Contributed paper

THE IMPACT OF GENOTYPE BY ENVIRONMENT INTERACTION ON BREEDING VALUES FOR 150-DAY WEIGHT IN KATAHDIN SHEEP IN MEXICO

L. De La Cruz[1,2], S.F. Walkom[1] and A.A. Swan[1]

[1] Animal Genetics Breeding Unit*, University of New England, Armidale, NSW, 2350 Australia
[1,2] National Institute of Forestry, Agriculture and Livestock Research in Mexico. Pachuca of Soto, Hidalgo, Mexico.

SUMMARY

With the objective of evaluating the impact of genotype by environment interaction (G by E) on breeding values for 150-day weight in Mexican Katahdin sheep, data from a total of 41,323 lambs, the progeny of 1,862 sires, were used to estimate genetic correlations between seven regional environments representing the majority of Mexico. Estimates of heritability within environments ranged from 0.24 ± 0.04 for Pacific Central to 0.42 ± 0.11 for North East. Genetic correlations across environments averaged 0.51 across all pairs, ranging from 0.07 ± 0.61 to 0.86 ± 0.27, indicating the presence of G by E interaction. A validation study predicted progeny performance within each environment with and without sire by flock effects from sire breeding values (EBVs) calculated from single trait BLUP analyses of data in the remaining environments. The regression of offspring performance on sire EBV were predictable across environments, but at lower levels than the expected value, in the absence of G by E, of 0.5. Fitting sire by flock improved the predictability with the regression coefficient increasing from 0.31 to 0.36.

INTRODUCTION

Katahdin sheep are a composite breed, developed in Maine, USA from crosses between hair and wool breeds (Wildeus 1997). This breed is low maintenance, highly prolific, does not require shearing, and is relatively resistant to internal parasites (Vanimisetti *et al.* 2007). The Katahdin breed plays an important role in the Mexican sheep industry as a maternal breed and makes a major contribution to sheep meat production in the country. Currently Katahdin sheep are dispersed across the diverse environments of Mexico, and have been dominating stud book registrations, with 87,807 animals in the database of the Mexican National Ovinocultores Union (UNO). However, recording of performance is limited, which has affected the development of any systematic breeding programs to improve the breed through selection. The objective of this study was to evaluate performance for 150-day weight, with a particular focus on the importance of G by E across the diverse Mexican environments where the breed is represented.

MATERIALS AND METHODS

Data structure. Records from 41,323 lambs, the progeny of 1,862 sires and 15,340 dams, were used to conduct genetic analyses of the performance for 150-day weight (W150) in Katahdin sheep across seven environments: North Central (NC), North East (NE), Pacific Central (PC), Central (C), Gulf Central (GC), Pacific South, (PS), and South East (SE). A summary of the number of animals, sires and dams represented at each environment is shown in Table 1.

---

* A joint venture of NSW Department of Primary Industries and the University of New England

Table 1. Data structure and descriptive statistics for 150-day weight (W150, kg) recorded in Katahdin sheep for all of Mexico and by environment

Component	Mexico	NC	NE	PC	C	GC	PS	SE
Number of records	41,323	4,839	3,118	10,977	12,206	2,140	5,153	2,890
Mean	40.9	42.1	38.1	44.1	41.9	40.8	31.9	42.1
Standard deviation	8.8	9.5	7.3	7.8	8.6	7.2	6.2	7.1
Variation coefficient (%)	21.5	22.6	19.2	17.7	20.5	17.7	19.5	16.9
Number of dams	15,340	1,751	1,264	3,871	4,545	787	1,977	1,272
Number of sires	1,862	350	176	556	734	151	161	215

NC: North Central; NE: North East; PC: Pacific Central; C:Central; GC: Gulf Central; PS: Pacific South; SE: South East.

Statistical analysis. Fixed effects fitted were sex (males, females), birth type and rearing type (single, twin, and triplet) and age of dam in years (factor: eight levels). Age at measurement was included as a linear regression. Variance components and heritabilities for each environment were estimated using univariate sire model. Random effects included sire genetic (G), maternal permanent environmental (PE) and contemporary group (CG), which was defined by flock of birth, year of birth and season. Genetic correlations across environments were estimated using bivariate analyses between each pair of environments. Bivariate models included the same fixed and random effects, with all analysis performed using ASReml software (Gilmour *et al.* 2018). The general bi-variate form of the variance structures including the residual term (R) was:

$$G=\begin{bmatrix} \sigma_{s_i}^2 & \sigma_{s_{ij}} \\ \sigma_{s_{ji}} & \sigma_{s_j}^2 \end{bmatrix};$$

$$PE=\begin{bmatrix} \sigma_{m_i}^2 & 0 \\ 0 & \sigma_{m_j}^2 \end{bmatrix}; \quad R=\begin{bmatrix} \sigma_{e_i}^2 & 0 \\ 0 & \sigma_{e_j}^2 \end{bmatrix}; \quad CG=\begin{bmatrix} \sigma_{cg_i}^2 & 0 \\ 0 & \sigma_{cg_j}^2 \end{bmatrix}$$

Genetic correlations between environments i and j were derived from the G matrix as: $r_{g_{ij}} = \sigma_{s_{ij}}/(\sigma_{s_i}\sigma_{s_j})$. The CG, PE, and R matrixes have a diagonal structure because contemporary groups, dams, and progeny can only be represented in a single environment. Heritabilities were estimated for environment i as $h_i^2 = 4\sigma_{s_i}^2/\sigma_{P_i}^2$, with $\sigma_{P_i}^2$ the phenotypic variance calculated as the sum of all components, excluding CG.

Validation analyses were used to study the impact of ignoring G by E correlations in the genetic evaluation analyses. For each of the seven environments defined as "targets", we estimated EBVs from "training" data combining the other six environments using a single trait animal model including and excluding the sire by flock interaction, not considering G by E effects (apart from any variation explained by sire by flock effects). Adjusted progeny performance in the target environment was then regressed on sire EBVs from the training analysis, i.e. for sires with progeny in both target and training data. The expected value of this regression in the absence of G by E is 0.5, and regressions of adjusted offspring performance on sire breeding values were calculated from linear models including the fixed effects of sex, birth type, rearing type, age of the dam and age at measurement (W150), along with contemporary groups treated as a random effect.

RESULTS AND DISCUSSION

The mean weights at W150 for NC, NE, PC, C, GC, PS and SE were 42.1, 38.1, 44.1, 41.8, 40.7, 31.9 and 42.1 kg, respectively (Table 1). The highest level of performance was observed for Pacific Central (PC) and the lowest for Pacific South (PS). The number of the records shows the distribution

of this breed across the country, with the largest numbers located in the central area (Central, Pacific Central) and the least in the Gulf Central.

Genetic parameters used to calculate EBVs for the validation study were based on analyses of the whole data set. Parameters estimated from the animal model without sire by flock were 26.4 ± 0.22 for the phenotypic variance, 37.2 ± 1.27 for the CG variance, with corresponding heritability and maternal permanent environment ratios of 0.14 ± 0.01 and 0.06 ± 0.01. For the model with sire by flock fitted phenotypic variance was 27.3 ± 0.25, CG was 34.7 ± 1.23, with heritability, maternal permanent environment and sire by flock ratios of 0.11 ± 0.01, 0.06 ± 0.01 and 0.07 ± 0.01 respectively.

Sire model variance components within environments are presented in Table 2. Estimates of heritability ranged between 0.25 ± 0.04 to 0.42 ± 0.11, averaging 0.32 across all regions. Lower heritability estimates (0.20 ± 0.02) have previously been reported in Katahdin lambs weighed at approximately 120 days of age (Ngere *et al.* 2017). Estimates of the ratio of maternal permanent environment effects ranged from 0.05 ± 0.01 to 0.10 ± 0.01 (averaging 0.08).

Table 2. Estimates of phenotypic variance (σ^2_p), heritability (h^2), maternal permanent environmental effects (m^2) for 150-day weight (W150, kg) within environments in Katahdin sheep

Component	NC	NE	PC	C	GC	PS	SE
σ^2_p	30.6 ± 0.77	23.7 ± 0.86	31.4 ± 0.52	28.3 ± 0.45	20.9 ± 0.78	12.7 ± 0.36	20.1 ± 0.71
h^2	0.31 ± 0.07	0.42 ± 0.11	0.24 ± 0.04	0.25 ± 0.04	0.28 ± 0.10	0.33 ± 0.09	0.42 ± 0.10
m^2	0.10 ± 0.01	0.08 ± 0.02	0.08 ± 0.01	0.09 ± 0.01	0.08 ± 0.02	0.05 ± 0.01	0.08 ± 0.02

The highest number of common sires and their progeny were between Pacific Central and Central environments, where more linkage is facilitated by greater sharing of genetic material between farms located relatively closely together (Table 3).

Genetic correlations across environments were positive, ranging between 0.07 ± 0.61 and 0.86 ± 0.27, with an average weighted by the inverse of standard errors of 0.51 (Table 3). The precision of estimates was highly variable, driven by differences in the number of common sires and progeny between regions. However, given the average correlation of 0.51 and considerable variation around the average, there is evidence that G by E interaction will affect selection across regions. We speculate that these interactions could have both biological and industry structural origins. The "biological" refers to true genetic adaptation to diversity in environmental conditions while the "industry structural" refers to the structure of the breed in Mexico into a large number of different breeding associations within regions that have different aims and ways of operating. Development of a single genetic evaluation system across environments would help to address this structural issue.

The validation analysis implemented an evaluation model across environments without assuming G by E interaction i.e., a single trait for all environments. The model was tested with and without sire by flock interactions, which would in theory partly correct G by E interactions. Regression of offspring performance in each environment on sire breeding values calculated in the other environments (Table 4) showed that breeding values were predictable across environments but at a level lower than the expectation of 0.5 in the absence of G by E. Encouragingly, fitting sire by flock improved the weighted average regression (predictability) from 0.31 to 0.36. The reductions in predictability relative to the expectation of 0.5 were slightly higher than proportional to the average genetic correlation between environments: a regression of 0.31 would suggest a genetic correlation of 0.62 (0.31/0.5) and a regression of 0.36 a correspondingly higher correlation of 0.72.

Several environments in Table 4 showed good levels of predictability from data recorded in other environments, including NC, NE, C, and GC, whereas predictability in other environments was

lower, including PC, SE, and PS. The two latter regions were smaller and characterised by poor connectivity with other regions. Improved linkage would be beneficial for these regions.

Table 3. Number of common sires (and progeny) in upper triangle, genetic correlations for 150-day weight (W150, kg) between environments in lower triangle

Environment	NC	NE	PC	C	GC	PS	SE
NC		30 (842)	96 (4,741)	85 (3,552)	28 (723)	21 (937)	32 (845)
NE	0.86 ± 0.27		39 (1,777)	39 (1,633)	19 (352)	11 (395)	17 (245)
PC	0.48 ± 0.19	0.21 ± 0.37		147 (7,511)	44 (1,796)	31 (1,447)	45 (2,213)
C	0.84 ± 0.17	0.19 ± 0.33	0.41 ± 0.18		50 (198)	29 (1,630)	46 (2,110)
GC	0.56 ± 0.37	0.64 ± 0.40	0.50 ± 0.31	0.31 ± 0.32		16 (463)	30 (754)
PS	0.74 ± 0.28	0.31 ± 0.58	0.77 ± 0.27	0.35 ± 0.33	0.07 ± 0.61		15 (250)
SE	0.72 ± 0.31	0.77 ± 0.33	0.12 ± 0.34	0.23 ± 0.34	0.38 ± 0.47	0.45 ± 0.52	

Table 4. Number of sires and progeny used to calculate regression coefficients of sire breeding values from training data on progeny performance for 150-day weight in validation data from each environment with and without sire by flock interaction fitted in the training data

Environment	Sires	Progeny	Regression coefficient	
			No sire flock	Sire flock
NC	124	1689	0.57 ± 0.09	0.64 ± 0.11
NE	58	691	0.32 ± 0.10	0.35 ± 0.11
PC	198	5189	0.17 ± 0.06	0.25 ± 0.07
C	201	4335	0.38 ± 0.06	0.42 ± 0.06
GC	69	599	0.53 ± 0.18	0.55 ± 0.20
PS	39	738	0.15 ± 0.09	0.23 ± 0.13
SE	61	920	0.06 ± 0.13	0.11 ± 0.14
Average			0.31	0.36

CONCLUSIONS

While this study showed evidence of G by E interaction across regions of Mexico, there was still evidence of predictability of breeding values across regions, albeit at a lower level than expected in the absence of G by E. Development of a national genetic evaluation system for the Katahdin breed in Mexico will stimulate greater linkage between associations, allowing breeders to benefit from across-flock selection even if there is a biological component of G by E in some circumstances.

ACKNOWLEDGMENTS

REFERENCES

Gilmour A.R., Gogel B.J, Cullis B.R, Thompson R. and Butler D. (2018) ASReml User Guide Release 4.0. VSN International Ltd, Hemel Hempstead, UK

Ngere L., Burke J., Notter D. and Morgan J. (2017) *J. Anim. Sc.* **95:** 3396.

Vanimisetti H.B., Notter D.R. and Kuehn L.A. (2007) *J. Anim. Sc.* **85:** 60.

Wildeus S. (1997) *J. Anim. Sc.* **75:** 630.