSS) et
indice de gracilité sous-sternale (IGs) chez femelles. Rev

Bourzat D, Souvenir-Zafindrajaona P, Lauvergne JJ and Zueh V
(1993) Comparaison morpho-biométrique dès chèvres au
Nord Cameroun et au Tchad. Rev d’Élevage et de Méd Vét


Associate Editor: Bertram Brenig