ABSTRACT

Interaction of genotype with environment was studied with 10,780 Spanish production records (daughters of 210 Holstein-Friesian bulls) and two subsets of United States data (800,821 records with 1170 sires and 762,152 records with 1186 sires). Only 74 bulls had daughter records in both Spanish and United States data. Genetic and phenotypic (co)variances and correlations and heritability for milk and fat yields were estimated both within country and between countries with countries considered as separate traits (joint analysis). (Co)variance components were estimated with a REML procedure. Heritability estimates for milk and fat in the Spanish population (.16 and .14, within country; .12 and .09, joint) were smaller than for United States data (means of .33 and .31, within country; .26 and .24, joint). Genetic and phenotypic correlations between milk and fat within country were higher for Spain (.94 and .91) than for United States data (means of .66 and .81). Genetic correlation between countries averaged .81 for milk and .50 for fat. Rankings of bulls for milk yield are expected to be similar in Spain and the United States, although a scaling effect on predicted values is expected from different genetic variances in each country. Ranking of bulls for fat yield may be quite different.

INTRODUCTION

In the last 20 yr, dairy cattle management in Europe has been oriented toward increased milk production per animal. Importation of semen and live animals from highly specialized breeds of dairy cattle (mainly North American Holsteins) has been one element in this process. Spain has participated in this general trend. Calcedo (2) reported that 60% of all inseminations of Friesian cows in Spain used Holstein-Friesian semen.

A concern about introduction of foreign semen is the possible existence of an interaction of genotype with environment that could cause a change in ranking of bulls in different countries. Numerous studies have dealt with genotype by environment interactions in dairy cattle. Most (4, 6, 7, 11, 12, 25) showed no such interaction of practical importance for different definitions of environment within countries. The relatively large interac-
tion between sire and amount of grain fed in summer detected by Mao and Burnside (12) with Canadian data is an exception. Danell’s (4) summary of the literature reported significant interactions between feeding regimens and genotypes in experiments with twins. Correlations among predicted breeding values in different countries have been found to be above .70 in studies dealing with European and North American populations (3, 17, 22) and in some studies involving North and Central American data (13, 14). For North and Central American data, Romero (20) and Abubakar et al. (1) found correlations among breeding values of bulls in different countries to be low to intermediate (.26 to .50).

The purpose of this study was to estimate variance and covariance components for milk and fat yields in Spain and the US and to estimate the genetic correlation between performance of daughters of bulls in Spain and the US by considering milk production in the two countries as separate traits.

MATERIALS AND METHODS

Data

Data were provided by USDA and the Ministry of Agriculture of Spain. The original USDA data set of 4,409,545 records included all 305-d, mature equivalent, first lactation records of US Holstein cows sired by bulls that also had daughters in the Spanish data set and all their herdmates for calving years 1967 through 1983. All records included milk and fat yields, and all had sire identification. Duplicate records were eliminated. Seasons were defined as December through April and May through November. Sires were divided into five groups based on year of entry into AI service from 1951 to 1975 to account for selection over time.

Because the size of the original data set and the number of sires was too great for available computing capabilities, records were discarded 1) if the record was not in a herd-year-season (HYS) class in which at least 1 of the 74 bulls in common for the US and Spain was represented; 2) if the cow was sired by a bull that did not have more than 25 daughters; 3) if the cow was not the daughter of a registered AI sire; or 4) if the record was in a HYS class with only one sire represented. After edits, the US data set consisted of 2,001,435 records with 4569 sires represented. To reduce the data set further to a computationally manageable size, two samples were chosen so that they included records of all daughters of the 74 sires in common plus records of cows sired by about an additional 1000 randomly selected sires. After eliminating records in HYS classes with only one sire represented, the two samples contained 800,821 records with 1170 sires and 762,152 records with 1186 sires.

Data provided by the Ministry of Agriculture of Spain consisted of 10,780 records produced from 1975 through 1983 by first lactation cows sired by 210 Holstein-Friesian bulls. All records included 305-d milk and fat yields. Short records were extended to 305 d. Information included herd code, calving date, calving age, lactation length, region where the record was produced (North or South), and sire origin (European, North American, or unknown) and birth date. Yields, dates, calving ages, and lactation lengths were required to be within the following ranges: milk yield from 2500 to 10,000 kg, fat percentage from 2.5 to 5%, calving age from 20 to 40 mo, and minimum lactation length of 120 d. Mean milk and fat yields for Spanish data and both US samples are in Table 1. Calving ages in the final data set ranged from 21 to 38 mo and were divided into four groups. Two seasons

<table>
<thead>
<tr>
<th></th>
<th>Spain</th>
<th>United States</th>
</tr>
</thead>
</table>
| Milk yield, kg | 4982 | 7709 Sample 1  
| Fat yield, kg  | 171  | 278 Sample 2  
| Number of records | 10,780 | 800,821 Sample 1 
|                   |       | 762,152 Sample 2 |

TABLE 1. Mean milk and fat yields for Spanish data and two United States samples.
were defined (October through March and April through September). Sire birth year ranged from 1957 to 1975 and was divided into four groups.

Models

Linear models were used in all analyses:

\[ y = X\beta + Zs + e \]

where \( y \) is the vector of observations; \( \beta, s, \) and \( e \) are vectors of fixed, sire, and residual effects; and \( X \) and \( Z \) are known matrices that associate elements of \( \beta \) and \( s \) with \( y \). The vectors \( s \) and \( e \) were considered to be jointly normally distributed.

The model for US data was:

\[ Y_{ijkl} = h_i + f_j + S_{jk} + e_{ijkl} \]

where \( Y_{ijkl} \) is milk or fat yield of cow \( l \) in HYS \( i \) and sired by bull \( k \) in group \( j \); \( h_i \) is effect of HYS \( i \) (146,765 HYS for sample 1; 142,676 HYS for sample 2); \( f_j \) is effect of sire group \( j \); \( S_{jk} \) is effect of sire \( k \) within sire group \( j \) (1170 sires for sample 1; 1186 sires for sample 2); and \( e_{ijkl} \) is residual.

The model for Spanish data was:

\[ Y_{ijklmnopq} = h_{ij} + b_{jl} + e_{lm} + s_n + t_{jo} + u_{ko} + v_{lo} + w_{mo} + d_{op} + e_{ijklmnopq} \]

where \( Y_{ijklmnopq} \) is milk or fat yield of cow \( q \) calving at age \( 1 \) in region \( m \), herd \( i \), year \( j \), and season \( k \) and sired by bull \( n \) with origin \( o \) and birth year \( p \); \( h_{ij} \) is effect of herd-year \( ij \) (2068 herd-years); \( b_{jl} \) is interaction effect between calving year \( j \) and calving age \( 1 \); \( e_{lm} \) is interaction effect between calving age \( 1 \) and region \( m \); \( s_n \) is effect of sire \( n \); \( t_{jo} \) is interaction effect between calving year \( j \) and sire origin \( o \); \( u_{ko} \) is interaction effect between calving season \( k \) and sire origin \( o \); \( v_{lo} \) is interaction effect between calving age \( 1 \) and sire origin \( o \); \( w_{mo} \) is interaction effect between region \( m \) and sire origin \( o \); \( d_{op} \) is interaction effect between sire origin \( o \) and birth year \( p \); and \( e_{ijklmnopq} \) is residual. This more complicated model was used because results of a study by L. Gomez Raya et al. (1989, personal communication) with a similar sample of the Spanish Friesian population indicated that all these effects were significant.

Estimation of Variance Components

To estimate variances of \( s \) and \( e \), REML was used by following an expectation-maximization (EM) algorithm as described by Dempster et al. (5) for the exponential family of densities. In this study, the vector of \( s \) and \( e \) (\( x^t = [s^t; e^t] \)) was considered to be complete data, and the vector of “error contrasts” (\( K'y \) where \( K \) is a matrix that transforms \( y \) into a vector of deviations from herdmates’ yield) (16) was considered to be incomplete data.

Two sets of analyses were performed. The first was estimation of genetic (4 times sire) and phenotypic (sire plus residual) variances and covariances between milk and fat within each country. The second yielded estimates for genetic and phenotypic variances and covariances for milk and fat yields in each country with countries considered as separate traits.

Analysis Within Country for Milk and Fat

Milk and fat yields within each country were considered jointly for estimation of variances and covariances. The factors that affected milk and fat yields were assumed to be the same, and the following model was used:

\[
\begin{bmatrix}
    y_1 \\
    y_2
\end{bmatrix} =
\begin{bmatrix}
    X & 0 \\
    0 & X
\end{bmatrix}
\begin{bmatrix}
    \beta_1 \\
    \beta_2
\end{bmatrix} +
\begin{bmatrix}
    Z & 0 \\
    0 & Z
\end{bmatrix}
\begin{bmatrix}
    s_1 \\
    s_2
\end{bmatrix} +
\begin{bmatrix}
    e_1 \\
    e_2
\end{bmatrix}
\]

where \( y \) is the vector of observations for the yield trait (1 = milk; 2 = fat); \( X \) and \( Z \) are known incidence matrices for fixed and sire effects; \( \beta, s, \) and \( e \) are vectors of fixed, sire, and residual effects associated with the yield trait.

Distribution of

\[
\begin{bmatrix}
    s \\
    e
\end{bmatrix}
\]
was assumed to be multivariate normal with mean

\[
\begin{bmatrix}
0 \\
0
\end{bmatrix}
\]

and variance

\[
\begin{bmatrix}
G & 0 \\
0 & R
\end{bmatrix}
\]

where \( G = \text{Var}(s) = G_0 \otimes A \) and \( R = \text{Var}(e) R_0 \otimes I \)

with \( A = \) numerator relationship matrix among sires, \( I = \) identity matrix, \( \otimes \) indicates direct (Kronecker) product operation

\[
G_0 = \begin{bmatrix}
g_{11} & g_{12} \\
g_{21} & g_{22}
\end{bmatrix}
\]

and

\[
R_0 = \begin{bmatrix}
r_{11} & r_{12} \\
r_{21} & r_{22}
\end{bmatrix}
\]

A series of transformations were performed to simplify computations (8). These transformations involved a canonical transformation (23) to allow computation with single-trait systems, decomposition of the relationship matrix as \( A = LL^t \) (19) to eliminate \( A^{-1} \) in the quadratics in \( s \) and in the traces, and tridiagonalization of the coefficient matrix as described by Lawlor (10).

**Joint Analysis of Spanish and United States Yield Traits**

Spanish and US milk yields were analyzed jointly to estimate genetic correlation between milk yields in the two countries. A linear model also was used for this analysis. Milk yields in the two countries were considered as different traits and analyzed in a multiple-trait context so that estimates for genetic and residual variances within country and genetic covariance between genetic expression in the two countries could be obtained. The following model was used:

\[
y_1 = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}
\]

where \( y \) is the vector of observations for milk yield in the country (1 = US; 2 = Spain); \( X \) and \( Z \) are known incidence matrices relating the observation vectors to fixed and sire effects; and \( \beta, s, \) and \( e \) are vectors of fixed, sire, and residual effects associated with the country. The \( Z \) contain null columns corresponding to sires with no daughters in that country because the same sires were included in \( s \) for both countries even though most of them did not have daughters in both countries. The assumptions for \( s \) and \( e \) were the same as before except that

\[
R = \text{Var}(e) = \begin{bmatrix} r_{11}I_{N_1} & 0 \\ 0 & r_{22}I_{N_2} \end{bmatrix}
\]

An analogous joint analysis was performed for fat yield.

**RESULTS AND DISCUSSION**

**Analyses Within Country**

Estimates of genetic and phenotypic (co)variance components after 100 rounds of iteration are in Table 2. Estimates of both genetic and phenotypic variances were considerably smaller for Spain than for the US.

Heritability estimates for milk and fat in the Spanish population (.16 and .14) were smaller than estimates for US data sets (Table 2). However, Spanish estimates agree with results of L. Gomez Raya et al. (1989, personal communication), who applied Hender-
TABLE 2. Estimates of variance components, heritabilities, and genetic and phenotypic correlations for milk and fat yields from analyses within country for Spain and two United States samples.

<table>
<thead>
<tr>
<th></th>
<th>Spain</th>
<th>United States</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sample 1</td>
<td>Sample 2</td>
</tr>
<tr>
<td>Genetic (co)variances, kg²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>108,608</td>
<td>496,015</td>
<td>629,929</td>
</tr>
<tr>
<td>Fat yield</td>
<td>102</td>
<td>568</td>
<td>656</td>
</tr>
<tr>
<td>Milk-fat covariance</td>
<td>3147</td>
<td>10,564</td>
<td>14,058</td>
</tr>
<tr>
<td>Phenotypic (co)variances, kg²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>676,175</td>
<td>1,691,576</td>
<td>1,711,458</td>
</tr>
<tr>
<td>Fat yield</td>
<td>752</td>
<td>1993</td>
<td>2004</td>
</tr>
<tr>
<td>Milk-fat covariance</td>
<td>20,487</td>
<td>47,038</td>
<td>47,573</td>
</tr>
<tr>
<td>Heritability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>.16</td>
<td>.29</td>
<td>.37</td>
</tr>
<tr>
<td>Fat yield</td>
<td>.14</td>
<td>.28</td>
<td>.33</td>
</tr>
<tr>
<td>Correlation milk-fat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic</td>
<td>.94</td>
<td>.63</td>
<td>.69</td>
</tr>
<tr>
<td>Phenotypic</td>
<td>.91</td>
<td>.81</td>
<td>.81</td>
</tr>
</tbody>
</table>

son's method 3 to estimate variance components with a single-trait model for a similar sample of the Spanish population. Their heritability estimates were .12 for milk yield and .11 for fat yield. The slightly larger estimates in this study may be the result of including relationships among sires. Smaller genetic variances and heritabilities for Spain than for the US also agree with results found by several researchers (4, 10; L. Gomez Raya et al., 1989, personal communication), who observed an increase of genetic variance and, in some cases, heritability as yield increased. Mean yield in Spain was approximately two-thirds of that in the US (Table 1). Heritability estimates for US data sets (.29 and .28 for milk and fat yields in Sample 1 and .37 and .33 for milk and fat yields in Sample 2) were in the range of previous heritability estimates for the US (6, 21).

A possible explanation for the smaller heritability estimates for the Spanish population is that environmental conditions related to management limited expression of genetic potential of superior cows and, therefore, restricted differences in yields among animals due to genetic value. A higher disease incidence and less accurate pedigree recording in Spain than in the US also could have contributed to smaller estimates of genetic variance in Spain.

Correlations between milk and fat yields are in Table 2. Correlations between Spanish yield traits were higher than for US samples (.94 and .91 for genetic and phenotypic correlations in Spain and .66 and .81 for genetic and phenotypic correlations for mean of the two US samples).

Because computational costs for iteration were inexpensive, 100 rounds were performed. However, most parameters did not change significantly after round 30. As a check that the maximum found was global rather than local, different starting values were used, and all led to the same final results. These starting values included negative covariances, larger genetic than phenotypic (co)variances, and identity matrices for G₀ and R₀.

Joint Analyses

Estimates of genetic and phenotypic variance components are in Table 3 for the Spanish data analyzed with each of the two US samples. Estimates for genetic variances in this analysis were consistently smaller for both countries and for both milk and fat yields than were estimates from within-country analyses. For US data sets, the decrease in estimates of the genetic component might be expected if smaller genetic differences among bulls in the Spanish data
TABLE 3. Estimates of variance and covariance components, heritabilities, and genetic and phenotypic correlations for milk and fat yields from joint analysis of Spanish and United States data sets.

<table>
<thead>
<tr>
<th></th>
<th>Spain-US sample 1</th>
<th></th>
<th>Spain-US sample 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spain</td>
<td>US</td>
<td>Spain</td>
<td>US</td>
</tr>
<tr>
<td>Genetic variances, kg²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>79,177</td>
<td>431,575</td>
<td>88,182</td>
<td>466,231</td>
</tr>
<tr>
<td>Fat yield</td>
<td>66</td>
<td>449</td>
<td>68</td>
<td>477</td>
</tr>
<tr>
<td>Genetic covariances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Spain, US)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>145,526</td>
<td>165,455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat yield</td>
<td>84</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotypic variances, kg²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>674,221</td>
<td>1,694,106</td>
<td>676,255</td>
<td>1,698,973</td>
</tr>
<tr>
<td>Fat yield</td>
<td>748</td>
<td>1993</td>
<td>748</td>
<td>1998</td>
</tr>
<tr>
<td>Heritability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>.12</td>
<td>.25</td>
<td>.13</td>
<td>.27</td>
</tr>
<tr>
<td>Fat yield</td>
<td>.09</td>
<td>.23</td>
<td>.09</td>
<td>.24</td>
</tr>
<tr>
<td>Genetic correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Spain, US)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk yield</td>
<td>.79</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat yield</td>
<td>.49</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Heritabilities for milk yield at successive rounds of iteration for Spain-US sample 1 joint analysis (□ = Spain; △ = US sample 1) and Spain-US sample 2 joint analysis (○ = Spain; * = US sample 2).
ESTIMATION OF GENETIC PARAMETERS

Figure 2. Heritabilities for fat yield at successive rounds of iteration for Spain-US sample 1 joint analysis (□ = Spain; Δ = US sample 1) and Spain-US sample 2 joint analysis (◊ = Spain; * = US sample 2).

could reduce estimated differences among those bulls in US data because of the positive genetic correlation between countries. For the same reason, estimated differences of Spanish bulls might be expected to increase because of larger differences among estimated breeding values for the data set; however, this was not observed. The decrease in the sire variance component for the Spanish data set did not have a clear explanation. Simulation studies by Walter and Mao (24) and Pollak et al. (18) and Meyer and Thompson's (15) study with field data found that estimates of genetic variances tended to decrease if correlated traits with selection on one trait were analyzed as single traits rather than as a multiple trait. Magnitude of decrease was related to selection intensity and correlation between traits. In this study, milk and fat were highly correlated in the Spanish data, although selection intensity was not expected to be high. In contrast to Walter and Mao's simulation (24) for which a decrease in estimates of residual variance components was reported for single-trait as compared with multiple-trait analyses, estimates of residual components for the Spanish data did not decrease if analyzing milk and fat yields independently rather than jointly but instead increased slightly. Estimates of phenotypic variances were similar for both within-country analyses and for joint analyses. As a consequence of the decrease in estimates of genetic variance but not in estimates of phenotypic variance, heritability estimates decreased from within-country to joint analyses (Tables 2 and 3).

Genetic correlations between countries for milk and fat yields are in Table 3. Correlation for milk yield is relatively high (.79 for the Spain-US sample 1 analysis and .82 for the Spain-US sample 2 analysis), which indicates that bulls tended to rank similarly in both countries (although a scaling effect is ex-
expected because of the differences in genetic variances). Correlations between countries for fat yield were considerably smaller (.49 for the Spain-US sample 1 analysis and .50 for the Spain-US sample 2 analysis) than correlations for milk. The two countries' different genetic correlations between milk and fat yields (.94 for Spain vs. .63 and .69 for US samples) might explain partially the differences in genetic correlation between countries. In addition, fat yield is calculated from fat percentage of the milk and, therefore, may be subject to more measurement errors than is milk yield.

Figures 1 through 3 illustrate the iterative process to convergence that led to estimates of heritabilities for milk yield (Figure 1) and fat yield (Figure 2) and genetic correlations between countries (Figure 3) for the joint analyses. Fewer rounds of iteration (from 24 to 51) were performed for joint than for within-country analyses because of higher cost per round. Size of the Frobenius norm (computed as square root of sum of squares of elements of the matrix of differences in variance components between successive rounds) relative to estimates of variance components (which was smaller than .1% at stopping point) was the criterion used in deciding when convergence had been reached. Because of the extremely slow rate of convergence, especially for genetic variance for the Spanish data and genetic covariance between countries, the procedures described by Laird et al. (9) was used to speed convergence by predicting expected estimates of parameters of interest after an infinite number of rounds. Discontinuities in Figures 1 through 3 represent points of estimate projection.

**SUMMARY AND CONCLUSIONS**

Estimates of genetic components of variance as well as heritability estimates for both
milk and fat yields in Spain were relatively small. Possible causes of the low estimates related to the data set include accuracy of pedigree recording, sampling, and size of the data set. Additional factors not related to the data set that may have caused small estimates are management conditions that limit expression of genetic potential of superior cows. Further investigation of causes of the small estimates of genetic variance is recommended, including studies with larger data sets or those on quality of paternity identification or genetic variance in different management conditions. If these additional studies also show evidence of a small genetic variance in the Spanish dairy cattle population, a change in the parameters currently used (heritability of .25 for milk) in sire evaluations might be necessary.

Genetic and phenotypic correlations between milk and fat yields were close to 1.00 in Spain but somewhat smaller for the United States. Rankings for US bulls based on milk or fat yields would not be expected to be the same, whereas Spanish rankings for the two traits should be nearly identical. Rankings of bulls in Spain and the US are expected to be similar for milk yield but different for fat yield. Further investigation is necessary to determine causes of the smaller correlation between countries for fat yield, especially as related to import of bulls for their fat value.

Low estimates for genetic variance and relatively high genetic correlations between countries for milk yield suggest that importation of semen from the US Holstein population might enhance genetic progress for milk yield in the Spanish Friesian population. Although use of US semen likely would increase milk and fat yields, economic costs and benefits should be compared with those from improving management conditions.

ACKNOWLEDGMENTS

This research was supported in part by a grant from the US-Spain Joint Committee for Scientific and Technological Cooperation. Computing facilities were provided by the Northeast Dairy Records Processing Laboratory, Ithaca, NY, and the Cornell National Supercomputer Facility, Ithaca, NY, which is funded by the National Science Foundation, New York State, and IBM Corporation. Assistance by S.M. Hubbard in technical improvement of the manuscript is acknowledged gratefully.

REFERENCES


