GUEST AUTHORS

Genetic Evaluation of Criollo Cattle and Their Crossbreds in Colombia

Mauricio A. Elzo[1],
Carlos Manrique P.[2]
and Gustavo Ossa S.[3]

ABSTRACT

Colombia has seven Criollo breeds of cattle adapted to tropical conditions, with a total number of purebreds of about 18.000, and about 50.000 cattle if Criollo crossbreds are also considered. The 1986 Livestock Development Project suggested increasing their numbers and developing animal improvement programs at a farm level. These goals can be achieved through the development of a national genetic evaluation program involving all Criollo breeds and their crossbreds. The basic population and data structures that would allow the realization of a national genetic evaluation are presented and discussed in terms of the Colombian situation. Current unibreed and multibreed linear methodology, and computational procedures, are discussed, and computer software programs are suggested. A general strategy to develop a National Genetic Evaluation Program is outlined, and specific points discussed in further detail from a Colombian perspective. The implementation of a large national genetic evaluation program will not only benefit the Criollo breeds and their crossbreds, but it can serve as a model for the genetic evaluation of all cattle breeds and their crossbreds in Colombia.

Key words: cattle, criollo, genetic evaluation, crossbreeding, variance components.

1. Department of Animal Science, University of Florida, Gainesville (Florida), 32611-0910, U.S.A.; 2. Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional de Colombia, Santa Fe de Bogotá, Colombia; 3. Centro de Investigación Turipaná, CORPOICA, Apartado Postal 602, Montería (Córdoba), Colombia.

RESUMEN

Título: **Evaluación genética del ganado criollo y sus cruces en Colombia.**

Colombia posee siete razas de bovinos criollos adaptadas a las condiciones tropicales; el número total de animales puros es menor de 18.000, aunque si se consideran los animales criollos cruzados, dicho censo asciende a 50.000 animales aproximadamente. El Proyecto de Desarrollo Ganadero de 1986 sugirió aumentar su número y desarrollar planes de mejoramiento a nivel de finca. Estos objetivos se pueden lograr mediante un programa de evaluación genética nacional que incluya todas las razas criollas y sus cruces. El presente artículo presenta las estructuras básicas de las poblaciones y de los datos, a partir de los cuales es posible realizar una evaluación genética nacional, la cual se analiza en términos de la actual situación colombiana. Se discuten además, métodos lineales y procedimientos computacionales uniraciales y multiraciales para su manejo, y se sugiere el uso de programas de computación. Así mismo, se plantea una estrategia general para desarrollar dicho Programa Nacional de Evaluación Genética, enfatizando en algunos aspectos específicos desde el punto de vista colombiano. La implementación de dicho programa beneficiaría la conservación y el desarrollo de las razas criollas y de sus cruces, además de servir como modelo para desarrollar de manera sistemática la evaluación genética de los bovinos colombianos.

Palabras claves: ganado bovino, criollo, evaluación genética, cruzamientos, componentes de varianza.

INTRODUCTION

COLOMBIA is currently pursuing a Livestock Modernization Plan with the collaboration of producers and governmental institutions. One of the specific objectives of this plan is to improve reproductive and productive traits through the utilization of cattle adapted to tropical conditions. As part of its strategy to achieve this goal, the Livestock Modernization Plan: 1) suggests the development of animal improvement programs at a farm level, and 2) encourages the specialization of researchers in areas of critical importance to allow access and use of advanced methodologies (Afanador, 1996). Colombia is in a good position to achieve these goals because: 1) the country has several Criollo breeds of cattle, although small in number, well adapted to their tropical environment, and 2) it has a sizable number of capable researchers in various research centers and universities. From a genetic improvement standpoint, what is needed is a coordinated effort among producers, government officials, and researchers to develop the infrastructure to genetically evaluate animals of the various Criollo breeds for their use in straightbreeding and crossbreeding. Thus, the objectives of this work are: 1) to present population and data structures that allow the prediction on production of the animal genetic effects in farm conditions, 2) to describe current genetic evaluation methods and computational procedures, and 3) to suggest a strategy for development, in medium term, a national genetic evaluation program for the Criollo breeds and crossbreds in Colombia.

Population and Data Structures for Genetic Predictions under Farm Conditions

Types of Populations

A population is defined as a group of animals (males and females) that interbreed. Populations can be unibreed and multibreed. Unibreeed populations contain animals of a single breed. For example, seven unibreed Criollo populations exist in Colombia: Blanco Orejinegro

(BON), Casanare (C), Chino Santanderiano (CS), Costeño Con Cuernos (CCC), Hartón del Valle (HDV), Romosinuano (R) and San Martinero (SM). Multibreed populations are formed by animals of several breeds and crossbred groups. For example, a two-breed RZ (R x Zebú) multibreed population was formed in the Turipaná Investigation Center (Córdoba region) in the 1980 s. This RZ multibreed population was composed of purebred R and Z as well as RZ, 3/4R1/4Z, 1/2R1/2Z, 1/4R3/4Z and several other crossbred groups of other R and Z fractions. This was a complex multibreed population. A much simpler multibreed population would be one formed by purebreds and first generation crossbreds only (e.g., R, Z, and 1/2R1/2Z). Several simple multibreed populations of this type were to be generated by the mating plans for purebred Criollo bulls of the seven Colombian breeds above menctioned and suggested in the 1986 Livestock Development Project. This plan called for the mating of purebred Criollo bulls to Criollo and to Zebu dams (Bejarano *et al.*, 1986). Criollo cattle will, in all likelihood, continue to be used for intrabreed and interbreed mating programs in Colombia in the foreseeable future. Thus, both types of populations will be considered here.

Number of Animals

The number of purebred and crossbred Criollo cattle in Colombia is small. The 1986 livestock development plan of Criollo breeds (Bejarano *et al.*, 1986) reported the existence of about 15,000 purebred and upgraded dams and 3,000 purebred bulls of the seven Colombian Criollo breeds (BON, C, CS, CCC, HDV, R, and SM). When crossbred Criollo cattle were included, the total number increased to approximately 50,000 animals. Probably a smaller number of purebred and crossbred Criollo cattle exist today. These population numbers are too small to obtain accurate predictions of genetic values for a large number of parents. It would be desirable to increase the number of animals per Criollo breed as much as economically feasible for their use in straightbreeding and crossbreeding programs. On the other hand, the current number of Criollo cattle per breed should not be a deterrent to conducting the necessary research for development national genetic evaluation programs for these populations. Research in genetic evaluation procedures under Colombian conditions would help determine the distribution and amount of data that could be feasibly collected under field conditions, and the accuracies to be expected from such data sets.

Plans to increase the number of cattle per Criollo breed have existed in Colombia for at least 10 years. The 1986 Livestock Development Project based on Criollo breeds suggested the use of multiplication herds and upgrading to increase the numbers of the seven Colombian Criollo breeds (Bejarano *et al.*, 1986). Unfortunately, this plan has yet to materialize. These two options, upgrading and multiplication herds, may still be the most feasible options to increase the numbers of Criollo cattle under current Colombian conditions. In particular, upgrading would be greatly helped by the use of artificial insemination. A third option that might be considered, provided that sufficient funds and trained personnel were available, would be superovulation and embryo transfer. A joint research project between CORPOICA, the University of Antioquia, and Texas A&M University will preserve 400 embryos per breed of all the Colombian Criollo breeds (Estrada, 1996).

Contemporary Groups

A contemporary group is a group of animals that allows a fair comparison among animals of interest. For example, contemporary groups for birth weight (BW), weaning weight (WW), and postweaning gain (PG) are usually defined as follows. Contemporary groups for BW are formed by all calves that: 1) were born in the same farm and the same season, 2) are of the same sex, and 3) came from dams in the same management group. Contemporary groups for WW include all calves that: 1) belong to the same BW contemporary group, 2) were in the same preweaning feeding and management group, and 3) were weaned on the same date. Contemporary groups for PG consider calves that: 1) were in the same WW contemporary group, and 2) were in the same postweaning feeding and management group. It should be noted that PG contemporary groups are nested within WW contemporary groups, which in turn are nested within BW contemporary groups.

The only difference between contemporary groups in unibreed and in multibreed populations is that unibreed contemporary groups contain calves of a single breed only, whereas multibreed contemporary groups include calves of various breeds and crossbred groups.

Connected Data Sets

A data set with linked contemporary groups that permits the comparison of any pair of animals is a connected data set. A connected data set is the basis of a genetic evaluation system. Connections across contemporary groups are created by using reference animals (e.g., sires) in several contemporary groups. Reference animals need *not* be in all comparison groups, but they must create enough connections across contemporary groups such that the comparison between any two evaluated animals is possible. Artificial insemination would be a very important tool to quickly achieve a high degree of connectedness across contemporary groups.

Records in a connected data set should contain the following information: 1) identification of the calf, its sire and dam; if available, its maternal grandsire; 2) the breed composition of each calf and its ancestors; 3) the codes for each fixed effect (e.g., contemporary group, sex of calf, age of dam), and 4) the measurement of each trait being evaluated (adjusted if required).

Pedigree Data Set

The pedigree data set must include identification numbers of all evaluated animals and their immediate ancestors (sire, dam, maternal grandsire), their birth dates, and for multibreed populations, the breed composition of each animal and its ancestors. The animals and ancestors included in the pedigree file will depend on the model chosen for the analysis. For example, if an Animal Model (AM, see definition page 19) is used, all calves, sires, and dams must be included in the pedigree file, but only sires, maternal grandsires, and their male ancestors are needed for a Sire-Maternal Grandsire Model (SMM, see definition page 19).

The pedigree file will be used to construct the inverse of the covariance matrix of additive genetic effects (unibreed population; Henderson, 1975 a,b, and 1976 a,b; Quaas, 1976), and the covariance matrices of additive and nonadditive genetic effects (multibreed population; Elzo, 1996 and 1990 a,b), using sets of computational rules.

Traits and Genetic Effects

The traits considered in the evaluation will depend on the objectives of the evaluation, and the capability of the software. For sire summaries in general, the main limitation is the set of traits that can be recorded in the field. Thus, the usual reproduction and survival traits (e.g., age at first calving, calving interval, survival after 3 days, survival to weaning and to yearling), growth traits (BW, WW, PG), and carcass

Figure 1. Hato de Chino Santandereano

Figure 2. Vaca Blanco Orejinegro

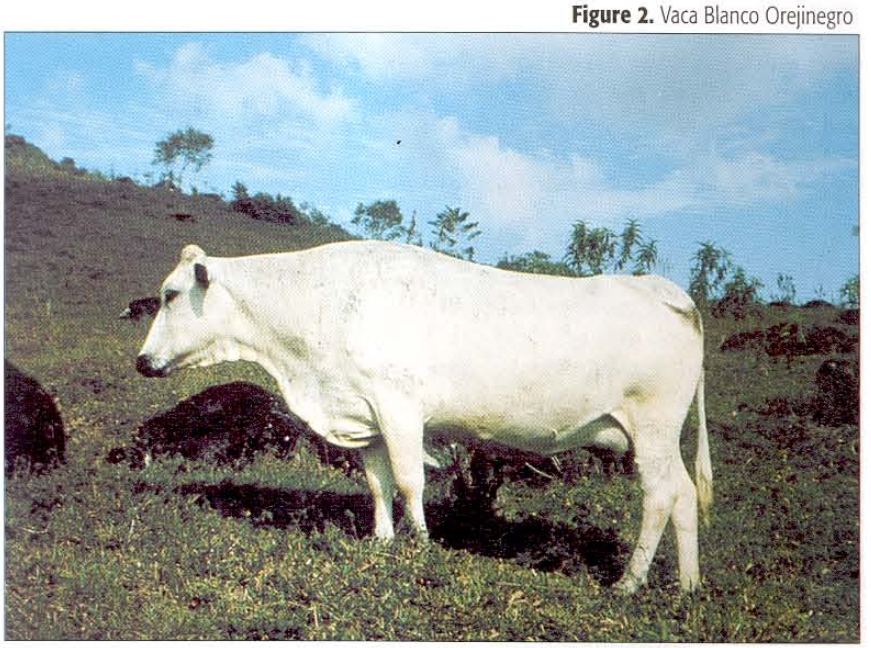

Figure 3. Toro Casanareño

Figure 4. Chino Santandereano

Figure 5a. Torete Costeño con cuernos

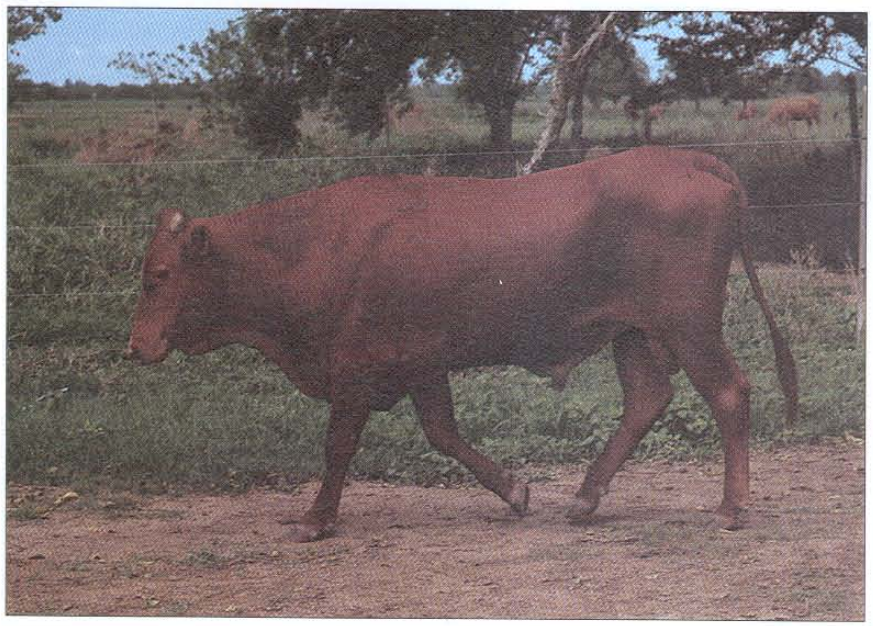

Figure 6. Vaca Costeña con cuernos

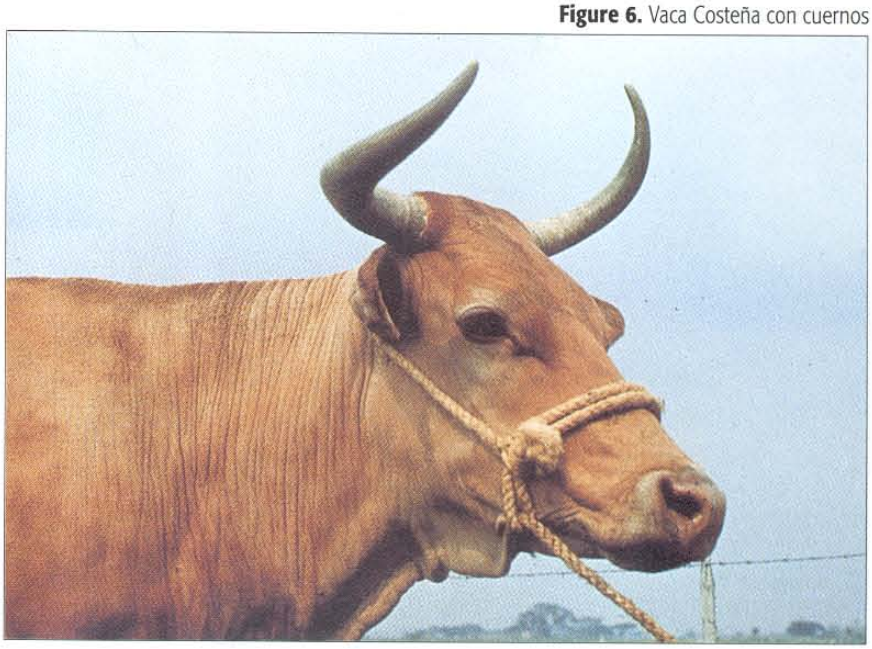

Figure 7. Novilla Hartón del Valle

Figure 8. Toro Romosinuano

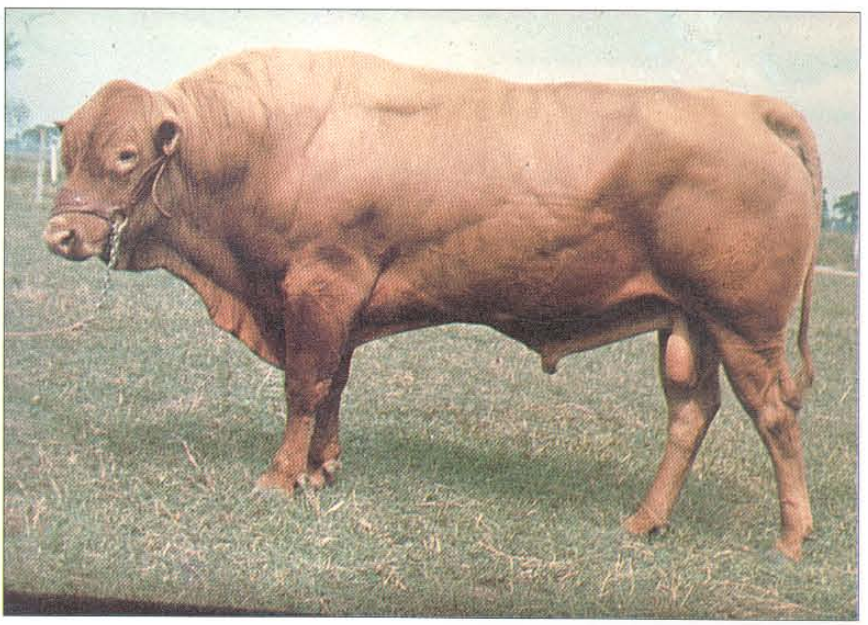

Figure 9. Toro Sanmartinero

traits (e.g., hot carcass weight, fat over the longissimus dorsi, marbling score, shear force) might be the first batch of traits that can be evaluated. As the evaluation system matures over time, either more traits, or a different set of traits that better reflect the prevailing economic interests at that time, might be considered.

The types of genetic effects considered per trait will vary according to the type and the level of inheritance of a trait and the type of population animals belong to. For traits that are lowly heritable, both additive and nonadditive effects may need to be considered, particularly interbreed nonadditive genetic effects in multibreed populations. Further, if the dam affects the outcome of a trait (e.g., WW), then maternal effects will be considered in addition to direct genetic effects. Direct genetic effects are the result of an animal's own genes. Maternal effects are caused by the influence of the dam over the calf (e.g., the effect of maternal milk over WW). From a statistical standpoint, each effect within a trait will be treated as a different trait. Thus, a trait influenced by direct and maternal effects (e.g., WW) in a multibreed population may need to consider up to four embedded traits (direct and maternal additive and nonadditive genetic effects).

Assuming that a suitable software package is available, the genetic effects for each trait considered in the model will depend on: 1) the previous knowledge about the inheritance of a trait; for example, a model for a highly heritable post-weaning trait (e.g., PG) or a carcass trait (e.g., hot carcass weight) will probably consider direct genetic effects, but no maternal effects, and 2) the type of population where animals are being evaluated; for instance, a model for traits of medium to low heritability (e.g., BW, WW) in a multibreed population may include nonadditive genetic effects, in addition to additive genetic effects, if the structure of the data set permits such predictions.

Genetic Evaluation Methods and Computational Procedures

Genetic Evaluation Methods

The genetic evaluation tests currently used to evaluate animals in the field can be broadly grouped in two categories: linear and nonlinear methods. Linear methods are better suited to predict genetic values of animals for traits that show continuous responses (e.g., BW, WW, PG) than for traits that have categorical responses (e.g., calving ease, calf survival). If traits have categorical responses, then nonlinear methods are preferred to linear ones. The method of choice to predict genetic values of animals for continuous traits is called Best Linear Unbiased Prediction (BLUP; Henderson *et al.*, 1959; Henderson, 1973; Pollak and Quaas, 1980; Elzo, 1983; Elzo and Famula, 1985; Arnold *et al.*, 1992), whereas Bayesian methods are preferred for nonlinear traits (Foulley *et al.*, 1983; Gianola and Foulley, 1983; Gianola and Fernando, 1986; Hoeschele *et al.*, 1995). Detailed description of nonlinear models and procedures is beyond the scope of this paper. Thus, only linear models will be discussed.

Best Linear Unbiased Predictor (BLUP) and Mixed Model Equations

The BLUP predictor minimizes the error variance of a prediction within the group of linear unbiased predictors. BLUP computing using the original formula is not an easy task. Thus, C. R. Henderson *et al.* (1959) came up with a set of mixed model equations (MME) that yielded not only BLUP of random effects, but also generalized least squares (GLS) of fixed effects in the model. The MME became, and still are, the procedure of choice for computing BLUP.

Genetic Models

From a practical standpoint, the purpose of a genetic evaluation mo-del is to account genetic and environmental effects in a population as thoroughly as possible such that the predictions of the genetic animal values are the best that can be computed given the available data. This purpose applies to any kind of mo-del and any type of population.

The genetic evaluation models can be classified according to the type of animal included in the model as follows:

1) Animal Model (AM): a model that contains all animals in a vector of random genetic effects. The AM predicts expected breeding values (EBV) and it accounts for all relationships among animals in the data set. These EBV must subsequently be divided by TWO to compute the expected progeny differences (EPD) that currently are published in sire summaries.

2) Reduced Animal Model (RAM): a model that explains nonparents in terms of their parental genetic effects and their random Mendelian factors, thus leaving only parents in vector of random genetic effects. The RAM also predicts EBV, and it accounts for all relationships. The RAM yields the same predictions as the AM. The purpose of the RAM is to save the storage space when computing EBV in large data sets.

3) Sire-Dam Model (SDM): a model that contains only parents in the vector of random genetic effects. The SDM predicts EPD, and it accounts for relationships among parents only.

4) Sire-Maternal Grandsire Model (SMM): a model that contains only sires and maternal grandsires in the vector of random genetic effects. The SMM predicts EPD, and it accounts for relationships among sires and maternal grandsires only.

5) Sire Model (SM): a prediction model that contains only sires in the vector of random genetic effects. The SM predicts EPD, and it accounts for relationships among sires only.

The models SDM, SMM, and SM are *approximations* to the AM. Their predictions are less accurate than those of the AM (or RAM) because require less information than the AM. Although it may be desirable to any use to perform an AM in order to predict genetic values, computing constraints may force the utilization of an approximation model.

All these genetic models can be either unibreed or multibreed depending on the type of population animals they are going to be applied. By definition, unibreed models include additive and nonadditive genetic effects from a single breed only. Multibreed models are an extension of unibreed models in that they consider both intrabreed and interbreed additive and nonadditive genetic effects.

Genetic Evaluation Model and Assumptions

A genetic evaluation model could be represented, in general form, as follows:

RECORD
=
CONTEMPORARY GROUP (FIXED)
+
OTHER FIXED ENVIRONMENTAL EFFECTS (e.g., age of dam)
+
GROUP(FIXED) GENETIC EFFECTS
+
RANDOM GENETIC EFFECTS
+
PERMANENT ENVIRONMENTAL EFFECTS
+
RESIDUAL

The assumptions of this model are as follows:

1) The expected value of a record is equal to the sum of the fixed effects in the model: contemporary group (assumed here to include an overall mean) + other fixed environmental effects + group genetic effects.

2) The variance of a record is equal to the variance of the random genetic effects + the variance of the residual. The random genetic effects, the permanent environmental effects and the residual are assumed to be uncorrelated.

This general model applies to all unibreed and multibreed models described above (AM, RAM, SDM, SMM, SM).

Group genetic effects. Group genetic effects are usually constructed when animals being evaluated have incomplete pedigree information. In a multibreed model group genetic effects could be formed for both additive and nonadditive genetic effects. The preferred grouping strategy is the accumulated grouping which incorporates information of ancestor groups by exploiting the structure of the additive relationship matrix (Thompson, 1979; Quaas, 1988; Westell *et al.*, 1988). In models with genetic groups, random genetic effects are computed as deviations from genetic group effects. Group genetic effects are more likely to be needed in field data sets where parents come from many different sources where the level of accuracy of record keeping varies widely. If only a few herds with well-kept pedigree records supply most what bulls needed for straightbreeding and crossbreeding, then genetic groups may not be needed because most (if not all) of the important relationships among evaluated animals will be accounted for. This is the situation with some of the Colombian Criollo breeds (e.g., R, CCC). It is also the case for some of the available multibreed data sets. For example, the SMM model used for the genetic evaluation of the multibreed RZ data set (1980-1992) of Turipaná did not require the use of genetic groups (Elzo *et al.*, 1996) because all relevant relationships among sires and maternal grandsires were accounted for.

Random genetic effects. Depending on the trait, additive direct and maternal additive and nonadditive genetic effects can be considered in the model. By the usual assumptions of the infinitesimal theory of quantitative inheritance, additive and nonadditive genetic effects are independent (Kempthorne, 1955). Direct and maternal effects, on the other hand, will usually be correlated. Thus, the structure of the covariance matrix for each animal is block diagonal with two blocks: one for additive direct and maternal genetic covariances among all the traits being evalua-ted, and another block for the nonadditive direct and maternal covariances among the traits in the analysis. For example, consider the data set of the RZ multibreed population of Turipaná, and assume that, i) a multiple-trait AM for two traits: BW and WW will be used in the evaluation, and ii) only additive and nonadditive (interbreed [RZ] intralocus interactions) are important. Here, 1) both the additive and the nonadditive genetic covariance blocks of each animal in the model will be 2 ´ 2, and an animal's complete block-diagonal covariance matrix will be 4 ´ 4, and 2) the size of covariance matrix of all random effects will be 4n x 4n, where n is the number of animals in the evaluation.

Permanent and residual environmental effects. Permanent environmental effects are considered when there are several records per animal, where all these records can be assumed to be repeated measures of the *same* trait. There could be direct and maternal permanent. Permanent environmental effects are usually assumed to be uncorrelated to genetic effects; they are also uncorrelated across animals. Thus, the covariance matrix of permanent environmental effects is block diagonal where each block contains covariances among permanent environmental effects for each animal.

Residual effects contain genetic and (or) environmental random effects not explicitly accounted for in the model. Residual effects are assumed to be uncorrelated among animals. Thus, the covariance matrix of residual effects is block diagonal, with one block per animal.

Genetic Base

If genetic predictions are obtained for national genetic evaluation summaries, then a factor that needs to be considered is a genetic base. The choice of an appropriate genetic base will be particularly important for genetic evaluations in multibreed populations because of the potential economic implications of comparing animals of different breeds and crossbred groups.

A genetic base is formed by a group of animals whose mean predicted genetic value is the mean from which the predicted genetic values of all animals in the population are deviated. A genetic base can be floating or fixed. A floating genetic base is determined automatically by the genetic evaluation procedure; thus, its composition changes each time animals or data are added to or deleted from a genetic evaluation. A fixed genetic base, on the other hand, is formed by a specific group of animals, and it remains unchanged for a specific period of time (e.g., 10 years).

The criteria used to define the group of animals for a fixed genetic base will depend on the type of population. In unibreed populations, a time factor (e.g., year of entry into the stud) may be sufficient. In multibreed populations, breed composition and time may be needed. For example, in a RZ multibreed population composed of straightbred R and Z, and crossbred groups of these two breeds, a fixed genetic base could be defined by all R animals that entered the stud the first year data were recorded.

Computational Procedures and Computer Programs

Currently all computational algorithms use the MME to obtain the BLUP of the random effects in the model either by direct or by iterative procedures. Direct procedures usually involve some kind of factorization (Golub and Van Loan, 1983; Quaas, 1984) of the left hand side of the MME (LHS), which can be used to compute solutions and to obtain the elements of the inverse of LHS needed to compute standard errors from predicted genetic values. Iterative procedures usually use one of the following procedures alone or in combination: Gauss-Seidel, Block Gauss-Seidel, Jacobi, or Successive Over-Relaxation (Golub and Van Loan, 1983; Van Vleck and Dwyer, 1985a,b). Iterations may or may not require constructing the LHS. The preferred algorithm to solve very large sets of MME is one that iterates on the data provided (Schaeffer and Kennedy, 1986), and it does not require the construction of the LHS.

The best two free-of-charge unibreed computer packages, that can be run on personal computers, are: i) MTDFREML (Boldman *et al.*, 1993), and ii) DMU (Jensen and Madsen, 1993). Both programs can estimate covariance components in addition to predicting genetic values under a great variety of unibreed genetic models. There is a research computer system for multibreed populations resulting from two base breeds (LREM3; Elzo, 1996b) that predict genetic values and estimate covariance components for several traits and effects using various multibreed genetic models.

The basic steps involved in the computation of genetic predictions using one of the above packages are as follows:

1) Preparation of input animal files. Users need to prepare a data file and a pedigree file. The data file contains the identification of animals with records, the identification number of each environmental fixed effect in the model, the identification number of the ancestors of the animal with records, and the animal's records for all traits. The pedigree file contains the number of each animal and its ancestors. In addition to this information, the LREM3 multibreed program requires the breed composition of each animal and its ancestors in the data file and in the pedigree file.

2) Preparation of input parameter files. These files specify the type of model to be run (e.g., AM), effects to be considered per trait, variances and covariances among all traits and effects in the model, and output files.

3) Execution of the programs needed for the desired objective. The set of programs to be executed may depend on whether the objective is genetic evaluation only, estimation of covariance components only, or estimation of covariance components *and* genetic evaluation. These programs can be executed separately, or sequentially through a script file. A script file is an executable file that contains commands that will be executed by operating system in the order they were specified.

If one of the objectives of the study were to estimate covariances, this would be specified in the parameter files; then, the specified covariances would be taken as initial values by the iterative procedures used to estimate these covariances. The MTDFREML and the DMU packages use a derivative-free algorithm to estimate covariance components within a breed (Smith and Graser, 1986). The LREM3 system uses an EM algorithm (Dempster *et al.*, 1977; Elzo, 1996b) to estimate multibreed covariance components.

Strategy to Develop a National Genetic Evaluation Program for the Colombian Criollo Breeds

The development of a national genetic evaluation program for the Criollo breeds in Colombia should be an *integral part* of the Livestock Modernization Plan. A *task force* formed by breeders and producers, government officials, and genetic researchers would probably need to be created to develop the necessary policies and to organize the research and development work. A *genetic evaluation center* should be created in order to provide stability and continuity to the genetic evaluation effort.

General Strategy

A strategy to develop a genetic evaluation program for the Colombian Criollo breeds could consider the following secuential phases:

A. Research Phase:

1) Obtain a current inventory of the animals and historical data sets available.

2) Determine the level of connectedness of various data sets within and across Criollo breeds and crossbred groups among Criollo breeds and between Criollo breeds and other breeds (e.g., RZ crossbred groups in Turipaná).

3) Agree on a set of traits of economic importance whose measurements are feasible to be collected in the field.

4) Agree on a uniform set of data recording formats for all traits of interest and for all breeds.

5) Agree on a uniform set of adjustment formulas for traits that require such additional adjustments prior to their use in genetic evaluation.

6) Define a tentative mating plan for each breed (intrabreed only, or intrabreed and interbreed), and the subpopulations (unibreed and multibreed) created by these mating schemes.

7) Agree on genetic models and procedures to be used in the genetic evaluation.

8) Agree on a set of reference bulls to ensure that the defined unibreed and multibreed subpopulations are connected.

9) Agree on a *genetic base* from which all genetic evaluations will be deviated.

10) Define the computer resources (hardware and software) for: i) the editing and preparation the data files; ii) the computation of the genetic predictions; and, iii) the publication of the genetic evaluation summaries.

B. Developmental Phase:

1) Organize a reliable *in-farm* data recording scheme.

2) Assign responsibilities for the various tasks to be accomplished during the genetic evaluation process to the various organizations and individuals involved. Examples of tasks are: i) gathering and preliminary editing of the data; ii) final editing of the data, construction of the evaluation data and pedigree files; iii) conducting the genetic evaluation; and, iv) publishing the summaries of predicted genetic values.

3) Conduct research to improve the genetic evaluation procedures and the efficiency of the genetic evaluation program.

Specific Considerations

The main areas involved in the general strategy are discussed in more detail below.

Inventory of animals and historical data sets. This inventory should cover *all* types of data sets where Criollo cattle were used intrabreed, crossbred, experimental, field data sets kept by breeders and commercial producers, and pedigree information. This inventory: i) will be used to create the inventory data and pedigree files used to determine connectedness within and across breed groups (purebred *and* crossbred), and to identify traits in common across the component individual data sets, ii) will create a knowledge base upon which future decisions will be based upon (e.g., common recording formats, adjustment procedures, mating plans), and iii) will be the first step in the creation of a national data base for Criollo breeds and their crossbred groups.

Traits. A national genetic evaluation should include all traits (reproductive, growth, carcass) of economic importance that were recorded in the field data sets. The initial inventory files will probably have only a limited set of traits that were measured in all contemporary groups. The traits most likely to be available will be growth traits (e.g., BW, WW). There could also be some reproductive traits that could be either directly available or that could be recovered from the existing files (e.g., survival from birth to weaning, age at first calving, calving interval). The least recorded traits are likely to be carcass traits (e.g., hot carcass weight, area of the muscle *longissimus dorsi*, marbling score, tendernes or slices force); and their only source might be experimental data sets.

Thus, the first set of traits to be considered for a national genetic evaluation of Criollo breeds is a set of growth traits. Reproduction traits would be considered next. The evaluation of sires (and dams) for carcass traits will probably need to wait until a national carcass recording scheme is in place.

Mating Plans and Subpopulations. Mating decisions are made by breeders and commercial producers. The role of the participating animal breeders will largely be the identification of unibreed and multibreed subpopulations created by these mating decisions, and the identification of bulls to help connect contempo-

rary groups for genetic evaluation purposes. The most important factor when identifying these unibreed and multibreed subpopulations is the identification of contemporary groups. The key factor that determines whether a contemporary group is unibreed or multibreed is the breed composition of the animals in them. Multibreed contemporary groups must have animals representing *several* breeds and (or) crossbred groups. It is *not* necessary to have *all* base breeds and crossbred groups in *every* multibreed contemporary group. However, it would be desirable that reference sires of several breed compositions were represented in multibreed contemporary groups. The connected unibreed and multibreed subpopulations found in the inventory data file will reflect past mating decisions based on a great variety of traits ranging from purely experimental goals to purely commercial goals. Mating plans will probably be narrower if only breeder and commercial producer goals are agreed upon. These mating plans will be the ones that define the first set of subpopulations whose data sets will be used for national genetic evaluations.

Genetic Models and Procedures. Several models will need to be explored during the research phase. The type of models tested will primarily depend on the defined subpopulation types (i.e., unibreed or multibreed) and on the pattern of missing data (some effects may be completely confounded). Regardless of the shortcomings that inventory files may have, a clearer vision should emerge of the environmental and genetic effects to be included in models for the defined unibreed and multibreed subpopulations, and the additional data and links across contemporary groups needed. The model chosen for each subpopulation will surely stablish an significant compromise between accuracy of prediction and practicality. However, this should not be a major concern. These initial models will evolve into more complete (and complicated) ones as the data sets expand in future years.

Genetic Bases. There will be a different genetic base (floating or fixed) for each separate unibreed and multibreed subpopulation. A choice of a genetic base should be made during the research phase. Perhaps all subpopulations should choose the *same* type of genetic base (floating or fixed). The advantage of floating genetic base is that the computer program automatically defines the base. However, the chosen genetic base may not be the most suitable for the publication of the genetic predictions, particularly for the multibreed subpopulations. Additional computations will be necessary to deviate the predicted genetic values from the *mean genetic value of specific group of animals*, effectively turning this floating genetic base into a fixed genetic base. A major advantage of a fixed genetic base is that it provides stability for comparisons over a period of time. In addition, genetic predictions of animals that were born, or that became parents, in subsequent years will give breeders and producers information on genetic trends in their herds.

Computer Hardware. A dedicated set of computers is essential for the implementation of national genetic evaluation programs. Various types of computers and printers, and a redundant disk storage system will be needed. Microcomputers could be used for data entry and editing, whereas powerful microcomputers or UNIX workstations will be required for the computation of genetic predictions. The speed and storage capabilities of the microcomputers have increased dramatically in recent years. Perhaps powerful microcomputers (e.g., a 200 megahertz Pentium Pro) with appropriate amounts of memory (e.g., 128 megabytes of RAM), and disk space (e.g., 10 gigabytes) may be enough for the computation of genetic predictions during the research phase. On the other hand, if these machines were also used for the evaluation of the larger Zebu population in Colombia, then perhaps more powerful UNIX workstations (e.g., IBM RS6000) with more memory and disk space would be needed. Server-type microcomputers and UNIX workstations with a redundant disk storage system will minimize data loss due to an hardware failure and should be considered as an option. Impact printers for production runs and laser printers for high-quality printouts would be needed. In general, current *and* future computing needs and hardware update schedules should play a major role in hardware purchases.

Computer Software. Various types of software will be needed for specific stages of the production of genetic prediction summaries.

1) Data entry and initial editing of data files. Perhaps standard spreadsheet and data base management programs could be used here. To eliminate those data file formatting problems, it would be most advisable to choose a single software system that can create standard ASCII files for the next stage of the genetic evaluation process.

2) Final editing of data files, and construction of input data and pedigree files for the genetic prediction programs. These programs are very specific to the populations being evaluated. Thus, specialized software will need to be written (e.g., in FORTRAN) and the necessary compilers will need to be purchased. These computer programs will be the ones that define contemporary groups and all other environmental fixed effect subclasses, adjust trait measurements, and assign sequential numbers to animals in the evaluation.

3) Computation of genetic predictions. Computer programs for this stage can be general or specialized. General programs are more flexible but less efficient than specialized programs. The unibreed computer packages MTDFREML and DMU are general computer programs. The DMU program is more flexible and complete, but more difficult to learn than the MTDFREML. The multibreed computer system LREM3 can currently handle a variety of genetic models for *two-breed* multibreed populations only. This type of research system is continuously evolving. An earlier version of this system (RCLREM) was used to compute covariance components and genetic predictions of sires used between 1980 and 1992 in the RZ multibreed herd at Turipaná (Elzo *et al.*, 1996b). The structure of the multibreed subpopulations were defined in the research phase will determine if changes and additions to the LREM3 system will need to be made.

4) Publication of genetic prediction summaries. Distribution of the genetic prediction summaries can be done electronically or in print. Electronic distribution of genetic evaluation files could be done by one (or more) of the following means: diskettes, downloading files from a dedicated computer in the genetic evaluation center, and(or) accessing these files directly in a national genetic evaluation homepage. These last two options would require an excellent telephone or network communications. The software required would be web communication and editing programs. Printed distribution of genetic evaluations will probably be more expensive than electronic distribution. A desktop publishing software package and a laser printer could be used to produce high quality genetic prediction summaries.

In-Farm Recording Scheme. An in-farm recording program is the *core* of a national genetic evaluation program. Thus, the cooperation of breeders and producers as well as geneticists in participating experi-

mental stations is of paramount importance. Furthermore, the quality and quantity of field information will determine the accuracy of the prediction of genetic values. The in-farm recording scheme should be *realistic* and *flexible*. It is important that participating breeders, commercial producers, and geneticists collaborate in the design of recording forms. The initial set of forms should be as simple as possible; additional items could be added later on.

Distribution of Responsibilities. The genetic evaluation strategy outlined here is a two-tier system: 1) collection of data: breeders, commercial producers, experimental stations; and, 2) genetic evaluation and publication of genetic prediction summaries: genetic evaluation center.

Other options are certainly possible. For example, the publication of intrabreed evaluations might be the responsibility of the individual Criollo breed associations.

Developmental Research. The procedures discussed here were only linear. As indicated previously, nonlinear methods are favored over linear ones for categorical traits. Thus, implementation of nonlinear systems should be considered in future stages of this program. Additional traits may want to be recorded in future years, requiring modifications to the in-farm recording program. Editing programs will need continuous updating because they are data-set dependent, and additional data will be added every year. The structure and connectedness of population may change over time, thus, their genetic evaluation systems should change accordingly. Changes like these will need to be made in a planned and organized manner. It will be the responsibility of *all* the participating groups (breeders, producers, and geneticists) to identify problems, to propose creative solutions, and to implement the necessary changes to the genetic evaluation system.

Implications

The existence of seven distinct Criollo breeds places Colombia in a very favorable position in terms of its capability for supplying the world with *Bos taurus* germplasm adapted to tropical and subtropical environments. To realize this potential, Colombia needs to increase the numbers of cattle in each one of these Criollo breeds. A way to encourage this growth is to increase the economic value of Criollo cattle. One way to increase the value of individual animals (especially the bulls) is to compute predictions of their additive genetic values, and their combining abilities in various crossbreeding schemes. The development of a national genetic evaluation program involving the seven Criollo breeds and their crossbreds will not only achieve this goal, but it will create the infrastructure for a national sire evaluations for all cattle breeds and their crossbreds in Colombia.

AKNOWLEDGEMENTS

The senior author thanks Corpoica for the opportunity to participate in the Third Iberoamerican Congress of Criollo and Native Breeds, which allowed him to participate in many invaluable discussions with the coauthors of this paper, and with G. Martínez, C. Tobón, O. Acosta, and F. Gómez. Additionally, the senior author would like to thank G. Martínez for making available a copy of the colombian 1986 Livestock Development Project.

BIBLIOGRAPHY

Afanador, G. 1996. Plan Estratégico de Modernización Tecnológica de la Ganadería Colombiana. Revista CORPOICA (1)1:52-54.

Arnold, J. W., Bertrand, J. K. and Benyshek, L. L. 1992. Animal model for genetic evaluation of multibreed data. J. Anim. Sci. 3322-3332.

Bejarano, A., Hernández, G. and Rico, G. 1986. Proyecto de desarrollo ganadero con base en el uso de las razas criollas y colombianas. Publ. Misc. No. 628-ISSN-534-5391. Bogotá, Colombia. 70 pp.

Boldman, K. G., Kriese, L. A., Van Vleck, L. D. and Kachman, S. D. 1993. A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances. USDA-ARS. Clay Center, Nebraska, U.S.A.

Dempster, A. P., Laird, N. M. and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 38:1-38.

Elzo, M. A. 1983. Multibreed sire evaluation within and across countries. Ph. D. Dissertation. Univ. California, Davis, U.S.A.

Elzo, M. A. 1990a. Recursive procedures to compute the inverse of the multiple trait additive genetic covariance matrix in inbred and noninbred multibreed populations. J. Anim. Sci. 68:1215-1228.

Elzo, M. A. 1990b. Covariances among sire by breed group of dam interaction effects in multibreed sire evaluation procedures. J. Anim. Sci. 68:4079-4099.

Elzo, M. A. 1996a. Unconstrained procedures for the estimation of positive definite covariance matrices using restricted maximum likelihood in multibreed populations. J. Anim. Sci. 74:317-328.

Elzo, M. A. 1996b. LREM3: A computer program to predict genetic values and to estimate covariance components in two-breed multibreed populations. University of Florida, Gainesville, U.S.A.

Elzo, M. A. and Famula, T. R. 1985. Multibreed sire evaluation procedures within a country. J. Anim. Sci. 60:942-952.

Elzo, M. A., Manrique, C., Ossa, G. and Acosta, O. 1996. Uso de modelos multiraciales en la evaluación genética de bovinos en sistemas de producción tropicales. 12 pp. *En*: Proc. Tercer Congreso Iberoamericano de Razas Criollas y Nativas. CORPOICA, Bogotá, Colombia.

Estrada, L. 1996. Aplicación de la biotecnología animal para la preservación, caracterización y utilización del germoplasma bovino Criollo de Colombia. 13 pp. *En*: Proc. Tercer Congreso Iberoamericano de Razas Criollas y Nativas. CORPOICA, Bogotá, Colombia.

Foulley, J. L., Gianola, D. and Thompson, R. 1983. Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight, and pelvic opening. Genet. Sel. Evol. 15:401-424.

Golub, G.H. and Van Loan, C.F. 1989. Matrix Computations 3. 2nd edition. John Hopkins University Press, Baltimore, MA, U.S.A.

Gianola, D. and Fernando, R. 1986. Bayesian methods in animal breeding theory. J. Anim. Sci. 63:217-244.

Gianola, D. and Foulley, J. L. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:201-224.

Henderson, C. R. 1973. Sire evaluation and genetic trends. pp 10-41. *In*: Proc. Anim. Breed. Genet. Symp. in Honor of Dr. Jay L. Lush. ASAS and ADSA, Champaign, Illinois, U.S.A.

Henderson, C. R. 1975 a. Rapid method for computing the inverse of a relationship matrix. J. Dairy Sci. 58:1727-1730.

Henderson, C. R. 1975 b. Inverse of a matrix of relationships due to sires and maternal grandsires. J. Dairy Sci. 58:1917-1921.

Henderson, C. R. 1976 a. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.

Henderson, C. R. 1976 b. Inverse of a matrix of relationships due to sires and maternal grandsires in an inbred population. J. Dairy Sci. 59:1585-1588.

Henderson, C. R., Kempthorne, O., Searle, S. R. and Von Krosigk, C. M. 1959. The estimation of environmental and genetic trends from records subject to culling. Biometrics 15:192-218.

Hoeschele, I., Tier, B. and Graser, H. U. 1995. Multiple-trait genetic evaluation for one policotomous trait and several

continuous traits with missing data and unequal models. J. Anim. Sci. 73:1609-1627.

Jensen, J. and Madsen, P. 1993. A user's guide to DMU. A package for analyzing multivariate mixed models. National Inst. Anim. Sci., Res. Ctr. Foulum. Tjele, Denmark.

Kempthorne, O. 1955. The theoretical values of correlations between relatives. Genetics 40:153-167.

Quaas, R. L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics 32:949-953.

Quaas, R.L. 1984. Linear Prediction. *In*: BLUP School Handbook. Use of Mixed Models for Prediction and for stimation of covariance components. Feb. 5-7, Animal Genetics and Breeding Unit, University of New England, New South Wales, Australia. 158 pp.

Quaas, R. L. 1988. Additive genetic model with group and relationships. J. Dairy Sci. 71:1338-1345.

Schaeffer, L. R. and Kennedy, B. W. 1986. Computing strategies for solving mixed model equations. J. Dairy Sci. 69:575-579.

Smith, S. P. and H. U. Graser. 1986. Estimating variance components in a class of mixed models by restricted maximum likelihood. J. Dairy Sci. 69:1156-1165.

Thompson, R. 1979. Sire evaluation. Biometrics 35:339-353.

Van Vleck, L. D. and L. V. Cundiff. 1996. Data from designed experiments for across-breed adjustments: A discussion. *In*: Development and Implementation of Statistical Analyses and Computer Strategies for National Cattle Evaluation. WRCC-100, Brainerd, Minnesota, U.S.A.

Van Vleck, L. D. and Dwyer, D. J. 1985a. Successive overrelaxation, block iteration, and method of conjugate gradients for solving equations for multiple trait evaluation of sires. J. Dairy Sci. 68:760-767.

Van Vleck, L. D. and Dwyer, D. J. 1985b. Comparison of iterative procedures for solving equations for sire evaluation. J. Dairy Sci. 68:1006-1014.

Westell, R. A., Quaas, R. L. and Van Vleck, L. D. 1988. Genetic groups in an animal model. J. Dairy Sci. 71:1310-1318.

Wilson, D. E., Berger, P. J., and Wilham, R. L. 1986. Estimates of beef growth trait variances and heritabilities determined from field records. J. Anim. Sci. 63:386-394.