Extensive genome-wide association analyses identify genotype-by-environment interactions of growth traits in Simmental cattle

Camila U. Braz[1], Troy N. Rowan[1,2], Robert D. Schnabel[1,2,3], Jared E. Decker[1,2,3*]

[1] Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA.
[2] Genetics Area Program, University of Missouri, Columbia, MO 65211, USA.
[3] Informatics Institute, University of Missouri, Columbia, MO 65211, USA.

*Corresponding Author: deckerje@missouri.edu

Abstract

We identify genotype-by-environment interactions (GxE) for birth, weaning and yearling weights (BW, WW, and YW, respectively) using ~835,000 SNPs in ~13,500 Simmental cattle by conducting: direct GxE genome-wide association analyses using continuous environmental variables (minimum, mean, maximum and mean dew point temperatures; elevation; precipitation; and minimum and maximum vapor pressure deficit), and in combination (U.S. ecoregions); and variance-heterogeneity genome-wide association (vGWA) analyses, indicative of interactions, using residuals adjusted for additive, dominance, and epistatic relationships. Genotype-by-environment interactions contributed to 10%, 4%, and 3% of the phenotypic variance of BW, WW, and YW, respectively. Genes were related to response to stimulus, nitrogen compound, gene expression, development and metabolic processes. Twenty-two vQTL (difference in variance between genotypes) were detected. Some vQTL were enriched with GxE effects while one vQTL was also a QTL (difference in the mean between genotypes). This study reveals the importance to investigate alternative approaches using genomic information to identify loci contributing to genetic control of complex traits. Further, we provide evidence for extensive GxE interactions in Simmental cattle.

Introduction

Cattle are reared over a wide range of climatic and topographic regions across the United States, contributing to genotype-by-environment (GxE) interactions and leading to poorly adapted cattle, which negatively affects the profitability of the beef cattle production system (Kolmodin et al., 2002; Howard et al., 2014). GxE interactions refer to a variable response of genotypes to changes in the environment, resulting in different outcomes for animals and their offspring in different environments than those in which they were raised or selected (Falconer et al. 1996; Kolmodin et al., 2002). Therefore, understanding the genetic basis of GxE interaction is crucial to improve the environmental adaptation and sustainability of beef cattle production. This understanding will better drive selection and mating decisions, and consequently, improve beef production efficiency (Guidolin et al., 2012).

Studies have reported the existence of GxE interaction for growth traits in beef cattle. Williams et al. (2012) found evidence for GxE interaction for weaning weight and post-weaning gain in beef cattle reared in different altitudes. Bradford et al. (2016) reported GxE interaction for weaning weight when comparing groups of US Angus cattle during and without heat stress. Potential genes involved in GxE interaction when comparing US geographic regions were identified by Smith et al. (2019) using Gelbvieh cattle. These studies suggest that the same trait may be modeled by different sets of genes across divergent environments (Falconer & Mackay, 1996). However, a limited amount of research has been done on identifying genes or biological pathways directly interacting with environmental factors in beef cattle.

In this study we searched for QTLs enriched for GxE interaction for birth weight (BW), weaning weight (WW), and yearling weight (YW) in Simmental cattle, using univariate and multivariate linear mixed models through three approaches: i) conducting direct GxE genome-

wide association (GxE GWA) analyses using continuous environmental variables, such as minimum, mean, maximum and dew point temperatures; elevation; precipitation; and minimum and maximum vapor pressure deficit; ii) performing GxE GWA analyses using United States ecoregions as environmental factors; and iii) carrying out variance-heterogeneity genome-wide association (vGWA) analyses, as indicators of interactions, using residuals accounting for additive, dominance and epistasis effects. To our knowledge, this is the first study investigating the interaction of genome-wide SNP markers with a vast assortment of environmental variables and searching for SNPs controlling phenotypic variance using a large beef cattle data set. Our results have the potential to provide novel insights into how the environment influences complex traits and adaptation in beef cattle and potentially across mammals.

Materials and Methods

Phenotype and genotype data

Data from Simmental males and females born between 1975 and 2016 were transferred to the University of Missouri from the American Simmental Association. The traits studied were birth weight (BW), 205-d adjusted weaning weight (WW) and 365-d adjusted yearling weight (YW), with the traits also adjusted for age of dam as per breed association practice. The animals were assigned to contemporary groups (CG), defined by the combination of farm, and season and year of birth.

The DNA samples were genotyped with various low density assays and were imputed to the combination of Illumina BovineHD (Illumina, San Diego, CA) and the GeneSeek Genomic Profiler F250 (GeneSeek, Lincoln, NE) according to Rowan et al. (2019). The genotype filtering was performed to remove non-autosomal variants, as well as both variants and individuals with

call rate less than 0.90, using PLINK 1.9 (Purcell et al., 2007). SNP positions were based on the ARS-UCD1.2 Bovine reference genome assembly (Rosen et al., 2018). The genotypes were then phased using Eagle 2.4 (Das et al., 2016) and imputed using Minimac3 (Loh et al., 2016). The final imputed genotype data contained a total of 835,947 bi-allelic variants with imputation accuracy of 99.6% (Rowan et al., 2019). Variants with minor allele frequency less than 0.01 were removed using PLINK 1.9 for further analysis, remaining 710,202, 709,054, and 706,771 SNP markers for 13,427, 11,847, and 8,546 animals for BW, WW and YW, respectively.

Eight environmental continuous variables were used for the GxE GWA analyses, including minimum, mean, maximum and mean dew point temperatures; elevation; precipitation; and minimum and maximum vapor pressure deficit, which were drawn from the PRISM climate dataset (http://prism.oregonstate.edu). The United States was partitioned into nine regions (Figure 1A) based on similar topographic and environmental conditions (ecoregions), including mean temperature, elevation and precipitation information, using k-means clustering implemented in '*RStoolbox*' R package (Leutner and Horning, 2017, R Core Team, 2013). The optimal ecoregions were identified using '*pamk*' function from the R package '*fpc*' (Hennig, 2013). The ecoregions were named Desert (D), High Plains (HP), Arid Prairie (AP), Forested Mountains (FM), Upper Midwest & Northeast (UN), Southeast (SE), Rainforest (R), Cold Desert (CD), and Fescue Belt (FB). The animals were assigned to the ecoregions based on the zip code of their breeder (Figure 1B), and only ecoregions with more than 200 records were analyzed. Thus, animals located in the Rainforest ecoregion were removed from the dataset, and those that were reared in the Desert and Arid Prairie ecoregions were analyzed together (DAP) due to their similar environmental conditions. No animal record was assigned to the Cold Desert ecoregion. A complete description

of the dataset, including weight records and environmental variable measurements is shown in Table 1.

Variance components

Variance components for BW, WW, and YW were estimated using multi-component restricted maximum likelihood (REML) approach, implemented in GCTA software (Yang et al., 2010), in univariate linear mixed models that included only additive effects (MA), or additive, dominance and epistasis effects (MADE) as follows:

$$y = Xb + Kc + Z_a a + e \text{ (MA)}$$

$$y = Xb + Kc + Z_a a + Z_d d + Z_{ep} ep + e \text{ (MADE)}$$

where y is the vector of phenotypes; b is the vector of fixed effect (sex); c is the vector for random CG effects; a is the vector of random additive genetic effects; d is the vector of random dominance effects; ep is the vector of random epistatic effects; and e is the vector of random residuals. Z matrices are incidence matrices relating observations to animals. Assuming that a ~ N(0, $G\sigma_a^2$), d ~ N(0, $D\sigma_d^2$), ep ~ N(0, $E\sigma_{ep}^2$), e ~ N(0, $I\sigma_e^2$), where σ_a^2, σ_d^2, σ_{ep}^2, σ_e^2 are the additive, dominance, epistatic, and residual variances, respectively. K is the CG information; I is the identity matrix; and G, D, and E are the additive, dominance and epistatic genetic relationship matrices, respectively. The genomic (G) and the dominance (D) relationship matrices were calculated using SNP marker information according to Zhu et al. (2015) as: $G = W_G W_G'/N$, and $D = W_D W_D'/N$, where $w_{G(i)} = \left(x_{G(i)} - 2p_i\right)/\sqrt{2p_i(1-p_i)}$ and $w_{D(i)} = \left(x'_{D(i)} - 2p_i^2\right)/\lfloor 2p_i(1-p_i)\rfloor$, with p_i being the allele frequency at SNP i; and N is the number of SNPs. According to Henderson (1985) the epistatic relationship matrix (E) can be derived from the additive genomic relationship matrix

as: $E = G\#G$, where # denotes the Hadamard product operation. For WW, the random maternal permanent environmental effect was also included in the model.

The variance of GxE interaction was also estimated considering the ecoregions as an environmental factor and treated as random effect in the model in order to estimate the amount of phenotypes that are explained by the GxE interaction among the ecoregions.

Genome-wide association analyses

Phenotypes were pre-adjusted for fixed effect of sex, estimated using '*--reml-est-fix*' option, and random effect of CG was predicted by the BLUP method using '*--reml-pred-rand*' flag using single-trait animal model implemented in GCTA software. The random maternal permanent environmental effect was also included in the WW adjustment.

Univariate linear mixed model analyses were performed for the pre-adjusted BW, WW or YW using GEMMA software (Zhou et al., 2012) as follows:

$$y^* = 1\mu + X\beta + u + e;$$

where y^* is an *n*-vector of pre-adjusted phenotypes; μ is the overall mean; X is the incidence matrix of genotypes; β is the genotype effects; u is an *n*-vector of random additive genetic effects; and e is an *n*-vector of random residual. Assuming that $u \sim MVN_n(0, GV_g)$ and $e \sim MVN_n(0, IV_e)$, G is the *n* x *n* genomic relationship matrix (GRM); V_g is the residual additive genetic variance; V_e is the residual variance component; I is an *n* x *n* identity matrix. MVN_n denotes the *n*-dimensional multivariate normal distribution. GRM were estimated using the standardized genotypes of 710,202, 709,054, and 706,771 SNP markers for 13,427, 11,847, and 8,546 animals for BW, WW and YW, respectively.

We also fitted multivariate linear mixed model for the pre-adjusted BW, WW and YW using GEMMA in the following form: $Y^* = 1\mu + X\beta + U + E$; where Y^* is an n x d matrix of d pre-adjusted phenotypes for n individuals; μ is the overall mean; X is the incidence matrix of genotypes; β is a d vector of the genotype effects for the d phenotypes; U is an n x d matrix of random additive genetic effects; and E is an n x d matrix of random residuals. $U \sim MN_{nxd}(0, G, V_g)$ and $E \sim MN_{nxd}(0, I_{nxn}, V_e)$, G is the n x n genomic relationship matrix (GRM); V_g is the d x d matrix of additive genetic variance component; V_e is a d x d matrix of the residual variance component; I is an n x n identity matrix. MN_{nxd} denotes the n x d matrix normal distribution with mean 0, row covariance matrix G or I (n x n), and column variance matrix V_g or V_e (d x d). For all GWA analyses performed, single-marker *P*-values were used to generate Manhattan plots using '*manhattan*' function implemented in '*qqman*' R package (Turner, 2018).

Genotype-by-environment GWA analyses

The GxE GWA analyses were performed using GEMMA through two approaches: using eight environmental variables (i.e. minimum, mean, maximum and mean dew point temperatures; elevation; precipitation; and minimum and maximum vapor pressure deficit) separately; and using the ecoregions as environmental factors, where each ecoregion was compared against the total dataset using 0 and 1 dummy coding. In the GxE GWA analyses, for each SNP in turn, GEMMA fits a linear mixed model that controls both the SNP main effect and environmental main effect, while tested for the interaction effect and controlling for population stratification, evaluating the alternative hypothesis (H_1: $x_j\beta_{ij} \neq 0$) against the null hypothesis (H_0: $x_j\beta_{ij} = 0$) for each interaction, therefore the resulting *P*-values correspond to the significance of the GxE interaction

(Smith et al., 2019). This model can be described as: $y^* = 1\mu + x_i\beta_i + x_j\beta_j + x_i * x_j\beta_{ij} + u + e$; where y^* is an *n*-vector of pre-adjusted phenotype; x_i is an *n*-vector of the genotypes of the SNP i; β_i is the effect size of the SNP i; x_j is an *n*-vector of environmental variable j; β_j is the fixed effect of the environmental variable j; β_{ij} is the interaction effects between the genotypes of the SNP i and the environmental variable j; u is an *n*-vector of random additive genetic effects; and e is an *n*-vector of random residual. GxE GWA analyses using multivariate linear mixed model were also performed.

Variance-heterogeneity GWA analyses

We performed vGWA analyses to detect locus affecting the difference in variance between genotypes (vQTL). Residual effects (e) from the MA and MADE models were standardized to z score by the rank-based inverse-normal transformation and squared (z^2), which is a measure of phenotypic variance (Yang et al., 2012; Wang et al., 2019). The vGWA analyses were performed for BW z^2, WW z^2, and YW z^2 using univariate and multivariate linear mixed models implemented in GEMMA software as denoted in "Genome-wide association analyses" section, where y^* is an *n*-vector of z^2; and β_i is the effect size of the SNP i on z^2. Linkage disequilibrium (LD) analyses were performed using PLINK 1.9 (Purcell et al., 2007). All vQTL were defined by SNP markers with r^2 greater than 0.6.

Functional enrichment analyses

We explored 10 kb and 100 kb sequence windows that flanked the significant SNPs ($P <$ 1e-05) to scan for genes located in their vicinity and to identify possible regulatory elements, respectively. Enrichment analysis based on Gene Ontology (GO) and KEGG pathways were

performed using ShinyGO v0.60 (Ge and Jung, 2018), in which the analyses were adjusted for FDR of 5%. GO annotations were also retrieved from the AmiGO browser (Carbon et al., 2009).

Results

Variance component estimates

Contemporary groups (CG) were included in the model as random effect since 19% of individuals analyzed belong to CG with only a single animal (Ugarte et al., 1992). The amount of BW, WW, and YW variation explained by CG effect was 19%, 53%, and 61%, respectively, showing that the effect of farm management on the animal body weight increases at older ages. After adjusting the phenotypes for the CG and sex effects, we estimated the contribution of additive, dominance, epistasis and GxE variances over the phenotypic variances using SNP marker information. GxE variances contribute in 10.1%, 3.8%, and 2.8% to the BW, WW, and YW variances, respectively, using the ecoregions as environmental factor, suggesting that, opposite of the CG effects there is a decrease of the GxE effect on the animal body weight throughout life.

Models accounting only for additive effects (MA) or for additive, dominance and epistasis effects (MADE), based on SNP marker information, were analyzed in order to verify the amount of the phenotypic variances are possibly explained by additive and non-additive genetic effects (Table 2). Overall, the heritabilities (proportion of additive genetic variance in phenotypic variance) were greater than the ratio of the variance explained by the non-additive genetic effects to the phenotypic variances. When non-additive genetic effects were included in model MADE, the heritabilities for BW and WW decreased compared with that of MA model. This result suggests that when dominance and epistasis effects are not included in the model, e.g. MA, the variation due to non-additive genetic effects is distributed to other variance components and may lead to

overestimation of heritability (Palucci et al., 2007; Su et al., 2012). However, with the inclusion of dominance and epistasis effects in the model, the heritability value of YW did not change. Proportions of dominance variances in phenotypic variances for WW and YW were low (2% and 5%), whereas no dominance variances was estimated for BW. Epistatic variances accounted for 25%, 12% and 11% of the phenotypic variances for BW, WW, and YW, respectively, showing that epistasis has a substantial contribution to phenotypic variation for body weight traits, mainly for BW, in Simmental cattle.

Genome-wide association analyses

The GWA analyses identified major QTL on BTA 6, 7, 14, and 20 for BW, WW and YW (Supplementary Figure 1), showing that these genomic regions are very important for the development and growth of the animals during their whole life. Genes that harbor those locations are involved in growth and development (*IBSP*, *BMPR1B*, *SLIT2*, *PLAG1*, *CHD7*, *STC2*, *MSX2*, *SOX17*, *SH3PXD2B*, *SNAI2*, *PPDPFL*, *PDLIM5*), hormone activity (*SLIT2*, *LYN*, *CHD7*, *CRH*, *STC2*), and MAPK signaling (*LYN*, *DUSP1*, *MOS*) processes. In addition, significant SNPs ($P <$ 1e-5) were identified on BTA 3 (53.6 Mb) and BTA17 (73 Mb) for BW; and on BTA 1 (81.1-81.3 Mb), BTA 2 (116.1 Mb), and BTA 5 (96.7 Mb) for WW. The GWA analyses using multivariate model detected QTL in the same positions that the univariate GWA analyses did on BTA 5, 6, 7, 14, 17, and 20. In addition, the multivariate GWA analysis also identified significant signals ($P <$ 1e-5) on BTA 4 (24.5 Mb), 9 (57.9 Mb), 13 (63.3 Mb), 19 (50.5 Mb), and 26 (45.4 Mb).

Genotype-by-environment GWA analyses

The GxE GWA analyses were performed using minimum, mean, maximum and mean dew point temperatures; elevation; precipitation; and minimum and maximum vapor pressure deficit as environmental variables. The correlations between the environmental variables are displayed in Supplementary Table 1. The temperature variables have high correlation between them (0.82 to 0.98); precipitation showed positive correlation with the temperatures (0.57 to 0.86); minimum vapor pressure deficit showed correlation with maximum vapor pressure deficit (0.54); on the other hand, elevation have negative correlation with the temperatures (-0.33 to -0.76) and precipitation (-0.72). Mean dew point temperature was the environmental variable that most interacted with genotypes for BW (941 significant GxE SNPs, $P < 1e\text{-}5$), followed by minimum temperature with 756 SNPs ($P < 1e\text{-}5$) enriched with GxE effects (Table 3). Consequently, the ecoregion that identified more GxE interaction for BW was the Desert and Arid Prairie followed by Forested Mountains, ecoregions with high and low temperatures, respectively. Likewise Forested Mountains has the lowest dew point temperature among all ecoregions. Elevation was found to be the environmental variable that most contribute to GxE interaction for WW and YW, in which 28 and 15 GxE SNPs were identified ($P < 1e\text{-}5$), respectively. In agreement, Forested Mountain, the ecoregion with the highest elevation, showed more SNP-by-environment interaction for WW and YW (31 and 49 GxE SNPs, $P < 1e\text{-}5$, respectively).

The range of the GxE SNP effects identified by the environmental variables and the ecoregions are shown in Table 3. Minimum vapor deficit pressure was the environmental variable that showed the highest GxE SNP effects for all traits varying from -1.399 to 1.831 kg for BW, -3.254 to -3.242 kg for WW, and -6.615 kg for the unique GxE SNP detected for YW. GxE SNP effects for the ecoregions were considerably larger than those for environmental variables. One plausible explanation for this result is that when analyzing ecoregions much more than climate and

topographic information are taken into account including the combination of them, as well as local pathogens and forages. The highest GxE SNP effects were observed for Desert and Arid Prairie ecoregion (-8.243 to 6.043 kg for BW, -11.570 to 6.456 for WW, and -16.450 to 23.620 kg for YW), followed by Forested Mountains, with GxE SNP effects varying from -3.135 to 1.215, -6.785 to 4.737, and -9.397 to 17.690 kg for BW, WW, and YW, respectively.

No SNP with significant GxE effects was identified for all traits, indicating that different genes enriched with GxE effects influence the animal weight over the time. However, using the multivariate GWA analyses 193 GxE SNPs (P < 1e-5) were detected (Table 3). Only one GxE SNP in common was detected interacting with two different ecoregions, showing that the ecoregions have very specific GxE interaction and the traits are modeled by different sets of genes across divergent environments. The SNP *rs133638007*, located on BTA 18 at 61,721,281 bp position, was identified with GxE effects for BW in the Desert & Arid Prairie ecoregion and using multivariate analysis for Forested Mountains ecoregion. The T allele increase BW in 1.06 kg in the Desert & Arid Prairie ecoregion, on the other hand, results from the multivariate analysis showed that the T allele, when interacting with Forested Mountains ecoregion decrease BW in 0.62 kg.

Many SNP markers interact with multiple environmental variables/factors according to the GxE GWA analysis. As expected, many of these SNPs are located on BTA 6 demonstrating the importance of this chromosome for GxE effects. E.g., the SNP *rs134507148* (BTA 6 at 30,910,038 bp) interacts with elevation, precipitation, and mean and minimum temperatures for BW, as well as with Forested Mountains ecoregion using multivariate analysis with effect of -0.60, 1.68, and 2.32 kg on BW, WW, and YW, respectively. Another SNP, *rs110128085* (BTA6 at 37,521,235 bp), interacts with elevation, precipitation, minimum, maximum, mean and dew point

temperatures, and also with Forested Mountains ecoregion for BW. Using multivariate model, *rs110128085* also showed interaction with minimum, mean and dew point temperatures and Forested Mountains ecoregion (-0.48, 2.74, 3.32 kg, for BW, WW and YW, respectively). GxE interaction between the SNP *rs109865961* (BTA 6 at 65,667,126 bp) and mean, minimum, maximum and dew point temperatures for BW were verified, as well as for Fescue Belt ecoregion using univariate BW (-0.49 kg) and multivariate analysis (-0.66, 0.16, -2.48 kg for BW, WW and YW, respectively). The SNP *rs111030267* (82,186,837 bp) interacts with mean, minimum, maximum and dew point temperatures, and maximum vapor pressure deficit for BW. In addition, this SNP showed interaction with Upper Midwest & Northeast ecoregion for BW (0.54 kg). GxE interaction was verified in the SNP *rs133964846* (35,157,409 bp) with precipitation, mean, minimum, maximum and dew point temperatures, and with Southeast ecoregion for BW (0.71 kg). The SNP *rs135398062* (47,833,835 bp) interacts with precipitation, minimum and dew point temperatures, as well as with Fescue Belt ecoregion (0.78 kg).

On BTA 5, GxE interaction was observed between the SNP *rs137064164* (4,161,661 bp) and mean, dew point and minimum temperatures for BW. When analyzing in the multivariate mo dels, *rs137064164* showed interaction with precipitation and with the Southeast ecoregion (-1.79, -2.57, -3.89 kg for BW, WW and YW, respectively). The SNP *rs43195878* (BTA 5 at 88,761,789 bp) interacts with elevation, precipitation and mean, maximum, minimum and dew point tempera tures for BW. We verified GxE interaction between the SNP *rs134867512* (BTA 10 at 56,735,84 5 bp) and elevation and with Forested mountains (8.19 kg) for YW and for BW, WW, and YW w hen analyzed in a multivariate model (0.31, 2.21, 8.31 kg, respectively). The SNP *rs110081354* (BTA 18 at 55,824,967 bp) interacts with Southeast ecoregion for BW (-0.83 kg) and also for mea n, minimum and dew point temperatures for BW. On BTA 21, the SNP located at 27,759,568 bp

(*rs378480293*) interacts with Southeast ecoregion (0.74 kg), mean, maximum, minimum and dew point temperatures for BW.

In total, GxE GWA analysis detected 2,319 SNPs (P < 1e-5) interacting with environmental variables and/or in combination (ecoregions). The GxE SNPs were detected on all chromosomes whereas the majority located on BTA 6 (Figure 2A). Consequently, more genes interacting with environment were also identified on BTA 6 when scanning 10 kb region near the significant SNPs (Figure 2B). Exploring 100 kb region flanking the significant SNPs, BTA 18 was the chromosome with the largest number of candidate genes interacting with environments.

The top 30 enriched GO terms for the genes located at 10 kb or 100 kb from the significant SNPs (P < 1e-05) are summarized in Table 4. In both analyses, the most significant GO term was response to stimulus (GO:0050896), and its direct descendants GO terms cellular response to stimulus (GO:0051716) and response to stress (GO:0006950) were also enriched. Many biological processes involved in metabolism, development, localization, gene expression, transport, signaling, and cell adhesion were also enriched. The top 30 enriched pathways for the genes located at 10kb and 100 kb from the significant SNPs (P < 1e-05) are reported in Table 5. Both analyses revealed the metabolic pathway (bta01100) as the most significantly enriched pathway. The analyses also enriched pathways related to cell proliferation, differentiation, growth, survival, immune and inflammatory responses, and mechanisms of neurotransmission in central nervous system. In addition, circadian entrainment (bta04713), a pathway related to environmental adaptation, was also enriched.

Figures 3 and 4 provide the percentage of candidate genes for each ecoregion that participate in the biological processes and pathways enriched, respectively, using genes located at 100 kb from the significant SNPs (P < 1e-05). The percentage of candidate genes, located at 10 kb

from the significant SNPs, for each ecoregion involved in the enriched biological processes and pathways are displayed in the Supplementary Figures 2 and 3, respectively. Genes associated with GxE in each ecoregion where involved in all enriched biological process, expect for Fescue Belt, which did not have any associated genes with the response to stress biological process annotation (GO:0006950). The majority of the candidate genes with GxE effects for Upper Midwest & Northeast, High Plains, Forested Mountains, and Desert & Arid Prairie ecoregions participate in the single-organism metabolic process (GO:0044710), while for Southeast ecoregion response to stimulus (GO:0050896) was the most representative GO term. Response to stimulus (GO:0050896), single-multicellular organism (GO:0044707), multicellular organism (GO:0032501), single-organism developmental (GO:0044767), developmental (GO:0032502), anatomical structure development (GO:0048856), and multicellular organismal development (GO:007275) processes showed equal representation for Fescue Belt ecoregion. Many pathways are affected by candidate genes with GxE effects for the ecoregions Desert & Arid Prairie, Forested Mountains, Southeast, and High Plains. However, genes with GxE effects for Fescue Belt and Upper Midwest & Northeast influence in specific pathways, whereas those involved in mechanisms of neurotransmission in central nervous system, as GABAergic synapse (bta04727), serotonergic synapse (bta04726) and neuroactive ligand-receptor interaction (bta04080) were the most representative pathways for Fescue Belt ecoregion.

Comparing GxE GWA with GWA analyses results, we verified that 312 SNPs with additive effects on mean trait between genotypes (QTL) were also enriched with GxE effects. Some SNPs showed GxE effects even greater than their additive effects on the traits. E.g., the SNP *rs136523580* (BTA 6 at 30,358,905) has additive effect of -0.43 kg and GxE effects of -0.67 kg in the Forested Mountains ecoregion for BW. Additive effects of the SNP *rs110620934* (BTA 20 at

6,208,301 bp) was 0.46 kg, but its SNP-by-minimum vapor pressure deficit interaction effect was 0.65 kg for BW. The SNP *rs109868461* (BTA 6, at 36,902,704 bp) has additive effects of -0.69 kg, although it has GxE effects of 1.11 kg for BW in the Southeast ecoregion. Such QTL and GxE QTL were located on BTA 6 (*ABCG2, ADGRA3, BMPR1B, bta-mir-218-1, DCAF16, FAM13A, FAM184B, GPRIN3, GRID2, HERC3, HERC5, HERC6, HPGDS, IBSP, KCNIP4, LAP3, LCORL, MED28, MEPE, NAP1L5, NCAPG, PCDH7, PDLIM5, PIGY, PKD2, PPM1K, SLIT2, SMARCAD1*), BTA 7 (*bta-mir-3660*) and BTA 20 (*ATP6V0E1, bta-mir-584-6, CCSER1, ERGIC1, RPL26L1*).

As pointed out above, we identified major QTL for the growth traits on BTA 6 where 531 SNPs (from 26.8 to 48.8 Mb) were detected out of 948 total significant SNPs using multivariate analysis. Likewise, several SNPs interacting with Forested Mountains ecoregion, elevation, precipitation and, mainly, with the temperatures were also detected on BTA 6. Given the importance of BTA 6 underlying the growth traits, we repeated the multivariate GWA analysis including some SNPs as fixed effects in the model in order to identify SNPs accounting for QTL effect. Results indicated that there are many QTL in this region due to the fact that even including six SNPs (not in LD) simultaneously in the model (*rs109849093*, *rs464458177*, *rs110305942*, *rs43459303*, *rs109278547*, *rs137209027* located at 37,211,057; 37,234,136; 37,418,164; 37,695,352; 38,139,495; 41,084,765 bp, respectively), a small number of SNPs were still associated (six out of 531 SNPs).

Most of the SNPs enriched with GxE effects were identified for BW in this study. As BW is an indicator of intrauterine growth (Ferrell, 1993), we investigated whether the environment detected interacting with those SNPs could be reflecting the intrauterine environment of the dam. To do this we adjusted BW for random maternal affect (MAM) including it in the model and

compared the results with those from the GxE GWA analysis without removing maternal effect (MA). Results showed that 876 SNPs were not associated with GxE effects when accounting for the effect of dam, suggesting that those SNPs are interacting with the uterine environment. Pathway and GO enrichment analyses using the genes located near these regions revealed, in particular, metabolic, insulin signaling, starch and sucrose metabolism, and cGMP-PKG signaling pathways; and response to stimulus, blastocyst development, and regulation of signaling processes.

Variance-heterogeneity GWA analyses

The proportion of variance in phenotypes explained by SNP markers in GEMMA (PVE), also called ”chip heritability”, for BW z^2, WW z^2 and YW z^2 was 0.049 ± 0.008, 0.047 ± 0.008, and 0.059 ± 0.011, respectively, showing that phenotypic variance is heritable. In total, 44 SNPs with effect on z^2 (P < 1e-05) were identified on BTA 3, 5, 8, 10, 11, 13, 14, 16, 17, 18, 21, 27, and 29 (Table 6), which are putative 24 vQTL (based on LD analysis; data not shown) being seven vQTL for BW, four vQTL for WW, and eleven vQTL for YW. The multivariate analysis identified five vQTL, of which two vQTL (BTA 8 at 4 Mb; and BTA 29 at 39 Mb) were not identified by the univariate analyses.

Nine vQTL were only significant using z^2 from MA (three vQTL for BW, three vQTL for WW, two vQTL for YW, and one vQTL using multivariate analysis), suggesting that these regions may have gene-by-gene interactions (epistasis effects). Using z^2 from MADE, 15 vQTL were detected (four vQTL for BW; one vQTL for WW; nine vQTL for YW; and four vQTL using multivariate analyses, of which three vQTL were also identified by the univariate analyses), which are supposed to be enriched with GxE effects**.**

Five vQTL were also identified by the GxE GWAS analysis. vQTL for BW on BTA 13 at 35.4 Mb showed interaction with minimum vapor pressure deficit also for BW. The vQTL for YW on BTA 16 at 36.1 showed interaction with precipitation for BW; and with mean, minimum and dew point temperatures for BW and using multivariate analysis. In addition, this locus showed to be enriched with GxE effects for Forested Mountains ecoregion using multivariate analysis. The vQTL for BW on BTA 17 at 57.6 Mb showed interaction with dew point temperature and Southeast ecoregion also for BW. The vQTL for YW on BTA 17 at 70.5 Mb showed interaction with maximum vapor pressure deficit for BW. The vQTL for BW on BTA 10 at 22.3 Mb showed interaction with minimum vapor pressure deficit for BW. However this vQTL was not significant for BW using z^2 from MADE (P = 5e-5).

Comparing vGWA and GWA analyses results, only the vQTL on BTA 14 at 23.1-23.3 Mb was also found as a QTL for BW, indicating that this locus affects both mean and the variance of the trait. This results show that SNPs influencing the variance of the trait (vQTL) are more enriched with GxE effects compared to those with QTL effects.

Funding

This study was supported by Agriculture and Food Research Initiative Competitive, Grant No. 2016-68004-24827, from the USDA National Institute of Food and Agriculture.

Acknowledgment

The authors would like to acknowledge the American Simmental Association for providing phenotypic and genotypic data.

References

1. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, et al. AmiGO: online access to ontology and annotation data. *Bioinformatics.* 2009;25:288-9.

2. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48:1284-7.

3. Falconer DS, Mackay, TFC. Introduction to quantitative genetics, 4th ed. Longman Group. Essex; 1996, 464 pp.

4. Ferrell CL. Factors Influencing Fetal Growth and Birth Weight in Cattle. Roman L. Hruska U.S. Meat Animal Research Center. 1993, 132.

5. Guidolin DGF, Buzanskas ME, Ramos SB, Venturini GC, Lobo RB, Paz CCP, et al. Genotype-environment interaction for post-weaning traits in Nellore beef cattle. *Anim Prod Sci.* 2012;52:975-80.

6. Howard JT, Kachman SD, Snelling WM, Pollak EJ, Ciobanu DC, Kuehn LA, et al. Beef cattle body temperature during climatic stress: a genome-wide association study. *Int J Biometeorol*. 2014;58:1665.

7. Leutner B, Horning N. RStoolbox: Tools for Remote Sensing Data Analysis. R package version 0.1.10.; https://cran.r-project.org/package=RStoolbox. 2017.

8. Hennig C. fpc: Flexible Procedures for Clustering. R package version 2.1-11.1.; https://CRAN.R-project.org/package=fpc. 2013.

9. Loh P, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet.* Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2016;48:1443.

10. Palucci V, Schaeffer LR, Miglior F, Osborne V. Non-additive genetic effects for fertility traits in Canadian Holstein cattle. *Genet Sel Evol*. 2007;39:181-93.

11. PRISM Climate Group. Oregon State University, http://prism.oregonstate.edu, 2011.

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-75.

13. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013; URL https://www.R-project.org/.

14. Kolmodin R, Strandberg E, Madsen P, Jensen J, Jorjani H. Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. *Acta Agric Scand A Anim Sci*. 2002;52:11-24.

15. Rosen B, Bickhart D, Schnabel R, Koren S, Elsik C, Zimin A, et al. Modernizing the Bovine Reference Genome Assembly. Proceedings of the World Congress on Genetics Applied to Livestock Production. *World Congress on Genetics Applied to Livestock Production*; 2018; Molecular Genetics 3:802.

16. Smith JL, Wilson ML, Nilson SM, Rowan TN, Oldeschulte DL, Schnabel RD, Decker JE, et al. Genome-wide association and genotype by environment interactions for growth traits in U.S. Gelbvieh cattle. *BMC Genomics*. 2019;20:926.

17. Steven Xijin Ge, Dongmin Jung. ShinyGO: a graphical enrichment tool for ani-mals and plants

18. bioRxiv 315150; 2018. doi: https://doi.org/10.1101/315150

19. Su G, Christensen OF, Ostersen T, Henryon M, Lund MS. Estimating Additive and Non-Additive Genetic Variances and Predicting Genetic Merits Using Genome-Wide Dense Single Nucleotide Polymorphism Markers. *PLoS ONE*. 2012;7:e45293.

20. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Softw*. 2018;3:731.

21. Ugarte E, Alenda R, Carabaño MJ. Fixed or Random Contemporary Groups in Genetic Evaluations. *J Dairy Sci*. 1992;75:269-78.

22. Wang H, Zhang F, Zeng J, Wu Y, Kemper KE, Xue A. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Sci Adv*. 2019;5:eaaw3538.

23. Williams JL, Bertrand JK, Misztal I, Łukaszewicz M. Genotype by environment interaction for growth due to altitude in United States Angus cattle. *J Anim Sci*. 2012;90:2152-8.

24. Yang J, Loos RJ, Powell JE, Medland SE, Speliotes EK, Chasman DI, et al. FTO genotype is associated with phenotypic variability of body mass index. *Nature*. 2012;490:267-72.

25. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010;42:565-9.

26. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44:821-4.

Table 1. Summary statistics for growth traits and environmental variables across the US ecoregions.

Traits/Env.Variables	SE	HP	FM	FB	UN	DA	Total
Birth Weight (kg)							
N	1151	3271	2397	3924	2173	511	13,427
Mean	35.5	39.9	38.7	37.1	38.9	37.2	38.2
SD	4.4	4.9	4.8	4.4	4.5	5.2	4.8
Weaning Weight (kg)							
N	873	3283	2329	3025	1883	454	11,847
Mean	303.6	313.0	287.1	296.8	315.7	289.5	302.6
SD	42.2	44.0	38.6	46.2	43.3	42.0	44.6
Yearling Weight (kg)							
N	533	2525	1999	1926	1293	270	8,546
Mean	491.5	551.4	512.0	484.9	542.5	486.0	520.1
SD	78.2	81.1	71.5	89.7	81.5	71.6	85.1
Multivariate							
N	533	2307	1992	1917	1278	267	8294
Elevation (m)							
Mean	170.2	681.8	1587.0	341.6	372.0	914.8	747.5
SD	89.9	215.1	307.5	147.5	65.9	496.7	517.2
Precipitation (ml)							
Mean	1289.8	420.7	390.6	946.0	749.2	497.6	643.8
SD	188.1	118.4	90.3	155.3	111.7	150.3	329.5
Mean Temperature (^oC)							
Mean	17.1	7.2	6.8	12.3	7.3	14.1	9.9
SD	1.6	1.8	1.4	1.3	1.4	3.9	3.8
Min. Temperature (^oC)							
Mean	11.1	0.60	-0.80	6.1	1.8	7.4	3.3
SD	1.8	1.6	1.4	1.2	1.4	4.9	4.0
Max. Temperature (^oC)							
Mean	23.7	14.2	14.5	18.3	13.3	22.8	16.4
SD	1.6	2.1	1.8	1.5	1.5	3.1	3.7
Mean Dew point Temperature (^oC)							
Mean	11.5	0.2	-2.2	6.3	2.5	5.8	3.0
SD	1.7	1.5	1.0	1.2	1.9	5.2	4.3
Min. Vapor Pressure Deficit (hPs)							
Mean	1.0	1.3	1.4	1.1	0.9	1.7	1.2
SD	0.5	0.4	0.3	0.2	0.2	0.6	0.4
Max. Vapor Pressure Deficit (hPs)							
Mean	17.0	13.0	13.7	13.2	10.0	21.0	13.3
SD	1.7	2.1	1.8	1.6	1.1	2.6	2.9

Env = environmental; Max = maximum; Min = Minimum; Ecoregions: SE = Southeast, HP = High Plains, FM = Forested Mountains, FB = Fescue Belt, UN = Upper Midwest & Northeast, DA = Desert & Arid Plains.

Table 2. Proportion of genetic variance based on the additive genomic model (MA), and additive, dominance and epistasis genomic model (MADE) and their standard errors for each trait.

Trait	MA	MADE		
	σ_A^2/σ_P^2	σ_A^2/σ_P^2	σ_D^2/σ_P^2	σ_{Ep}^2/σ_P^2
Birth weight	0.37 ± 0.01	0.33 ± 0.01	0.00 ± 0.01	0.25 ± 0.03
Weaning weight	0.30 ± 0.01	0.28 ± 0.02	0.02 ± 0.02	0.12 ± 0.04
Yearling weight	0.40 ± 0.01	0.40 ± 0.02	0.05 ± 0.02	0.11 ± 0.05

σ_A^2/σ_P^2, σ_D^2/σ_P^2, σ_{Ep}^2/σ_P^2 are proportion of additive, dominance, and epistasis variances to total phenotypic variance, respectively.

Table 3. Number of significant SNPs (*P* < 1e-05) and magnitude of the allele substitution effects by environmental variables/factors for each trait.

Env. Variables/Factor	Birth Weight		Weaning Weight		Yearling weight		Multivariate	Total
	N	ASE	N	ASE	N	ASE	N	N
Elevation	142	-0.001 – 0.001	28	-0.008 – 0.002	15	-0.005 – 0.006	50	216
Precipitation	338	-0.001 – 0.002	4	-0.004 – -0.004	1	-0.015	12	349
Mean Temperature	555	-0.170 – 0.201	1	0.491	4	-1.240 – 1.704	69	611
Minimum Temperature	756	-0.159 – 0.212	6	-0.868 – 0.439	-	-	88	823
Maximum Temperature	408	-0.175 – 0.207	1	0.426	4	-1.267 – 1.848	23	427
Mean Dew point Temperature	941	-0.178 – 0.168	4	-0.516 – 0.409	1	-1.105	94	1018
Minimum Vapor Pressure Deficit	152	-1.399 – 1.831	2	-3.254 – -3.242	1	-6.615	10	165
Maximum Vapor Pressure Deficit	224	-0.213 – 0.262	10	-0.546 – 0.556	1	1.078	17	236
Southeast	54	-2.078 – 1.649	-	-	2	-11.120 – -10.500	7	59
High Plains	40	-1.381 – 1.132	16	-5.353 – 3.930	4	-22.900 – 6.722	23	63
Forested Mountains	90	-3.135 – 1.215	31	-6.785 – 4.737	49	-9.397 – 17.690	148	221
Fescue Belt	47	-2.145 – 0.953	1	-2.943	1	18.530	11	56
Upper Midwest and Northeast	24	-0.709 – 0.872	-	-	1	8.181	10	33
Desert and Arid Prairie	125	-8.243 – 6.043	8	-11.570 – 6.456	6	-16.450 – 23.620	*	125
Total	1923		101		78		353	2319

Env = Environmental; N = number of significant SNPs; ASE = allele substitution effects from GEMMA software (kg).

Table 4. Top 30 enriched gene ontology (GO) terms for candidate genes located at 10 kb and 100 kb sequence window from the significant SNPs (P < 1e-05).

GO Term	GO name	10 kb window		100 kb window	
		N	P-value$_{adj}$	N	P-value$_{adj}$
GO:0050896	Response to stimulus	164	1.9e-11	452	5.7e-33
GO:0006807	Nitrogen compound metabolic process	178	1.4e-08	541	6.7e-28
GO:0019222	Regulation of metabolic process	107	1.5e-09	335	2.2e-27
GO:1901360	Organic cyclic compound metabolic process	85	4.7e-07	294	3.4e-27
GO:0032501	Multicellular organismal process	150	2.3e-07	404	5.2e-27
GO:0034641	Cellular nitrogen compound metabolic process	92	3.2e-08	311	6.7e-27
GO:0044707	Single-multicellular organism process	150	5.0e-07	404	1.3e-26
GO:0051716	Cellular response to stimulus	-	-	326	1.7e-25
GO:0080090	Regulation of primary metabolic process	97	1.4e-07	312	1.8e-25
GO:0048518	Positive regulation of biological process	105	7.6e-10	309	2.9e-25
GO:0060255	Regulation of macromolecule metabolic process	-	-	312	1.0e-24
GO:0031323	Regulation of cellular metabolic process	98	2.3e-07	315	1.4e-24
GO:0046483	Heterocycle metabolic process	81	5.7e-07	277	2.1e-24
GO:0044710	Single-organism metabolic process	210	1.5e-09	618	3.7e-24
GO:0006725	Cellular aromatic compound metabolic process	-	-	280	1.4e-23
GO:0071840	Cellular component organization or biogenesis	119	9.1e-10	280	2.4e-22
GO:0051179	Localization	128	6.6e-10	337	4.4e-22
GO:0019538	Protein metabolic process	118	3.7e-09	334	8.6e-22
GO:0010467	Gene expression	77	5.4e-07	252	8.9e-22
GO:0006139	Nucleobase-containing compound metabolic process	-	-	271	9.7e-22
GO:0032502	Developmental process	-	-	339	1.9e-21
GO:0044767	Single-organism developmental process	-	-	339	6.1e-21

GO:0007154	Cell communication	-	-	328	1.4e-20
GO:0006950	Response to stress	-	-	176	1.7e-20
GO:0016043	Cellular component organization	118	1.8e-08	338	2.1e-20
GO:0007275	Multicellular organismal development	-	-	281	5.3e-20
GO:0051234	Establishment of localization	101	6.6e-10	264	9.5e-20
GO:0048519	Negative regulation of biological process	102	4.0e-08	280	2.1e-19
GO:0048856	Anatomical structure development	-	-	309	3.7e-19
GO:0023052	Signaling	-	-	323	7.8e-19
GO:1902578	Single-organism localization	128	1.2e-09	-	-
GO:0006811	Ion transport	53	2.0e-09	-	-
GO:0006810	Transport	97	3.8e-09	-	-
GO:0044765	Single-organism transport	97	3.2e-08	-	-
GO:0048523	Negative regulation of cellular process	91	3.2e-08	-	-
GO:0007155	Cell adhesion	29	1.8e-07	-	-
GO:0022610	Biological adhesion	30	1.8e-07	-	-
GO:0044267	Cellular protein metabolic process	95	2.1e-07	-	-
GO:0009893	Positive regulation of metabolic process	53	2.3e-07	-	-
GO:0043412	Macromolecule modification	82	2.6e-07	-	-
GO:0030001	Metal ion transport	30	6.8e-07	-	-

N = number of candidate genes; $P\text{-value}_{adj}$ = *P*-value adjusted for FDR of 5%.

Table 5. Top 30 enriched pathways for candidate genes located at 10 kb and 100 kb sequence window from the significant SNPs ($P < 1e-05$).

Pathway entry	Pathway name	10 kb window		100 kb window	
		N	P-value$_{adj}$	N	P-value$_{adj}$
bta01100	Metabolic pathways	36	1.8e-10	82	1.5e-15
bta05032	Morphine addiction	10	2.4e-08	18	7.3e-12
bta04151	PI3K-Akt signaling pathway	11	1.4e-03	32	6.9e-11
bta04723	Retrograde endocannabinoid signaling	10	4.6e-06	18	2.2e-09
bta04510	Focal adhesion	10	1.3e-04	21	2.5e-09
bta05200	Pathways in cancer	-	-	44	5.9e-09
bta04727	GABAergic synapse	9	1.5e-06	15	8.5e-09
bta05033	Nicotine addiction	7	1.5e-06	11	1.2e-08
bta05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	8	3.3e-06	13	4.2e-08
bta04141	Protein processing in endoplasmic reticulum	-	-	18	5.8e-08
bta04014	Ras signaling pathway	9	5.0e-03	23	8.0e-08
bta03013	RNA transport	-	-	20	8.0e-08
bta04010	MAPK signaling pathway	13	5.4e-03	29	2.9e-07
bta04912	GnRH signaling pathway	-	-	12	2.4e-06
bta04012	ErbB signaling pathway	-	-	12	2.9e-06
bta05169	Epstein-Barr virus infection	5	5.8e-03	23	6.1e-06
bta04724	Glutamatergic synapse	7	5.7e-04	13	7.3e-06
bta04725	Cholinergic synapse	6	3.2e-03	13	7.7e-06
bta04921	Oxytocin signaling pathway	-	-	15	8.7e-06
bta05410	Hypertrophic cardiomyopathy (HCM)	7	1.0e-04	11	9.2e-06
bta04726	Serotonergic synapse	7	1.1e-04	13	1.0e-05
bta04664	Fc epsilon RI signaling pathway	-	-	10	1.2e-05

bta04080	Neuroactive ligand-receptor interaction	11	1.3e-03	25	1.5e-05
bta05152	Tuberculosis	-	-	16	2.4e-05
bta05160	Hepatitis C	-	-	10	2.5e-05
bta04390	Hippo signaling pathway	-	-	13	2.9e-05
bta04730	Long-term depression	-	-	10	3.5e-05
bta03030	DNA replication	-	-	6	6.9e-05
bta05161	Hepatitis B	-	-	15	6.9e-05
bta04722	Neurotrophin signaling pathway	-	-	12	6.9e-05
bta05414	Dilated cardiomyopathy	7	1.3e-04	-	-
bta05416	Viral myocarditis	3	2.2e-03	-	-
bta04514	Cell adhesion molecules (CAMs)	5	2.5e-03	-	-
bta04144	Endocytosis	9	2.5e-03	-	-
bta04976	Bile secretion	5	3.1e-03	-	-
bta00770	Pantothenate and CoA biosynthesis	2	3.8e-03	-	-
bta05330	Allograft rejection	2	3.8e-03	-	-
bta04060	Cytokine-cytokine receptor interaction	7	3.9e-03	-	-
bta00785	Lipoic acid metabolism	2	3.9e-03	-	-
bta04728	Dopaminergic synapse	6	4.7e-03	-	-
bta04512	ECM-receptor interaction	2	5.0e-03	-	-
bta05320	Autoimmune thyroid disease	4	5.3e-03	-	-
bta00510	N-Glycan biosynthesis	4	5.8e-03	-	-
bta04713	Circadian entrainment	5	6.7e-03	-	-

N = number of candidate genes; $P\text{-value}_{adj}$ = *P*-value adjusted for FDR of 5%.

Table 6. Significant vQTL (*P* < 1e-5) for growth traits when using residuals adjusted for additive effects (MA) or for additive, dominance and epistasis effects (MADE).

Trait	SNP	BTA	Position	Freq	MA		MADE		Genes
					Effect (kg)	*P*-value	Effect (kg)	*P*-value	
BW	rs110323698	10	22,312,496	0.127	0.026 ± 1.6e-04	7.07e-06	-	-	*DAD1, ABHD4*
	rs137529355	13	35,496,312	0.278	-	-	-0.020 ± 9.3e-05	6.04e-06	*LYZL1*
	rs133623280	14	17,005,811	0.458	-0.019 ± 7.9e-05	2.94e-06	-0.018 ± 7.9e-05	6.69e-06	*ZHX2*
	rs109136768	14	17,007,761	0.418	0.018 ± 8.6e-05	9.60e-06	-	-	
	rs43026996	14	17,166,118	0.344	0.022 ± 1.1e-04	3.51e-06	0.021 ± 1.1e-04	6.67e-06	
	rs43186199	14	17,183,088	0.347	0.022 ± 1.1e-04	3.10e-06	0.021 ± 1.1e-04		
	rs42796317	14	17,195,903	0.356	0.021 ± 1.1e-04	9.31e-06	-	-	
	rs109515648	14	23,128,784	0.376	0.024 ± 1.4e-04	6.04e-06	-	-	*TMEM68, TGS1, LYN, RPS20, MOS, PLAG1, CHCHD7, SDR16C5*
	rs137303549	14	23,298,062	0.388	0.024 ± 1.4e-04	6.48e-06	-	-	
	rs134215421	14	23,329,375	0.385	0.025 ± 1.4e-04	2.49e-06	-	-	
	rs109815800	14	23,338,890	0.384	0.024 ± 1.4e-04	5.38e-06	-	-	
	rs41843599	17	57,685,345	0.144	-	-	0.026 ± 1.7e-04	6.42e-06	*NOS1, KSR2, FBXO21, TESC, FBXW8*
	rs41843601	17	57,685,935	0.146	0.026 ± 1.6e-04	6.60e-06	0.026 ± 1.6e-04	5.62e-06	
	rs109209365	17	57,699,496	0.141	-	-	0.026 ± 1.7e-04	8.79e-06	
	rs41847363	17	57,975,119	0.146	0.026 ± 1.6e-04	7.91e-06	0.026 ± 1.6e-04	9.43e-06	
	rs134957935	21	51,465,082	0.188	-0.022 ± 1.2e-04	9.21e-06	-	-	*LRFN5*
WW	rs109967828	3	106,206,442	0.169	0.025 ± 1.6e-04	9.18e-06	-	-	*MFSD2A, MYCL, TRIT1, BMP8B*
	rs43657596	11	7,293,721	0.133	0.027 ± 1.8e-04	7.17e-06	-	-	*IL18RAP, SLC9A4, SLC9A2*
	rs134174267	11	7,304,992	0.101	0.035 ± 2.4e-04	5.08e-07	0.033 ± 2.4e-04	1.78e-06	
	rs109287854	18	37,372,422	0.219	0.022 ± 1.2e-04	8.65e-06	-	-	-
	rs110803856	27	40,193,292	0.209	0.023 ± 1.3e-04	8.88e-06	-	-	*TOP2B, RARB*

YW	rs135719485[£]	5	115,845,266	0.336	-0.025 ± 1.3e-04	1.93e-06	-0.026 ± 1.3e-04	4.19e-07	*RIBC2, FBLN1, SMC1B*
	rs133130907[¥]	14	74,597,382	0.092	0.044 ± 3.1e-04	7.42e-08	0.044 ± 3.2e-04	4.08e-08	*MMP16*
	rs41802941	16	36,184,280	0.037	0.057 ± 7.7e-04	6.73e-06	-	-	*XCL2, XCL1, DPT*
	rs41802957	16	36,200,777	0.050	0.050 ± 5.8e-04	4.27e-06	0.050 ± 5.8e-04	4.86e-06	
	rs137574197	16	36,231,107	0.050	0.051 ± 5.9e-04	3.36e-06	0.051 ± 5.9e-04	4.02e-06	
	rs109835794	16	36,242,339	0.052	0.050 ± 5.6e-04	3.09e-06	0.050 ± 5.7e-04	3.46e-06	
	rs134575841	17	44,811,665	0.432	-	-	0.022 ± 1.2e-04	8.59e-06	*GALNT9*
	rs111011858	17	70,514,528	0.010	0.103 ± 2.5e-03	5.18e-06	0.107 ± 2.5e-03	2.53e-06	*YWHAH, SLC5A1, DEPDC5, SLC5A4*
	rs110473197	17	70,523,020	0.010	0.103 ± 2.5e-03	5.08e-06	0.107 ± 2.5e-03	2.50e-06	
	rs109010785	17	70,527,754	0.010	0.103 ± 2.5e-03	5.08e-06	0.107 ± 2.5e-03	2.50e-06	
	rs108982594	17	70,540,863	0.010	0.103 ± 2.5e-03	5.08e-06	0.107 ± 2.5e-03	2.50e-06	
	rs378352090	18	62,939,763	0.149	-	-	0.034 ± 2.8e-04	7.85e-06	*TTYH1, LENG8, LENG9, LAIR1, CDC42EP5, RPS9, TSEN34, MBOAT7, TMC4, LENG1, CNOT3, PRPF31, TFPT, NDUFA3, OSCAR, TARM1*
	rs383417308	18	63,229,878	0.155	-	-	0.034 ± 2.8e-04	7.42e-06	
	rs433152283	18	63,230,047	0.155	-	-	0.034 ± 2.8e-04	6.22e-06	
	rs43727853	27	33,796,650	0.460	-0.023 ± 1.3e-04	7.00e-06	-	-	*TACC1*
	rs110205571	27	34,027,900	0.281	0.024 ± 1.4e-04	5.33e-06	-	-	*PLEKHA2*
	rs110909628	27	35,715,673	0.223	-	-	0.026 ± 1.6e-04	8.82e-06	*ZMAT4*
	rs137312252[*]	27	35,899,682	0.097	0.041 ± 3.2e-04	3.42e-07	0.042 ± 3.2e-04	2.68e-07	*ZMAT4*
	rs135719206[*]	27	35,902,981	0.092	0.042 ± 3.4e-04	6.24e-07	0.042 ± 3.4e-04	4.34e-07	
	rs134473626[*]	27	35,906,532	0.095	0.042 ± 3.3e-04	2.64e-07	0.042 ± 3.3e-04	2.51e-07	
	rs136265144	29	48,072,601	0.254	-0.027 ± 1.5e-04	1.90e-06	-0.027 ± 1.5e-04	1.40e-06	*SHANK2*
MV	rs135719485[£]	5	115,845,266	0.334	-	-	-	8.47e-06	*RIBC2, FBLN1, SMC1B*
	rs109142386	8	4,822,842	0.258	-	-	-	7.78e-06	*GALNTL6*
	rs133130907[¥]	14	74,597,382	0.091	-	9.98e-07	-	8.47e-06	*MMP16*
	rs137312252[*]	27	35,899,682	0.098	-	2.04e-06	-	3.77e-07	*ZMAT4*

rs135719206*	27	35,902,981	0.092	-	1.06e-06	-	3.45e-06	
rs134473626*	27	35,906,532	0.096	-	1.81e-06	-	3.10e-06	
rs135379559	29	39,441,489	0.099	-	2.83e-06	-	-	*PAG6, PAG11, LOC528815*

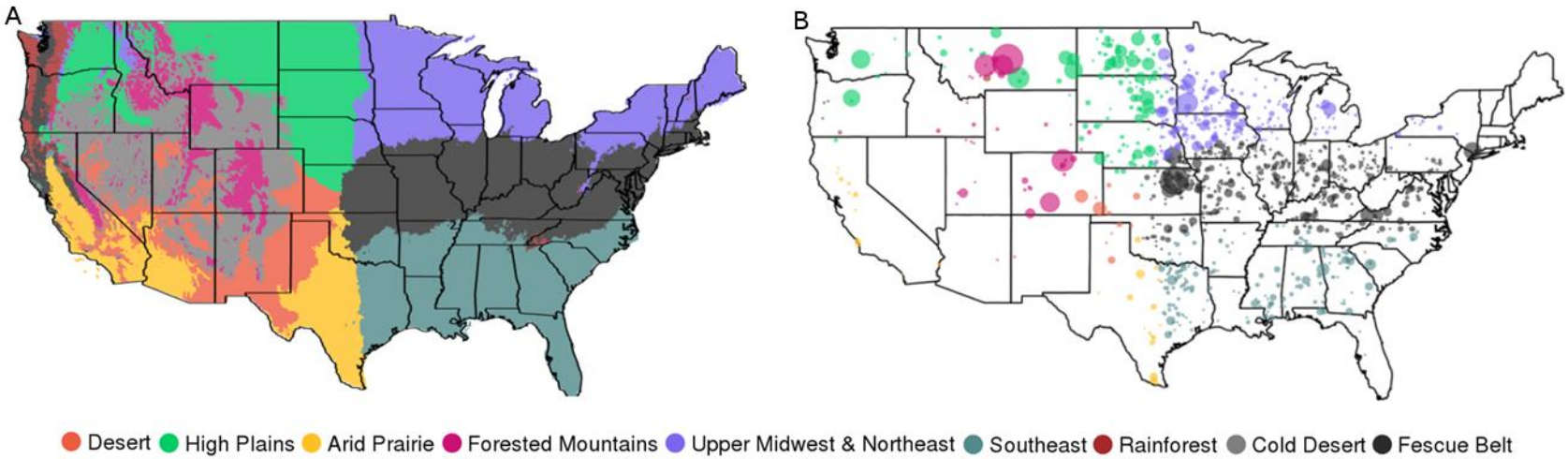

Figure 1. Boundaries for ecoregion assignments in the United States. (*A*) United States partitioned into nine ecoregions based on similar topographic and environmental conditions. (*B*) Location of beef farms for which data was retrieved.

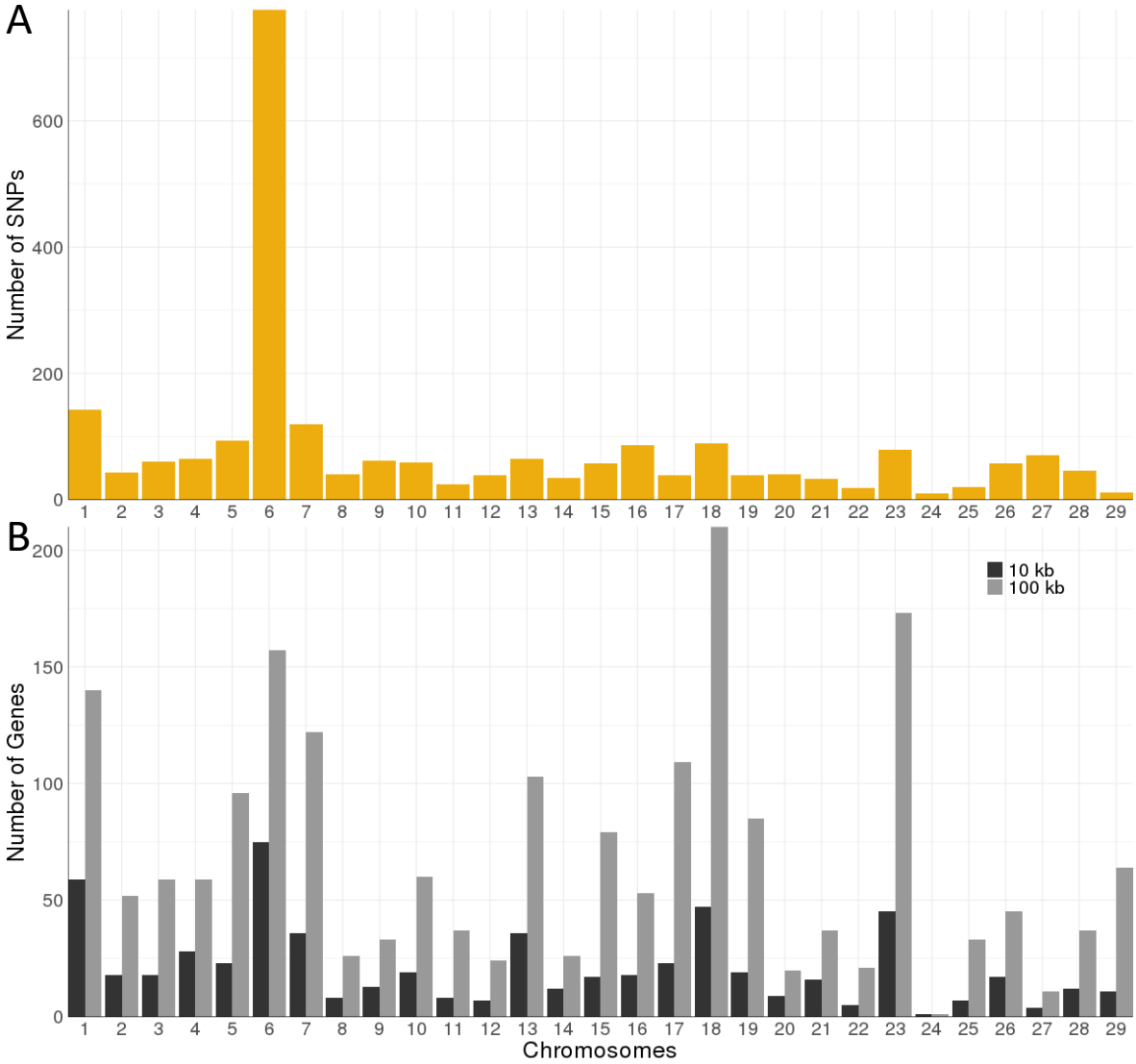

Figure 2. Distribution of genotype-by-environment genome-wide association results by chromosomes. (*A*) Number of significant SNP (P < 1e-5) enriched with genotype-by-environment effects. (*B*) Number of candidate genes located less than 10 kb or 100 kb from the significant SNPs.

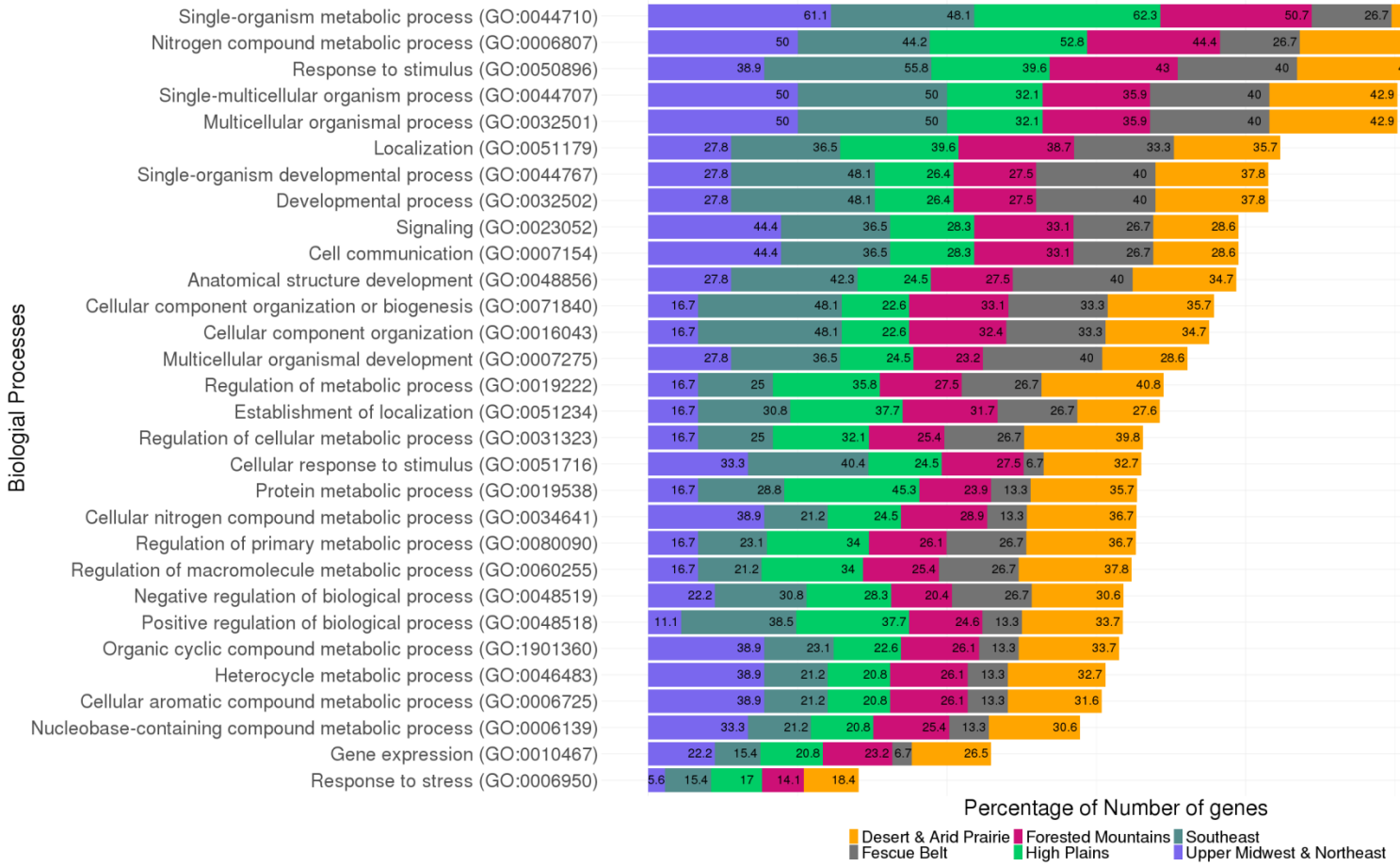

Figure 3. Percentage of candidate genes with genotype-by-environment effects for each ecoregion that participate in the enriched gene ontology (GO) terms. Genes were located 100 kb from significant SNPs (P < 1e-5).

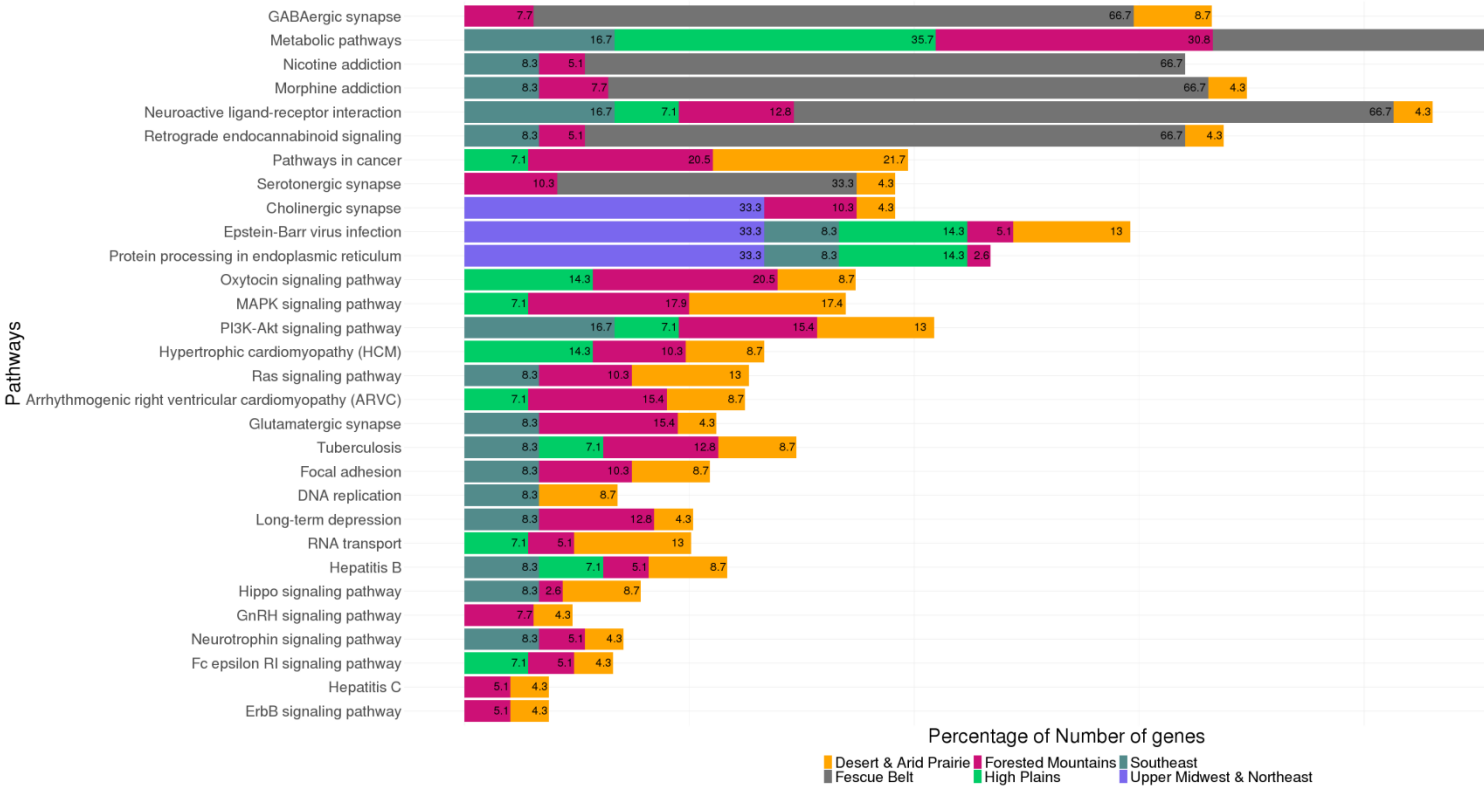

Figure 4. Percentage of candidate genes with genotype-by-environment effects for each ecoregion that participate in the enriched biological pathways. Genes were located 100 kb from significant SNPs (P < 1e-5).