ANIMAL GENETICS AND GENOMICS

Random regression of Hereford percentage intramuscular fat on geographical coordinates

Jose S. Delgadillo Liberona,[†,1] John M. Langdon,[†,1] Andy D. Herring,[†]
Harvey D. Blackburn,[‡] Scott E. Speidel,[‖] Stacy Sanders,[$] and David G. Riley[†,2]

[†]Department of Animal Science, Texas A&M University, College Station, TX 77843, [‡]USDA-ARS, National Animal Germplasm Program, Fort Collins, CO 80521, [‖]Department of Animal Science, Colorado State University, Fort Collins, CO 80523, [$]American Hereford Association, Kansas City, MO 64153

[1]These authors contributed equally to this work as co-first authors.

[2]Corresponding author: david-riley@tamu.edu

Abstract

Accounting for genotype–environment interactions may improve genetic prediction and parameter estimation. The objective was to use random regression analyses to estimate variances and thereby heritability for intramuscular fat (**IMF**) across longitude and latitude coordinates within the continental United States. Records from the American Hereford Association (n = 169,440) were used. Analyses were first conducted using the continental United States in its entirety, and then as subdivided into two or four regions. Data were analyzed with an animal model, and linear and quadratic random regressions of additive genetic merit on longitude or latitude as covariate (separately). Subdivided data were analyzed with linear random regressions unique to regions. Regions were North and South separated at 40°N latitude, or West and East separated at 99°W longitude using longitude or latitude as covariate, respectively. Further subdivision to four regions included additional boundaries of 44.46° and 36.46°N latitude and 104.55° and 92.22°W longitude. The estimated heritability of IMF from the traditional model was 0.19 ± 0.004. Without regional subdivision of data, quadratic random regression had the best fit for the data based on likelihood ratio tests using longitude or latitude as covariate (P < 0.01). Estimates of heritability from quadratic random regression on latitude ranged from 0.12 in the South to a high of 0.27 at the extreme Northern latitude. Estimates of heritability from quadratic random regression on longitude ranged from 0.17 in the middle of the parameter space (corresponding to the central United States) to 0.37; higher estimates were noted at the extremes, that is, the far West and East longitudes. Random regression analyses of data divided into regions were conducted with a linear coefficient, as increasing to a quadratic polynomial was never accomplished. Results from random regression on latitude in the East region were similar to results from analyses without regions (h^2 ranged from 0.09 to 0.32); however, estimates of heritability in the West region had a lower range from South to North (0.14 to 0.27). Estimates of heritability from random regression on longitude with data divided into two regions were similar to those from analyses that did not include region. Estimates in the South region were somewhat lower and had a lower range (0.15 to 0.31) than those from the North region (0.19 to 0.47). When data were further subdivided, estimation of only a subset of covariances among random regression coefficients was possible, that is, within-region covariances of intercept and linear terms (latitude); those and covariances between all linear random regression coefficients were estimated when longitude was the covariate. Results from random regression analyses of data with four regions modeled produced very high estimates of heritability in low latitudes in the furthest West and high latitudes in the furthest East region, with approximate difference of 0.3 and 0.2 between estimates in the two West regions and the two East regions, respectively. Results from random regression on longitude indicated higher estimates of heritability in North region, especially at the furthest East longitudes of the most Northern region. There appeared to be substantial additive genetic variance differences, as well as estimates of heritability,

that correspond to different geographical environments as modeled by random regressions on within-region latitude or longitude coordinates.

Key words: genetic prediction, Hereford, intramuscular fat, random regression

Introduction

Genetic merit is likely contextual, and it may be beneficial to predict location-specific genetic merit for economically important traits. Burns et al. (1979) found genotype–environment interactions (G × E) in Hereford cattle for multiple traits including birth weight, preweaning gain, estimated 205 d weight, and body condition score. Fennewald et al. (2017) identified G × E for weight at birth and weaning in different regions within the United States for Red Angus. Genetic variation within breed subpopulations reared in different environments impacts heritability (Blackburn et al., 2017) and can lead to over- or underestimate breeding values if predictions are done without accounting for environment. Notter et al. (1992), Hayes et al. (2016), and MacNeil et al. (2017) have supported the use of G × E in beef cattle genetic evaluation, but G × E has not been employed in U.S. National Cattle Evaluation. Not accounting for G × E could lower the rate of genetic change for traits. The American Hereford Association has records to account for G × E in genetic merit predictions. Among relevant traits fit for this analysis strategy, intramuscular fat (**IMF**) impacts beef quality, and its improvement in postnatal life via nutrition is governed by the genetic potential of the breed (Pethick et al., 2004; Hocquette et al., 2010).

Random regression procedures make it possible to model G × E and attain a greater level of precision for parameter estimates and genetic merit predictions (Cardoso et al., 2012), as well as to estimate parameters across environment gradients instead of just for a given set of environments (Santana et al., 2015). Accounting for geographical location would help producers select sires more fit to the environment in their operations. The objective of this study was to estimate genetic parameters for IMF in American Hereford using random regressions across latitude or longitude coordinates within continental United States.

Materials and Methods

Records

Records were provided by the American Hereford Association. Using open source databases (http://federalgovernmentzipcodes.us/), longitude and latitude coordinates were obtained for each IMF record using the U.S. Postal Service zip code of the ranch listed as the breeder. Contemporary groups were defined by the American Hereford Association, combining information related to herd, sex, management group, and birth date of the animals. Records without an associated zip code or with no contemporary group assignment were removed. Records greater than the mean +4 SD or less than the mean −4 SD were considered outliers and removed. After editing, the final data included 169,440 IMF records. The pedigree included 227,902 ancestors.

Statistical Analyses

Analysis assessed the benefits of using linear or quadratic random regressions of IMF on latitude or longitude coordinates versus traditional animal models in genetic parameters estimation, accounting for United States as a unique geographic location. Linear random regressions were evaluated subdividing the United States into two and four regions, in order to identify the impact of an increase in geographical subdivision of the country over genetic parameters estimation. Regional subdivision was performed as an attempt to assess the joint influence of latitude and longitude on parameter estimation, as interactions of random regressions were not accomplished. Random regressions (either linear or quadratic) were modeled using Legendre polynomials. Distributions of records across the United States subdivided in four regions for latitude and longitude analyses are presented in Figures 1 and 2, respectively.

The animal model followed this form:

$$y = X\beta + Zu + Wc + e$$

where y is a vector of IMF records, β is a vector of estimated fixed effects for a linear regression on longitude or latitude coordinate, u is a vector of random additive genetic effects, c is a vector of random contemporary group effects, e is a vector of residuals, and X, Z, and W are incidence matrices relating observations in y to values in β, u, and c.

Expectation for the components in the random vector was equal to vectors of 0, with variance–covariance structure:

$$Var \begin{bmatrix} u \\ c \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & R \end{bmatrix}$$

in which $G = A\sigma_a^2$, where A is the numerator relationship matrix constructed with the pedigree information, and σ_a^2 is the additive genetic variance; $C = I\sigma_c^2$, where I is an identity matrix and σ_c^2 is the contemporary group variance; $R = I\sigma_e^2$, and σ_e^2 is the residual variance.

The linear and quadratic random regression models followed the general form:

$$y = X\beta + Qu + Wc + e$$

in which y, β, c, and e vectors are as described for the animal model, and u is a vector of random regression coefficients for additive genetic effects. X and W are incidence matrices as described for the animal model, and Q is the design matrix containing the longitude or latitude coordinates as covariates, and relates the IMF records in y to the additive genetic random regression coefficients in u. The number of columns in the Q matrix is equal to the order of the random regression (two or three for the linear and quadratic random regressions, respectively). Expectation of the random vectors is a vector of 0. The variance–covariance structure is:

$$Var \begin{bmatrix} u \\ c \\ e \end{bmatrix} = \begin{bmatrix} A \otimes G & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & R \end{bmatrix}$$
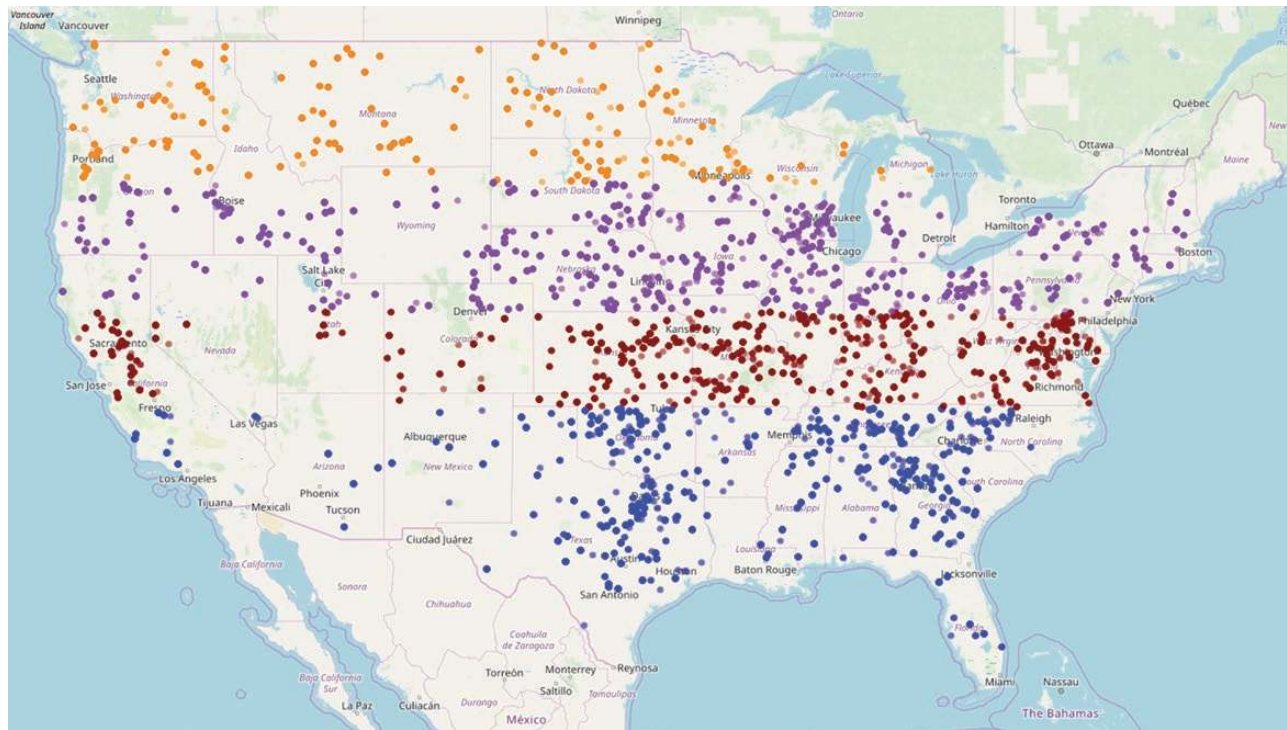
Figure 1. Distribution of intramuscular fat (**IMF**) records across continental United States and according to regional subdivision of the United States into North 1 (yellow, *n* = 42,403), North 2 (purple, *n* = 51,785), South 1 (red, *n* = 42,927), and South 2 (blue, *n* = 32,325) regions.
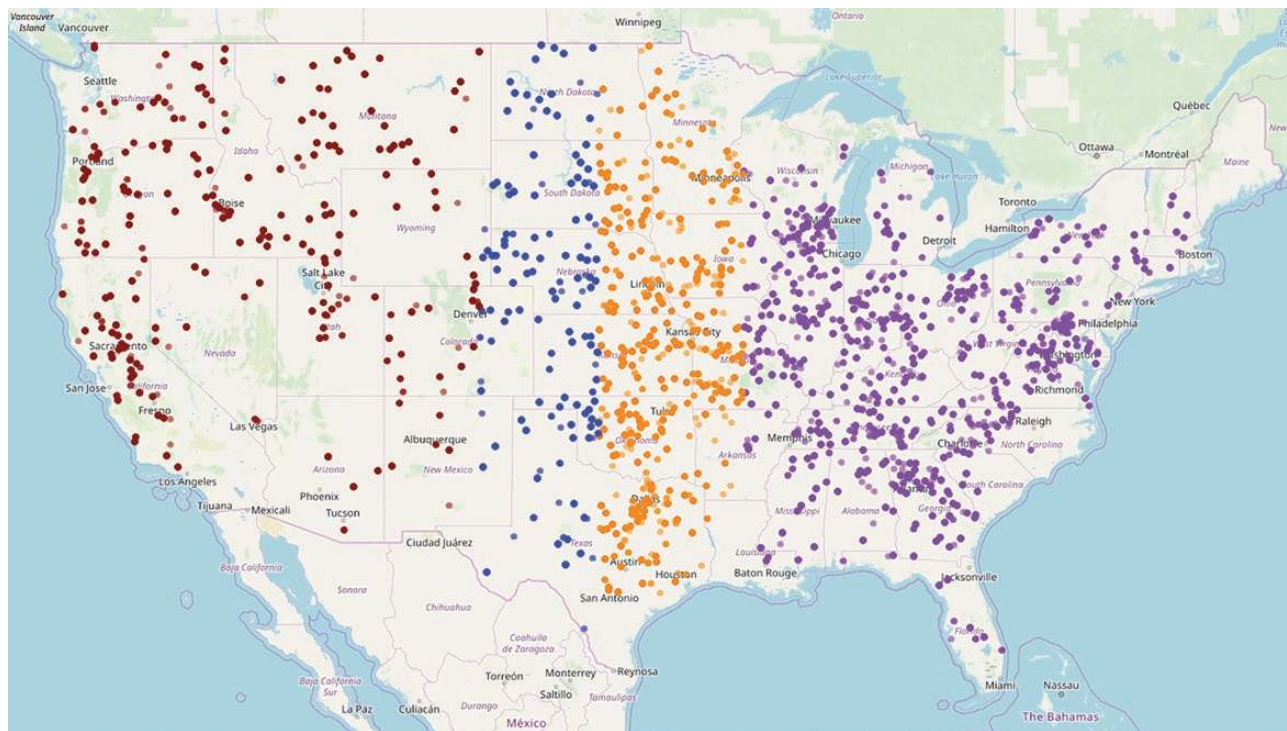

Figure 2. Distribution of IMF records across continental United States and according to regional subdivision of the United States into West 1 (red, *n* = 47,151), West 2 (blue, *n* = 37,949), East 1 (yellow, *n* = 46,427), and East 2 (purple, *n* = 37,913) regions.

where **A** is the numerator relationship matrix, **G** is the covariance matrix of additive genetic regression coefficients with an order equal to the polynomial modeled; **C** and **R** are matrices as described for the animal model. The **G** matrix used in the random regression models included the estimation of the variances and covariances of the intercept and the regression coefficients.

Models for analyses of regionally subdivided data included a linear random regression of IMF on longitude or latitude coordinates unique by region (i.e., one linear random regression nested within each region in the model). Analyses of data subdivided into two regions included unique random regressions within North and South (longitude) or West and East (latitude) regions. Analyses of data divided into four regions included random regressions within North 1, North 2, South 1, and South 2 (longitude; Figure 1) and West 1, West 2, East 1, and East 2 (latitude; Figure 2) regions. Regional boundaries were designated at specific longitudes or latitudes which were chosen to evaluate an overall even land territory across regions, and to keep similar numbers of records in each region. The North and South regions were delineated at 40°N latitude, with 94,188 and 75,252 records, respectively. The boundary between West and East regions was set at 99°W longitude, with 84,340 and 85,100 records, respectively. There were 11.1% and 10.7% of the total sires (9,815) that had progeny with records in both subregions for longitude and latitude, respectively. Boundaries for analyses of random regression on longitude within four regions were set at 44.46°N (between North 1 and North 2), 40°N (between North 2 and South 1), and 36.46°N (between South 1 and South 2). Corresponding numbers of records were 42,403, 51,785, 42,927, and 32,325 for the North 1, North 2, South 1, and South 2 regions, respectively. The regional boundaries for random regression on latitude were set at 104.55°W (between West 1 and West 2), 99°W (between West 2 and East 1), and 92.22°W longitude (between East 1 and East 2), with 47,151, 37,949, 46,427, and 37,913 records represented in the West 1, West 2, East 1, and East 2 regions, respectively. Of the 9,815 sires with progeny that had records, 14.9% and 14.1% had progeny in more than one region (of the four regions in analyses with random regression on longitude or latitude, respectively).

The analyses of unique random regressions per region were the same as the first set of random regression models with the same expectations for first and second moments. They differed in that Q is the incidence matrix containing the longitude or latitude coordinates covariate nested within regions and relates the IMF records in y to the additive genetic random regression coefficients in u for each region in the model. The number of columns in the Q matrix was associated with the order of the random regression amplified by the number of modeled regions (four for the two-region subdivision, and eight for four-region subdivision). The strategy for G matrix estimation was to first estimate variances of coefficients with all other covariances fixed at 0, and then incrementally add covariance components for estimation while holding previously estimated parameters constant; various analyses were attempted varying the set of parameters held constant. Analyses were repeated as necessary with estimation of previously fixed components. Non-estimable covariance components were fixed to zero and all other parameters were simultaneously estimated in final models. For analyses with two regions modeled, the **G** matrix included the estimation of variances for the intercept and the linear regression coefficient from both regions, as well as all covariance components between intercepts and linear regression coefficients from those two regions. In analyses with four regions, the **G** matrix included the variances for intercepts and linear regression coefficients from each region, as well as the covariance between the intercept and the linear regression coefficient within each region. Estimation of other covariance components in the **G** matrix were prioritized in this order: (i) covariances between each pair of linear regression coefficients across regions, (ii)

covariances between each pair of intercept coefficients across regions, and (iii) all other covariances.

Likelihood ratio tests were conducted for analyses of the data without regional subdivisions. The two-region models were not subsets of the four-region models, and larger models did not have additional terms that could be evaluated with likelihood ratio tests. Analyses were conducted using ASReml (Gilmour et al., 2009), and the Texas A&M University High Performance Research Computing Service.

From random regression analyses, gradients of heritability for IMF across longitude or latitude coordinates were estimated using variance component estimates (Schaeffer, 2016).

Results from linear random regression analyses without regional subdivision were selected to assess the potential effects of G × E on genetic merit predictions for IMF, as this modeling strategy is the most feasible approach to be included in early stages for genetic improvement strategies in Hereford cattle. Breeding values were estimated across latitude and longitude coordinates. Ranks of predicted breeding values were compared at different longitudes and latitudes by estimation of Spearman correlation coefficients.

Results

Fixed Effects

The fixed regression on latitude coordinates was an important model component ($P < 0.001$). The corresponding fixed regression on longitude coordinates was significant only in the quadratic random regression model analysis. Nevertheless, the fixed linear regression was kept in the animal and linear random regression models that utilized longitude as covariate to permit likelihood-ratio tests.

The fixed effect of region and the linear random regression of IMF on latitude nested within region were significant ($P < 0.001$) when data were subdivided into either two or four regions. When longitude was used as covariate in the model, the previous effects were only significant ($P < 0.001$) when data were subdivided into four regions.

These effects were kept in random regression analyses regardless of significance in order to facilitate model comparison.

Across-Region Random Regression

Results from the likelihood ratio test between each pair of the three continental U.S. models (animal model, linear random regression model, and quadratic random regression model), using either latitude or longitude coordinates as covariate, indicated that the quadratic random regression model had the better fit for these data ($P < 0.001$).

Heritability estimated using the animal model with either latitude or longitude coordinates as covariate was low (0.19 ± 0.004). Variances estimated for linear and quadratic random regression coefficients (latitude as a covariate) are shown in Tables 1 and 2, respectively. Heritability estimated with linear and quadratic random regression parameters resulted in similar ranges (0.08 to 0.27, linear; 0.12 to 0.27, quadratic). Plotted estimates of heritability from linear random regression on latitude appeared to increase from South to North (Figure 3). Estimates produced from the quadratic random regression were lower than those from the linear random regression in Southern latitudes, but higher in middle latitudes.

The variances estimated for the random regression coefficients using longitude as a covariate are shown in Tables 1 and 2. These yielded estimates of heritability ranging from 0.17

**Table 1.** Estimates of variance from linear random regression analyses[1]

	Contemporary group	Intercept	Linear	Residual
Latitude covariate				
Contemporary group	**0.38 ± 0.004**			
Intercept		**0.25 ± 0.006**	0.58 ± 0.056	
Linear		0.04 ± 0.002	**0.02 ± 0.003**	
Residual				**0.23 ± 0.002**
Longitude covariate				
Contemporary group	**0.37 ± 0.004**			
Intercept		**0.25 ± 0.006**	0.02 ± 0.018	
Linear		0.00 ± 0.003	**0.09 ± 0.005**	
Residual				**0.24 ± 0.002**

[1]Variances are in bold type. Covariances are below that diagonal and correlation coefficients are above. Covariances of contemporary group and residual with other terms were assumed to be 0.

Table 2. Estimates of variance from quadratic random regression analyses[1]

	CG	Intercept	Linear	Quadratic	Residual
Latitude covariate					
Contemporary group	**0.38 ± 0.004**				
Intercept		**0.20 ± 0.008**	0.57 ± 0.045	-0.47 ± 0.097	
Linear		0.05 ± 0.002	**0.04 ± 0.006**	-0.10 ± 0.11	
Quadratic		−0.02 ± 0.003	0.00 ± 0.003	**0.01 ± 0.004**	
Residual					**0.23 ± 0.002**
Longitude covariate					
Contemporary group	**0.37 ± 0.004**				
Intercept		**0.26 ± 0.010**	0.05 ± 0.024	0.19 ± 0.053	
Linear		0.01 ± 0.003	**0.05 ± 0.009**	-0.46 ± 0.089	
Quadratic		0.02 ± 0.005	−0.02 ± 0.003	**0.03 ± 0.004**	
Residual					**0.23 ± 0.002**

[1]Variances are in bold type. Covariances are below that diagonal and correlation coefficients are above. Covariances of contemporary group and residual with other terms were assumed to be 0.
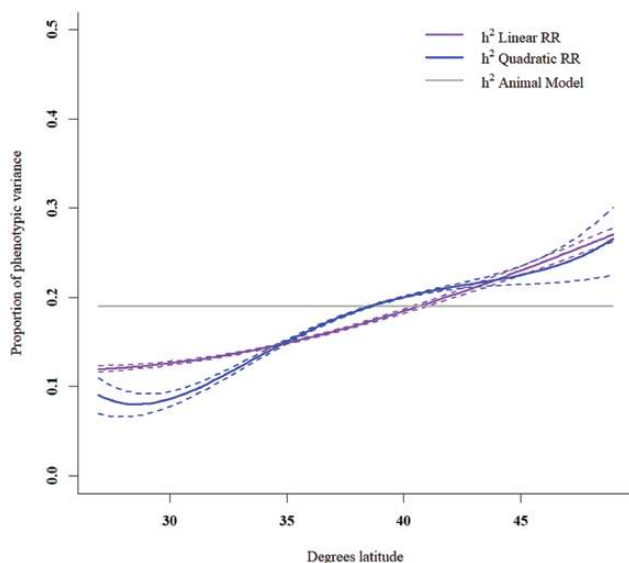
Figure 3. Estimates of heritability for IMF from the animal model, linear, and quadratic random regression of IMF on latitude (dashed lines indicate ± 1 SE).

only a linear random regression was fairly symmetric and positively parabolic, generally smooth, and indicated greater estimates of heritability at the two extremes of longitudinal coordinates (farthest West, and farthest East), with a minimum heritability in the middle of the regression. The curve of heritability estimates from analyses with a quadratic random regression was asymmetric, with a single inflection less centered and positioned closer to the West than for the linear random regression (Figure 4).

Unique Random Regressions by Region

The continental United States was subdivided into two major regions (North and South or West and East for latitude or longitude coordinates as covariate, respectively), with an overall similar number of records and geographical area covered. This approach made it possible to estimate variances for random intercepts and random regression linear coefficients and covariance components in analyses with latitude coordinates as was covariate (Table 3) and longitude coordinates as covariate (Table 4). Increasing the order of the random regression to a quadratic polynomial (intercept, linear, and quadratic) was never accomplished when data were divided into regions when the additional covariances were included in the model. This may be a consequence of differing data support available in the same latitudes (or longitudes) of different regions. That is, the parameter spaces of the covariate were different for the different regions, and this is especially the case in the four-region analyses. For instance, consider as an example the records in Florida (Figures 1 and 2); there are few (or none) records in
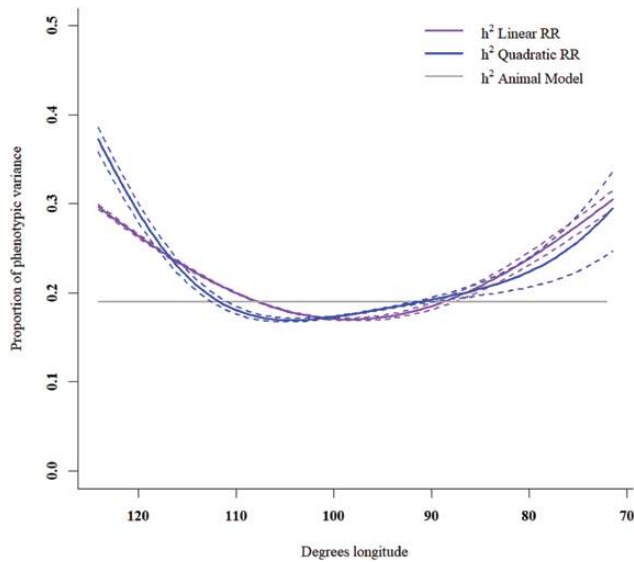
to 0.30 (linear random regression only), and from 0.17 to 0.37 using the quadratic random regression (Figure 4). The curve of plotted estimates of heritability from the analysis that included

Figure 4. Estimates of heritability for IMF from the animal model, linear, and quadratic random regression of IMF on longitude (dashed lines indicate ± 1 SE).

Table 3. Linear random regression (co)variance components estimates for IMF on latitude in two regions[1,2]

	β_0 East	β_1 East	β_0 West	β_1 West
β_0 East	**0.27 ± 0.007**	0.77 ± 0.076	0.91 ± 0.031	0.37 ± 0.081
β_1 East	0.06 ± 0.004	**0.03 ± 0.005**	0.62 ± 0.115	0.41 ± 0.215
β_0 West	0.22 ± 0.009	0.05 ± 0.008	**0.22 ± 0.008**	0.34 ± 0.065
β_1 West	0.04 ± 0.009	0.01 ± 0.007	0.03 ± 0.004	**0.04 ± 0.007**

[1]Variances are on diagonal and in bold type. Covariances are below that diagonal and correlation coefficients are above.
[2]Data were divided into West and East regions at 99°W longitude.

Table 4. Linear random regression (co)variance components estimates for IMF on longitude in two regions[1,2]

	β_0 North	β_1 North	β_0 South	β_1 South
β_0 North	**0.32 ± 0.008**	0.32 ± 0.020	0.78 ± 0.038	0.18 ± 0.085
β_1 North	0.07 ± 0.007	**0.17 ± 0.011**	0.06 ± 0.070	0.11 ± 0.148
β_0 South	0.20 ± 0.011	0.01 ± 0.013	**0.21 ± 0.006**	−0.03 ± 0.027
β_1 South	0.03 ± 0.016	0.02 ± 0.020	−0.01 ± 0.004	**0.11 ± 0.008**

[1]Variances are on diagonal and in bold type. Covariances are below that diagonal and correlation coefficients are above.
[2]Data were divided into North and South regions at 40°N latitude.

the other regions that had corresponding latitudes. Even in latitudes where there are records in each of the regions, some regions have much less support of records (see the Southwest region including New Mexico, Arizona, Nevada, and Southern California). Variances of intercepts were similar in magnitude in random regressions of IMF on latitude or longitude; however, the variances of the linear random regression coefficients on longitude (Table 4) were from two to five times larger than those from regression on latitude (Table 3). Correspondence among random regression coefficients (latitude) was positive and large, as correlations of those components ranged from 0.34 to 0.91 (Table 3); much lower correspondence of random regression coefficients (longitude) is shown in Table 4. As an exception,

the correlation of intercept coefficients for the North and South regions was large (r = 0.78 ± 0.04).

Random regression of IMF on latitude resulted in estimation of only variances of regression coefficients and within-region covariances of coefficients (Table 5). Intercept variances were similar in the central regions (West 2 and East 1) and between one half and two-thirds of the magnitude of those on the extremes (West 1 and East 2). The West 1 region (furthest West) differed substantially from the other regions as the variance estimates of the linear regression coefficient was two to three times larger than all other regional variances of that coefficient. The West 1 region had a negative genetic correlation between the intercept and linear regression coefficients; all other regions were large and positive.

A larger number of parameters was estimated in analyses of random regression on longitude in four regions (Table 6). Variances estimated for the intercept and linear regression terms were largest in the Northernmost region (North 1) and progressively smaller in each region to the South, which was similar to the pattern of variances from analyses of data in two regions (Table 4). Within-region correlations between the intercept and linear terms were large and positive in both North regions, but of low magnitude in both South regions, and that in the Southernmost region (South 2) was negative (Table 4). Across-region correlations between the linear regression coefficient terms were large and positive for North 1 with North 2 and South 1 with South 2. That for North 2 with South 1 (these are adjacent regions) was less than half the magnitude. Correlations of linear regression coefficients of North 1 with South 1 and North 2 with South 2 did not differ from 0. The correlation between linear regression coefficients from the extreme regions (North 1 with South 2) was large and negative.

Random regression analyses of subdivided data indicated reasonably similar estimates of heritability for IMF at Northern latitudes (Figures 5 and 6). However, at Southern latitudes, modeling distinct random regressions in the Western U.S. resulted in much higher estimates of heritability in the Westernmost region (West 1; Figure 5). Heritability estimates determined within the East and West regions ranged from 0.09 to 0.32, and 0.14 to 0.27, respectively. Heritability estimates in the West 1 and West 2 regions ranged from 0.21 to 0.46 and 0.09 to 0.26, respectively. Heritability estimates in the East 1 and East 2 regions ranged from 0.13 to 0.29 and 0.05 to 0.48, respectively. The curves of estimates of heritability obtained for the West, West 1, and West 2 regions differed in shape as well as in minimum and maximum values; particularly the curve obtained within the West 1 region, which had a shape more similar to a parabola in comparison to curves from the West and West 2 regions, which appear linear. The minimum heritability value within the West 1 region was nearer to the center of the latitude coordinates evaluated (41°N), unlike the minimum value determined within the West, and West 2 regions, which was closer to the farther South coordinate (32.9 and 26.9°N, respectively). These results could also be in at least part a consequence of differing quantities of data at the same values of the covariate in different regions.

Random regression analyses of data subdivided into two regions resulted in patterns of heritability of somewhat greater estimates in the Northern region (Figures 7 and 8), especially at the Easternmost longitudes. Those estimates ranged from 0.19 to 0.47 and 0.15 to 0.31, for the Northern and Southern regions. Analyses of data subdivided into four regions resulted in again larger estimates of heritability for IMF in Eastern longitudes,

Table 5. Linear random regression (co)variance components estimates for IMF on latitude in four regions[1, 2, 3]

	β_0 East 1	β_1 East 1	β_0 East 2	β_1 East 2	β_0 West 1	β_1 West 1	β_0 West 2	β_1 West 2
β_0 East 1	**0.24 ± 0.008**	0.49 ± 0.078						
β_1 East 1	0.04 ± 0.004	**0.03 ± 0.008**						
β_0 East 2			**0.37 ± 0.010**	0.92 ± 0.072				
β_1 East 2			0.15 ± 0.008	**0.07 ± 0.014**				
β_0 West 1					**0.35 ± 0.015**	−0.31 ± 0.044		
β_1 West 1					−0.07 ± 0.015	**0.14 ± 0.020**		
β_0 West 2							**0.19 ± 0.009**	0.64 ± 0.126
β_1 West 2							0.04 ± 0.004	**0.03 ± 0.008**

[1]Absence of a value indicates that estimation of this parameter was not accomplished and was fixed to 0.
[2]Variances are on diagonal and in bold type. Covariances are below that diagonal and correlation coefficients are above.
[3]Lower numbers indicate regions further West. Boundaries separating the four regions were 104.55°W (between West 1 and West 2), 99°W (between West 2 and East 1), and 92.22°W (between East 1 and East 2).

Table 6. Linear random regression (co)variance components estimates for IMF on longitude in four regions[1, 2, 3]

|	β_0 North 1	β_1 North 1	β_0 North 2	β_1 North 2	β_0 South 1	β_1 South 1	β_0 South 2	β_1 South 2
β_0 North 1	**0.42 ± 0.014**	0.46 ± 0.033						
β_1 North 1	0.17 ± 0.022	**0.33 ± 0.033**		0.54 ± 0.094		0.02 ± 0.147		−0.63 ± 0.167
β_0 North 2			**0.30 ± 0.008**	0.49 ± 0.022				
β_1 North 2		0.13 ± 0.022	0.11 ± 0.007	**0.16 ± 0.013**		0.21 ± 0.143		−0.11 ± 0.202
β_0 South 1					**0.25 ± 0.008**	0.13 ± 0.030		
β_1 South 1		0.01 ± 0.033		0.03 ± 0.022	0.02 ± 0.006	**0.15 ± 0.010**		0.49 ± 0.202
β_0 South 2							**0.16 ± 0.008**	−0.11 ± 0.057
β_1 South 2		−0.10 ± 0.029		−0.01 ± 0.023		0.05 ± 0.023	-0.01 ± 0.007	**0.08 ± 0.016**

[1]Absence of a value indicates that estimation of this parameter was not accomplished and was fixed to 0.
[2]Variances are on diagonal and in bold type. Covariances are below that diagonal and correlation coefficients are above.
[3]Lower numbers indicate regions further North. Boundaries separating the four regions were 44.46°N (between North 1 and North 2), 40°N (between North 2 and South 1), and 36.46°N (between South 1 and South 2).
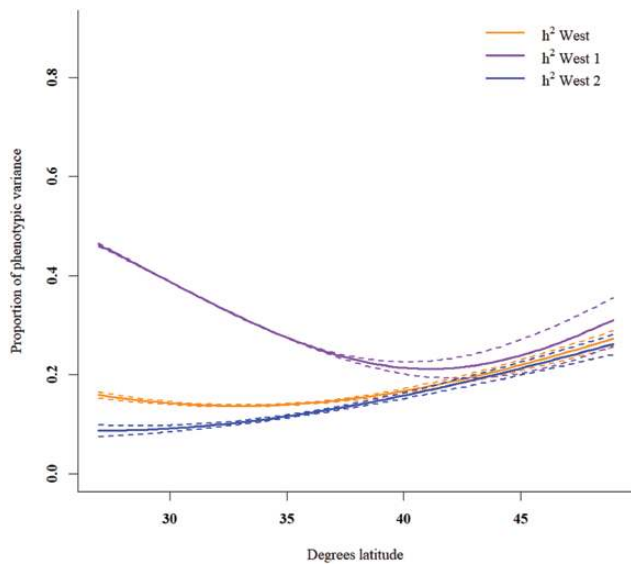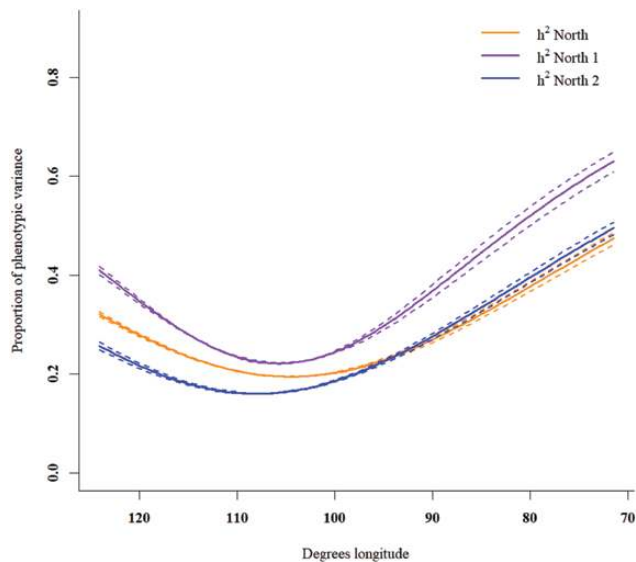

Figure 5. Estimates of heritability from linear random regression of IMF on latitude within the West region (data divided into West and East regions at 99°W), and within West 1 (furthest West subregion with boundary at 104.55°W) and West 2 regions (dashed lines indicate ± 1 SE).
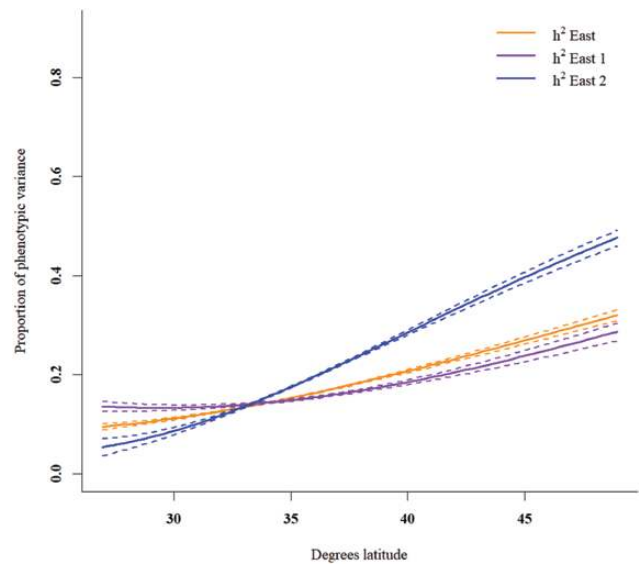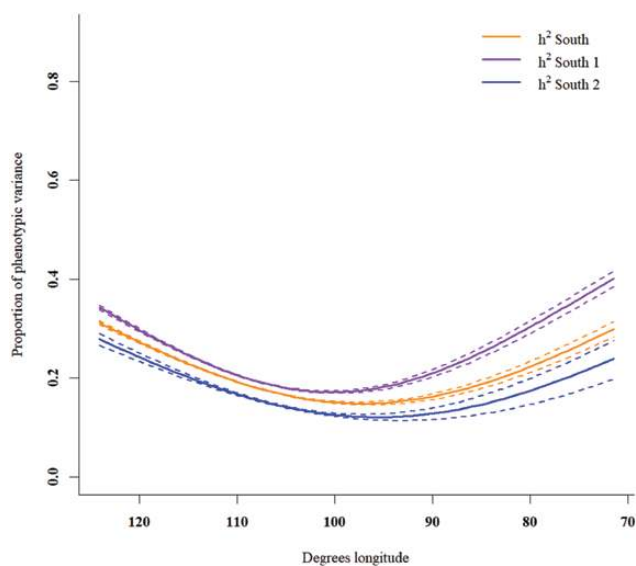
Figure 6. Estimates of heritability from linear random regression of IMF on latitude within the East region (data divided into West and East regions at 99°W), and within East 1 (furthest West subregion with boundary at 92.22°W) and East 2 regions (dashed lines indicate ± 1 SE).

especially in the Northernmost region (North 1); or in other words, successively lower estimates in each region Southward. Those estimates ranged from 0.22 to 0.63 (North 1), from 0.16 to 0.50 (North 2; Figure 7), from 0.17 to 0.40 (South 1), and from 0.12 to 0.28 (South 2; Figure 8). The change in heritability estimates across longitude coordinates was similar in trajectory, where from West to East the heritability exhibited a decrease in value, achieving a minimum at longitude coordinates 104, 105, and 108° W for the North, North 1, and North 2 regions, respectively. With decreasing longitude (moving Eastward), heritability estimates increased until achieving a maximum value at 71°W for North, North 1, and North 2 regions (Figure 7). Heritability estimates

Figure 7. Estimates of heritability from linear random regression of IMF on longitude within the North region (data divided into North and South regions at 40°N), and within North 1 (furthest North subregion with boundary at 44.46° N) and North 2 regions (dashed lines indicate ± 1 SE)

Figure 8. Estimates of heritability from linear random regression of IMF on longitude within the South region (data divided into North and South regions at 40°N), and within South 1 (furthest North subregion with boundary at 36.46°N) and South 2 regions (dashed lines indicate ± 1 SE).

across longitude coordinates also decreased in the South, South 1, and South 2 regions (Figure 8); but reach a minimum value near the center of the evaluated longitude coordinates, and higher values at the extremes of the respective curves. Differences at the extreme covariate values were less pronounced in the South (Figure 8) than those in Northern regions.

Rank changes of breeding values across pairs of latitude and longitude coordinates under evaluation were minimal, as indicated by large (> 0.9) Spearman rank correlation coefficients.

Discussion

Random regression methodology may provide an effective way to model the geographic–environmental complex. Likelihood ratio tests across models with data not subdivided into regions, using either latitude or longitude, indicated that the quadratic random regression better fit the data in comparison to the linear random regression and animal model. However, the variance component estimates from the quadratic random regression model were low in comparison to estimates from the linear random regression using either latitude or longitude. Additionally, the extremes of the heritability curves estimated across latitude or longitude have larger standard errors when a quadratic random regression is modeled in comparison to the linear random regression. All this considered, maybe the use of a quadratic random regression is not necessarily the best alternative to model IMF with the current database, and more accurate estimations could be done using the linear random regression.

Heritability estimates using random regression of IMF on latitude coordinates (both linear and quadratic regressions) were greater in the Northern and lower in the farthest South latitudes of the United States. Random regression of IMF on longitude coordinates yielded heritability estimates that were lowest in the middle section of the country, and largest in the far West. Evaluation of linear random regression within four rather than two regions did not substantially impact the shapes of heritability curves from regions or subdivided regions; but the minimum and maximum values for IMF heritability were influenced. In contrast, more restricted region size influenced the rate of change in the estimates of heritability across latitudes, especially in the Western half of the country. In those analyses, the maximum estimate of heritability in the furthest East region of 0.48 was noticeably greater than the other East subdivided region (0.32) and the overall (not subdivided) East region (0.29). The maximum estimate of heritability in the coastal West region was larger than the inland West region and the overall West region (0.46, 0.27, and 0.26, respectively), and this was the case for the minimum estimates of heritability in those same regions (0.21 vs. 0.14 and 0.09, respectively). Results from random regression analyses on longitude and latitude jointly suggest that the greater heritability for IMF can be found in the coastal areas of the Northern United States, and the lowest values are found in the central South.

The range of IMF heritability estimated with the random regression models was similar to results out of traditional (not random regression) analyses (0.26 to 0.42) in Hereford cattle (0.26; Moser, 2006; 0.42; Su et al., 2017), suggesting that results are representative of the U.S. Hereford population.

This current work suggests that the environment where animals were evaluated may impact the additive genetic estimates. Sire or other genetic components as an interaction with environment may merit inclusion in genetic evaluation

(Bertrand et al., 1985, 1987). Genotype–environment interactions also have been studied for birth weight and weaning weight in Red Angus cattle using random regressions that evaluated sire's progeny within regions that differed in temperature and humidity indices (Fennewald et al., 2017). Those authors concluded that genotype–environment interactions, although present, did not influence the rank of sire predictions of genetic merit for those traits. There were substantially different regional heritability estimates (not random regression) for Red Angus birth weight (0.00 and 0.46, in the area near the Gulf Coast and the Upper Great Plains, respectively) and weaning weight (0.05 and 0.41, Gulf Coast and desert subregions; Fennewald et al., 2017).

The use of either longitude or latitude coordinates as an environmental gradient is a proxy for a combination of ambient and nutrition conditions. Those appeared to be most severe when considering covariance (of coefficients) differences from North to South and would at first consideration suggest differences in forage species which are being consumed by the cattle or in environment temperature, especially considering that temperatures tend to increase with increased proximity to the equator (at lower latitudes farther South), and decrease when farther from the equator (at higher latitudes farther North). However, the large differences in estimates of variances from West to East were noteworthy, especially differences between the farther West region and the rest of the regions, which may be associated to the lower annual precipitations observed in the Western regions of the country, especially in the farther West region (NOAA, 2019). On the other hand, observed similarity in IMF heritability estimates in regions closer together in proximity could have been associated to a greater degree of environmental connectedness than between regions farther apart.

Whether modeled through linear or quadratic random regressions, estimates of heritability (without modeling region) were lower in latitudes farther South (Figure 3). Linear random regressions on latitude supported this with an exception of in the far West (West 1; Figure 5) in which estimates of heritability at lower latitudes were higher. North to South differences were also evident in the random regressions on longitude. Estimates of heritability were lower in the central (90 to 110° longitude) United States as indicated by either linear or quadratic random regression results (Figure 4). Random regression analysis within regions (linear random regression only) produced similar results (Figures 7 and 8) except that heritability estimates were much higher in the lowest longitudes of all regions, but especially the furthest North (Figure 7). Estimated correlation coefficients from random regression within regions supported positive correspondence of East–West regions in comparison to those from the North and South regions (0.41 vs. 0.11, respectively). This appropriately suggests greater difference between the environmental factors from Northern and Southern sections of the country. Adjacent regions often had positive correlations, and regions farther apart had negative correlations which became more negative with distance; that for the Northernmost and Southernmost was –0.63.

When evaluating the effects of G × E interaction on genetic merit predictions for IMF, no substantial changes in ranks of animals across latitude and longitude coordinates were observed. Fennewald et al. (2017) evaluated G × E effects over birth weight and weaning weight in Red Angus cattle using random regression methodology, finding that ranking among sires was similar across environments. In dairy cattle, Kolmodin et al. (2002) used random regression methodology to assess G × E interactions effects over production and fertility traits,

identifying just a small change in the ranking of sires across environments, with exception of the extremes points of the environmental gradient. G × E interactions have been often attributed to changes in scale (McDaniel and Corley, 1967; Stanton et al., 1991; Cromie et al., 1998). Landscape genetics (Manel et al., 2003; Manel and Holderegger, 2013) represents a formidable opportunity to describe populations in a spatial context. Especially if combined with genomic information (Storfer et al., 2018), issues of local adaptation through detection of candidate genes under selection alter the G × E study paradigm, and this field may be the next appropriate focus of livestock selection genetics.

These results show the feasibility of using random regressions to account for genotype–environment interactions in genetic merit predictions with Hereford cattle. Furthermore, the use of this type of strategy would make it possible to select sires based on their location-specific genetic merit instead of an overall average across the country. This would lead to a more efficient genetic improvement of IMF, where the potential improvement per generation will depend on the genetic variability within specific geographic locations.

Considering the findings out of this project, the use of random regressions represents a potential better tool to select sires according to the specific environments where their operations are located. Nevertheless, further studies are needed in order to assess this methodology in additional economically relevant traits for this breed.

Implications

Results from the current work indicate that random regressions reveal differences in additive genetic variance and heritability for IMF in American Hereford cattle in different geographical locations within the continental United States. A traditional animal model may not adequately account for the possible change in genetic variability for economically relevant traits in livestock production systems across environments. The incorporation of random regression methodology in genetic merit predictions has potential for the American Hereford Association evaluations, as it accommodates a more flexible and appropriate characterization of additive genetic variance contingent upon environmental differences. However, further research needs to be done in order to confirm that current methodology is rightfully accounting for all major sources of environmental effects influencing genotypes.

Acknowledgments

Literature Cited

Bertrand, J. K., P. J. Berger, and R. L. Willham. 1985. Sire × environment interactions in beef cattle weaning weight field data. *J. Anim. Sci.* **60**:1396–1402. doi:10.2527/jas1985.6061396x

Bertrand, J. K., J. D. Hough, and L. L. Benyshek. 1987. Sire × environment interactions and genetic correlations of sire progeny performance across regions in dam-adjusted field data. *J. Anim. Sci.* **64**:77–82. doi:10.2527/jas1987.64177x

Blackburn, H. D., B. Krehbiel, S. A. Ericsson, C. Wilson, A. R. Caetano, and S. R. Paiva. 2017. A fine structure genetic analysis evaluating ecoregional adaptability of a *Bos taurus* breed (Hereford). *PLoS One* **12**:e0176474. doi:10.1371/journal.pone.0176474

Burfening, P. J., D. D. Kress, and R. L. Friedrich. 1982. Sire × region of United States and herd interactions for calving ease and birth weight. *J. Anim. Sci.* **55**:765–770. doi:10.2527/jas1982.554765x

Burns, W. C., M. Koger, W. T. Butts, O. F. Pahnish, and R. L. Blackwell. 1979. Genotype by environment interaction in Hereford cattle: II. Birth and weaning traits. *J. Anim. Sci.* **49**:403–409. doi:10.2527/jas1979.492403x

Cardoso, F. F., and R. J. Tempelman. 2012. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *J. Anim. Sci.* **90**:2130–2141. doi:10.2527/jas.2011-4333

Cromie, A. R., D. L. Kelleher, F. J. Gordon, and M. Rath. 1998. Genotype by environment interaction for milk production traits in Holstein Friesian dairy cattle in Ireland. *Interbull Bull.* **17**:100–104.

Fennewald, D. J., R. L. Weaber, and W. R. Lamberson. 2017. Genotype by environment interactions for growth in Red Angus. *J. Anim. Sci.* **95**:538–544. doi:10.2527/jas.2016.0846

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. *ASReml user guide release 3.0*. Hemel Hempstead, UK: VSN International Ltd..

Hassen, A., D. E. Wilson, and G. H. Rouse. 2003. Estimation of genetic parameters for ultrasound-predicted percentage of intramuscular fat in Angus cattle using random regression models. *J. Anim. Sci.* **81**:35–45. doi:10.2527/2003.81135x

Hayes, B. J., H. D. Daetwyler, and M. E. Goddard. 2016. Models for genome × environment interaction: examples in livestock. *Crop. Sci.* **56**:1–9. doi:10.2135/cropsci2015.07.0451

Hocquette, J. F., F. Gondret, E. Baéza, F. Médale, C. Jurie, and D. W. Pethick. 2010. Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. *Animal* **4**:303–319. doi:10.1017/S1751731109991091

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agric. Scand.* **52**:11–24. doi:10.1080/09064700252806380

MacNeil, M. D., F. F. Cardoso, and E. Hay. 2017. Genotype by environment interaction effects in genetic evaluation of preweaning gain for Line 1 Hereford cattle from Miles City, Montana. *J. Anim. Sci.* **95**:3833–3838. doi:10.2527/jas2017.1829

Manel, S., and R. Holderegger. 2013. Ten years of landscape genetics. *Trends Ecol. Evol.* **28**:614–621. doi:10.1016/j.tree.2013.05.012

Manel, S., M. K. Schwarz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends. Ecol. Evol.* **18**:189–197. doi:10.1016/S0169-5347(03)00008-9

McDaniel, B. T., and E. L. Corley. 1967. Relationships between sire evaluations at different herdmate levels. *J. Dairy Sci.* **50**:735–741. doi:10.3168/jds.S0022-0302(67)87503-9

Moser, D. 2006. Performance Update. Available from https://hereford.org/wp-content/uploads/2017/02/issue-archive/0706_PerformanceUpdate.pdf – [accessed November 5, 2018].

NOAA. 2019. National Temperature and Precipitation Maps. Available from https://www.ncdc.noaa.gov/temp-and-precip/us-maps/12/201812?products[]=nationalpcpnrank&products[]=regionalpcpnrank&products[]=statewidepcpnrank&products[]=divisionalpcpnrank#us-maps-select – [accessed August 14, 2019].

Notter, D. R., B. Tier, and K. Meyer. 1992. Sire x herd interactions for weaning weight in beef cattle. *J. Anim. Sci.* **70**:2359–2365. doi:10.2527/1992.7082359x

Pethick, D. W., G. S. Harper, and V. H. Oddy. 2004. Growth, development and nutritional manipulation of marbling in cattle: a review. *Aust. J. Exp. Agric.* **44**:705–715. doi:10.1071/EA02165

Santana, M. L., Jr, J. P. Eler, A. B. Bignardi, A. Menéndez-Buxadera, F. F. Cardoso, and J. B. Ferraz. 2015. Multi-trait linear reaction norm model to describe the pattern of phenotypic expression of some economic traits in beef cattle across a range of environments. *J. Appl. Genet.* **56**:219–229. doi:10.1007/s13353-014-0242-9

Schaeffer, L. R. 2016. Random regression models. Available from http://animalbiosciences.uoguelph.ca/~lrs/BOOKS/rrmbook.pdf – [accessed November 5, 2018].

Stanton, T. L., R. W. Blake, R. L. Quaas, L. D. Van Vleck, and M. J. Carabaño. 1991. Genotype by environment interaction for Holstein milk yield in Colombia, Mexico, and Puerto Rico. *J. Dairy Sci.* **74**:1700–1714. doi:10.3168/jds.S0022-0302(91)78333-1

Storfer, A., A. Patton, and A. K. Fraik. 2018. Navigating the interface between landscape genetics and landscape genomics. *Front. Genet.* **9**:68. doi:10.3389/fgene.2018.00068

Su, H., B. Golden, L. Hyde, S. Sanders, and D. Garrick. 2017. Genetic parameters for carcass and ultrasound traits in Hereford and admixed Simmental beef cattle: accuracy of evaluating carcass traits. *J. Anim. Sci.* **95**:4718–4727. doi:10.2527/jas2017.1865

Williams, J. L., M. Lukaszewicz, J. K. Bertrand, and I. Misztal. 2012. Genotype by region and season interactions on weaning weight in United States Angus cattle. *J. Anim. Sci.* **90**:3368–3374. doi:10.2527/jas.2011-4276.