GENOMICS OF SEASONAL HAIR SHEDDING AND ECOREGION-SPECIFIC

GROWTH TO IDENTIFY ENVIRONMENTALLY-ADAPTED BEEF CATTLE

_____

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

_____

In Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

_____

by

HARLY JANE DURBIN

Dr. Jared Decker, Dissertation Supervisor

December 2020

APPROVAL PAGE

The undersigned, appointed by the Dean of the Graduate School, have

examined the dissertation entitled:

GENOMICS OF SEASONAL HAIR SHEDDING AND ECO-REGION SPECIFIC

GROWTH TO IDENTIFY ENVIRONMENTALLY-ADAPTED BEEF CATTLE

Presented by Harly Jane Durbin, a candidate for the degree of Doctor of

Philosophy, and hereby certify that in their opinion it is worthy of acceptance.

_____

Dr. Jared E. Decker, Animal Sciences, UMC

_____

Dr. Robert D. Schnabel, Animal Sciences, UMC

_____

Dr. Jeremy F. Taylor, Animal Sciences, UMC

_____

Dr. Elizabeth G. King, Biological Sciences, UMC

_____

Dr. Stephen P. Miller, Angus Genetics, Inc.

DEDICATION

Momma – thank you for teaching me compassion and also how to approach life with a sense of humor. You do such an incredible amount for other people and I strive to be as selfless every day. Daddy – thank you for encouraging me to think independently and for teaching me how to stand up for myself. Thank you also for the years of FFA (also known sometimes as Father Feeds Animal) and getting up early to haul heifers to jackpot shows. Darby – thank you for holding me to a high standard. I often think there was some mix-up and that you were meant to be the older sister for how much you hold me up. Mimi and Pawpaw – thank you for your unconditional support and love. I cherish the time we spend together and the impact it's had on making me who I am, from "Pooh stories" with Mimi to going out to feed cows with Pawpaw to weekend of getting spoiled (and loaded up on canned food) during college. Kitty – thank you for encouraging my curiosity. I think of all people that helped raised me, you've had the biggest impact on how I see the world and on my personality (I'll leave it up to everyone else to decide if that's a good or a bad thing). I love you all!

Lauren Ciernia – I truly treasure our friendship and cannot state how much important your support has been in keeping me afloat, especially during the first part of grad school. I'm so lucky to have such a great hype-man and I can't wait to know you for the rest of our lives.

Dr. Troy Rowan – I truly don't know how I would have done it without you. You're my best friend and favorite co-author, but hopefully you already know that. Thank

you for keeping me grounded but encouraging me to think outside the box… I

love you so much and can't wait for our future collaborations.

ACKNOWLEDGEMENTS

First and foremost I have my advisor, Dr. Jared Decker, to thank for getting me to get to this point. I appreciate most of all your constant willingness and patience to do things in the way that works best for the way my brain works. Thank you for helping me to become an independent thinker and researcher!

Thank you also to Dr. Bob Schnabel for among many other things, changing the way that I approach problems (and endowing me with a healthy amount of skepticism!). I really appreciate you advocating for my success and your encouragement over the last 4.5 years. I've been so lucky to know that I can come to you with a question and that you will always take the time to give thoughtful guidance.

Interning at the American Angus Association was among the best decisions I made during my Ph.D., almost entirely because of Dr. Steve Miller's mentorship. Steve, thank you for always letting me work through (sometimes hare-brained) ideas while still guiding me in the right direction. You had such a big impact on my self-confidence and how I see myself as a scientist and I'm very grateful. Thank you to everyone else who made my time at AGI so great, including Kelli Retallick, Dan Moser, Duc Lu, Lou-Ann Adams, Jason Kenyon, and the Lunch Bunch.

A huge thanks to Helen Yampara-Iquise, without whom I would still be wading through sample cards and trying to update metadata information. Your help and

you to the undergraduate student workers and interns who helped with organization and accession of DNA samples from 2016-2019:

- Natalie Barr

- Sarah Van Ausdal

- Brian Arisman

- Aaron Mott

  You all really took a load off of my shoulders! Thank you also to the MU Extension and research farm teams that took the time to help out with the hair shedding project, be it via scoring, advice, or otherwise:

- Eldon Cole, the original master of hair shedding

- MU Thompson Research Center, including Jon Schreffler, Stoney, Charles, and Jacob

- MU Southwest Research Center, including David Cope, Steve Stamate, and Matt Massie

- Erin Larimore

- David Mallory

  I have so many other mentors and allies to thank for my success, including but not limited to Dr. Jerry Taylor, Dr. Libby King, Dr. Jim Amos-Landgraf, Becki

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SUPPLEMENTARY FIGURES

LIST OF SUPPLEMENTARY TABLES

GENOMICS OF SEASONAL HAIR SHEDDING AND ECOREGION-SPECIFIC

GROWTH TO IDENTIFY ENVIRONMENTALLY-ADAPTED BEEF CATTLE

Harly Jane Durbin

Dr. Jared Decker, Dissertation Supervisor

ABSTRACT

Recently, interest has increased in augmenting current national scale cattle evaluations with precision genetic predictions tailored to specific environmental conditions. Some efforts to develop environmentally-aware predictions have focused on the use of novel phenotypes and others on the incorporation of genotype-by-environment interactions (GxE) to existing methodologies. Cattle and other mammal species molt thick winter coats at the beginning of summer in order to prepare for the oncoming stress of warmer weather. In warm climates, cattle that shed their winter coat earlier and more completely have an adaptive advantage over later-shedding herd-mates, and previous work has demonstrated the relationship between seasonal coat shedding and production traits. Using a novel trait (early summer hair shedding score) we develop a genetic evaluation for heat tolerance. We find that hair shedding score is moderately heritable and controlled by genomic loci involved in light sensing and metabolism. Additionally, we explore the degree to which GxE interactions across discrete ecoregions affect pre-weaning growth in American Angus cattle. We find evidence for GxE in the maternal but not direct effect of weaning weight, particularly in heat-stressed environments. Together, these efforts will help beef cattle breeders match genetics to the environmental conditions in which they are best suited.

CHAPTER 1

DEVELOPMENT OF A GENETIC EVALUATION FOR HAIR SHEDDING IN AMERICAN ANGUS CATTLE TO IMPROVE THERMOTOLERANCE

Abstract

Background

Heat stress and fescue toxicosis caused by ingesting tall fescue infected with the endophytic fungus *Epichloë coenophiala* represent two of the most prevalent stressors to beef cattle in the United States and cost the beef industry millions of dollars each year. The rate at which a beef cow sheds her winter coat early in the summer is an indicator of adaptation to heat and an economically relevant trait in temperate or subtropical parts of the world. Furthermore, research suggests that early-summer hair shedding may reflect tolerance to fescue toxicosis, since vasoconstriction induced by fescue toxicosis limits the ability of an animal to shed its winter coat. Both heat stress and fescue toxicosis reduce profitability partly via indirect maternal effects on calf weaning weight. Here, we developed parameters for routine genetic evaluation of hair shedding score in American Angus cattle, and identified genomic loci associated with variation in hair shedding score via genome-wide association analysis (GWAA).

Results

Hair shedding score was moderately heritable (h^2 = 0.34 to 0.40), with different repeatability estimates between cattle grazing versus not grazing endophyte-infected tall fescue. Our results suggest modestly negative genetic and phenotypic correlations between a dam's hair shedding score (lower score is earlier shedding) and the weaning weight of her calf, which is one metric of performance. Together, these results indicate that economic gains can be made by using hair shedding score breeding values to select for heat-tolerant cattle. GWAA identified 176 variants significant at FDR < 0.05. Functional enrichment analyses using genes that were located within 50 kb of these variants identified pathways involved in keratin formation, prolactin signaling, host-virus interaction, and other biological processes.

Conclusions

This work contributes to a continuing trend in the development of genetic evaluations for environmental adaptation. Our results will aid beef cattle producers in selecting more sustainable and climate-adapted cattle, as well as enable the development of similar routine genetic evaluations in other breeds.

Background

At the beginning of the summer, many mammalian species molt thick winter coats in response to changing day length in order to prepare for warmer temperatures [1–6]. There is evidence of quantitative variation in the rate and timing of this yearly shedding across taxa [7,8], including cattle [9]. In warm

climates, cattle that shed their winter coat earlier and more completely have an adaptive advantage over later-shedding herd-mates. Late-shedding cattle will need to partition energy that could have gone towards growth and production towards overcoming heat stress [10]. Economic losses attributable to heat stress cost the U.S. beef cattle industry more than $360 million each year in 2003 [11], which equates to ~ $518 million in 2020 after adjustment for inflation. In the cow-calf sector, a portion of this economic impact is a result of lowered calf weaning weights caused by reduced dam productivity [12]. However, there is currently no national-scale genetic evaluation for heat tolerance. In the United States, much of the beef herd that is at risk of heat stress is also at risk for fescue toxicosis. Tall fescue (*Lolium arundinaceum*) is the most widely available forage in the United States [13], thanks in part to its symbiotic relationship with the endophytic fungus *Epichloë coenophiala*. *E coenophiala* produces ergot alkaloids that benefit the forage by increasing drought tolerance and pathogen resistance [14], but negatively impact livestock to varying degrees. In cattle, one side-effect of fescue toxicosis is peripheral vasoconstriction, which reduces the animal's ability to dissipate heat. The ergot alkaloids that cause fescue toxicosis also disrupt the hair follicle growth cycle, which interferes with hair coat shedding and, in turn, further increases the potential for heat stress [15]. Therefore, effective early-summer hair shedding while grazing endophyte-infected (hereafter referred to as "toxic") tall fescue may also be an indicator of tolerance to fescue toxicosis. One way to mitigate heat stress is through introgression of beneficial alleles from tropically-adapted breeds [16]. However, this can take many generations and

may come at the cost of other production traits. An alternative strategy is the exploitation of standing genetic variation in the population of interest. Recently, interest has grown in augmenting national genetic evaluations with predictions of regional adaptability and suitability [17–19], particularly by using novel traits [20]. Here, we develop parameters for a prototype national genetic evaluation of hair shedding in American Angus cattle, a novel trait that directly influences cattle's ability to dissipate heat. To assess one potential impact of such an evaluation on beef cattle producers, we also demonstrate the relationship between dam hair shedding score and the weaning weight of her calf. This evaluation will aid beef cattle producers in heat-stressed regions in the selection of more sustainable cattle.

Methods

Data

All data originated from purebred cattle registered in the American Angus Association (AAA) and commercial cattle enrolled in the AAA Breed Improvement Record program. Phenotypic data comprised hair shedding scores recorded by beef cattle producers enrolled in the Mizzou Hair Shedding Project (MU data) between 2016 and 2019 in combination with hair shedding scores collected by technicians in 2011, 2012, 2018, and 2019 as part of Angus Foundation-funded projects at Mississippi State University and North Carolina State University (AGI data). Across all years and both datasets, scores were recorded on one day between April 17th and June 30th in the late spring or early summer, with most

4

scores recorded in mid- to late-May. Hair shedding was evaluated using a 1 to 5 visual appraisal scale, where 1 was 0% dead winter coat remaining and 5 was 100% winter coat remaining based on the systems developed by Turner and Schleger [21] and Gray et al. [22] (Fig. 1). While there is variation in the onset of hair shedding between individuals, cattle and other mammals tend to shed from the head towards the tail and from the topline towards the legs [2,8,23].

Records were removed when the breeder-reported sex of an animal did not match the sex recorded in the AAA pedigree. Hair shedding scores that originated from male animals comprised less than 1% of the dataset and only female records were retained. Age classifications were assigned to each record based on age in days determined by the AAA-recorded birth date and the date the hair shedding score was recorded. Similar to the system used in the Beef Improvement Federation (BIF) Guidelines for age-of-dam classification [24], age classifications were defined as to , where is the age classification and is days. Records where the breeder-reported age in years differed from the calculated age classification by more than two years and records from animals younger than 275 days-of-age were removed. When no calving season was reported, it was imputed using the most recent natural birth calving date available in the AAA database prior to the recorded score. When no natural birth calving dates were available, calving season was imputed using the animal's own birth date. In the AGI data, some animals were scored by multiple scoring technicians on the same day. In these cases, phenotypes recorded on the same animal and the same day were averaged. In the MU data, participating producers were asked to

5

report whether or not (yes or no) animals grazed toxic fescue during the spring of the recording year. Grazing status was not explicitly recorded in the AGI data, but animals scored in Texas were assumed not to have grazed toxic fescue. This resulted in 14,839 scores in the combined, filtered dataset. Among the 8,619 individuals included, 49% had between 2 and 6 years of data. Most data came from herds in the Southeast and Fescue Belt (Fig. 2). The mean hair shedding score was slightly higher in the AGI data (μ = 3.10; n = 6,374) compared to the MU data (μ = 2.65; n = 8,465), but the standard deviation was identical in both datasets (σ = 1.15).

Genotypes and imputation

Genotypes for 3,898 of the 8,619 animals were imputed to a union marker set (= 233,246) of the GGP-F250 genotyping chip and various commercial assays using FImpute v.3.0 [25]. The commercial assays were those used in routine genotyping of Angus cattle for genomic selection purposes, which include ~ 50 K markers or a lower density panel that can be imputed to ~ 50 K with sufficient accuracy. Although FImpute provides the capacity to infer the genotypes of un-genotyped animals based on information from relatives, markers were imputed only for genotyped individuals. Prior to imputation, markers with a GenCall score lower than 0.15 were set to missing and individuals with Mendelian error rates higher than 2% had their parents set to missing in the pedigree. The GGP-F250 was designed to genotype functional variants and thus has more variants at low minor allele frequencies [26]. Therefore, no minor allele frequency filter was

applied during or after imputation beyond the removal of monomorphic SNPs.

Animals and markers with call rates lower than 85% were removed. The resulting

marker set consisted of 174,894 autosomal variants.

Construction of the blended relationship matrix H^{-1}

In single-step genomic best linear unbiased prediction (BLUP) as used in the

AAA National Cattle Evaluation (NCE), relationships between individuals are

represented in the matrix **H⁻¹** which is a blended form of the genomic and additive

relationship matrices [27], allowing information from both genotyped and non-

genotyped animals to be used. **H⁻¹** calculated as:

$$\mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where **A⁻¹** represents the inverted pedigree relationship matrix traditionally used

to represent relationships, **A⁻¹₂₂** represents the inverted pedigree relationship

matrix for the subset of animals with genotypes available, and **G⁻¹_w** is the inverted

genomic relationship matrix. The genomic relationship matrix was calculated

using the VanRaden method [28] and was blended with **A₂₂** using the default

weight of 0.05 using the preGSf90 program [29]. In all subsequent models

including a random genetic effect, **H⁻¹** was constructed using the 3-generation

pedigree (in total, 17,652 animals; 1,987 distinct sires and 9,509 distinct dams) in

combination with imputed genotypes.

Effect of age on hair shedding score and contemporary group definition

Understanding how and which environmental factors shape phenotypic variation enables the development of more appropriate contemporary group definitions during genetic evaluation. In order to quantify the effect of animal age on hair shedding score, we fitted age as a categorical fixed effect in a repeated records animal model. Age categories were defined in three ways. First, age in years was fit (i.e. all possible values between 1 and 16). Second, ages were grouped as 1, 2, 3, or other ("four-class model"). Third, age groups were defined according to the guidelines set forth by the BIF for age-of-dam effects on birth weight and weaning weight (i.e., 2, 3, 4, 5-9, 10, 11, 12, 13+; [24]) plus yearlings ("BIF model"). The four-class model and the BIF model were each compared against a null model with no age effect included using Akaike's Information Criterion (AIC) and likelihood ratio tests. In all three models with age classification fitted as a categorical fixed effect, classifications with fewer than five animals were excluded. These models are summarized below:

$$\mathbf{y} = \mathbf{X_1 b} + \mathbf{X_2 a} + \mathbf{Z_1 u} + \mathbf{Z_2 p} + \mathbf{e},$$

where $\mathbf{y}$ is a vector of hair shedding scores; $\mathbf{b}$ is a vector of contemporary group effects for each hair shedding score, with contemporary group defined as farm ID, year scored, calving season, score group, and toxic fescue grazing status; $\mathbf{a}$ is a vector of age classification effects for each individual (based on age-in-years, BIF classifications, or the four age classes); $\mathbf{u}$ is the random additive genetic effect with $\mathbf{u} \sim N(0, \mathbf{H}\sigma_a^2)$; $\mathbf{p}$ is the random permanent environment effect with

$\mathbf{p} \sim \mathrm{N}(0, \mathbf{I}\sigma_{\mathrm{pe}}^2);$ is the random residual with $\mathbf{e} \sim \mathrm{N}(0, \mathbf{I}\sigma_{\mathrm{pe}}^2);$ and $\mathbf{X_1}$, $\mathbf{X_2}$, $\mathbf{Z_1}$, and $\mathbf{Z_2}$ are incidence matrices relating the elements of $\mathbf{y}$ to $\mathbf{b}$, $\mathbf{a}$, $\mathbf{u}$, and $\mathbf{p}$, respectively.

Effect of toxic fescue grazing status on hair shedding

Cattle reared in heat-stressed regions but not exposed to endophyte-infected fescue demonstrate similar benefits from early summer hair shedding, but it is unclear if the biological mechanisms that govern hair shedding under fescue toxicosis and heat stress alone are the same. This could have implications for routine genetic evaluation, as it might require that some hair shedding score observations be treated as a separate trait. In order to clarify the relationship between hair shedding score while grazing toxic fescue versus while not grazing toxic fescue, we calculated the covariance and genetic correlation between hair shedding score grazing toxic fescue and not grazing toxic fescue using the bivariate repeated records animal model below:

$$\mathbf{y_t} = \mathbf{X_t b_t} + \mathbf{Z_{1t} u_t} + \mathbf{Z_{2t} p_t} + \mathbf{e_t},$$

where $\mathbf{y}$ is a vector hair shedding scores and $\mathbf{t}$ is toxic fescue grazing status (yes or no); $\mathbf{b}$ is a vector of contemporary group effects for each hair shedding score, with contemporary groups defined as farm ID, year scored, calving season, score group, and age class (yearling, 2-year-old, 3-year-old, or other; based on the results of the age classification analyses above); $\mathbf{u}$ is the additive genetic effect

and $\mathrm{Var}(\mathbf{u}) = \begin{bmatrix} \sigma_{\mathrm{uYes}}^2 & \sigma_{\mathrm{uYes,uNo}} \\ \sigma_{\mathrm{uNo,uYes}} & \sigma_{\mathrm{uNo}}^2 \end{bmatrix} \otimes \mathbf{H};$ $\mathbf{p}$ is the permanent environment effect

and $\mathrm{Var}(\mathbf{p}) = \begin{bmatrix} \sigma^2_{pYes} & 0 \\ 0 & \sigma^2_{pNo} \end{bmatrix} \otimes \mathbf{I}$; $\mathbf{e}$ is the random residual and $\mathrm{Var}(\mathbf{e}) =$

$\begin{bmatrix} \sigma^2_{eYes} & \sigma_{eYes,eNo} \\ \sigma_{eNo,eYes} & \sigma^2_{eNo} \end{bmatrix} \otimes \mathbf{I}$; and $\mathbf{X}$, $\mathbf{Z_1}$, and $\mathbf{Z_2}$ are incidence matrices relating the

elements of $\mathbf{y}$ to $\mathbf{b}$, $\mathbf{u}$ and $\mathbf{p}$, respectively.

In addition, we fitted a univariate model with toxic fescue grazing status included

as a categorical fixed effect. The goal of this model was to quantify the effect of

reported toxic fescue grazing status on hair shedding score:

$$\mathbf{y} = \mathbf{X_1 b} + \mathbf{X_2 f} + \mathbf{Z_1 u} + \mathbf{Z_2 p} + \mathbf{e},$$

where $\mathbf{y}$ is a vector of hair shedding scores; $\mathbf{b}$ is a vector of contemporary group

effects, defined in the same way as the univariate model above; $\mathbf{f}$ is the toxic

fescue status effect (yes or no); $\mathbf{u}$ is the additive genetic effect with

$\mathbf{u} \sim N(0, \mathbf{H}\sigma^2_a)$; $\mathbf{p}$ is the permanent environment effect with $\mathbf{p} \sim N(0, \mathbf{I}\sigma^2_{pe})$; $\mathbf{e}$ is the

random residual with $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2_e)$; and $\mathbf{X_1}$, $\mathbf{X_2}$, $\mathbf{Z_1}$, $\mathbf{Z_2}$, are incidence matrices

relating the elements of $\mathbf{y}$ to $\mathbf{b}$, $\mathbf{f}$, $\mathbf{u}$, and $\mathbf{p}$ respectively.

In both models, only females with known toxic fescue grazing status were

retained for analysis. Contemporary groups with fewer than five animals or no

variation were discarded, resulting in 5,832 observations from cattle grazing toxic

fescue and 4,197 observations from cattle not grazing toxic fescue. Three

hundred ninety-six animals had observations over multiple years both grazing

and not grazing toxic fescue.

Genetic parameters, breeding values, and estimated bias

Variance components, heritability, repeatability, and breeding values were estimated using the univariate repeated records animal model below implemented in AIREMLF90 [29].

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1 u} + \mathbf{Z_2 p} + \mathbf{e},$$

where $\mathbf{y}$ is a vector of hair shedding scores; $\mathbf{b}$ is the contemporary group effect; $\mathbf{u}$ is the additive genetic effect with $\mathbf{u} \sim N(0, \mathbf{H}\sigma_a^2)$; $\mathbf{p}$ is the permanent environment effect with $\mathbf{p} \sim N(0, \mathbf{I}\sigma_{pe}^2)$; $\mathbf{e}$ is the random residual with $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$; and $\mathbf{X}$, $\mathbf{Z_1}$, and $\mathbf{Z_2}$ are incidence matrices relating the elements of $\mathbf{y}$ to $\mathbf{b}$, $\mathbf{u}$, and $\mathbf{p}$, respectively.

The definition of contemporary groups used in this final prediction model was informed by the results of the age classification and toxic fescue grazing status analyses above. It included a combination of farm, year scored, calving season (spring or fall), toxic fescue grazing status (yes or no), age group (yearling, 2-year-old, 3-year-old, or other), and score group. In herds where cattle were scored for hair shedding over more than one day, the score group was determined using a 7-day sliding window to maximize the number of animals per contemporary group. In the future, it will be recommended that producers score all cattle for hair shedding within a week of one another to maximize the size of contemporary groups. Although yearling heifers have not yet experienced the stress of pregnancy, calving season/birth season is a good proxy for

management group in the absence of breeder-reported codes. Therefore, "calving season" was included in the contemporary group definition for all animals regardless of reproductive status. Contemporary groups with fewer than five animals or no variation were dropped. This resulted in 14,438 total scores from 8,449 animals in 395 contemporary groups.

In order to evaluate model bias, we estimated breeding values in ten separate iterations, excluding all phenotypes for a randomly selected 25% of animals. These "partial" breeding values were then compared to breeding values obtained via the "whole" model including all possible information using the "LR method" parameters proposed by Legarra and Reverter [30]. First, we calculated the absolute difference between whole breeding values and partial breeding values for the validation set, or animals whose phenotypes were excluded ($d^v_{w,p}$) and the reference set, or animals whose phenotypes were not excluded ($d^r_{w,p}$). The expectation of this value is zero in the absence of bias, where bias is introduced by incorrect estimation of the genetic trend. Next, we regressed whole breeding values on partial breeding values for both validation ($b^v_{w,p}$) and reference ($b^r_{w,p}$) sets. In this model, deviations of the slope from 1 are suggestive of dispersion. Finally, we calculated the correlation between partial and whole breeding values ($\rho_{p,w} = \frac{cov(\hat{\mu}_p, \hat{\mu}_w)}{\sqrt{var(\hat{\mu}_p, \hat{\mu}_w)}}$) within the validation and reference sets, where the correlation within the validation set ($\rho^v_{p,w}$) is a metric of prediction accuracy.

Weaning weight

The effects of heat stress on pre-weaning growth are well characterized in cattle. Heat stress impacts calf performance most severely via reduced milk production in the dam [12]. Fescue toxicosis induces reduced milk production in a similar fashion [31]. Therefore, we quantified the phenotypic and genetic correlations between hair shedding score and weaning weight.

Weaning weight phenotypes and contemporary group designations came from the weekly growth run of the AAA national cattle evaluation (NCE). Prior to entering the NCE, phenotypes were adjusted for age-of-dam effects as used in Angus's weekly NCE and to 205 days-of-age. Weaning weight data were retrieved for: (1) own weaning weight of cows with at least one hair shedding score recorded, (2) all of cow's recorded calves, (3) cow's weaning weight contemporary group peers, and (4) all of their recorded calves' weaning weight contemporary group peers. Weaning weights from animals born via embryo transfer and contemporary groups with fewer than five animals or no variation were excluded, resulting in 40,794 total weaning weight and 14,039 total hair shedding score records. Of the 45,420 phenotyped animals retained for analysis, 3,850 had both a recorded weaning weight and at least one hair shedding score. Furthermore, 6,448 dams had both hair shedding scores and calf weights recorded in at least one year (n = 9,092 score/weight pairs).

Conceivably, environmental factors that affect a dam's hair shedding performance could also affect the direct weaning weight of her calf and her

maternal effect on the calf's growth, creating a residual covariance between the two traits. In order to reflect this covariance, a bivariate model was fitted in which a direct hair shedding score effect was modeled for the cow, a direct weaning weight effect was modeled for the calf, and a maternal weaning weight effect was modeled for the cow. In practice, this model was implemented by fitting a maternal genetic effect for hair shedding, no direct genetic effect of hair shedding (no genetic effect of the calf on the hair shedding score of its dam), and direct and maternal genetic effects for weaning weight. This model created a direct tie between a dam's hair shedding score and the calf she weaned that year, which reflects more accurately the relationship of interest and is similar to models used to assess the correlations between weaning weight and actual milk yield [32]. For cows with a hair shedding score but no calf weaning weight reported during the scoring year, a "dummy calf" with a weaning weight set to missing and unknown sire was created. This model was fitted three separate times: once including only dams explicitly recorded to have been grazing toxic fescue, once including only dams explicitly recorded to have not been grazing toxic fescue, and once with all available data.

$$
\begin{bmatrix} \mathbf{y_{HS}} \\ \mathbf{y_{WW}} \end{bmatrix} = \begin{bmatrix} \mathbf{X_{HS}} & 0 \\ 0 & \mathbf{X_{WW}} \end{bmatrix} \begin{bmatrix} \mathbf{b_{HS}} \\ \mathbf{b_{WW}} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Z_{1_{WW}}} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{u_{WW}} \end{bmatrix} + \begin{bmatrix} \mathbf{Z_{2_{HS}}} & 0 \\ 0 & \mathbf{Z_{2_{WW}}} \end{bmatrix} \begin{bmatrix} \mathbf{m_{HS}} \\ \mathbf{m_{WW}} \end{bmatrix} +
$$

$$
\begin{bmatrix} \mathbf{Z_{3_{HS}}} & 0 \\ 0 & \mathbf{Z_{3_{WW}}} \end{bmatrix} \begin{bmatrix} \mathbf{mpe_{HS}} \\ \mathbf{mpe_{WW}} \end{bmatrix} + \begin{bmatrix} \mathbf{e_{HS}} \\ \mathbf{e_{WW}} \end{bmatrix},
$$

where $\mathbf{y_t}$ is the phenotype and t is the trait (hair shedding score (HS) or weaning weight (WW)); $\mathbf{b_t}$ is the contemporary group effect; $\mathbf{u_t}$ is the calf genetic effect (fit

only for weaning weight) and $\mathrm{Var}(\mathbf{u}) = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2_{uWW} \end{bmatrix} \otimes \mathbf{H}$, where σ^2_{uWW} represents

the genetic variance for the calf direct effect of weaning weight; $\mathbf{m_t}$ is the cow

genetic effect and $\mathrm{Var}(\mathbf{m}) = \begin{bmatrix} \sigma^2_{mHS} & \sigma_{mHS,mWW} \\ \sigma_{mWW,mHS} & \sigma^2_{mWW} \end{bmatrix} \otimes \mathbf{H}$, where σ^2_{mHS}

represents the genetic variance for hair shedding and σ^2_{mWW} represents the

genetic variance for the cow maternal effect of weaning weight; $\mathbf{mpe_t}$ is the cow

permanent environment effect and $\mathrm{Var}(\mathbf{mpe}) =$

$\begin{bmatrix} \sigma^2_{mpeHS} & \sigma_{mpeHS,mpeWW} \\ \sigma_{mpeWW,mpeHS} & \sigma^2_{mpeWW} \end{bmatrix} \otimes \mathbf{I}$, where σ^2_{mpeHS} represents the permanent

environmental variance for hair shedding and σ^2_{mpeWW} represents the permanent

environmental variance for the maternal effect of weaning weight; $\mathbf{e_t}$ is the

random residual and $\mathrm{Var}(\mathbf{e}) = \begin{bmatrix} \sigma^2_{eHS} & \sigma_{eHS,eWW} \\ \sigma_{eWW,eHS} & \sigma^2_{eWW} \end{bmatrix} \otimes \mathbf{I}$; and $\mathbf{X}$, $\mathbf{Z_1}$, $\mathbf{Z_2}$, and $\mathbf{Z_3}$,

are incidence matrices relating the elements of $\mathbf{y}$ to $\mathbf{b}$, $\mathbf{u}$, $\mathbf{m}$, and $\mathbf{mpe}$,

respectively.

We also evaluated the phenotypic relationship between dam hair shedding score

and calf weaning weight using the subset of 6,448 dams with both hair shedding

scores and calf weights recorded in at least one year. We did this by calculating

the estimated change in calf weaning weight as a function of dam hair shedding

score using four separate simple linear regression models. In the first two

models, unadjusted calf weaning weight was regressed on unadjusted dam hair

shedding score. Using weaning weight unadjusted for age in days captures

increased gain from an earlier birth date (older when weighed), which might be

an indicator of increased fertility for earlier shedding cows. In the other two models, 205-day, age-of-dam, and contemporary group solution adjusted calf weaning weight was regressed on un-adjusted dam hair shedding score. Both the unadjusted weaning weight and adjusted weaning models were fitted separately for all available data, dams explicitly recorded as grazing toxic fescue, and dams explicitly recorded as not grazing toxic fescue.

Genome-wide association

In order to evaluate the genetic architecture of hair shedding and identify variants that contribute to hair shedding score breeding values, we performed a single-SNP genome-wide association analysis using the SNP1101 v.1 software [33]. The breeding values calculated above using AIREMLF90 were de-regressed and used as the phenotype such that each of the 3,783 animals had one record. The de-regressed breeding values were weighted by their reliability $1 - \frac{PEV}{\sigma_a^2}$, where $PEV = (SE^2) * \sigma_e^2$ and σ_a^2 and σ_e^2 are the estimated additive genetic and residual variances for hair shedding score, respectively. Heritability was constrained to 0.40 and the genomic relatedness matrix used to control for family structure was calculated using the VanRaden method [28].

Using the UMD 3.1 bovine genome assembly [34] coordinates and annotations, we searched genes within 50 kb of SNPs with a genome-wide *q*-value lower than 0.05. The size of our search space was determined based on the density of our marker set, and the resulting gene list was used as input for cluster enrichment analysis within ClueGO v.2.5.6 [35]. KEGG pathways and biological process

gene ontologies with at least four associated genes were considered for search terms. We also searched for protein-protein interaction between genes in our list using STRING v.10 [36], considering co-expression, experimental data, and curated databases as active interaction sources.

Results

Effect of age on hair shedding score and contemporary group definition

The results of the age-in-years model suggest a non-linear effect of age with larger effect sizes in two-year-old, three-year-old cows, yearlings, and old cows relative to mature cows (Figure 1.3a). Both the BIF age class model and the four age class model had lower AIC values than the null model with no age effect (38912.38, 38906.17, and 38983.31 respectively). Likelihood ratio test results indicate a better fit of the four class model over the null ($-log_{10}(p)$ = 8.899) and no improvement in model fit using BIF age classes over four age classes ($-log_{10}(p)$ = 0). The power of contemporary grouping is undermined by over-parameterization, which can result in fewer animals per contemporary group. Therefore, we chose to classify age using the simpler four age class definition in all downstream analyses where contemporary group was fitted as a fixed effect in order to maximize contemporary group size.

Effect of toxic fescue grazing status on hair shedding

When treated as separate traits, hair shedding while grazing and not grazing toxic fescue had similar heritability estimates (Table 1.1) and a high genetic

17

correlation (r_g = 0.93). Furthermore, the Pearson correlation between breeding values grazing and not grazing toxic fescue was 0.99. The total phenotypic variation in hair shedding grazing toxic fescue was slightly higher than hair shedding not grazing toxic fescue, which suggests that reduced peripheral blood flow caused by fescue toxicosis is more detrimental to hair shedding than heat stress alone (Table 1.1). The fixed-effect model solutions support this conclusion (β_f = 0 vs. -0.59 hair shedding score units for grazing and not grazing toxic fescue, respectively). Furthermore, the estimated permanent environment effect (and therefore estimated repeatability, r) was much higher for hair shedding while grazing toxic fescue (Table 1.1).

Genetic parameters, breeding values, and estimated bias

Using all available data, the estimated narrow-sense heritability ($\frac{\sigma_a^2}{\sigma_a^2+\sigma_{pe}^2+\sigma_e^2}$) was 0.40 with an approximate standard error of 0.018. Likewise, the estimated repeatability ($\frac{\sigma_a^2+\sigma_{pe}^2}{\sigma_a^2+\sigma_{pe}^2+\sigma_e^2}$) was 0.44 with an approximate standard error of 0.012. These estimates are similar to those previously reported in Angus cattle based on pedigree relatedness [22].

Across ten iterations, $d^v_{p,w}$ averaged 0.25, ranging from 0 to 1.48. In the absence of bias introduced by incorrect estimation of the genetic trend, this value is expected to be zero. Estimates of $b^v_{p,w}$ ranged from 0.96 to 1.05, which suggests minimal dispersion of breeding values (Figure 1.4). Prediction accuracy ($\rho^v_{p,w}$) ranged from 0.70 to 0.73.

Relationship between hair shedding and weaning weight

All three bivariate models suggest a moderately negative genetic correlation between weaning weight and hair shedding score. In the model using all available data, the estimated r_g between the maternal component of weaning weight and hair shedding was -0.19 (Table 1.2). When the data were stratified by dam toxic fescue grazing status, this estimate increased slightly in magnitude for both grazing and not grazing toxic fescue (r_g = -0.25 and -0.28, respectively). For dams not grazing toxic fescue, the r_g between the direct and maternal effect of weaning weight fell near commonly reported estimates (r_g = -0.29; [37]) but was much higher for dams grazing toxic fescue (r_g = -0.63) and for all possible dams (r_g = -0.43) (Table 1.2). The r_g between the direct effect of weaning weight and hair shedding ranged from -0.10 (dams not grazing toxic fescue) to -0.03 (all possible data) to 0 (dams grazing toxic fescue).

In the simple linear models predicting unadjusted weaning weight from dam hair shedding score, unadjusted calf weaning weight was estimated to decrease by 1.30 kg with every unit increase in hair shedding score using all available data, by 3.22 kg for dams grazing toxic fescue and by 5.08 kg for dams not grazing toxic fescue. Slope estimates from the simple linear models predicting adjusted weaning weight from dam hair shedding score were more modest but also negative. Adjusted calf weaning weight was estimated to decrease by 1.45 kg using all possible data, by 2.47 kg among dams grazing toxic fescue, and by 1.11 kg among dams not grazing toxic fescue with every unit increase in hair shedding

score (Figure 1.5).

Genome-wide association analysis

We found 176 variants that passed the genome-wide false discovery rate threshold of 0.05 and 56 variants that passed the false discovery rate threshold of 0.01 (Figure 1.6). Of these 176 variants, 33% are on chromosome 5. Two hundred and six unique genes were found to be within 50 kb of significant variants. The two strongest associations were observed within *CEP290*. Perhaps interestingly, near our largest peak, we identified several members of the KRT gene family (*KRT1*, *KRT3*, *KRT4*, *KRT76*, *KRT77*, *KRT78*, and *KRT79*), which are involved in creating structural epithelial cells like hair.

We found significant enrichment (Benjamini-Hochberg corrected p-value < 0.05) for pathways involved in virus-host interaction, fat cell differentiation, prolactin signalling, cellular response to starvation, vasopressin-regulated water reabsorption, and other biological processes (Table 1.3). We also found more protein-protein interactions than expected (PPI enrichment p-value = 0.00462) and enrichment for PFAM protein domains "keratin type II head" (FDR = 8.89e-06), "somatotropin hormone family" (FDR = 8.09e-05), and "intermediate filament protein" (FDR = 0.00064).

Discussion

The expression of a phenotype is not always consistent across lifespan [38]. We found that the relationship between age and hair shedding is non-linear with

young cows, especially 2-year-old and 3-year-old cows, that displayed higher

hair shedding scores than their older herd mates. This is in line with

expectations, as young cows require increased energy expenditure associated

with continued growth [39] and the new stress of lactation [40]. To a lesser

extent, cows 10 to 13 years old tended to have higher hair shedding scores than

young animals. A similar U-shaped relationship between age and molt date was

reported in other ungulate species [8] and was reflected in the estimates of effect

size from the BIF age class model (Figure 1.2b). Cows are typically culled from

the herd or die after 10 to 11 years of age [41,42]. Thus, estimates of effects for

cows older than 12 years reflect a selected sample. However, the early shedding

estimates for these very old cows support early hair shedding as an important

characteristic of longevity, especially in heat-stressed environments.

Although our results suggest a high correlation between hair shedding score

breeding value across toxic fescue grazing status, we found a slightly higher

heritability estimate and much larger effect of permanent environment among

cattle grazing toxic fescue than those not grazing toxic fescue. Stress can

sometimes increase phenotypic variation [38], which could result in the higher

heritability observed among cattle grazing toxic fescue. Because repeatability is

the upper bound of broad sense heritability, the disparity found in permanent

environment estimates might be explained by a larger contribution of non-

additive genetic effects (i.e., epistatic and dominance effects) to variation in hair

shedding while grazing toxic fescue versus while not grazing toxic fescue. It is

also possible that certain permanent environmental effects (i.e., physiological

differences between the ability of animals to shed their winter hair) are manifest

when cattle graze infected tall fescue. Most likely, the increased estimate of the

permanent environment effect reflects the accumulation of physiological damage

from long-term fescue toxicosis. The medial layer of blood vessels tends to be

thickened in animals that suffer from fescue toxicosis, which Strickland et al. [43]

linked to hyperplasia of the smooth muscle. Repeated exposure to ergovaline

also increases venous contractile response, suggesting bioaccumulation [44].

Typically, measurements of the same trait across different environments that

result in genetic correlations r_g lower than 0.80 are considered "very different"

[45]. Hair shedding scores recorded while grazing toxic fescue versus while not

grazing toxic fescue have an r_g of 0.93, which suggests minimal re-ranking of

breeding values. However, the magnitude of the difference in permanent

environment effects found here may justify treating hair shedding grazing and not

grazing toxic fescue as separate traits in research studies that examine

physiological or non-additive genetic effects. For the implementation of the

American Angus NCE, we have chosen to minimize the effect of toxic grazing

status by including it in the definition of contemporary groups. Many biotic and

abiotic factors affect the prevalence of toxicity-inducing ergot alkaloids within

forage, including moisture, reproductive status, soil nitrogen, and most notably,

temperature [46]. Previous work suggests that animals must ingest a threshold

level of ergot alkaloids before fescue toxicosis symptoms become evident [47].

However, in these analyses, toxic fescue grazing was treated as a binary

producer-reported status in the absence of quantitative measures of ergot

alkaloid levels, which may affect the interpretation of our results. Furthermore, we did not account for the effect of grazing toxic fescue in previous years.

Our enrichment results identified pathways associated with prolactin signaling, which is a well-known modulator of seasonal hair shedding and hair growth as well as milk production [4]. In 2014, Littlejohn et al. identified mutations in prolactin (*PRL*) and its receptor (*PRLR*) that cause abnormal pelage, milk production, and thermoregulation phenotypes in cattle [48]. Furthermore, low serum prolactin level is often used as an indicator of fescue toxicosis [49]. Gray et al. [22] suggested that the negative relationship that they found between calf weaning weight and dam hair shedding was due in part to differences in serum prolactin level. Our results support this conclusion. While the genetic correlation found here using all possible data between a dam's hair shedding score and the weaning weight of her calf is moderate, it is nearly three times less than the previous estimate reported by Gray et al. [22] (r_g = -0.58), which was identical to the correlation reported by Turner and Schleger [21] for a calf's own hair shedding score and its post-weaning gain. This is likely due, in part, to our use of a slightly different phenotype. Turner and Schleger [21] used an expanded 7-point scoring system, whereas Gray et al. [22] used the same scoring system but categorized dams based on the month of the year that they first achieved a hair shedding score of 3 (about 50% shed; Figure 1.1c). We also considered the relationship between hair shedding score and the maternal effect of weaning weight rather than the direct effect of weaning weight. Another possibility could be confounding environmental effects. The relationship between dam hair

shedding score and calf weaning weight was also different across toxic fescue grazing statuses, and when toxic fescue grazing statuses are modelled separately the r_g between hair shedding score and the maternal effect of weaning weight increases relative to the r_g estimated using all data. This is similar to the results reported in Hoff et al. [50], where the accuracy of bovine respiratory disease (BRD) genomic prediction was higher when analysis of the data was done with data stratified by state than taken all together. The authors postulated that the discrepancy in prediction accuracy was likely due to the prevalence of different BRD-causing pathogens between environments [50]. Similarly, our results suggest that the relationship between hair shedding and other production traits may be environment- or context-specific.

In the four phenotypic regressions of calf weaning weight on dam hair shedding score, dam hair shedding while grazing toxic fescue was estimated to have the largest effect on adjusted weaning weight, but not on unadjusted weight. When contemporary grouping is fitted as a fixed effect in BLUP, the resulting contemporary group solution can be interpreted as a metric of environmental stress [51,52]. Larger contemporary group solutions indicate a greater advantage to the phenotype from the environment, including plane of nutrition and management practices. The disparity that we found between adjusted and unadjusted weaning weight results can be explained by smaller contemporary group solutions among calves whose dams grazed toxic fescue. Indeed, the mean contemporary group solution among calves whose dams did not graze toxic fescue was 20.75 kg higher than that of those whose dams did graze toxic

fescue (258.95 and 238.20 kg, respectively).

The negative genetic correlation that is often found between the maternal and direct genetic effects of weaning weight has puzzled researchers since the first large-scale national cattle evaluations, with some suggesting that it is an artefact and others that it reflects real biological phenomena [53]. We found that the magnitude of this genetic correlation varied across toxic fescue grazing statuses, with dams grazing toxic fescue showing a more negative correlation (-0.63) than dams not grazing toxic fescue (-0.29) (Table 1.2). There are several potential explanations for this result. First, the variation that we found in genetic correlations between maternal and direct weaning weight could result from sire-by-herd and sire-by year interactions [54]. These interactions can arise via multiple avenues, including genotype-by-environment interactions, selective data reporting, and preferential management of the progeny of certain sires. If this interaction were larger in certain herds, our estimates would be skewed. Alternatively, it is possible that our results reflect the effect of fescue toxicosis on dam nutrient partitioning. Our enrichment analysis identified multiple pathways involved in response to nutrient levels, response to starvation, and fat cell differentiation, which could support this conclusion. During the initiation of lactation, mammals draw upon their own energy reserves in order to meet increased metabolic demand [55,56], which implies genetic antagonism between maternal and direct weaning weight [37,53]. The nutrient partitioning process is influenced by stress. For example, Rhoads et al. [57] demonstrated that decreased feed intake explains only part of the reduction in milk yield found in

heat-stressed dairy cows, indicating further changes in metabolism and partitioning of nutrients in response to hyperthermia.

Although associations with a FDR less than 0.05 were detected on 20 of the 29 bovine autosomes and associations with an FDR less than 0.01 were found on seven chromosomes, one third of the associated variants were on chromosome 5. Among these, ten variants were located near or within members of the keratin gene family. In particular, *KRT1*, *KRT3*, *KRT4*, *KRT77*, *KRT78*, and *KRT79* form a protein-protein interaction network, the orthologs of which are co-expressed in other species during the formation of intermediate filament proteins. However, it is possible that significant variants near and within keratin genes are simply an artefact of extensive linkage disequilibrium (Figure 1.6b). Using the current sample, this result is difficult to disentangle.

The two most significant associations were both detected in *CEP290*. In humans, mutations in *CEP290* cause abnormal photoreceptors [58,59]. Photoreceptors affect an animal's ability to detect changes in seasons [60], and changes in photoreceptors could have large impacts on this function. Mutations in *CEP290* affect cilia formation, and are believed to interact with Bardet–Biedl syndrome (BBS) proteins [61]. Recently, *BBS1* was associated with local adaptation in Red Angus cattle [19]. Furthermore, the strength of these associations on chromosome 5 from 12 to 28 Mb could be due to multiple causal mutations [62] affecting multiple genes.

Many strategies have been proposed to phenotype heat stress in cattle. These

methods often require the use of specialized equipment and training (e.g., body temperature, respiration rate, heart rate, and sweating rate; see [63]), or at the very least increased labor cost. Therefore, routine collection of such "gold standard" phenotypes is currently limited to use in dairy cattle or in research settings. Early summer hair shedding scoring is minimally labor intensive, since cattle need not be physically handled or processed in order to be scored. Furthermore, accurate hair shedding scoring requires a relatively small time commitment and little to no training, making it an ideal candidate for genetic evaluation at a national scale. Automated sensing technologies present an opportunity to deeply phenotype animals at large scale [65], but are not currently accessible or easily implemented by the majority of beef producers. In the future, such "gold standard" measures of heat stress could be combined with routinely collected hair shedding scores to provide a comprehensive prediction of tolerance to heat stress, fescue toxicosis, or both.

Cattle produced in sub-tropical environments account for nearly 80% of the global beef herd [64]. However, exports to South America and Australia accounted for ~84% of 5,333,490 total units of beef semen exported from the United States in 2019, suggesting that selection decisions made in the U.S. beef herd still have an influence on the sustainability of beef production in the global south. Furthermore, Angus genetics accounted for 86% of total semen exports in 2019 (National Association of Animal Breeders, personal communication). At an international scale, a genetic evaluation for heat stress in American Angus cattle could enable global producers to exploit elite American genetics and make faster

genetic progress in production and meat quality traits while minimizing loss of environmental adaptability. Furthermore, because of the ease of phenotype collection, hair shedding scores can be collected in any temperate or subtropical environment and used in phenotypic or genetic selection for heat tolerance.

Conclusions

We developed a prototype genetic evaluation for early-summer hair shedding in American Angus cattle in order to enable genetic selection for heat tolerance. In agreement with previous research [21,22], we found that early summer hair shedding is moderately heritable. We also identified variants associated with biological pathways such as prolactin signaling, response to starvation, and keratin formation that contribute to genetic variation for hair shedding score. Weaning weight and hair shedding score appear to be negatively correlated. However, we found evidence for a greater impact of hair shedding score on performance for cows experiencing heat stress alone compared to cows grazing toxic fescue. Therefore, further investigation of the relationship between hair shedding and other symptoms of fescue toxicosis (such as reduced fertility) are warranted in order to determine the appropriateness of using hair shedding scores as an indicator trait for tolerance to fescue toxicosis. Exploration of the functional biology of hair shedding both on and off toxic fescue is also necessary. Finally, our results support the use of hair shedding scoring as a barometer of cow wellbeing in addition to other routinely collected phenotypes such as body condition score.

Figures
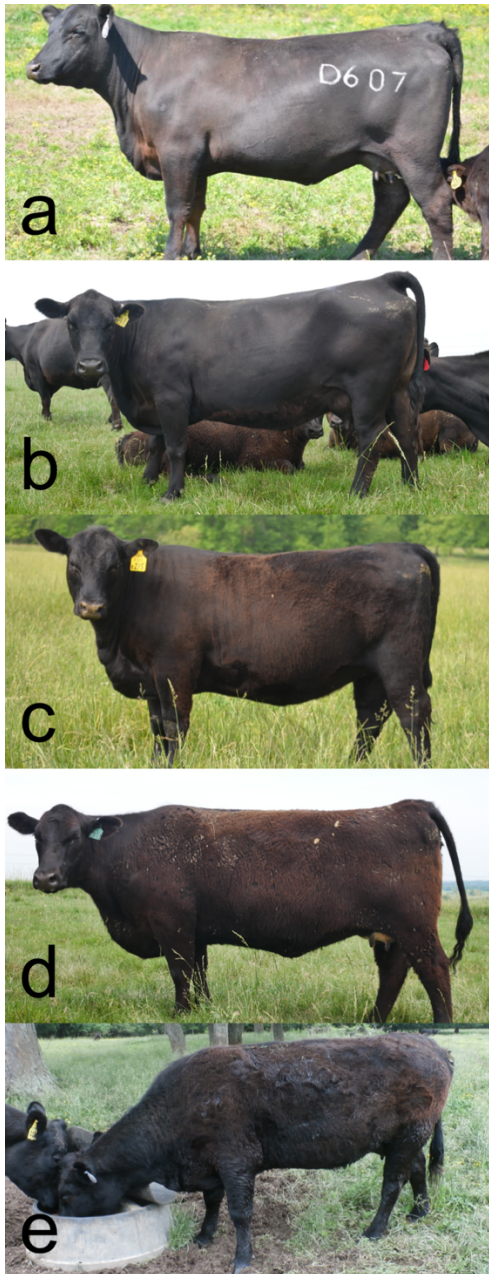
Figure 1.1. Hair shedding scoring system. Examples of the 1 to 5 visual appraisal hair shedding scoring system used in this research. (a) Score of 1, 0% dead winter coat remaining. (b) Score of 2, approximately 25% of winter coat remaining, typically observed on the lower hindquarter, flank and belly. (c) Score of 3, approximately 50% of winter coat remaining. (d) Score of 4, approximately 75% of winter coat remaining. Hair is typically removed from the head and neck first. (e) Score of 5, 100% winter coat remaining.
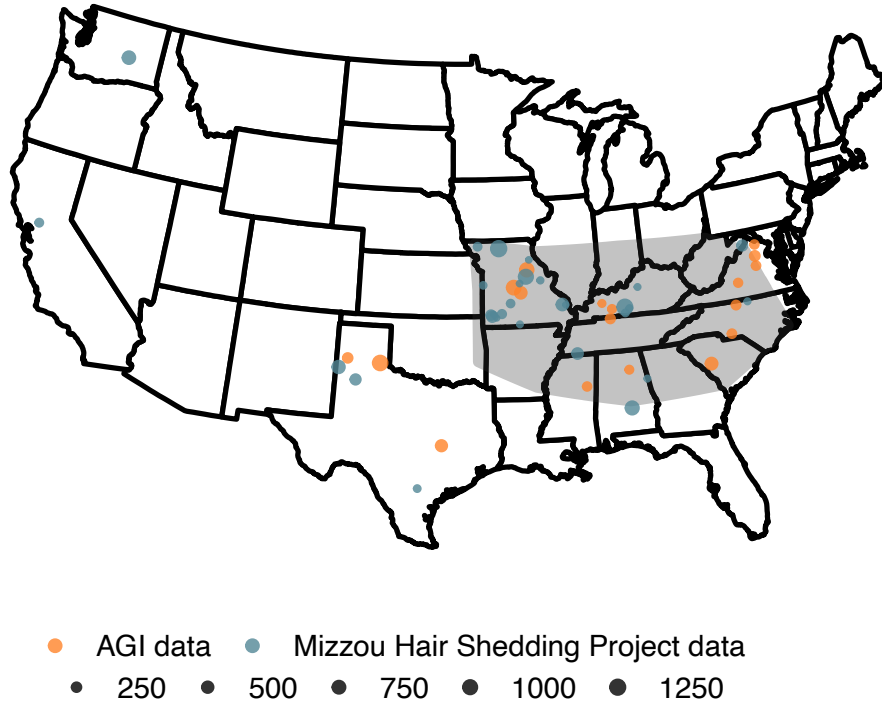
Figure 1.2. Geographic distribution of animals with hair shedding scores.
Hair shedding scores in both the AGI and MU datasets originated primarily from the South and the Fescue Belt. Here, the approximate location of the Fescue Belt is shaded in grey. Size of circles denotes the number of hair shedding scores recorded at that location. Farmers and ranchers in the MU dataset reported whether cattle grazed the predominant endophyte-infected fescue forage or a different forage species.
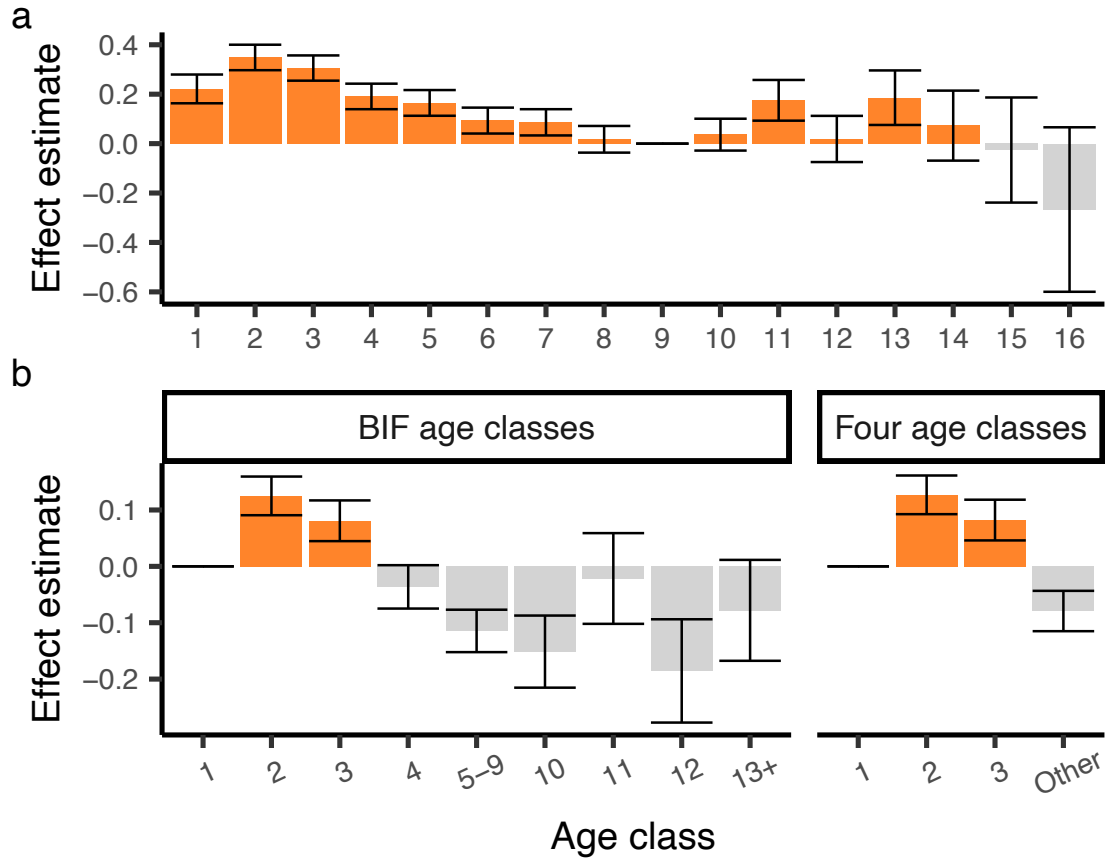
Figure 1.3. Estimates of the effect of age on hair shedding score. (a) The effect of age in years on hair shedding score appears to be non-linear and follows a U-shaped pattern. (b) Comparison of effect estimates using BIF age-of-dam classifications or four age classes. Error bars represent standard error. Age groups with at least five observations are plotted.
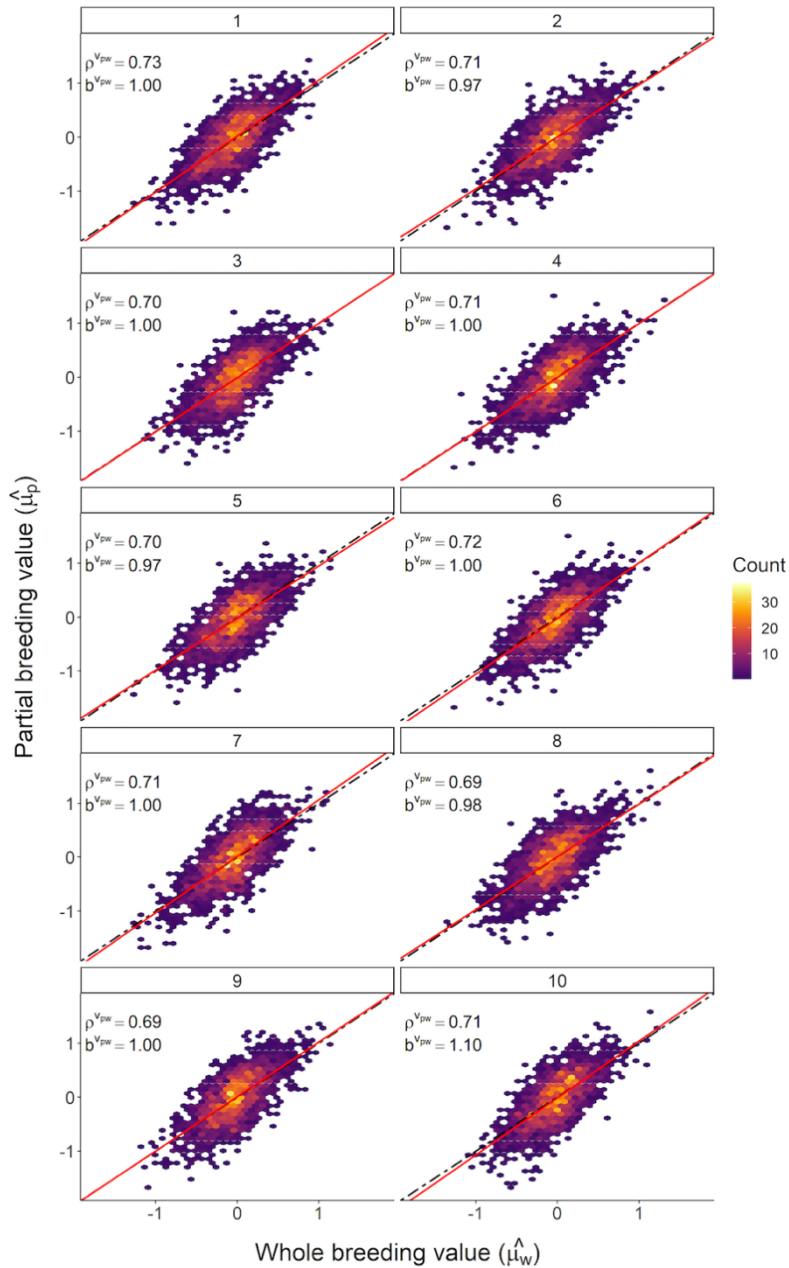
Figure 1.4. Linear regression evaluation of breeding values. Comparison of breeding values estimated using all available data ($\hat{\mu}_w$) and breeding values estimated using a reduced dataset ($\hat{\mu}_p$) across ten iterations within validation animals. The solid red line represents $b^v_{p,w}$ and the dotted black line represents the expectation of $b^v_{p,w} = 1$ in the absence of dispersion.
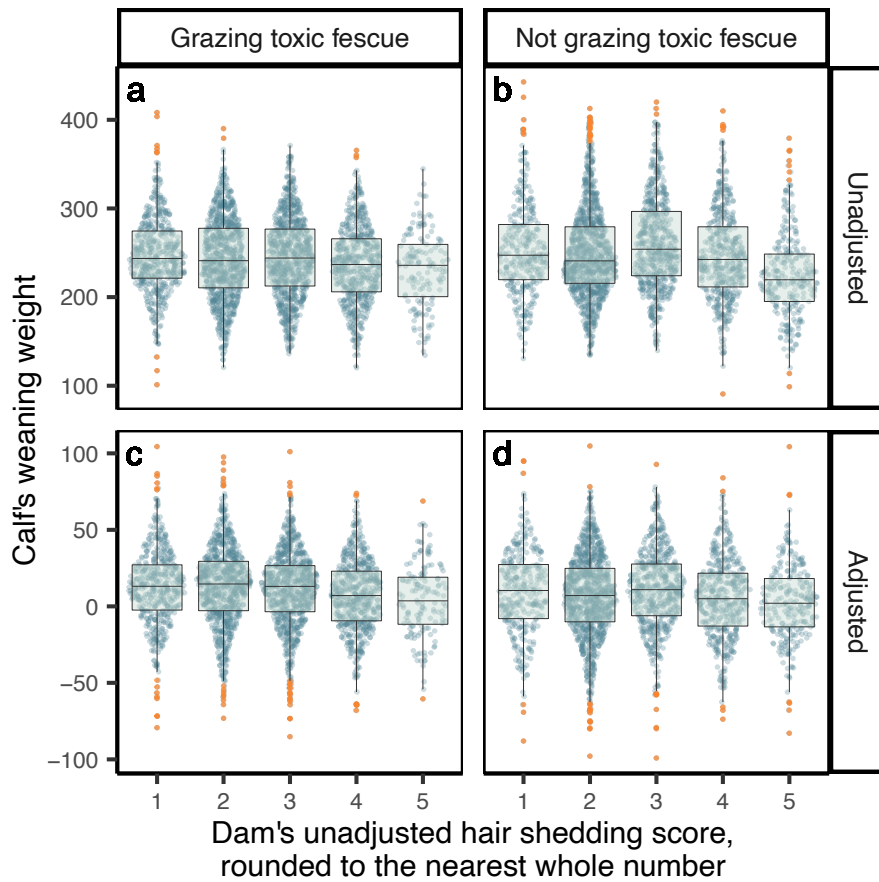
Figure 1.5. Comparison of dam's hair shedding score to the weaning weight of her calf. The effect of a dam's hair shedding score on the unadjusted (a, b) and adjusted (c, d) weaning weight of her calf with outlier weaning weights highlighted in orange. Regardless of fescue grazing status, there is very little difference in calf weaning weight between dams with hair shedding scores 1, 2, and 3.
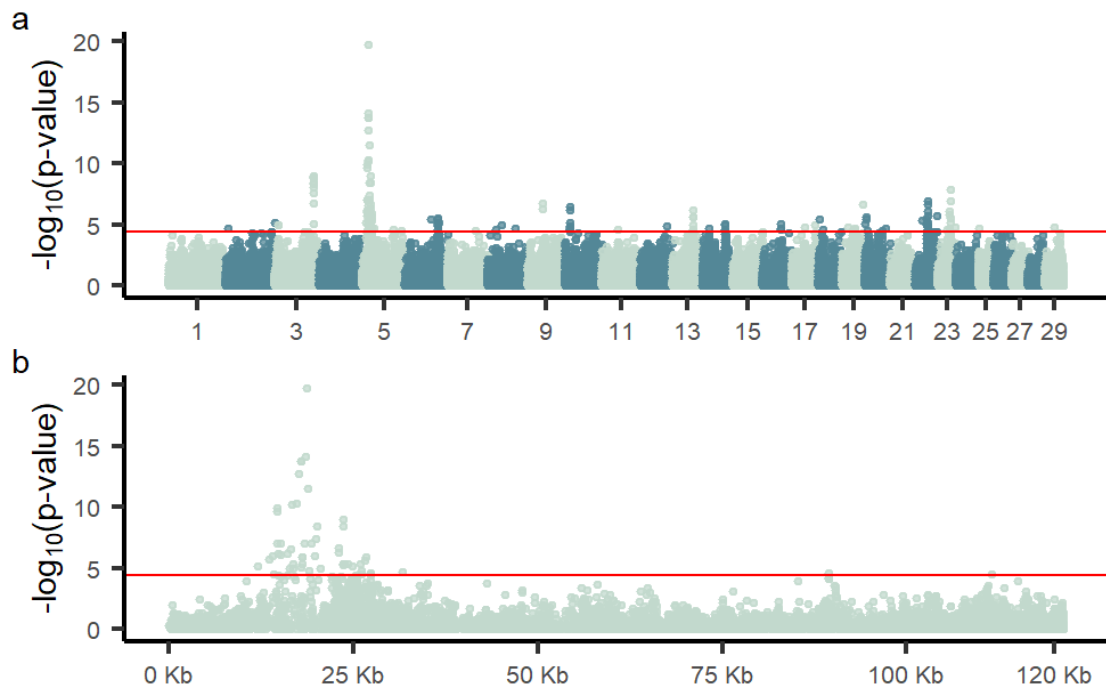
Figure 1.6. Manhattan plot of variants associated with hair shedding. Using de-regressed hair shedding score breeding values in SNP1101 sin regression, we found 176 variants that are significantly associated with hair shedding (FDR < 0.05, red line) (a). Of these 176 variants, 33% reside in a peak on chromosome 5 (b).

Tables

Table 1.1. Comparison of genetic parameters estimated using cattle grazing and not grazing toxic fescue.

	Bivariate model				Univariate model
	σ_P^2	σ_A^2	h^2	r	β_f
Grazing toxic fescue	0.90	0.38	0.40	0.45	0
Not grazing toxic fescue	0.95	0.30	0.34	0.34	-0.59 hair shedding score units

The estimated phenotypic variance (σ_P^2), additive genetic variance (σ_A^2), narrow sense heritability (h^2), and repeatability (r) from a bivariate model and fixed effect of grazing versus not grazing fescue from a univariate model. Additive genetic variance, heritability, and repeatability are higher for hair shedding recorded while grazing toxic fescue when treated as a different trait from hair shedding while not grazing toxic fescue. When fescue grazing status is fit as a fixed effect in a univariate model, the estimated effect of toxic fescue on hair shedding score (β_f) is also higher (i.e. later shedding animals).

Table 1.2. Estimated genetic correlations between dam hair shedding and calf weaning weight.

	Weaning weight (direct)	Weaning weight (maternal)
All available data		
Hair shedding	-0.03 (0.055)	-0.19 (0.066)
Weaning weight (direct)		-0.43 (0.050)
Grazing toxic fescue		
Hair shedding	0.01 (0.080)	-0.25 (0.104)
Weaning weight (direct)		-0.63 (0.071)
Not grazing toxic fescue		
Hair shedding	0.10 (0.091)	-0.28 (0.097)
Weaning weight (direct)		-0.29 (0.097)

Genetic correlation estimates between hair shedding, the direct effect of weaning weight, and the maternal effect of weaning weight vary across toxic fescue grazing statuses with approximated standard errors in parentheses.

Table 1.3 Terms significantly associated with genes within 50 kb of hair shedding GWAA variants with FDR < 0.05

Term	Ontology source	p-value	Associated genes
Modulation by virus of host morphology or physiology	GO BP	<0.001	ATG7, SMAD3, VAPB, ZC3H12A
Modification by symbiont of host morphology or physiology	GO BP	<0.001	ATG7, SMAD3, VAPB, ZC3H12A
Modification of morphology or physiology of other organism involved in symbiotic interaction	GO BP	0.004	ATG7, SMAD3, VAPB, ZC3H12A, ZNF502
Modification of morphology or physiology of other organism	GO BP	0.008	ATG7, SMAD3, VAPB, ZC3H12A, ZNF502
Interaction with host	GO BP	0.010	ATG7, SMAD3, VAPB, ZC3H12A, ZNF502
dsRNA fragmentation	GO BP	0.014	SNIP1, TARBP2, ZC3H12A
Production of small RNA involved in gene silencing by RNA	GO BP	0.014	SNIP1, TARBP2, ZC3H12A
Production of miRNAs involved in gene silencing by miRNA	GO BP	0.015	SNIP1, TARBP2, ZC3H12A
Gene silencing by miRNA	GO BP	0.020	SNIP1, TARBP2, ZC3H12A
Positive regulation of fat cell differentiation	GO BP	0.020	PRDM16, SH3PXD2B, ZC3H12A
Cellular response to extracellular stimulus	GO BP	0.021	AQP3, ATF4, ATG7, KLF10, ZC3H12A
Cellular response to dsRNA	GO BP	0.021	SNIP1, TARBP2, ZC3H12A
Vasopressin-regulated water reabsorption	KEGG	0.021	AQP3, LOC784058, RAB11A
Posttranscriptional gene silencing	GO BP	0.022	SNIP1, TARBP2, ZC3H12A
Regulation of viral genome replication	GO BP	0.022	TARBP2, VAPB, ZC3H12A
Posttranscriptional gene silencing by RNA	GO BP	0.022	SNIP1, TARBP2, ZC3H12A
Positive regulation of viral life cycle	GO BP	0.022	TARBP2, VAPB, ZNF502
Prolactin signaling pathway	KEGG	0.022	LOC100336962, PRP-VII, PRP14, PRP9
Regulation of viral life cycle	GO BP	0.022	TARBP2, VAPB, ZC3H12A, ZNF502
Cellular response to starvation	GO BP	0.023	ATF4, ATG7, KLF10, ZC3H12A
Regulation of mitochondrion organization	GO BP	0.024	MIEF1, PEMT, SNIP1, TRIAP1, WDR75
Positive regulation of mitochondrion organization	GO BP	0.024	MIEF1, PEMT, SNIP1, WDR75
Cellular response to nutrient levels	GO BP	0.024	AQP3, ATF4, ATG7, KLF10, ZC3H12A
Regulation of fat cell differentiation	GO BP	0.024	PRDM16, SH3PXD2B, SMAD3, ZC3H12A
Response to dsRNA	GO BP	0.024	IFNE, SNIP1, TARBP2, ZC3H12A

Regulation of viral process	GO BP	0.025	TARBP2, VAPB, ZC3H12A, ZNF502
Viral genome replication	GO BP	0.025	TARBP2, VAPB, ZC3H12A
Regulation of protein targeting to mitochondrion	GO BP	0.025	PEMT, SNIP1, WDR75
Gene silencing by RNA	GO BP	0.025	SNIP1, TARBP2, ZC3H12A
Osteoclast differentiation	GO BP	0.025	EPHA2, KLF10, OSTM1
Regulation of establishment of protein localization to mitochondrion	GO BP	0.025	PEMT, SNIP1, WDR75
Oxidative phosphorylation	GO BP	0.025	COX6A1, NDUFA12, TEFM
Response to starvation	GO BP	0.025	ATF4, ATG7, KLF10, ZC3H12A
Negative regulation of defense response	GO BP	0.026	KRT1, SMAD3, TARBP2, ZC3H12A
Positive regulation of viral process	GO BP	0.026	TARBP2, VAPB, ZNF502
Positive regulation of establishment of protein localization to mitochondrion	GO BP	0.026	PEMT, SNIP1, WDR75
Positive regulation of protein targeting to mitochondrion	GO BP	0.026	PEMT, SNIP1, WDR75
Negative regulation of inflammatory response	GO BP	0.036	KRT1, SMAD3, ZC3H12A
Ribosome biogenesis in eukaryotes	KEGG	0.039	GNL2, NOL6, WDR75
Regulation of protein targeting	GO BP	0.041	PEMT, SNIP1, WDR75

We find enrichment for pathways involved in virus-host interaction, response to starvation, prolactin signalling, and other biological processes. P-value are corrected for multiple testing using Benjamini-Hochberg methodology. Enrichments represent gene ontology biological process (GO BP) or Kyoto Encyclopedia of Genes and Genomes pathways (KEGG).

CHAPTER 2

GENOMIC LOCI INVOLVED IN SENSING ENVIRONMENTAL CUES AND METABOLISM AFFECT SEASONAL COAT SHEDDING IN BOS TAURUS AND BOS INDICUS CATTLE

Background

Most mammals replace their coat or molt either completely or incompletely at annual or bi-annual intervals as an adaptive response to seasonal and climatic variation [6]. In cattle, molting occurs annually in the late spring and early summer when thick winter coats are exchanged for short and sleek ones in preparation for warmer temperatures. Generally, the onset of seasonal shedding is driven by hormone cascades initiated by the hypothalamus–pituitary–gonadal axis in response to environmental cues such as day length and changes in temperature [65]. Among ungulates and other mammals, the effects of temperature and day length interact to induce seasonal molting. This interaction has never been explicitly demonstrated in cattle, although Yeates (1955) [9] showed that artificial manipulation of day length can be used to perturb the timing of hair coat shedding regardless of temperature, while Murray (1965) [66] found a moderate effect of temperature on hair coat shedding among cattle at similar latitudes.

The timing and completeness of molting is also influenced by variables intrinsic to the individual, including plane of nutrition, life stage, and social status [5,8,67]. In some species, inaccurate molt timing has a high fitness cost, and therefore phenotypic variation is limited [68]. In other species (including cattle), variation in molting has been documented within groups of contemporary individuals [21,69], suggesting genetic variation influences the ability to respond to environmental cues. Despite the extensive body of research exploring its biological basis in wild populations, domestic populations, and humans, very few studies have focused on the genetic basis of seasonal coat change (see [70,71]) and to our knowledge only Durbin et al., 2020 [72] has associated genetic variation with phenotypic variation in coat change. Though the consequences of a "mismatched" seasonal phenotype may be lower in domestic species like cattle compared to many wild populations, they can still have large impacts on productivity [10,11]. Previous work has demonstrated the impact of poor early summer hair shedding upon economically relevant traits such as growth and milk production in cattle. Here, we explore the genetic basis of variation that controls early summer hair shedding. We use a multi-breed, repeated records dataset of early summer hair shedding scores collected in a range of environments and latitudes to investigate how light, temperature, and nutrition interact with genomic loci to affect the degree of summer hair shedding in cattle.

Methods

Phenotypes

Hair shedding scores were collected over 9 years by 77 beef cattle producers and university groups. Hair shedding was classified on an integer 1-5 scale based on the systems developed by [22] and [21] as described in [72], where a score of 5 indicated 0% winter coat lost and a score of 1 indicated no remaining winter coat. Most herds were hair shedding scored once per year between mid-April and mid-June, but some groups chose to score cattle multiple times across the span of several months. This resulted in between 1 and 8 scores per animal per year. Most cattle were scored in at least two separate years (8,839 or 66.11% of all individuals; Figure S2.1b).When an animal's date of birth was available, its "age class" was calculated based on the date that the score was recorded. When no date of birth was available, the producer-provided integer age was used. Unreported score dates were assumed to be May 1 of the scoring year for the purposes of age class calculation. Age class was calculated as *(n*365d)-90d* to *(n+1)*365d-90d*, where *n* is the age classification and *d* is days. This means that animals that had not yet reached their first birthday could still be classified as yearlings and so on. Age class calculations were based on the Beef Improvement Federation age-of-dam definitions [24] as in [72], and hair shedding scores recorded on animals fewer than 275 (i.e., 365-90) days of age were excluded. Animals with differing sexes reported across multiple years were also excluded. Finally, hair shedding scores recorded on bulls and steers were

excluded as they comprised < 5% of the data, and work in other species

suggests the biological mechanisms underlying molting may be different between

sexes [8]. After filtering, 36,899 phenotypes from 13,364 cattle were retained for

analysis.

Genotypes and imputation

Array SNP genotypes were available for 10,511 phenotyped individuals and an

additional 1,049 relatives. These genotypes originated from multiple commercial

and research assays varying in density from 26,504 to 777,962 markers. Most

animals were genotyped with the GGP-F250, a research assay enriched for low-

frequency and putatively functional SNPs [26]. On an assay-by-assay basis,

markers with > 10% missing data and markers significantly deviating from Hardy-

Weinberg equilibrium ($p < 10^{-50}$) were set to missing. After site-level filtration,

samples with > 10% missing data were removed. The remaining genotypes for all

assays were merged by position, with discordant calls set to missing for

individuals genotyped with more than one assay. The merged genotypes were

then imputed to the union of the GGP-F250 and Illumina BovineHD assays using

a multi-breed reference panel and the two-step approach described in [26].

Finally, SNPs with a minor allele frequency below 1% were removed, resulting in

genotypes at 747,009 markers for 11,560 individuals.

Generation of the pedigree and relatedness matrices

Using records provided by various participating breed associations, a three-generation pedigree was constructed for registered animals with at least one phenotype retained for analysis. This pedigree was then supplemented with parentage information provided by project participants for un-registered and commercial animals with a registered sire and/or dam. To increase pedigree connectedness, the American Angus Association registration number was used for cross-registered American Angus individuals, sires, and dams with records in more than one breed association. Parentage was validated for genotyped animals with at least one genotyped parent using the SeekParentF90 program [29,73]. Based on imputation accuracy, the expected rate of genotyping error, and the distribution of Mendelian conflicts across all parent-progeny comparisons, parents found to have > 0.05% SNPs in Mendelian conflict were set to missing in pedigree. In total, 106 sires and 130 dams were excluded for 236 individuals. The final three-generation pedigree consisted of 13,221 un-phenotyped relatives in addition to the 13,364 phenotyped animals, with 6,733 unique sires and 17,954 unique dams. In order to take advantage of information from both genotyped and un-genotyped individuals, this pedigree was blended with genomic data to create the "hybrid" relationship matrix ($\mathbf{H^{-1}}$) [27]. $\mathbf{H^{-1}}$ is calculated as:

$$\mathbf{A^{-1}} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where A^{-1} represents the inversion of the numerator relationship matrix, A^{-1}_{22} represents A^{-1} subset to genotyped individuals, and G^{-1}_w represents the inverse of the genomic relatedness matrix (GRM) calculated using the VanRaden method [28]. In all models including a random effect of direct genetics, this matrix was used to represent relationships between individuals unless specified otherwise.

Estimation of breeding values and genetic parameters

<u>Full dataset</u>

Estimated breeding values (EBVs) and genetic parameters for hair shedding were first calculated using records from all available animals in the following repeated records animal model in AIREMLF90 [29].

$$y = X_1c + Z_1u + Z_2p + e$$

In this model, **y** represents a vector of hair shedding score phenotypes. Contemporary group effects are represented in the vector **c**, where X_1 is a matrix relating the elements of **c** to **y**. Contemporary groups were defined by the combination of herd ID, year, calving season (spring or fall), age group, toxic fescue grazing status, and score group. Based on the results of a model with age-in-years fit as a categorical fixed effect, age groups were defined as a) 1, b) 2-3, c) 4-9 or d) 10+. Grazing of tall fescue grass (*Lolium arundinaceum*) infected with the endophytic fungus *Epichloë coenophiala* has been shown to affect hair coat shedding in beef cattle [22,72]. Toxic fescue grazing status, or whether

cattle grazed endophyte-infected fescue in spring of the recording year, was reported by the participant as yes or no. Score group was used to account for differences in scoring dates within a herd and a year. In cases where an entire herd was not scored on the same day in a given year, records were assigned to a score group using a 5-day sliding window that maximized group size. Records from contemporary groups with fewer than 5 records were discarded.

We additionally tested a model which explicitly accounted for population structure by including the first two principal components from a PCA of all genotyped animals as fixed effects. Principal component analysis was conducted using all 11,560 individuals and 747,009 SNPs using EIGENSOFT smartPCA v.7.2.1 [74].

$$\mathbf{y} = \mathbf{X_1c} + \mathbf{pc1\beta_1} + \mathbf{pc2\beta_2} + \mathbf{Z_1u} + \mathbf{Z_2p} + \mathbf{e}$$

Only phenotypes from genotyped animals were included in this model. Otherwise, it was identical to the previous model besides the inclusion of principal components 1 and 2 as covariates. This model was compared to a model also identical to the previous model, except that only phenotypes from genotyped animals were included.

<u>Breed-specific datasets</u>

When performing ssGBLUP in crossbred populations, inclusion of data from both pure- and crossbred animals yields the highest predictive accuracy, assuming individuals have sufficiently similar genetic structure [75]. When individuals are not sufficiently similar, calculation of the GRM without accounting for differences

in allele frequencies between populations can result in inflated estimates of inbreeding. In turn, this can cause inflated EBVs and associated reliabilities for some individuals. Further, animal models assume that all individuals in the pedigree derive from the same founder individuals. Violation of this assumption (as in the case of multi-breed and cross-bred evaluations) can result in inflated estimates of the additive genetic variance [76]. Thus, we chose to replicate analyses completed in the full dataset in four breed-specific subsets with sufficient sample size for independent genetic evaluation. Calculation of within-breed EBVs also allowed us to search for differences in genetic parameters and architecture between breeds. The first through third datasets contained records from cattle registered with the American Angus Association (St. Joseph, MO; http://www.angus.org/), International Brangus Breeders Association (San Antonio, TX; https://gobrangus.com/), and American Hereford Association (Kansas City, MO; https://hereford.org/) respectively. The final dataset consisted of cattle registered with partner breed associations participating in the International Genetics Solutions (IGS) multi-breed evaluation (Bozeman, MT; https://www.internationalgeneticsolutions.com/). Further descriptions of these datasets can be found in Table 2.1. The three-generation pedigree and genotypes for associated animals were extracted for each dataset and used to construct the hybrid relatedness matrices $\mathbf{H^{-1}}$ (see *Generation of the pedigree and relatedness matrices* section).

The effects of temperature and photoperiod

Previous work in cattle has established roles for temperature and day length in initiating seasonal coat shedding [9,66]. To our knowledge, the effect of their interaction has never been explicitly examined. Understanding the biological mechanisms underpinning a trait can inform recommendations for standardized phenotyping and genetic evaluation protocols. Further, understanding the relative contributions of temperature vs. photoperiod could aid in identifying genetic variants that contribute to an animal's ability to respond to environmental cues.

Latitude and longitude coordinates were determined for each herd location using producer-provided addresses and the R [77] package {tidygeocoder} [78]. Based on these coordinates, the daily apparent high temperature, the sunrise time, and the sunset time were retrieved for the 30 days prior to each hair shedding scoring date with the {darksky} R package [79], which interfaces with the Apple Dark Sky API to query NOAA historical weather records. For each score date and geographic coordinate combination, the resulting 30-day range of apparent high temperatures was then averaged. Apparent temperature can be thought of as a proxy for heat stress, as it combines the effects of real temperature, relative humidity, and wind speed. Similarly, day lengths were calculated by subtracting the time of sunrise from the time of sunset, then averaged across the 30-day range to act as a proxy for light exposure prior to hair scoring. Next, we fit the following repeated records animal model using AIREMLF90 [29].

$$\mathbf{y} = \mathbf{X_1s} + \mathbf{X_2f} + \mathbf{X_3a} + \mathbf{X_4r} + \mathbf{t\beta_1} + \mathbf{l\beta_2} + \mathbf{Z_1u} + \mathbf{Z_2p} + \mathbf{e}$$

In this model, **y** is a vector of hair shedding score phenotypes; **s**, **f**, **a** and **r**, represent vectors of calving season, toxic fescue grazing status, age group, and year effects with matrices X_1, X_2, X_3, and X_4 relating observations to effects; β_1 represents the regression of **y** on mean apparent high temperature (**t**) and β_2 represents the regression of **y** on mean day length (**l**). The effect of farm or herd is confounded with the effect of latitude and by extension, both temperature and day length. Therefore, no herd effect was included.

Three additional models were also tested that were nearly identical to the base model above except for their inclusion of the temperature or day length variables. In two reduced models, only mean apparent high temperature or only mean day length were fit. In one expanded model, both temperature and day length were included plus an interaction effect, which was calculated by centering the individual variables then taking their product. All three of these models were compared to the base model using AIC and a likelihood ratio test.

Recommendations for genetic evaluations

In routine genetic prediction, additive and environmental variances are often partitioned by fitting a single contemporary group effect. For some traits, fitting an additional effect external to the contemporary group definition results in more accurate predictions despite the increased computational cost (i.e., the effect of age-of-dam fit for many maternally-influenced traits; [24]). For the purposes of large-scale genetic evaluations, it is of interest to know if the inclusion of additional environmental information provides a better fit than a simpler model

including only contemporary effect. To test this, we compared the base model

and four repeated records animal models similar to those explored in the

previous section. The first of these can be described as:

$$\mathbf{y} = \mathbf{X_1 c} + \mathbf{t\beta_1} + \mathbf{l\beta_2} + \mathbf{Z_1 u} + \mathbf{Z_2 p} + \mathbf{e}$$

In this model **y** represents a vector of hair shedding score phenotypes, and **c**

represents contemporary groups defined in the same way as the model

discussed in *Estimation of breeding values and genetic parameters*. Identical to

the models fit in the previous section, β_1 represents the regression of **y** on mean

apparent high temperature (**t**) and β_2 represents the regression of **y** on mean day

length (**l**). Two other models included only temperature or only day length

alongside the contemporary group effect. The final, expanded model included an

interaction between day length and temperature. In all models, records from

contemporary groups smaller than 5 animals were removed.

Genome-wide association

Deregression of breeding values and single-SNP regression

EBVs are an appealing pseudo-phenotype for further association studies as they

represent the estimated additive genetic merit of an individual with environmental

variance removed and combine repeated records into a single value. However,

failing to account for the heterogeneous variances between EBVs resulting from

the influence of familial data and in the case of repeated records traits, differing

numbers of phenotypes per individual, can result in decreased power and

increased false positive rate [80]. To take advantage of our repeated records,

genome-wide association analyses (GWAA) were performed using deregressed

breeding values (DEBVs).

First, reliabilities for EBVs were calculated as $1 - \frac{PEV}{(1+F)\sigma_a^2}$ [81], where PEV

represents the approximated prediction error variance and F represents the

pedigree-based inbreeding coefficient for the animal of interest calculated using

the R package {optiSel} [82]. Next, EBVs for genotyped animals were

deregressed and parent averages were removed using the method proposed by

[83], implemented in the {DRP} R package [84]. EBVs from animals with

unknown parentage were excluded. The resulting 9,865 DEBVs were used as

pseudo-phenotypes in SNP1101 single-SNP regression [33]. DEBVs were

weighted by (1/rel)-1, where rel is the DEBV's associated reliability with parent

information removed as calculated using the Garrick et al., 2009 [83] method.

These weights were used to construct the **R^{-1}** matrix of the mixed-model

equations. Covariance between records due to relatedness was accounted for

with a GRM constructed using the VanRaden method [28]. Five separate GWAA

were performed: one using DEBVs computed in the full dataset and the other

four using DEBVs computed in the breed-specific datasets.

Annotation and enrichment

 P-values for single-SNP associations in the full dataset were adjusted for false

discovery rate using the R package {qvalue} [85] and variants with -$log_{10}(q) > 1$

were retained for downstream analysis. In the breed-specific datasets, observed

p-values tracked closely with expected p-values, suggesting less power than in the full dataset. Therefore, we chose to determine significance in the breed-specific datasets using *p*- rather than *q*-values, and variants with *-log$_{10}$(p)* > 5 were retained.

Within each dataset, genes and annotated QTLs within 10 kb of significant variants were identified using the {GALLO} package [86], ARS-UCD1.2 bovine genome coordinates [87], and Animal QTLdb QTL annotations [88]. Enrichment analysis for annotated QTL within 10 kb of significant variants was also performed using the {GALLO} R package, and significant enrichments were determined using a Benjamini-Hochberg adjusted p-value threshold of 0.05.

Gene ontology enrichment analysis was performed for identified genes using the `gost` function provided in the {gprofiler2} package [89], which interfaces with the g:Profiler toolkit to query publicly available functional annotation databases. Significance values for functional enrichment results were corrected for multiple testing using the g:SCS algorithm, which is designed for hierarchically related, non-independent tests. Pathway enrichment analysis of significantly enriched functional terms was also performed using {gprofiler2}, then visualized using EnrichmentMap [90]. All pathways significant in the full dataset and any of the breed-specific datasets, plus unconnected terms significant in the full dataset were retained for visualization.

Results

Estimation of breeding values and genetic parameters

Principal component analysis showed that 60% of genetic variation was explained by differences between Hereford from Angus individuals (PC 1, 35%) and differences in *Bos taurus* versus *Bos indicus* ancestry (PC2, 25%; Figure S2.2). The remaining principal components explained much less variation, and therefore we chose to fit a model including eigenvectors for the first two principal components as covariates to evaluate the impact of explicitly accounting for population structure. A likelihood ratio comparing the models with and without principal components fit as fixed effects indicated that including PCs 1 and 2 provided a moderately better fit ($-log_{10}(p)$ = 3.04). However, this model required the exclusion of phenotypes from un-genotyped animals. To utilize all available information, we chose to consider results from the more basic model for downstream analyses.

Across the full and breed-specific datasets, narrow-sense heritability (h^2) and repeatability (r) were similar to parameters reported for American Angus cattle by Durbin et al., 2020 [72] and by Gray et al., 2011 [22] (Table 2.2). Estimated h^2 ranged from 0.32 (Hereford) to 0.41 (IGS) and estimated r ranged from 0.40 (Brangus, Hereford) to 0.48 (IGS). The additive genetic variance (σ^2_A) estimate in the full dataset fell within the range of estimates of σ^2_A in the breed-specific datasets, suggesting that this value was not inflated by inclusion of crossbred animals. The permanent environmental variance (σ^2_{PE}) accounted for 5-7% of

total variance with the exception of the Brangus dataset, for which σ^2_{PE} was essentially zero. The International Brangus Breeders Association did not begin participating in the project until 2018, and so 96% of Brangus animals had only one or two years of data. This might explain why no permanent environmental effect was estimated in the Brangus dataset.

In the full dataset, the median EBV was -0.02, ranging from -2.32 to 1.92. Though variation in EBVs largely overlapped between breeds, breeds recently selected for performance in the "show ring" (Shorthorn and Maine-Anjou) tended to have higher (i.e., less desirable) EBVs (Figure S2.3). Further, breeds with known *Bos indicus* ancestry (Brangus and Charolais; [91]) tended to have lower EBVs.

The effects of temperature and photoperiod

Mean hours of sunlight per day ranged from 10.89 to 15.41 hours, averaging 13.88 hours with a standard deviation of 0.74 while mean apparent temperature ranged from 4.23 to 39.33°C with a mean of 25.87 and standard deviation of 5.37. The base model including apparent temperature and day length provided a better fit over both the model with only apparent temperature ($-log_{10}(p)$ = 92.35) or day length ($-log_{10}(p)$ = 156.52), while the model including the interaction effect provided a better fit than the base model ($-log_{10}(p)$ = 3.42). The interaction model also had a lower AIC value than the base model and both of the reduced models (Table 2.3). The day length BLUE (best linear unbiased estimate or fixed effect solution) from this expanded model suggested that, on average, hair shedding

score decreases by 0.45 units for each hour increase in the mean hours of sunlight in the 30 days prior to scoring hair shedding. Further, hair shedding score was predicted to decrease by 0.07 units with every 1°C increase in the mean apparent high temperature for the 30 days prior to scoring.

Calving season, toxic fescue grazing status, and age group BLUEs from all four models are in Table S2.1. In general, BLUEs for grazing toxic fescue tended to be higher than BLUEs for not grazing toxic fescue and BLUEs for spring calving tended to be higher than BLUEs for fall calving. These results are congruent with trends reported by Durbin et al., 2020 [72]. Surprisingly, BLUEs for the oldest age group (cows aged 10 and up) were considerably more negative than BLUEs for mature cows (aged 4-9). In other mammals, patterns of seasonal coat shedding are typically "U-shaped" with age, with senescing animals in the last stage of their life typically shedding later than their younger counterparts (i.e. [8]). Since beef cows are typically culled from the herd or die between ages 11 and 12 [42], it's likely that the advantage predicted here for very old cows is reflective of selection allowing well-adapted individuals to remain productive later into their lives.

Recommendations for genetic evaluations

Based on a series of likelihood ratio tests and AIC comparisons, the base model including temperature and day length without an interaction effect provided the best fit to the data. However, the direction of the signs changed from negative to positive for all temperature and day length BLUEs relative to the models without

a contemporary group effect. Besides being incongruent with biological expectation, this is likely a sign of collinearity and suggests that including temperature or day length is redundant when contemporary groups are properly constructed. The combination of score group and farm ID captures these environmental conditions, and therefore we recommend that producers hair shedding score their entire herd on the same day in adequately sized score groups as are represented in our analyzed data.

Genome-wide association

Five-hundred fifty-seven variants in 24 peaks had $-log_{10}(q)$ values greater than 1 (Figure 2.2a) in the full dataset. Although twenty chromosomes contained at least one significant variant, over half were found on chromosome 5 as in Durbin et al., 2020 [72]. Peaks on chromosome five were also observed in the analyses of the Angus and IGS datasets, but not in the Hereford or Brangus datasets (Figure 2.3).

In the breed-specific datasets, 822, 20, 55, and 115 variants passed the significance threshold of $-log_{10}(p) > 5$ in the Angus, Brangus, Hereford, and IGS datasets, respectively. Five hundred fifty-three of the 1,522 total SNPs significant in at least 1 of the 5 GWAA resided within genes. Considering the GGP-F250 genotyping assay was designed to be enriched for genic and putatively functional variants [26], this is not surprising. However, variants within genes could also be tagging nearby regulatory elements. Most enriched gene ontologies and pathways were dataset-specific (Figure S2.4; Table S2.2; Table S2.3).

QTL enrichment analysis returned 6 significant terms with more than one

associated QTL in the full dataset (Table S2.2). An additional 48 QTL terms in

the full dataset had adjusted enrichment p-values > 0.05, including 6 terms

associated with growth or gain (Body weight (yearling), p-value = 0.133; Body

weight (birth), p-value = 0.207; Body weight (weaning), p-value = 0.606; Body

weight gain, p-value = 0.630; Body depth, p-value = 0.635; Average daily gain, p-

value = 0.650). "White spotting" was the most significantly enriched QTL term in

the full dataset (p-value = 3.8 x 10^{-27}) and the second most significantly enriched

term in the Angus dataset (p-value = 5.1 x 10^{-40}). Upon further examination, this

signal in the full dataset appeared to be driven by 51 SNPs within and near to the

MITF (*microphthalmia-associated transcription factor*) gene on chromosome 22,

a master regulator of melanocyte production that is highly conserved across

vertebrates [92–94]. Although American Angus cattle are required to be solid

black in order to qualify for registration (with the exception of a small allowances

for white spotting behind the navel; https://www.angus.org/Pub/brg.pdf), the QTL

term for "coat color" was significantly enriched in the dataset (Table S2.1).

"Non-return rate", or the proportion of females that conceive after artificial

insemination, was the top QTL enrichment term in the Angus dataset (Table

S2.1). "Interval from first to last insemination" was also significantly enriched in

the Brangus dataset (p-value = 0.018), though the signal was driven by a single

QTL.

Discussion

Results from the models quantifying the effects of external environmental variables on hair shedding score clearly point to roles for day length and temperature in regulating seasonal molt. However, the BLUEs for day length and temperature in the "base" model were nearly half the estimates from their respective reduced models, suggesting that the two variables are confounded. This makes it difficult to interpret the interaction effect estimate in the expanded model and to statistically disentangle the biological roles of photoperiod and temperature using the available data.

Gene ontology and QTL enrichment also support the role of light sensing and photoperiod in hair shedding score variation. For example, we found significant associations with the gene ontology term "long day photoperiodism" in the full and Angus GWAA results (Table S2.2). Further, variants associated with "negative regulation of testosterone secretion" and "positive regulation of testosterone secretion" were significantly enriched in the full and Angus datasets respectively (Table S2.2). Testosterone is part of the hormonal feedback loop that regulates seasonal coat change through its interaction with prolactin [95–97]. Additionally, significant variants on chromosome 23 in the full, Angus, and IGS datasets were within 10 kb of the *PRL* (prolactin) gene. Mutations in this gene have well-documented associations with abnormal milk production, hair coat, and thermoregulation [48].

QTL associated with "white spotting" were significantly enriched in the full and Angus datasets, driven by SNPs near and within *MITF* on BTA22. Perturbations to *MITF* are responsible for several auditory-pigmentary disorders across taxa, including Waardenburg syndrome [98] and Tietzs syndrome [99] in humans. In cattle, a family of German Fleckvieh cattle with bilateral deafness, incomplete albinism, and abnormal optical disks were found to harbor a missense mutation affecting the DNA binding domain of *MITF* [100]. Of particular interest, *MITF* is also essential for regulating the production of retinal pigment epithelial cells, which support the parts of the eye responsible for light sensing and color vision [101–103]. In cattle and other mammals, light stimulation in the eye activates the pineal gland, which is the main regulator of photoperiodic responses including coat molting [104]. "Coat color" was also enriched in the Angus dataset, despite the black coat color phenotype being fixed in American Angus cattle. The enrichment was determined to be driven by SNPs near 3 annotated QTL on chromosome 5 at ~ 15 Mb. The closest gene to these annotations was *ALX1*, a transcription factor that affects craniofacial development in which mutations cause severe facial dysplasias, including microphthalmia or abnormally small eyes [105]. *ALX1* has also previously been associated with facial markings in Fleckvieh cattle, but without an obvious mechanism [106]. These authors postulated that the association could have been caused by linkage with *KITLG*, a known modulator of coat color ~ 3 Mb away.

QTL associated with conception rate and other fertility phenotypes were significantly enriched in multiple datasets (Table S2.2). In wild populations, the

photoperiodic mechanisms regulating seasonal reproductive behavior and coat molting are closely intertwined (ferret [2]; badger [95]; mountain goat [8,96]; mink [97]). The same can be said for many livestock species [96]. Though cattle aren't strictly seasonal breeders, days with many hours of sunlight do result in earlier puberty in heifers and affect reproduction in other ways by increasing circulating prolactin [107–109]. In the future, investigation of the relationship between hair shedding and fertility could provide an interesting avenue of research. If variation in hair score and variation in fertility are both in part affected by variation in the ability to sense and respond to photoperiodic cues, it is possible that hair shedding score could inform breeding and management decisions as an indicator of fertility. Estimates of the genetic correlation between hair shedding score and fertility would help clarify this possibility.

Photoperiodic response is also associated with increased growth across taxa. When resources are more seasonally dependent, seasonal growth and nutrient partitioning is more extreme with higher adiposity during days with few hours of sunlight. Domestic animals are less dependent upon seasonal resources, but there is still evidence that growth, nutrient partitioning, and milk production respond to changing day length in livestock species [107,109–112]. We identified QTL associated with 18-month weight and mature weight in both the full and Angus datasets (Table S2.1). Independent of these associations, gene ontology terms were enriched for "growth hormone signaling" (Table S2.3), driven in part by variants on BTA19 near the *GH1* gene. Growth hormone acts directly on adipose tissue and forms a negative feedback loop with prolactin to regulate

photoperiodic response. During days with few hours of sunlight growth hormone signaling increases, while serum prolactin is elevated in response to days with many hours of sunlight. To our knowledge, an explicit relationship between photoperiod, growth hormone signaling, and seasonal coat shedding has not previously been demonstrated. However, our results suggest that they may be connected in cattle.

Multi-population datasets allow for the exploitation of heterogeneous patterns of LD and thus the mapping of association signals to more precise genomic intervals [113]. Here, we refined a previously published association with hair shedding score through the use of the full, multi-breed dataset. Fifty-nine percent (n = 306) of all significant variants in the full dataset were located on chromosome 5 (Figure 2.2a). Durbin et al. [72] found a similarly large association on chromosome 5 in American Angus cattle, but was unable to narrow down candidate genes. They theorized that the strength of the association, likely caused by extensive long-range linkage disequilibrium (LD), could contain multiple causal mutations affecting multiple genes [62]. In the full dataset, we find evidence for two separate associations between 14 and 24 Mb on chromosome 5 (Figure 2.2b). SNPs in the first peak fall within or near previously reported genes, including *LRRIQ1, TSPAN19,* and *CEP290,* whose action is involved in the biogenesis of the photoreceptor sensory cilia [114]. The lead SNP in this first peak and the SNP with the largest overall is located at BTA5:18,767,155. The closest genes to this SNP are an unannotated lncRNA and *KITLG*, ~ 320 kb downstream and ~ 414 kb upstream, respectively. *KITLG* and its receptor, *KIT*

have a multitude of roles across tissues, including in the retina. Recently, *KITLG* was shown to protect against retinal degenerative diseases by preventing photoreceptor death in a mouse model [115]. Also of interest, *KIT* and *KITLG* regulate the activity of *MITF*, although the mechanism by which this happens is unclear [94]. In the second association on this chromosome, the lead SNP and 14 flanking SNPs reside within the *CRADD* gene. *CRADD* was previously associated with male pattern baldness in humans [116] but is more often associated with height, body size, and morphology (https://www.ebi.ac.uk/gwas/genes/CRADD).

Only 47 SNPs were significant in more than one dataset, which is similar to results observed in other studies examining GWAA overlap between cattle breeds [117,118]. Though most variants found in domestic cattle are not breed-private, haplotype conservation is not consistent [119]. This means that even if causative QTL are shared between cattle populations, they may not be tagged by the same SNP in multiple populations. However, one might expect to find similarly enriched functional terms in pathways even in the absence of shared SNPs (i.e., [19] in studies of environmental adaptation) given a shared underlying mechanism between populations. We also found very few QTL, functional terms, or pathways enriched in more than one dataset (Table S2.1, Table S2.2, Figure S2.3). These results seem most likely to be a function of sampling and reduced power. Unfortunately, sub-setting the full dataset to "purebred" populations (with the exception of the IGS dataset) came at the cost of power in sample size, as nearly half of all records were derived from recently admixed animals with less

than ⅝ ancestry assigned to one breed based on pedigree estimates (Figure 2.1). As an extension, signal was only shared between datasets containing animals with common Angus ancestry; among the breed-specific datasets, Hereford was the only dataset free of introgression from the Angus breed. Surprisingly, the large association on chromosome 5 was not identified in Brangus, a *Bos indicus*-Angus advanced generation composite breed. Again, this is likely a function of power. In the future, more extensive hair shedding scoring of non-Angus animals would provide a more comprehensive picture of the genetic mechanisms influencing variation in hair shedding.

Estimated genetic parameters were generally in agreement across datasets (Table 2.2), with h^2 averaging ~ 0.37 and r averaging ~ 0.43 across all datasets. The estimated σ^2_A was much lower in Brangus and Hereford than any other dataset, which was also likely a function of sampling but might reflect smaller effective populations in these breeds.

In the future, several strategies could help to expand upon the results here. First, analyses aimed at clarifying the relationships between hair shedding score and temperature/day length would be helpful when creating recommendations for the implementation of national cattle evaluations for hair shedding. Based on evidence from other traits, it seems unlikely that the extent of hair shedding changes linearly with increased day length or temperature [52]. It is more likely that at some threshold, increased day length or temperature ceases to impact shedding or at least begins to have a reduced effect. Second, imputation of

genotypes to sequence level would enable further functional annotation and exploration of significantly associated regions. At the genotype density used here, causal variants are unlikely to have been directly assayed. Sequence-level data in combination with our multi-breed dataset could enable the refinement of causal variants to the base-pair level.

Data collected and maintained by non-professionals are often under-utilized, as they can introduce certain errors and biases. However, it can afford researchers an increased analytical power via vastly increased sample size when treated correctly. In a similar effort, Nowak et al. (2020) [120] recently quantified the effects of temperature and day length on molting in mountain goats using data and photographs collected by non-professionals. Similarly, we were able to explore the functional biology of a complex trait using farmer-sourced data. This work reinforces the utility of "citizen-science" type data collected by non-professionals as a powerful tool for studying complex trait biology.

Conclusions

We confirm once again that hair shedding is moderately heritable with consistent estimates of heritability and repeatability between datasets. Using a crossbred and multi-breed dataset, we were able to show that a previously published association found in American Angus cattle is likely driven by variants in multiple regions. However, collecting more data from non-Angus influenced cattle might provide a more comprehensive picture of the causal mutations driving variation in hair shedding across populations.

Together, these results point towards important roles of daylight sensing and temperature in regulating bovine seasonal hair coat shedding and provide compelling candidate regions for functional analyses. Particularly, there appears to be a clear relationship between variation in hair shedding and ocular function. Despite a vast body of research exploring the biological mechanisms regulating seasonal molting across the tree of life, to our knowledge there have been no previous studies of how genetic loci contribute to individual variation in seasonal molting. Additionally, the photoperiodic and light-sensing mechanisms regulating most seasonal phenotypes, including coat shedding, is largely shared (see Helm et al., 2013 [65] and references therein). Therefore, this work also provides an important stepping off point for research in other species.

Figures
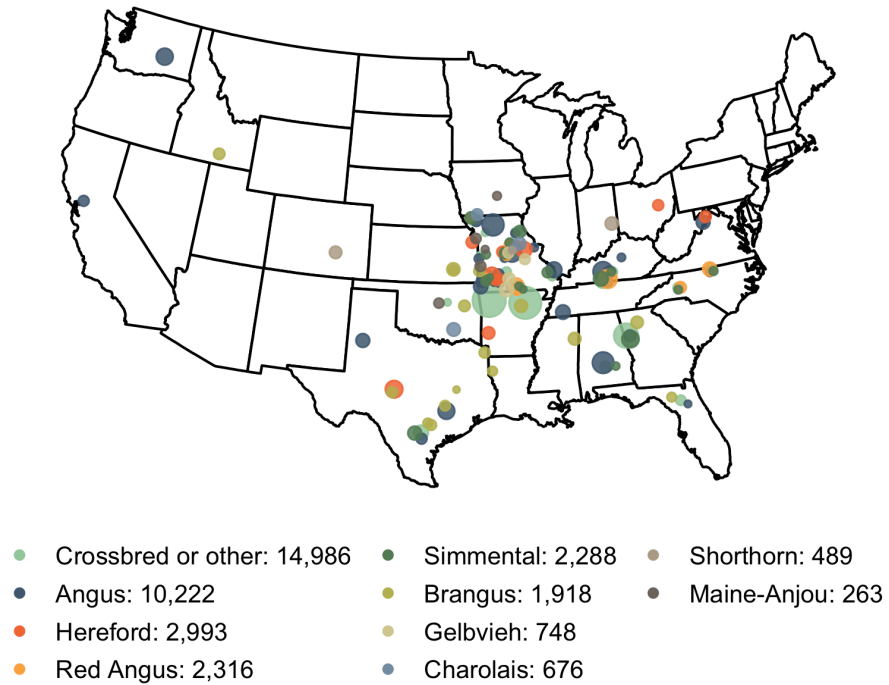
Figure 2.1. Counts of hair shedding score records by reported breed. Most phenotypes came from three breeds and were recorded in the Midwest or South. For the purposes of this map, Angus, Hereford, Red Angus, Simmental, and Gelbvieh animals with at least ⅝ ancestry assigned to the given breed based on pedigree estimates were included in that breed. Animals with unknown ancestry, less than ⅝ ancestry assigned to one breed, or of a breed not listed above were called "Crossbred or other".
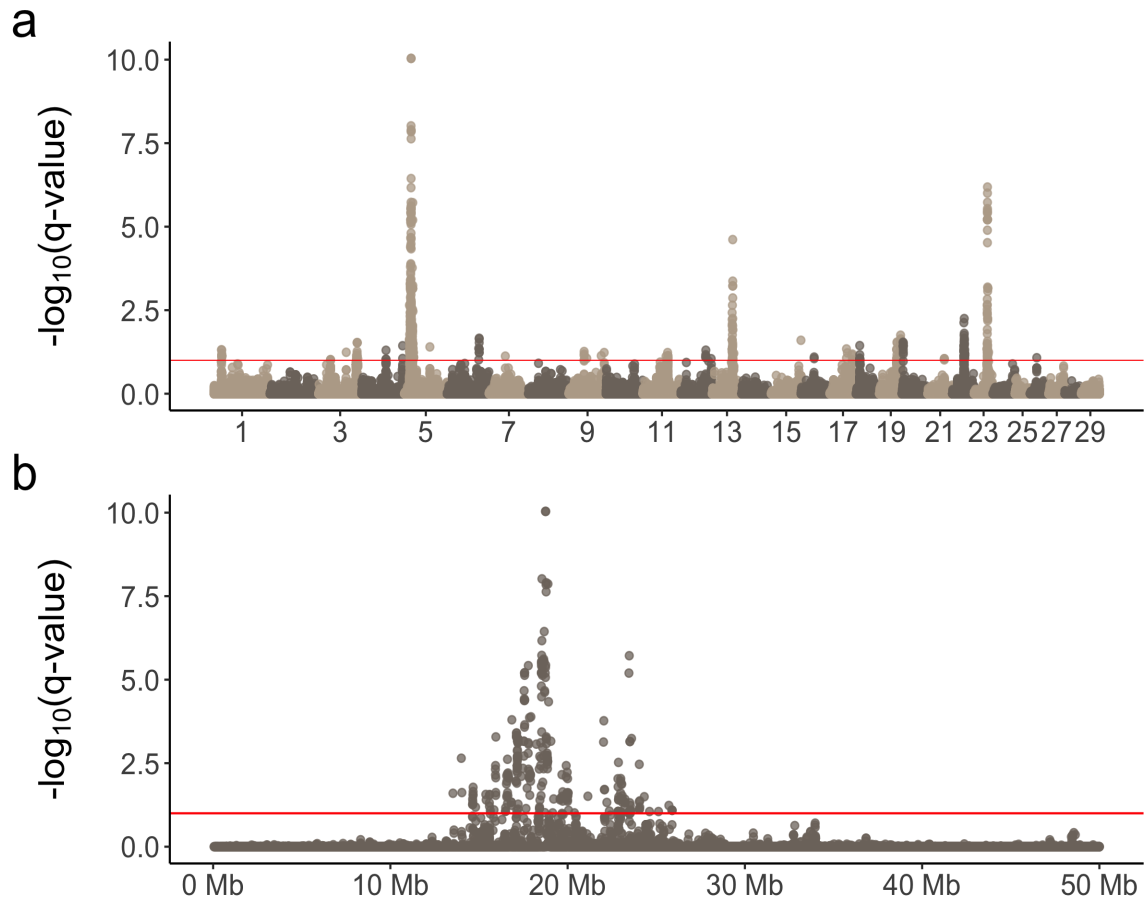
Figure 2.2. Manhattan plot of _-log₁₀(q)_ values in the full dataset (a) genome-wide and (b) on chromosome 5, truncated to 50 Mb. Red lines represent significance threshold of _-log₁₀(q)_ = 1.
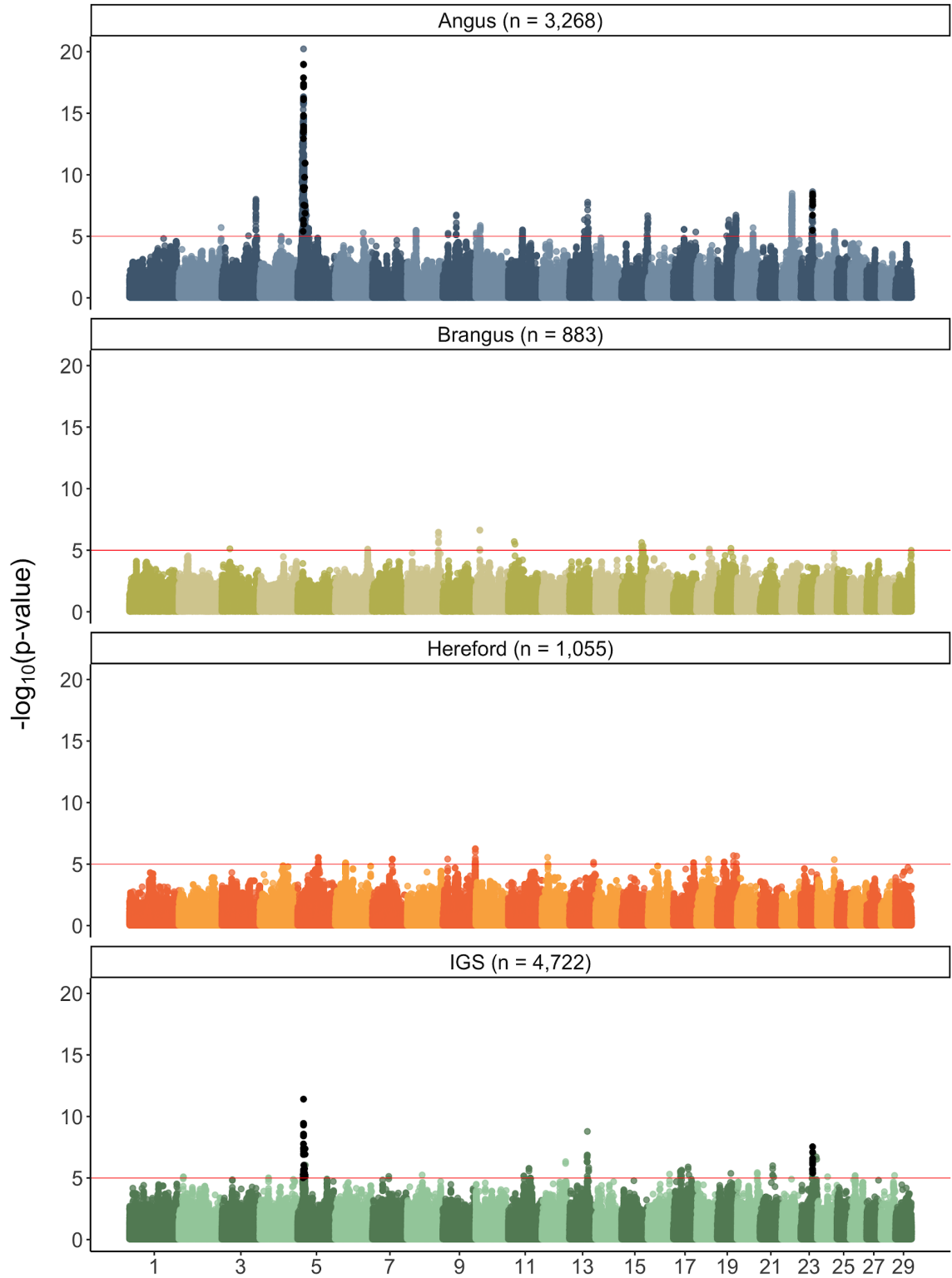
Figure 2.3. Manhattan plots of *-log₁₀(p)* values in the breed-specific datasets. Red lines represent a significance threshold of *-log₁₀(p)* = 5 and black points represent variants significant in more than one dataset.

Tables

Table 2.1. Descriptions of the breed-specific datasets.

Dataset	Description	N genotyped animals	N total phenotypes
Angus	Purebred Angus cattle registered in the American Angus Association and commercial Angus cattle enrolled in the Breed Improvement Record Program	3,286	8,674
Brangus	Brangus and Ultrablack cattle registered in the International Brangus Breeders Association	883	1,829
Hereford	Purebred Hereford cattle registered with the American Hereford Association	1,005	2,857
IGS	Purebred and crossbred cattle registered in one or more of the breed associations participating in the International Genetics Solutions multi-breed evaluation: American Simmental Association, Red Angus Association of America, American Shorthorn Association, American Gelbvieh Association	4,722	10,996

Table 2.2. Additive genetic, permanent environmental, and residual variance estimates as well as narrow-sense heritability and repeatability within the full dataset and each of the four breed-specific datasets.

Dataset	σ_A^2	σ_{PE}^2	σ_E^2	h^2	r
Angus	0.358	0.052	0.557	0.370 (0.022)	0.424 (0.014)
Brangus	0.267	0.000	0.403	0.400 (0.031)	0.400 (0.031)
Hereford	0.215	0.054	0.400	0.321 (0.044)	0.402 (0.024)
IGS	0.390	0.070	0.500	0.407 (0.020)	0.480 (0.012)
Full dataset	0.331	0.072	0.500	0.368 (0.012)	0.448 (0.008)

Approximated standard errors for heritability and repeatability estimates are in parentheses.

Table 2.3. BLUEs and AIC values from four increasingly complex models quantifying the effects of day length and temperature on hair shedding.

Model	Day length BLUE	Temperature BLUE	Day length*temperature BLUE	AIC
Day length + covariates	-0.830 (0.008)	-	-	94972.420
Temperature + covariates	-	-0.121 (0.001)	-	94380.790
Day length + temperature + covariates (base model)	-0.394 (0.013)	-0.074 (0.002)	-	93537.380
Day length + temperature + day length*temperature + covariates	-0.446 (0.016)	-0.072 (0.002)	-0.006 (0.001)	93509.710

CHAPTER 3

IDENTIFICATION OF GENOTYPE-BY-ENVIRONMENT INTERACTIONS ACROSS DISCRETE ECOREGIONS IN AMERICAN ANGUS CATTLE

Background

Compared to other major livestock species in the United States, beef cattle are raised in highly heterogeneous environments both in terms of management and in natural climatic conditions [121]. Often, selection occurs in breeding herds that are subject to more intensive management and less stressful environments than their descendants, perhaps especially descendants resulting from artificial insemination (AI). This can create the potential for genotype-by-environment interactions (GxE), whereby a genotype results in varying phenotypes depending on the environment in which it's expressed. GxE can arise via a multitude of avenues and creates the substrate upon which local adaptation acts.

Maternal effects occur when a dam's genotype influences the phenotype of her offspring beyond her additive contribution to the offspring's genotype [122], often through the environmental conditions she creates. However, the environment a dam creates for her offspring is usually a function of the environment in which she lives. In addition to direct GxE, "Maternal by environment interactions" can

arise when dams have differing responses to the environment they are subject to, thus creating differing environments for their offspring.

In cattle, GxE interactions have primarily been explored using random regression reaction norm models, in which each subject is regressed on a continuous environmental variable [18,123,124]. Subjects can be individuals or genotypes, but in studies of livestock species have typically been individuals. An intercept and slope are estimated for each subject, and the resulting line is interpreted as the degree of change in genetic potential across the environmental gradient. A slope of zero indicates robustness or no change in progeny performance across environments. Alternatively, GxE can be estimated using multivariate models, where observations on a trait made in multiple environments are fit simultaneously as potentially correlated traits. The genetic correlations between the trait expressed in different environments can then be estimated, which provides a measure of the extent of re-ranking of subjects for their additive genetic merit across environments. A genetic correlation of 1.00 indicates no GxE effect, whereas genetic correlations below 0.80 are typically thought of as indicating "very different" traits across environments [45,125]. Sometimes the environment of interest is best described by a continuous variable, in which case reaction norm models are more appropriate. However, environmental stress is often a function of many interacting variables. Discrete parameterization of environments can capture unmeasured stressors (i.e., local forages and pathogens) as well as their interactions.

Current large-scale national genetic evaluations provide predictions of genetic merit as the mean of an animal's progeny's performance across all environments in which they were generated. However, there is an increasing interest in augmenting current evaluations with predictions based on environmental suitability [18,20,72]. Though a large body of research surrounding GxE in animal agriculture has recently emerged, the magnitude of GxE effects have not been clearly elucidated and used to determine whether environmental specific national genetic evaluations are warranted. Here, we explore the extent to which GxE affects the direct and maternal components of weaning weight in American Angus cattle across discrete United States ecoregions.

Methods

National Cattle Evaluation phenotypes, breeding values, and contemporary group solutions

Weaning weight phenotypes recorded on registered and commercial Angus cattle between 1990 and May, 2019 were extracted from the May 13th weekly growth trait run of the American Angus Association (AAA) National Cattle Evaluation (NCE). Prior to entering the NCE, these phenotypes were adjusted in accordance with BIF guidelines for age-of-dam and actual age to 205-days. These pre-adjusted phenotypes were used in all downstream analyses and will hereafter be referred to as simply "weaning weights" or "phenotypes". For all phenotypes recorded within the continental United States, the weigh date, breeder zip code, and herd state (the state in which the phenotype was recorded)

were extracted from the AAA database. Phenotypes were then assigned to one of nine discrete ecoregions based on breeder zip code. Ecoregions were delineated by k-means clustering of 30-year normals for mean temperature, mean precipitation, and elevation recorded by the PRISM Climate Group [126] as in Rowan et al., 2020 [19]. To ensure accurate ecoregion assignment, phenotypes were discarded if the breeder zip code did not reside in the herd state. While our ecoregion assignments are helpful for characterizing ecoregion-specific stressors, some assignments were not useful for the goals of this project. In particular, some records assigned to the Fescue Belt ecoregion originated from locales where Kentucky-31 endophyte-infected tall fescue cannot be propagated (the western & eastern coasts and central plains). Records from California, Washington, Oregon, New York, New Jersey, Massachusetts, Maryland, Delaware, Connecticut, Rhode Island, Pennsylvania, Wisconsin, Michigan, and Nebraska assigned to the Fescue Belt were removed. Due to sample size disparity, records assigned to the Rainforest and Cold Desert ecoregions proposed by Rowan et al., 2020 [19] were also excluded. Finally, phenotypes from animals in contemporary groups with fewer than 5 animals were removed.

Some downstream analyses used estimated breeding values and contemporary group solutions produced by the weekly growth run of the AAA NCE. This dataset included 5,717,333 weaning weight measurements and associated breeding values for animals in 291,864 contemporary groups (Table 3.1; Figure 3.1).

Exploratory analyses

Contemporary group BLUEs

Contemporary group is often fit as a fixed effect during genetic evaluation to partition additive genetic from management and environmental effects. In the AAA NCE, weaning weight contemporary group members must have been born within 90 days of one another and must be of the same sex and birth type (natural service, artificial insemination, or embryo transfer). Further, they must be managed similarly at the same location and have been in the same birth weight contemporary group (https://www.angus.org/Performance/AHIR/PerfContempGrouping.aspx). Best linear unbiased estimates (BLUEs) of contemporary group effects can sometimes be used as a proxy metric to quantify the effects of the environment on a phenotype (i.e., [52,127]). We extracted contemporary groups BLUEs from the AAA NCE and then stratified and summarized them by ecoregion.

Artificial insemination sire inter-region variance

Wide-spread use of AI creates large half-sibling families distributed across a wide range of environments, which provides the opportunity to quantify sire performance across ecoregions. To explore the phenotypic variation of AI-sired progeny between discrete ecoregions, we asked:

1. For AI sires with calves born in multiple ecoregions, how much does the median weight of their calves vary from ecoregion to ecoregion?

2. Is the difference in the median weight of calves consistent between sires or are some sires less "robust", showing more variability in progeny phenotypes across ecoregions?

Sires with at least 50 calves in at least two unique contemporary groups in two or more ecoregions were retained for analysis (n = 1,710). The median number of ecoregions in which sires met these criteria was 3. One hundred ninety-seven of the sires had at least 50 recorded progeny in all seven ecoregions. For each sire-region combination, we compared the median weaning weight of the sire's progeny within each ecoregion to the median weaning weight of all of the sire's progeny in the dataset. We also compared the median weaning weight of calves in the High Plains to the median weaning weight of calves in other ecoregions for a subset of the AI sires (n = 1,522).

Calving season distribution

Bradford et al. [121] demonstrated that the weaning weight distributions of progeny of fall versus spring calved cows and the calving season that is most advantageous to growth varies across the United States. To verify that these findings were recapitulated in our data, we summarized the weaning weight distributions of fall versus spring born calves. Proxy birth dates for members of each contemporary group were calculated by subtracting 205 days from the

recorded weaning date. June 30th was used as the threshold between spring and fall calving seasons.

Genetic correlations between the High Plains and other ecoregions

Data and sampling

The genetic correlation between observations made upon the same trait in different environments provides an estimate of the degree of re-ranking of breeding values between environments and can be interpreted as the magnitude of GxE [45]. We chose to use the High Plains (Figure 3.1) as the constant in a series of bivariate variance components estimation analyses with one of the other six ecoregions.

For the High Plains and each of the six other comparison ecoregions, we sampled zip codes with at least 10 years of recorded data until a total of 100,000 +/- 500 animal records were sampled per ecoregion. This strategy increased the likelihood of sampling multiple offspring per dam, in turn increasing the accuracy of estimates of maternal genetic and maternal permanent environment effect estimates. Variance components were estimated over 10 iterations per ecoregion comparison (i.e., 60 total iterations).

Variance component estimation

The bivariate model used across all iterations and ecoregion pairs can be expressed as:

$$\begin{bmatrix} \mathbf{y_i} \\ \mathbf{y_j} \end{bmatrix} = \begin{bmatrix} \mathbf{X_i} & 0 \\ 0 & \mathbf{X_j} \end{bmatrix} \begin{bmatrix} \mathbf{b_i} \\ \mathbf{b_j} \end{bmatrix} + \begin{bmatrix} \mathbf{Z_{1_i}} & 0 \\ 0 & \mathbf{Z_{1_j}} \end{bmatrix} \begin{bmatrix} \mathbf{u_i} \\ \mathbf{u_j} \end{bmatrix} + \begin{bmatrix} \mathbf{Z_{2_i}} & 0 \\ 0 & \mathbf{Z_{2_j}} \end{bmatrix} \begin{bmatrix} \mathbf{m_i} \\ \mathbf{m_j} \end{bmatrix} + \begin{bmatrix} \mathbf{Z_{3_i}} & 0 \\ 0 & \mathbf{Z_{3_j}} \end{bmatrix} \begin{bmatrix} \mathbf{mpe_i} \\ \mathbf{mpe_j} \end{bmatrix} + \begin{bmatrix} \mathbf{e_i} \\ \mathbf{e_j} \end{bmatrix}$$

In this model, **y**, **b**, **u**, **m**, **mpe**, and **e** represent vectors of adjusted weaning weight phenotypes, fixed effects, random direct genetic effects, random maternal genetic effects, random maternal permanent environmental effects, and random environmental effects for ecoregions **i** (High Plains) and **j** (compared ecoregion). The incidence matrices **X**, $\mathbf{Z_1}$, $\mathbf{Z_2}$, and $\mathbf{Z_3}$ related records to their respective vectors of effects Variance components for this model were estimated using AIREMLF90 [29]. Within each iteration, the 3-generation pedigree was extracted for the ~ 200,000 sampled animals in the High Plains and the compared ecoregion. This pedigree was used to construct the inverse of the pedigree relationship matrix $\mathbf{A^{-1}}$, where **A** contains the covariances between phenotypes that are expected due to pedigree relatedness.

Calculation of breeding values and single-step GWAS

Based on results of the variance component estimation analyses, we chose to further explore genetic variation driving the direct and maternal effects of weaning weight in the High Plains, Southeast, and Fescue Belt.

First, phenotypes for 100,000 +/- 500 animals were sampled within each of the

High Plains, Southeast, and Fescue Belt using the zip code-based strategy

described in *Genetic correlations between the High Plains and other ecoregions*.

Three-generation pedigrees were then calculated within each ecoregion and

genotype data were extracted for all phenotyped animals with available

genotypes. The resulting High Plains, Southeast, and Fescue Belt datasets

contained 9,973, 9,872, and 9,906 genotyped animals respectively. We then

calculated breeding values using single-step genomic BLUP (ssGBLUP) [27] as

implemented in BLUPF90 [27] using the univariate model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1u} + \mathbf{Z_2m} + \mathbf{Z_3mpe} + \mathbf{e}$$

Vectors and matrices represent the same effects as in the previously described

bivariate model. ssGBLUP combines information from genotyped animals and

non-genotyped animals by replacing the inverse of the pedigree relationship

matrix $\mathbf{A^{-1}}$ with the inverse of a hybrid relationship matrix $\mathbf{H^{-1}}$, which is formed by

blending the portion of $\mathbf{A^{-1}}$ that corresponds to animals with genotypes with the

inverse genomic relationship matrix (GRM), calculated using the VanRaden

method [28]. The markers used to calculate the inverse GRM were those used in

the AAA weekly NCE and originated from commercial genotyping assays of

varying density that can be imputed to ~ 54 K with sufficient accuracy. Prior to

calculation of the inverse GRM, SNPs with a minor allele frequency lower than

1% and SNPs with > 10% missing data were excluded. Then, individuals with call

rates below 90% were excluded. Finally, we used the single-step GWAS method

(ssGWAS) described in Aguilar et al., 2019 [128] and implemented in POSTGSF90 [29] to back-solve for marker effects and their associated p-values.

<u>Annotation and enrichment</u>

First, we identified genes and annotated QTL [88] within 50 kb of variants with associated $-log_{10}(p) > 5$ using the {GALLO} R package [86] and UMD3.1 bovine assembly coordinates [34]. We next used the resulting QTL lists to perform QTL enrichment within ecoregions for the direct and maternal effects of weaning weight. QTL enrichment was also conducted using the {GALLO} package, and QTL terms with Benjamini-Hochberg adjusted enrichment p-values < 0.05 were retained for downstream analysis.

Results

Exploratory analyses

Season of calving was bimodal in the Desert, Southeast, Arid Prairie, and Fescue Belt but skewed towards spring calving in the High Plains, Forested Mountains, and Upper Midwest & Northeast (Figure 3.2). Largely, variation between ecoregions for both adjusted weaning weight and contemporary group BLUEs overlapped (Figure 3.2), but we found an 18.18 kg difference between the median adjusted weaning weight in the highest-ranked ecoregion (Upper Midwest & Northeast; 279.90 kg) and the lowest-ranked ecoregion (Southeast; 261.72 kg) (Figure 3.3a). The Southeast also ranked lowest in terms of contemporary group BLUEs, with a median solution (219.37 kg) 18.08 kg below

the median solution in the highest ranked ecoregion (High Plains; 237.35 kg) (Figure 3.3b).

Of 1,710 sires with at least 50 calves in multiple ecoregions, 311 had at least one ecoregion in which the median weaning weight of calves in that ecoregion was at least 20.00 kg lower than the median weaning weight of their calves across all ecoregions. Most frequently, this ecoregion was the Fescue Belt, occurring 40% more often than the second most frequently occurring ecoregion (Southeast). The median inter-region spread of weaning weights was 20.87 kg, though this varied considerably (SD = 15.17 kg) and 27 sires had inter-region spreads at least 3 SD above the median. Over 40% (n = 637) of the sires with at least 50 calves in both the High Plains and a comparison ecoregion had at least one comparison ecoregion in which the median weight of their calves was at least 20 kg lighter than the median weight of their High Plains calves. Again, this region was most often the Fescue Belt (n = 260) followed by the Southeast (n = 154; Figure 3.4).

Taken together, these results point to significant phenotypic variation in growth across discrete ecoregions, even between temperate ecoregions. Further, the environment of the High Plains appears to be more favorable than other ecoregions for growth, be it via management or natural phenomena.

Evidence for GxE based on estimated genetic correlations and heritabilities

For all compared ecoregion pairs, the mean direct genetic correlation between the High Plains and the comparison ecoregion (hereafter r_D) across iterations was between 0.85 and 0.87. Genetic correlations for the maternal effect of weaning weight (hereafter r_M) were much more variable between ecoregions. The mean of the 10 iteration r_M ranged from 0.77 (Southeast) to 0.86 (Desert and Arid Prairie). Further, r_M was much more variable across iterations within an ecoregion comparison. For example, the difference between the minimum r_M and maximum r_M were 0.29, 0.23, and 0.22 in the Fescue Belt, Southeast, and Upper Midwest & Northeast respectively. In the other 3 ecoregions, the range of r_M estimates across iterations were between 0.13 and 0.15, while the range of estimates averaged 0.15 across all ecoregion comparisons for r_D. The American Angus Association does not report a genetic correlation between the direct and maternal effects of weaning weight (r_{DM}), but estimates for the High Plains-comparison region r_{DM} (hereafter r_{DiMj}) fell near previously reported values ([37]; Table S3.1). However, within-region r_{DM} (hereafter r_{DiMi}) tended to be more negative (Table S3.1).

Heritabilities and estimates of the relative contribution of maternal permanent environmental effects to phenotype (c^2; $\frac{\sigma^2_{MPE}}{\sigma^2_{P}}$), were much more consistent than genetic correlations between iterations within a comparison. The mean estimated maternal effect heritability was similar to the estimate reported by the American Angus Association (h^2_M = 0.12; https://www.angus.org/Nce/Heritabilities.aspx) for

all ecoregions, with a slightly higher estimate in the Southeast (Table 3.3). The same was true for c^2, with again a slightly larger contribution estimated in the Southeast. Across all regions, direct effect heritability estimates tended to be slightly higher than but still near to the national estimate ($h^2_D = 0.28$).

Single-step GWAS

For the direct effect of weaning weight, associations were mostly shared across all 3 ssGWAS (Figure 3.5). Most significantly associated loci were located on chromosomes 7 and 20 with additional associations on chromosomes 14 and 15 in the Southeast and chromosomes 3 and 6 in the Fescue Belt. All these additional region-specific associations fell near known candidate genes for growth and body size in multiple species (*OR6N1*, *OR6K2*, *OR6K3*, *OR6K6,* and *SPTA1* on BTA3 (Table S3.2); *KCNIP4* and *SLIT2* on BTA6 [17,129]; *FBXO32* and *ATAD2* on BTA14 [130,131]; *METTL15* on BTA15 [132]). In all 3 regions, QTL enrichment analyses identified known large-effect growth QTL including those near *ARRDC3, ERGIC1* and *BNIP1* ([133,134]; Table S3.2).

Association signal was more heterogeneous between ecoregions for the maternal effect of weaning weight (Figure 3.6). Only one variant on BTA6 passed the significance threshold in the High Plains, while 18 and 20 variants had -$log_{10}(p)$ > 5 in the Southeast and the Fescue Belt respectively. The variant with the largest $log_{10}(p)$ for the Southeast maternal effect ssGWAS resided in a peak at the distal end of BTA23, ~ 30 kb upstream of *TFAP2A* (transcription factor AP-2α). In the ssGWAS of maternal effects in the Fescue Belt, the closest gene to

the lead SNP was *TIAM1.* Three QTL associated with milk composition were significantly enriched in the Fescue Belt but no QTL were significantly enriched in the Southeast.

Discussion

Overall, GxE interactions were predicted to have a larger effect on the maternal component of weaning weight than the direct component of weaning weight. Since a calf's plane of nutrition is almost entirely dependent upon its dam prior to weaning, this was expected. In most beef cattle genetic evaluations, the breeding value associated with the maternal effect of weaning weight is referred to as "maternal milk". However, maternal effects can shape weaning weight phenotype in a multitude of ways beyond milk production, including via prenatal environment and dam behavior [122]. Recent single-cell RNA sequencing analyses demonstrated the preferential expression of *TFAP2A*, the gene nearest the lead SNP in the Southeast maternal effect ssGWAS, in the trophectoderm of pre-implantation bovine blastocysts [135]. Additionally, *TIAM1* was previously associated with tropical adaptation in Senepol cattle [136]. Each of these results could be suggestive of GxE in the maternal effect of weaning weight for factors other than milk production.

Much of the U.S. cowherd is at risk for fescue toxicosis, caused by grazing tall fescue forage (*Lolium arundinaceum*) infected with the endophytic fungus *Epichloë coenophiala*. Both fescue toxicosis and heat stress reduce calf weaning weight via reduced dam milk production [12,31]. The low r_M we often found

between the High Plains and the Fescue Belt is likely reflective of the effects of fescue toxicosis on dam health.

"Cryptic genetic variation" is standing genetic variation that is typically not expressed or observed because it only modifies the expression of a phenotype in certain environments. Sometimes, environmental stress can result in higher estimates of additive genetic variance by exposing cryptic genetic variation [137–139]. For example, there are likely segregating variants that affect milk production when a dam is exposed to a low plane of nutrition, but not when her nutritional needs are met. Within the same ecoregion, it could be possible to sample many animals on a high plane of nutrition in one iteration then many animals on a low plane of nutrition in the next iteration. The iteration containing many animals on a low plane of nutrition would likely have a higher estimated additive genetic variance (and thus, likely a higher r_M) due to the expression of typically "cryptic" standing variation. In the absence of herd IDs, zip codes are an appropriate proxy for family cohorts. On the other hand, sampling animals by zip code could result in large breeding operations that may manage animals more or less intensively than other producers in their ecoregion having a disproportionately large influence on within iterate genetic correlation estimates. The example of cryptic genetic variation presented above could be expanded to include any number of within-region stressors to explain the high variation in r_D and r_M estimates observed across iterates within some ecoregions. For example, Bradford et al. 2016 [121] showed that fall-born calves tend to be heavier than spring-born calves in some parts of the United States. There may be cryptic

genetic variation that affects growth and milk production in one calving season

but not the other (i.e., higher temperatures in spring in the Southeast or presence

of toxic fescue in spring in the Fescue Belt). Thus, sampling of spring vs. fall

calving herds would result in varying estimates across iterates in ecoregions

where calving season is bimodal (Figure 3.2). Using the methods presented

here, it is not possible to disentangle the impact of GxE from genotype-by-

management (GxM) interactions or genotype-by-environment-by-management

interactions (GxExM). In the future, more explicit modelling of GxM and GxExM

will be helpful for assessing the relative roles of management and the broader

environment in exposing cryptic genetic variation.

Recently, Aldridge et al. 2020 [140] showed that (co)variances estimated using

the inverses of the traditional numerator relationship matrix **A** and the hybrid

relationship matrix **H** are very similar, suggesting that the less computationally

intensive method of variance component estimation using $\mathbf{A}^{-1}$ is appropriate.

However, this finding may not extend to estimates of (co)variances between

environments. Multivariate GxE models like those used here rely on half-sibling

relationships between each of the separate environments when $\mathbf{A}^{-1}$ is used to

capture pedigree relationships. When relationships are instead represented using

$\mathbf{H}^{-1}$ or the inverse of the genomic relatedness matrix $\mathbf{G}^{-1}$, a lesser degree of

relatedness is required to gain accurate estimates because genotype sharing

between animals unrelated by pedigree is more accurately represented. The

largest proportion of American Angus breeding operations reside in the High

Plains, evidenced by the number of registrations and the number of pedigree

connections of animals in the High Plains to animals in other ecoregions; most of the American Angus population can trace at least part of its recent ancestry to the High Plains. However, we made no explicit effort to screen for relatedness between the High Plains and the comparison ecoregion sample in our sampling scheme. This is another likely contributor to the large variation in (co)variance estimates across iterations within a comparison.

The consistently negative covariance between the direct and maternal genetic effects of weaning weight in beef cattle hints at an antagonism between selection for growth and selection for milk production [53]. In practice, this means that response to selection for growth can be hindered by maternal genetic effects [141]. Others have suggested that negative r_{DM} estimates are in part caused by multi-generational maternal effects [142]. In the classic example of multi-generational maternal effects, grand-dams with superior milking ability overfeed their daughters, causing "fatty udder syndrome". The daughter's milking ability is then inferior when she becomes a dam herself, introducing a genetic antagonism between growth and milking ability [143,144]. A similar balancing selective pressure could be imagined for dams experiencing environmental stress. In a study of American Angus cattle, Durbin et al. 2020 [72] found r_{DM} was lower when data were subset to dams who grazed endophyte-infected fescue and their calves compared to dams who grazed other forages and their calves. They suggested this result could indicate the influence of fescue toxicosis on nutrient partitioning in the dam. Variation in the magnitude of r_{DM} across ecoregions was

less dramatic here than in Durbin et al. 2020, but we similarly found more negative estimates in some ecoregions than others.

Conclusion

In agreement with previous research [18], we find little evidence for GxE in the direct effect of weaning weight between U.S. ecoregions using the 0.80 threshold suggested for "very different" environments by Falconer, 1952 [45]. However, even when considering temperate regions, we find genetic correlations consistently below 0.90. In the context of large-scale national genetic evaluations, this represents dramatic re-ranking and could have serious impacts on the selection decisions made by producers across ecoregions. We find strong support for GxE in the maternal effect of weaning weight, particularly in the Southeast and Fescue Belt ecoregions. Based on contemporary group BLUEs, local stressors in these two ecoregions also appear to have the most negative impact on weaning weight. The efficiency of selection in achieving the optimum value for a trait in all environments is determined by its across-environments genetic correlations, with progress slowing as genetic correlations approach one [145]. Therefore, the lower than expected genetic correlations presented here between some ecoregions should not be viewed as a deficit. Rather, they should be viewed as an opportunity for animal breeders to improve the sustainability of beef production through environmentally-aware genetic predictions.

Figures
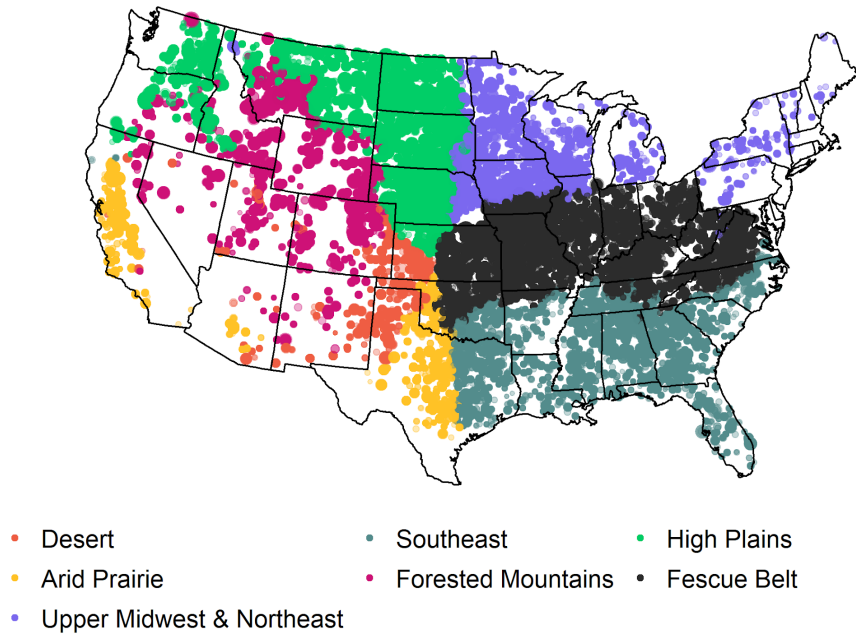
Figure 3.1. Geographic distribution of 5,717,333 weaning weight measurements in 291,864 contemporary groups samples across 7 U.S. ecoregions, recorded between 1990 and 2019. Points are scaled by the number of records.
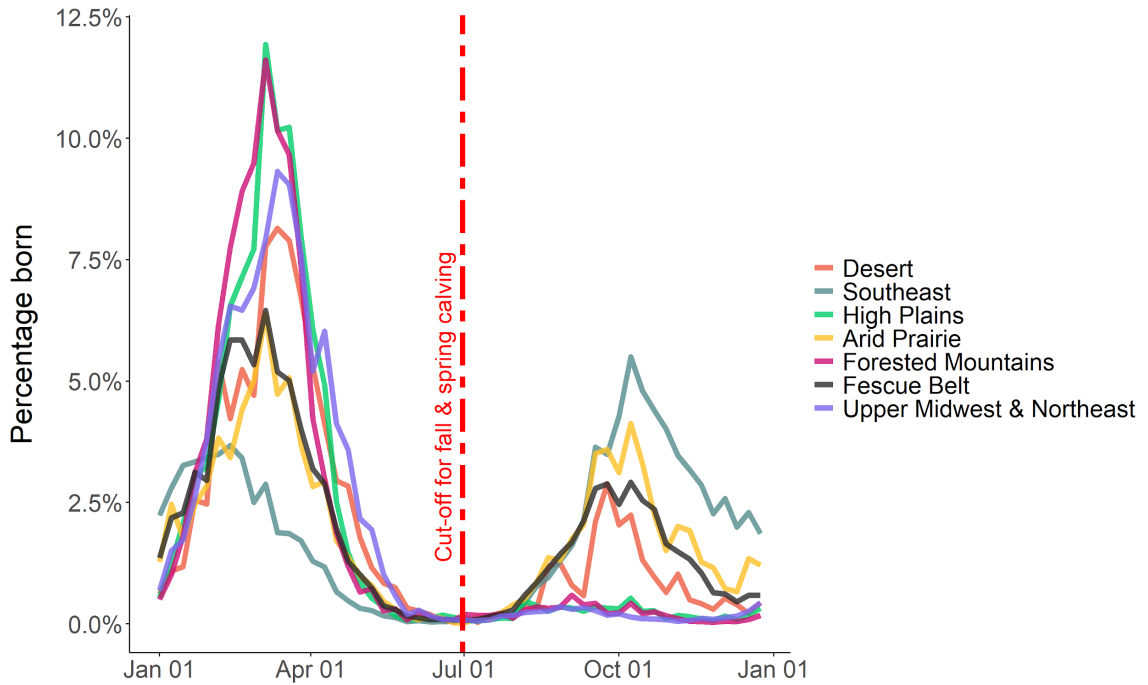
Figure 3.2. **Distribution of approximate birth date within the seven ecoregions based on weaning weight weigh-dates in 291,864 contemporary groups from 1990-2019.** Percentages on the X-axis represent the percentage of calves born on a given day relative to the total number of calves born within ecoregions.
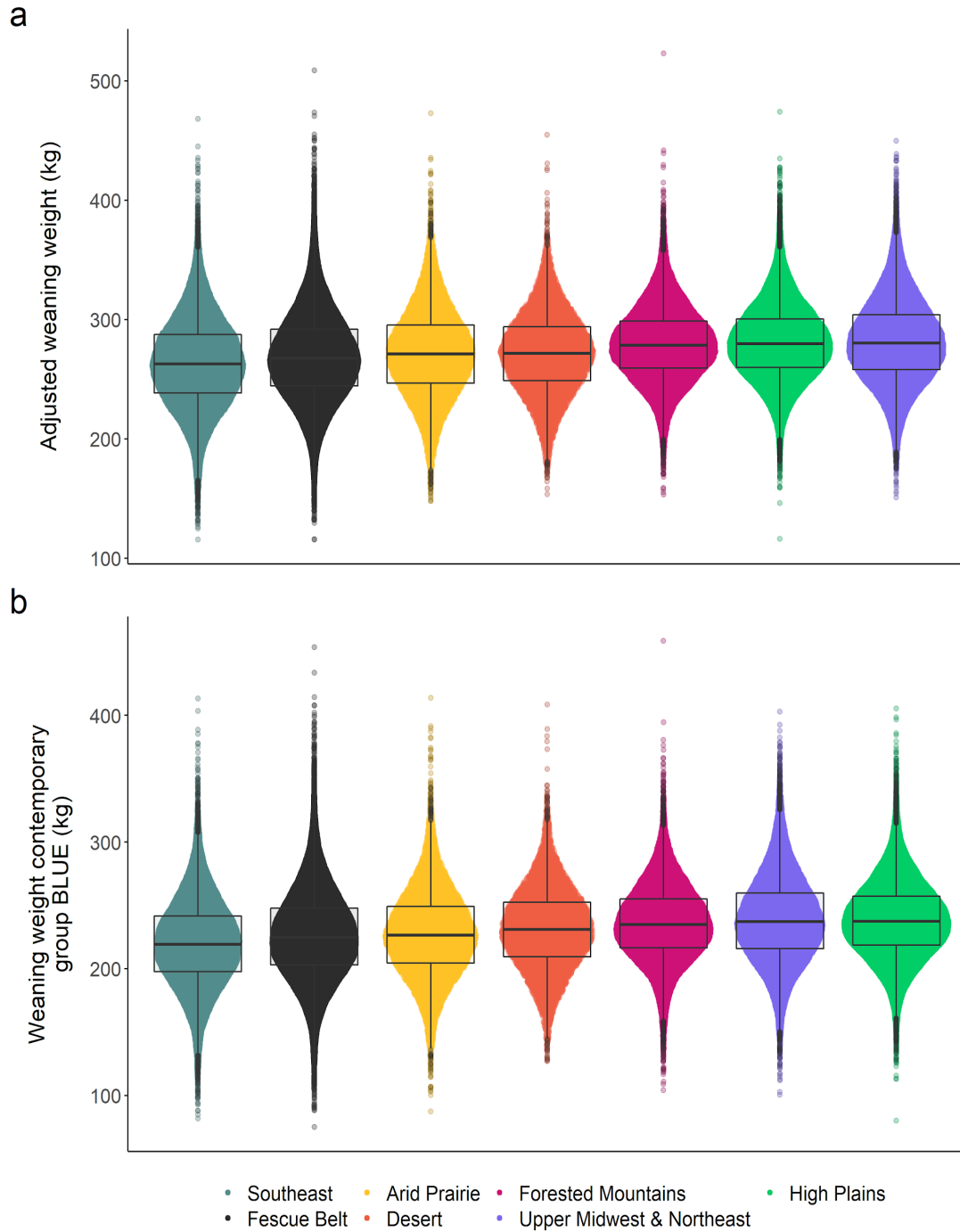
Figure 3.3. Distributions of (a) phenotypic weaning weights and (b) weaning weight contemporary group BLUEs. Hinges of box plots represent the first quartile, median, and third quartile while whiskers represent 1.5*IQR.
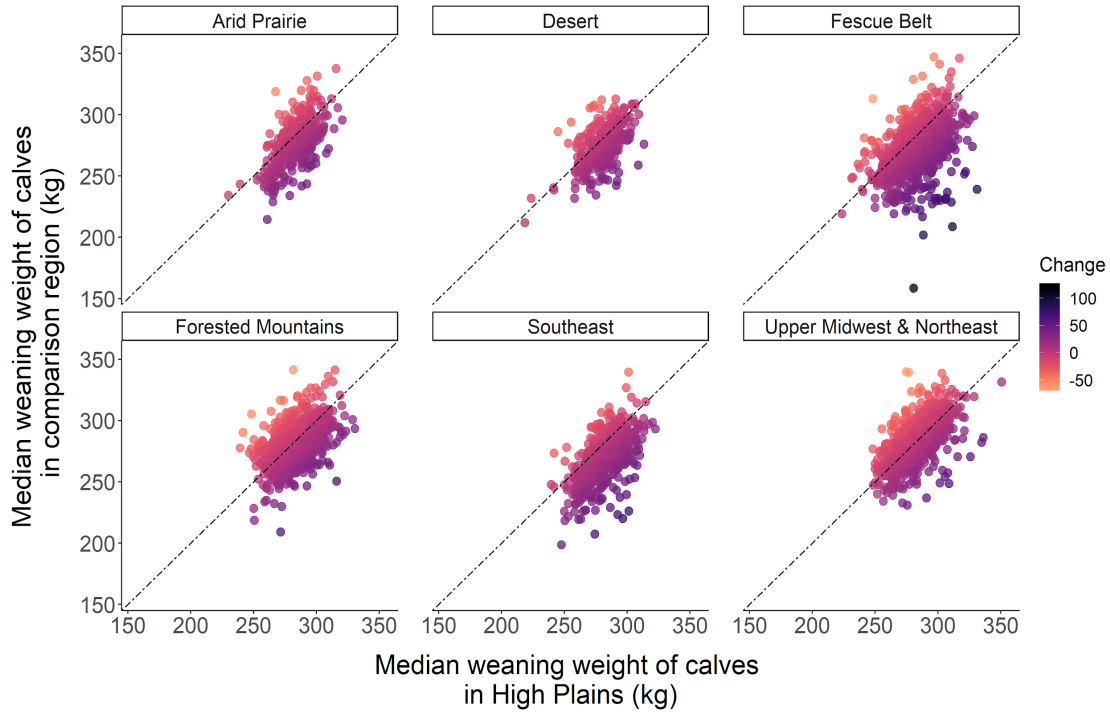
Figure 3.4. Difference between the median progeny weaning weight of calves born in the High Plains to the median progeny weaning weight of calves born in one of 6 comparison ecoregions for 1,522 AI sires. Each point represents an AI sire with at least 50 progeny in at least 2 contemporary groups in each of the High Plains and the comparison region. Most sires fall on or near the diagonal, indicating no difference between progeny performance in the High Plains and progeny performance in the comparison region. However, many AI sires have poor performance in the Fescue Belt compared to performance in the High Plains (many points below the diagonal line).
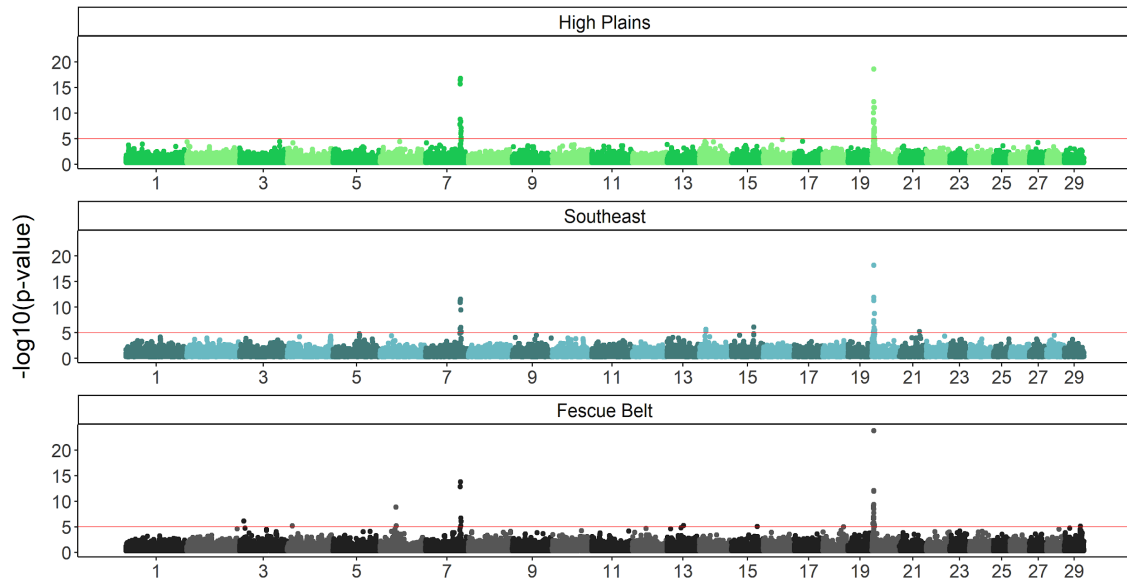
Figure 3.5. Manhattan plot of *-log₁₀(p)* values associated with the direct genetic effect of weaning weight in each of the High Plains, Southeast, and Fescue Belt. Red lines indicate significance thresholds of *-log₁₀(p)* = 5.
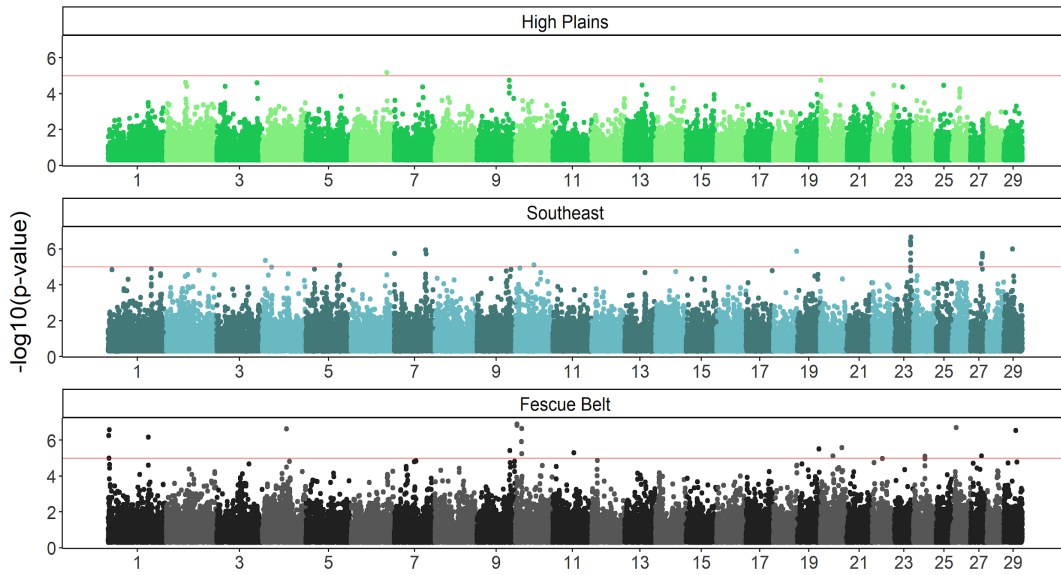
Figure 3.6. Manhattan plot of *-log$_{10}$(p)* values associated with the maternal genetic effect of weaning weight in each of the High Plains, Southeast, and Fescue Belt. Red lines indicate significance thresholds of *-log$_{10}$(p)* = 5.

Tables

Table 3.1. Number of records and contemporary groups, stratified by ecoregion.

Region	Number of records	Number of contemporary groups
Desert	165,057	8,908
Southeast	508,565	39,026
High Plains	2,075,979	69,002
Arid Prairie	208,689	13,180
Forested Mountains	696,033	26,184
Fescue Belt	1,462,959	101,806
Upper Midwest & Northeast	600,051	33,758

Table 3.2. Minimum, mean, and maximum genetic correlations for the direct and maternal effects of weaning weight between the High Plains and each of 6 comparison ecoregions across 10 iterations.

Direct effect

Comparison ecoregion	Minimum r_D	Mean r_D	Maximum r_D
Desert	0.80	0.86	0.91
Southeast	0.76	0.87	0.92
Arid Prairie	0.80	0.87	0.96
Forested Mountains	0.75	0.85	0.89
Fescue Belt	0.82	0.87	0.96
Upper Midwest & Northeast	0.71	0.85	0.95

Maternal effect

Comparison ecoregion	Minimum r_M	Mean r_M	Maximum r_M
Desert	0.80	0.86	0.95
Southeast	0.67	0.77	0.89
Arid Prairie	0.78	0.86	0.91
Forested Mountains	0.78	0.85	0.93
Fescue Belt	0.66	0.82	0.95
Upper Midwest & Northeast	0.72	0.84	0.95

Table 3.3. Minimum, mean, and maximum heritabilities for the direct and maternal effects of weaning weight within ecoregions.

Direct effect h^2_D

Comparison ecoregion	Minimum h^2_D	Mean h^2_D	Maximum h^2_D
Desert	0.32	0.34	0.35
Southeast	0.27	0.31	0.35
High Plains	0.24	0.30	0.41
Arid Prairie	0.31	0.33	0.35
Forested Mountains	0.22	0.27	0.33
Fescue Belt	0.27	0.30	0.33
Upper Midwest & Northeast	0.30	0.34	0.41

Maternal effect h^2_M

Comparison ecoregion	Minimum h^2_M	Mean h^2_M	Maximum h^2_M
Desert	0.11	0.12	0.13
Southeast	0.13	0.15	0.18
High Plains	0.10	0.12	0.15
Arid Prairie	0.11	0.12	0.13
Forested Mountains	0.09	0.12	0.15
Fescue Belt	0.10	0.12	0.14
Upper Midwest & Northeast	0.10	0.12	0.13

Maternal permanent environmental effect c^2

Comparison ecoregion	Minimum c^2	Mean c^2	Maximum c^2
Desert	0.11	0.11	0.12
Southeast	0.13	0.15	0.17
High Plains	0.07	0.10	0.11
Arid Prairie	0.12	0.12	0.13
Forested Mountains	0.10	0.11	0.13
Fescue Belt	0.12	0.13	0.15

Upper Midwest & Northeast	0.09	0.10	0.11

Estimates for the High Plains are across all 60 iterations while estimates in the other ecoregions are across 10 iterations.

SUPPLEMENTARY MATERIALS

Chapter 2 Supplementary Figures
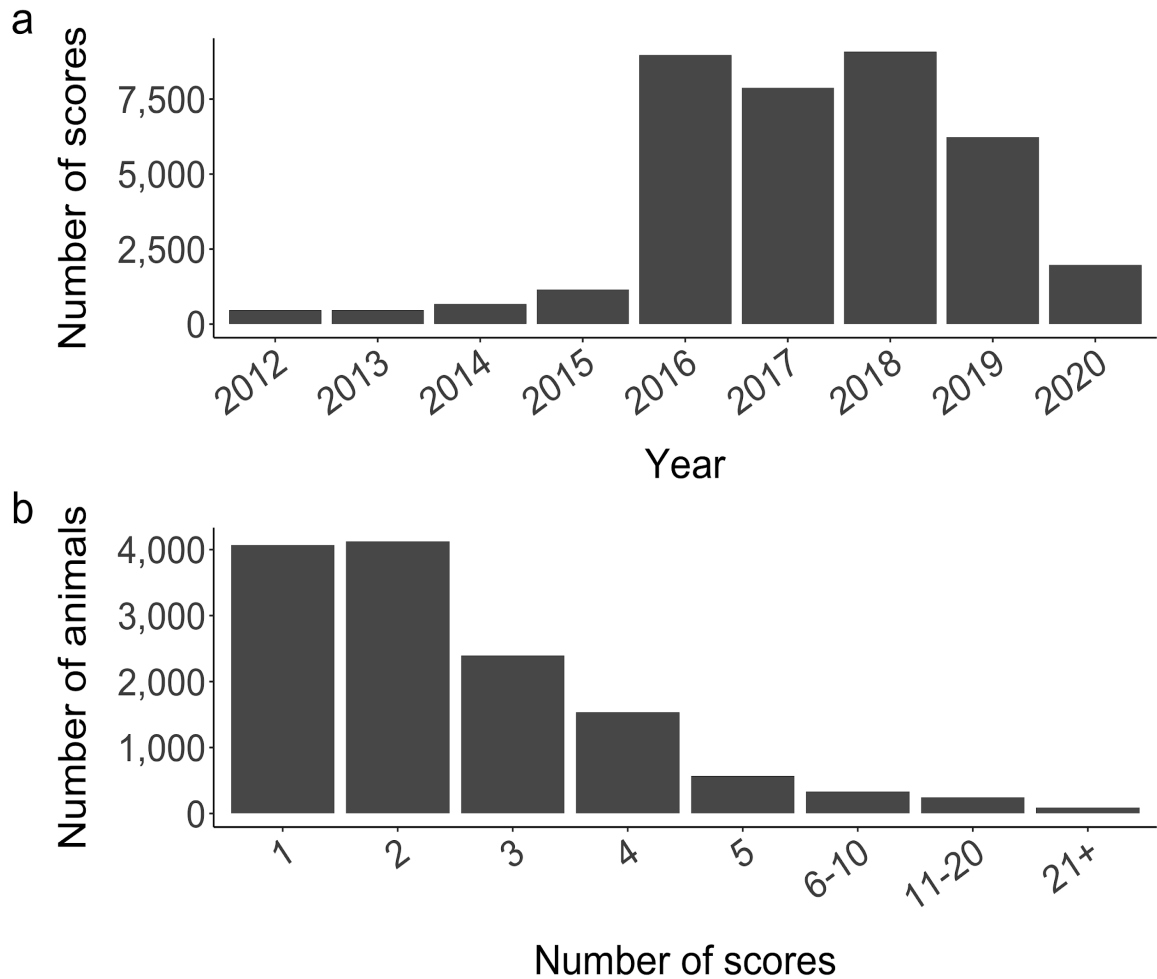
Figure S2.1. Counts of (a) hair shedding scores per year and (b) scores per animal across all years.
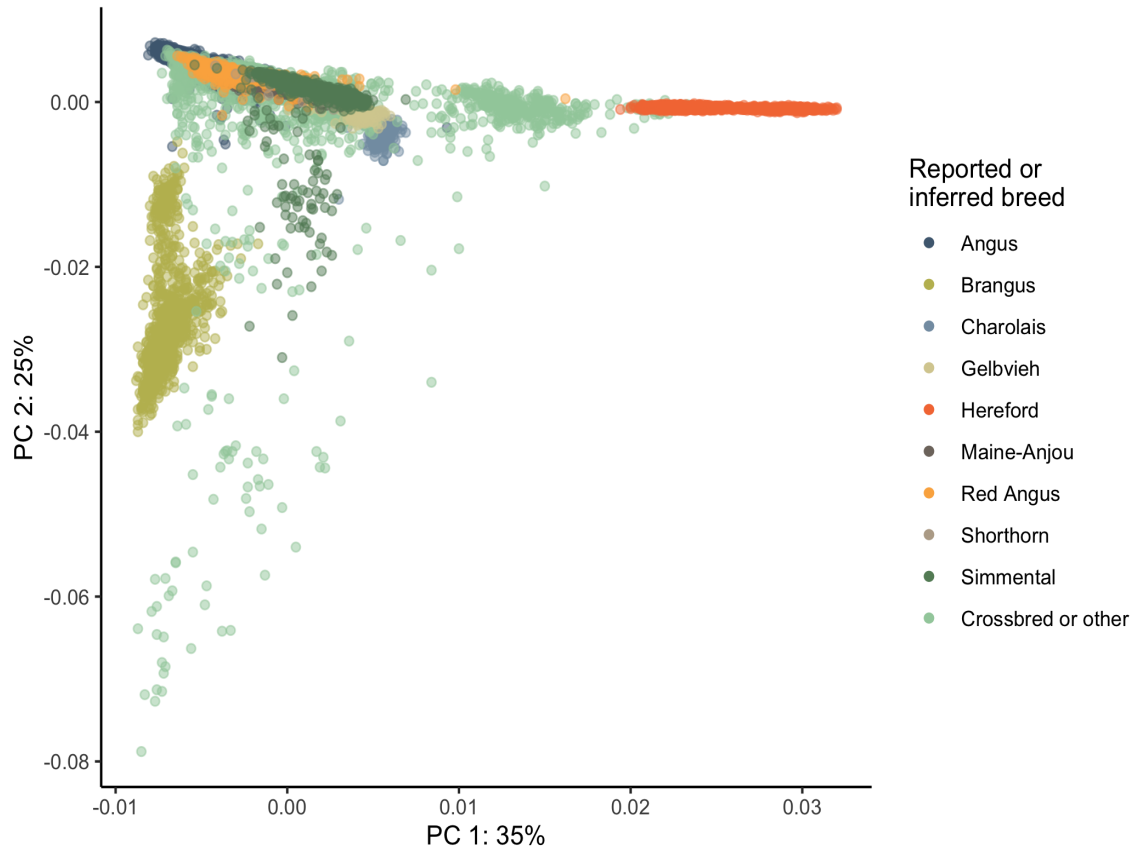
Figure S2.2. Principal components 1 and 2. For the purposes of this visualization, Angus, Hereford, Red Angus, Simmental, and Gelbvieh animals with at least ⅝ ancestry assigned to the given breed based on pedigree estimates were included in that breed. Animals with unknown ancestry, less than ⅝ ancestry assigned to one breed, or of a breed not listed above were called "Crossbred or other".
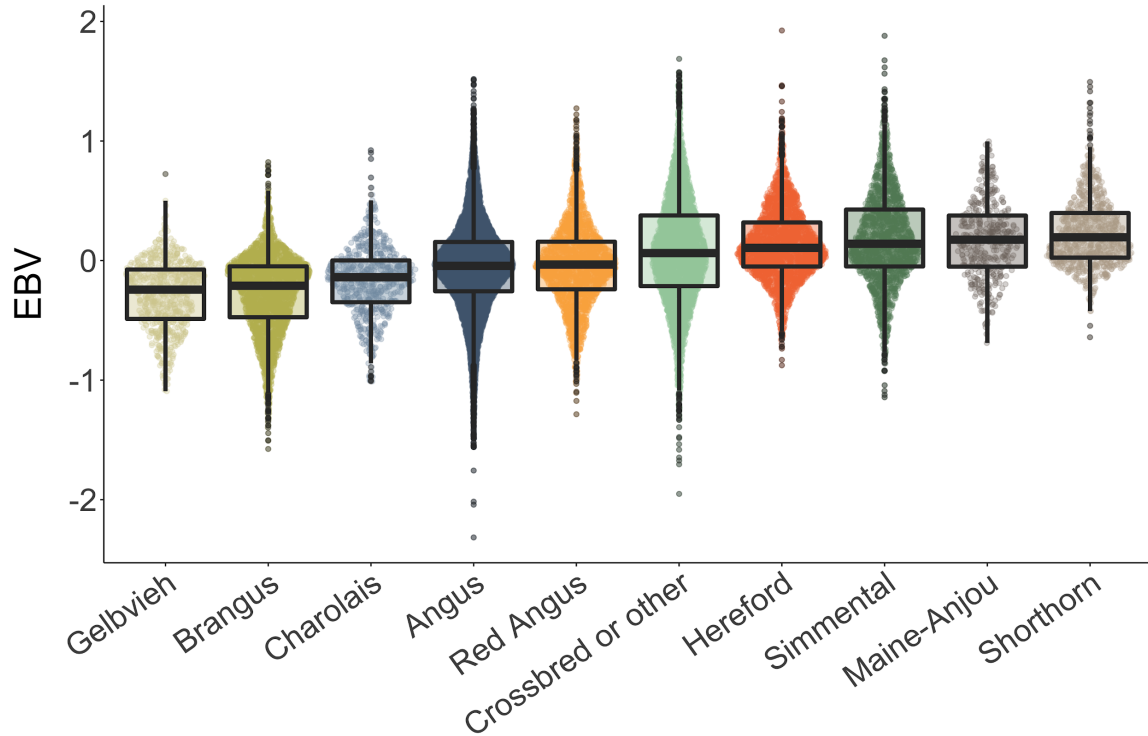
Figure S3.3. Comparison of EBVs from the full dataset by breed. For the purposes of this summarization, Angus, Hereford, Red Angus, Simmental, and Gelbvieh animals with at least ⅝ ancestry assigned to the given breed based on pedigree estimates were included in that breed. Animals with unknown ancestry, less than ⅝ ancestry assigned to one breed, or of a breed not listed above were called "Crossbred or other
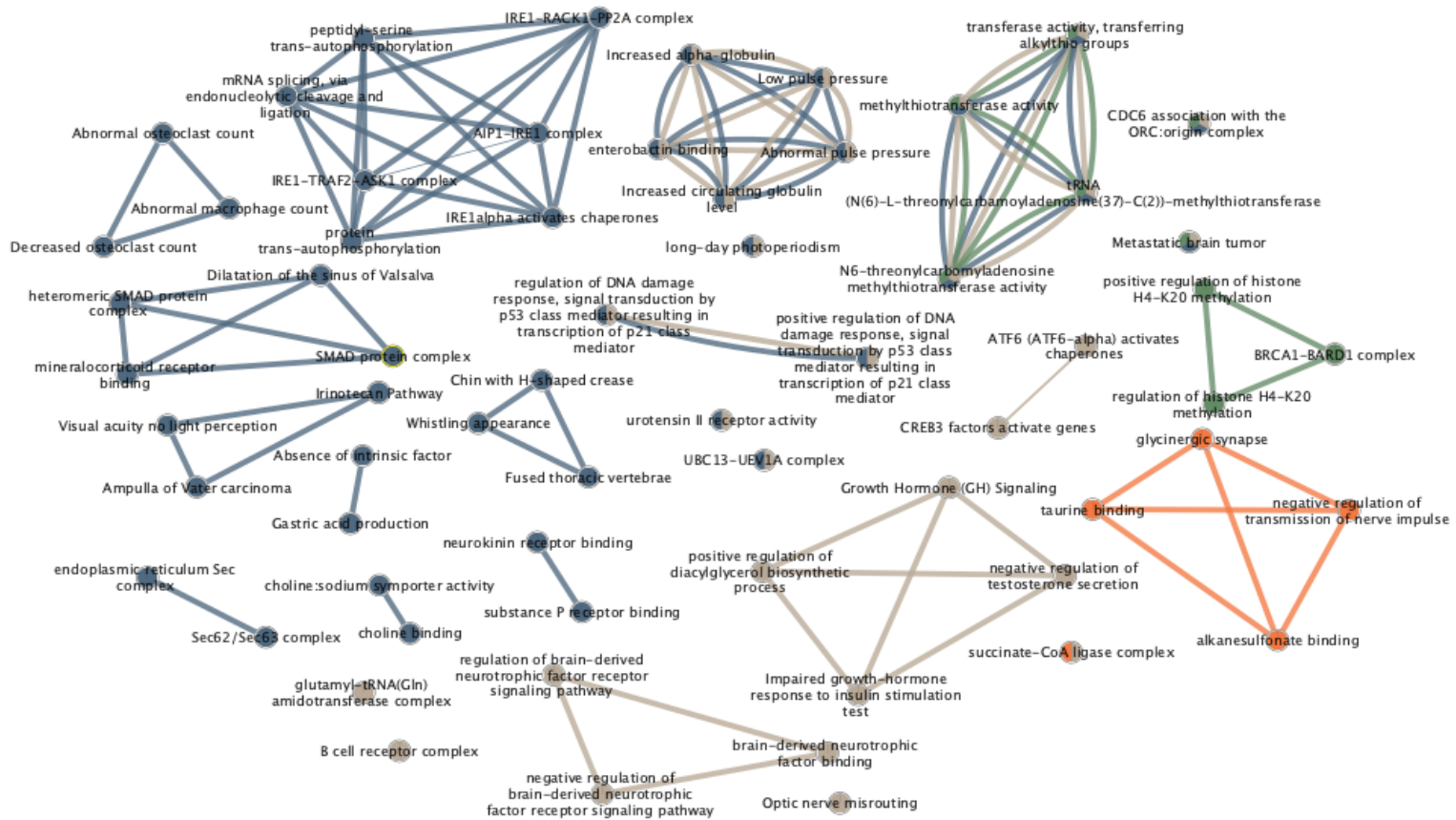
Figure S2.4. Pathway enrichment results visualized using EnrichmentMapper show that most identifiable pathways are dataset specific. Nodes represent gene sets corresponding to functional terms with adjusted p-values < 0.05 and edges represent overlap between gene sets. Edge width is determined by the degree of overlap between gene sets. Node and edge colors represent datasets where tan = full, blue = Angus, orange = Hereford, and green = IGS. No pathways were significantly enriched in the Brangus dataset. Unconnected nodes significant in only one breed-specific dataset were removed.

Chapter 2 Supplementary Tables

Table S2.1. BLUEs for toxic fescue grazing status and four increasingly complex models quantifying the effects of daily sunlight duration and temperature on hair shedding with approximated standard errors in parentheses.

Model	Fall calving	Not grazing toxic fescue	Age: 2-3 year olds	Age: 4-9 year olds	Age: 10+
Day length + covariates	-0.122 (0.020)	-0.222 (0.022)	0.091 (0.019)	-0.057 (0.021)	-0.206 (0.028)
Temperature + covariates	0.001 (0.019)	-0.055 (0.022)	0.106 (0.019)	-0.077 (0.021)	-0.231 (0.027)
Day length + temperature + covariates	-0.056 (0.019)	-0.124 (0.022)	0.100 (0.018)	-0.067 (0.021)	-0.215 (0.027)
Day length + temperature + day length*temperature + covariates	-0.061 (0.019)	-0.136 (0.022)	0.102 (0.018)	-0.061 (0.021)	-0.208 (0.027)

"Fall calving" BLUEs are relative to "spring calving" BLUE = 0, "not grazing toxic fescue" BLUEs are relative to "grazing toxic fescue" BLUE = 0, and age group BLUEs are relative to "age: yearling" BLUE = 0.

Table S2.2. Significantly enriched QTL terms in the full and breed-specific datasets.

Dataset	QTL	Adjusted p-value	N QTL in dataset	N QTL in database
Full	White spotting	3.84×10^{-27}	15	98
Full	Body weight (mature)	6.18×10^{-04}	4	161
Full	Body weight (18 months)	3.72×10^{-03}	2	21
Full	Milk tricosanoic acid content	3.97×10^{-03}	2	25
Full	Hoof and leg disorders	7.78×10^{-03}	2	39
Full	Eye area pigmentation	1.12×10^{-02}	2	51
AN	Non-return rate	2.63×10^{-67}	74	2213
AN	White spotting	5.10×10^{-40}	23	98
AN	Body weight (mature)	1.97×10^{-05}	6	161
AN	Coat color	7.31×10^{-04}	3	34
AN	Body weight (18 months)	1.10×10^{-02}	2	21
AN	Milk tricosanoic acid content	1.30×10^{-02}	2	25
AN	Hoof and leg disorders	2.71×10^{-02}	2	39
AN	Eye area pigmentation	3.21×10^{-02}	2	51
AN	Milk stearic acid content	3.21×10^{-02}	3	168
AN	Palmitoleic acid content	3.21×10^{-02}	2	46
AN	Myristoleic acid content	3.75×10^{-02}	2	68
AN	Residual feed intake	3.96×10^{-02}	4	414
HFD	Bovine respiratory disease susceptibility	3.64×10^{-02}	2	504
HFD	Gestation length	3.64×10^{-02}	2	591
IGS	Myristic acid content	2.05×10^{-02}	2	117

"N QTL in dataset" represents the number of independent QTL annotations present in Animal QTLdb within 10 kb of one or more significant SNPs.

Table S2.3. Significantly enriched functional terms based on genes within 10 kb of variants with *-log$_{10}$(q)* > 1 in the full dataset.

Term name	Adjusted p-value	Source	Significance in other datasets
UBC13-UEV1A complex	1.25E-02	GO cellular components	AN
tRNA (N(6)-L-threonylcarbamoyladenosine(37)-C(2))-methylthiotransferase	2.48E-02	GO molecular function	AN, IGS
N6-threonylcarbomyladenosine methylthiotransferase activity	2.48E-02	GO molecular function	AN, IGS
urotensin II receptor activity	2.48E-02	GO molecular function	AN
Metastatic brain tumor	2.49E-02	WikiPathways	AN, IGS
Growth Hormone (GH) Signaling	2.99E-02	WikiPathways	-
glutamyl-tRNA(Gln) amidotransferase complex	3.74E-02	GO cellular components	-
succinate-CoA ligase complex	3.74E-02	GO cellular components	HFD
ATF6 (ATF6-alpha) activates chaperones	3.75E-02	Reactome	-
negative regulation of brain-derived neurotrophic factor receptor signaling pathway	4.92E-02	GO biological processes	-
regulation of brain-derived neurotrophic factor receptor signaling pathway	4.92E-02	GO biological processes	-
negative regulation of testosterone secretion	4.92E-02	GO biological processes	-
positive regulation of diacylglycerol biosynthetic process	4.92E-02	GO biological processes	-
long-day photoperiodism	4.92E-02	GO biological processes	AN
regulation of DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	4.92E-02	GO biological processes	AN
positive regulation of DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	4.92E-02	GO biological processes	AN
enterobactin binding	4.96E-02	GO molecular function	-
methylthiotransferase activity	4.96E-02	GO molecular function	-

transferase activity, transferring alkylthio groups	4.96E-02	GO molecular function	-
brain-derived neurotrophic factor binding	4.96E-02	GO molecular function	-
B cell receptor complex	4.98E-02	GO cellular components	-
CDC6 association with the ORC:origin complex	4.99E-02	Reactome	-
CREB3 factors activate genes	4.99E-02	Reactome	-

Breed-specific datasets in which the term also had an adjusted p-value < 0.05 are indicated in the last column.

Chapter 3 Supplementary Tables

Table S3.1. Minimum, mean, and maximum genetic correlations for the direct and maternal effects of weaning weight between the High Plains and the comparison ecoregion as well as within ecoregion.

Weaning weight genetic correlation between direct effects in one ecoregion and maternal effects in a different ecoregion (r_{DiMj}), High Plains versus comparison ecoregion

Comparison ecoregion	Minimum r_{DiMj}	Mean r_{DiMj}	Maximum r_{DiMj}
Desert	-0.45	-0.31	-0.25
Southeast	-0.50	-0.38	-0.27
Arid Prairie	-0.33	-0.27	-0.14
Forested Mountains	-0.36	-0.21	-0.08
Fescue Belt	-0.47	-0.32	-0.21
Upper Midwest & Northeast	-0.40	-0.26	-0.08

Genetic correlation between direct and maternal effects of weaning weight (r_{DiMi}), within ecoregion

Comparison ecoregion	Minimum r_{DiMi}	Mean r_{DiMi}	Maximum r_{DiMi}
Desert	-0.41	-0.38	-0.35
Southeast	-0.50	-0.42	-0.38
High Plains	-0.47	-0.30	-0.14
Arid Prairie	-0.43	-0.39	-0.33
Forested Mountains	-0.38	-0.26	-0.17
Fescue Belt	-0.47	-0.38	-0.31
Upper Midwest & Northeast	-0.42	-0.36	-0.26

Estimates for the High Plains are across all 60 iterations while estimates in the other ecoregions are across 10 iterations.

Table S3.2. QTL terms significantly enriched within ecoregions and associated genes for the direct and maternal effects of weaning weight.

QTL term	Adjusted p-value	Ecoregion	Associated genes
Direct effect			
Birth index	0.000401	Fescue Belt	*ACTN4, ARRDC3, CAPN12, KPTN, MEIS3, NAPA, OR6K2, OR6K3, OR6K6, OR6N1, RF00100, SLC8A2, SPTA1*
Body weight (birth)	2.99E-05	Fescue Belt	*ARRDC3, ATP6V0E1, BNIP1, CREBRF, ENC1, ERGIC1, FAM184B, KCNIP4, LAP3, MED28, NEURL1B, NKX2-5, RF00001, RF00100, RPL26L1, SLIT2, STC2*
Body weight (weaning)	0.027859	Southeast	***ARRDC3, ATP6V0E1, BNIP1, bta-mir-584-6, CREBRF, ERGIC1, GFM2, HEXB, MSX2, NKX2-5, RF00001, RPL26L1, STC2***
Body weight (weaning)	0.021709	High Plains	***ARRDC3***, *ASNS*, ***ATP6V0E1, BNIP1, bta-mir-584-6, CREBRF, ENC1, ERGIC1, GFM2, HEXB***, *HIST3H2A, HIST3H2BB*, ***MSX2, NEURL1B, NKX2-5, RF00001***, *RF00322, RNF187*, ***RPL26L1, STC2***
Body weight (weaning)	1.9E-09	Fescue Belt	*ARHGEF2*, ***ARRDC3, ATP6V0E1, BNIP1, bta-mir-584-6, CREBRF***, *DYNC1I1*, ***ENC1, ERGIC1***, *FAM184B, KCNIP4, LAMTOR2, LAP3, MED28*, ***NEURL1B, NKX2-5***, *OR6K2, OR6K3, OR6K6, OR6N1, RAB25*, ***RF00001***, *RPL26L1, SLIT2, SPTA1, SSR2*, ***STC2***, *UBQLN4*
Body weight (yearling)	0.008516	Fescue Belt	*ARRDC3, ATP6V0E1, BNIP1, CREBRF, ERGIC1, FAM184B, KCNIP4, LAP3, MED28, NEURL1B, NKX2-5, RF00001, RPL26L1, SLIT2, STC2*
C22:1 fatty acid content	0.008885	Fescue Belt	*MPP7*
Carcass weight	6.53E-05	Fescue Belt	*ARRDC3, ATP6V0E1, BNIP1, BOD1, CREBRF, ERGIC1, FAM184B, LAP3, MED28, NEURL1B, NKX2-5, RF00001, RPL26L1, STC2*
Lignoceric acid content	0.013241	Fescue Belt	*VWA5B1*
Longissimus muscle area	0.041984	Fescue Belt	*ACTN4, ARRDC3, CAPN12, FAM184B, LAP3, MED28*
Yield grade	0.024483	High Plains	***ATP6V0E1, BNIP1, BOD1, CREBRF, ERGIC1, NEURL1B, NKX2-5, RF00001, RF00100, RPL26L1, STC2***
Yield grade	0.008516	Fescue Belt	***ATP6V0E1, BNIP1, BOD1, CREBRF, ERGIC1, NEURL1B, NKX2-5, RF00001, RF00100, RPL26L1, STC2***

Maternal effect			
Bovine viral diarrhea virus susceptibility	0.001161	High Plains	*LIN54, SEC31A, THAP9*
Calving index	0.014075	Fescue Belt	*HUNK, VIP*
Long-chain fatty acid content	0.03898	Fescue Belt	*MYO9A*
Medium-chain fatty acid content	0.03898	Fescue Belt	*MYO9A*
Shear force	0.005342	High Plains	*LIN54, SEC31A, THAP9*

Bolded gene names are shared across multiple ecoregions within a QTL term.

REFERENCES

1. Kay RNB, Ryder ML. Coat growth in Red deer (Cervus elaphus) exposed to a day-length cycle of six months duration. J Zool [Internet]. Wiley Online Library; 1978;185:505–10. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7998.1978.tb03348.x

2. Harvey NE, Macfarlane WV. The Effects of Day Length on the Coat-Shedding Cycles, Body Weight, and Reproduction of the Ferret. Aust Jnl Of Bio Sci [Internet]. CSIRO PUBLISHING; 1958 [cited 2019 Feb 28];11:187–99. Available from: http://www.publish.csiro.au/bi/BI9580187

3. Ryder ML. Coat structure and seasonal shedding in goats. Anim Sci [Internet]. Cambridge University Press; 1966 [cited 2019 Feb 28];8:289–302. Available from: https://www.cambridge.org/core/journals/animal-science/article/coat-structure-and-seasonal-shedding-in-goats/49915A9847CDD39DFD998FA885167475

4. Hazlerigg DG, Wagner GC. Seasonal photoperiodism in vertebrates: from coincidence to amplitude. Trends Endocrinol Metab [Internet]. 2006;17:83–91. Available from: http://dx.doi.org/10.1016/j.tem.2006.02.004

5. Heydon MJ, Milne JA, Brinklow BR, Loudon AS. Manipulating melatonin in red deer (Cervus elaphus): differences in the response to food restriction and lactation on the timing of the breeding season and prolactin-dependent pelage changes. J Exp Zool [Internet]. 1995;273:12–20. Available from: http://dx.doi.org/10.1002/jez.1402730103

6. Beltran RS, Burns JM, Breed GA. Convergence of biannual moulting strategies across birds and mammals. Proc Biol Sci [Internet]. 2018;285. Available from: http://dx.doi.org/10.1098/rspb.2018.0318

7. Matika O, Bishop SC, Pong-Wong R, Riggio V, Headon DJ. Genetic factors controlling wool shedding in a composite Easycare sheep flock. Anim Genet [Internet]. 2013;44:742–9. Available from: http://dx.doi.org/10.1111/age.12070

8. Déry F, Hamel S, Côté SD. Getting ready for the winter: Timing and determinants of molt in an alpine ungulate. Ecol Evol [Internet]. 2019;9:2920–32. Available from: http://dx.doi.org/10.1002/ece3.4970

9. Yeates NTM. Photoperiodicity in cattle. I. Seasonal changes in coat character and their importance in heat regulation. Aust J Agric Res [Internet]. CSIRO PUBLISHING; 1955 [cited 2019 Feb 28];6:891–902. Available from: http://www.publish.csiro.au/cp/AR9550891

10. Baumgard LH, Keating A, Ross JW, Rhoads RP. Effects of heat stress on the immune system, metabolism and nutrient partitioning: implications on reproductive success. Anim Reprod [Internet]. 2015;173–83. Available from: http://www.cbra.org.br/pages/publicacoes/rbra/v39n1/pag173-183%20(RB545).pdf

11. St-Pierre NR, Cobanov B, Schnitkey G. Economic Losses from Heat Stress by US Livestock Industries. J Dairy Sci [Internet]. 2003;86:E52–77. Available from: http://www.sciencedirect.com/science/article/pii/S0022030203740405

12. Berry IL, Shanklin MD, Johnson HD. Dairy Shelter Design Based on Milk Production Decline as Affected by Temperature and Humidity. Trans ASAE [Internet]. St. Joseph, MI: ASAE; 1964;7:329–0331. Available from: http://elibrary.asabe.org/abstract.asp?aid=40772&t=3

13. Hoveland CS, Fribourg HA, Hannaway DB, West CP. Origin and History [Internet]. Agronomy Monographs. 2009. Available from: http://dx.doi.org/10.2134/agronmonogr53.c1

14. Clay K. Fungal Endophytes of Grasses: A Defensive Mutualism between Plants and Fungi. Ecology [Internet]. 1988;69:10–6. Available from: http://doi.wiley.com/10.2307/1943155

15. Aiken GE, Klotz JL, Looper ML, Tabler SF, Schrick FN. Disrupted hair follicle activity in cattle grazing endophyte-infected tall fescue in the summer insulates core body temperatures. The Professional Animal Scientist [Internet]. 2011;27:336–43. Available from: http://www.sciencedirect.com/science/article/pii/S1080744615304976

16. Dikmen S, Khan FA, Huson HJ, Sonstegard TS, Moss JI, Dahl GE, et al. The SLICK hair locus derived from Senepol cattle confers thermotolerance to intensively managed lactating Holstein cows. J Dairy Sci [Internet]. 2014;97:5508–20. Available from: http://dx.doi.org/10.3168/jds.2014-8087

17. Smith JL, Wilson ML, Nilson SM, Rowan TN, Oldeschulte DL, Schnabel RD, et al. Genome-wide association and genotype by environment interactions for

growth traits in U.S. Gelbvieh cattle. BMC Genomics [Internet]. 2019;20:926. Available from: http://dx.doi.org/10.1186/s12864-019-6231-y

18. Bradford HL, Fragomeni BO, Bertrand JK, Lourenco DAL, Misztal I. Genetic evaluations for growth heat tolerance in Angus cattle. J Anim Sci [Internet]. 2016;94:4143–50. Available from: http://dx.doi.org/10.2527/jas.2016-0707

19. Rowan TN, Durbin HJ, Seabury CM, Schnabel RD, Decker JE. Powerful detection of polygenic selection and environmental adaptation in US beef cattle [Internet]. bioRxiv. 2020 [cited 2020 May 13]. p. 2020.03.11.988121. Available from: https://www.biorxiv.org/content/10.1101/2020.03.11.988121v3

20. Pauling RC, Speidel SE, Thomas MG, Holt TN, Enns RM. Evaluation of moderate to high elevation effects on pulmonary arterial pressure measures in Angus cattle. J Anim Sci [Internet]. 2018;96:3599–605. Available from: http://dx.doi.org/10.1093/jas/sky262

21. Turner HG, Schleger AV. The significance of coat type in cattle. Aust J Agric Res [Internet]. CSIRO PUBLISHING; 1960 [cited 2020 Feb 9];11:645–63. Available from: http://www.publish.csiro.au/CP/AR9600645

22. Gray KA, Smith T, Maltecca C, Overton P, Cassady JP. Differences in hair coat shedding, and effects on calf weaning weight and BCS among Angus dams. Livest Sci [Internet]. 2011 [cited 2019 Feb 10];140:68–71. Available from: https://www.researchgate.net/publication/251707867_Differences_in_hair_coat_s hedding_and_effects_on_calf_weaning_weight_and_BCS_among_Angus_dams

23. Nay T, Fraser AS. Growth of the mouse coat. III. Patterns of hair growth. Aust J Biol Sci [Internet]. 1954;7:361–7. Available from: http://dx.doi.org/10.1071/bi9540361

24. Cundiff LV, Van Vleck LD, Hohenboken WD, editors. Guidelines For Uniform Beef Improvement Programs [Internet]. Beef Improvement Federation; 2018. Report No.: 9. Available from: https://beefimprovement.org/wp-content/uploads/2018/03/BIFGuidelinesFinal_updated0318.pdf

25. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. BMC Genomics [Internet]. 2014;15:478. Available from: https://doi.org/10.1186/1471-2164-15-478

26. Rowan TN, Hoff JL, Crum TE, Taylor JF, Schnabel RD, Decker JE. A multi-breed reference panel and additional rare variants maximize imputation accuracy in cattle. Genet Sel Evol [Internet]. 2019;51:77. Available from: http://dx.doi.org/10.1186/s12711-019-0519-x

27. Aguilar I, Misztal I, Legarra A, Tsuruta S. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J Anim Breed Genet [Internet]. 2011;128:422–8. Available from: http://dx.doi.org/10.1111/j.1439-0388.2010.00912.x

28. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci [Internet]. 2008;91:4414–23. Available from: http://dx.doi.org/10.3168/jds.2007-0980

29. Misztal I, Tsuruta S, Lourenco D, Masuda Y, Aguilar I, Legarra A, et al. BLUPF90 family of programs [Internet]. University of Georgia; 2014. Available from: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf

30. Legarra A, Reverter A. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. Genet Sel Evol [Internet]. 2018;50:53. Available from: http://dx.doi.org/10.1186/s12711-018-0426-6

31. Peters CW, Grigsby KN, Aldrich CG, Paterson JA, Lipsey RJ, Kerley MS, et al. Performance, forage utilization, and ergovaline consumption by beef cows grazing endophyte fungus-infected tall fescue, endophyte fungus-free tall fescue, or orchardgrass pastures. J Anim Sci [Internet]. 1992;70:1550–61. Available from: http://dx.doi.org/10.2527/1992.7051550x

32. Miller SP, Wilton JW. Genetic relationships among direct and maternal components of milk yield and maternal weaning gain in a multibreed beef herd. J Anim Sci [Internet]. Narnia; 1999 [cited 2020 Feb 4];77:1155–61. Available from: https://academic.oup.com/jas/article-pdf/77/5/1155/22624609/1155.pdf

33. Sargolazei M. SNP1101 User's Guide Version 1.0. HiggsGene Solutions Inc.; 2014.

34. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol [Internet]. Springer; 2009;10:R42. Available from: http://dx.doi.org/10.1186/gb-

35. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics [Internet]. 2009;25:1091–3. Available from: http://dx.doi.org/10.1093/bioinformatics/btp101

36. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res [Internet]. 2015;43:D447-52. Available from: http://dx.doi.org/10.1093/nar/gku1003

37. Lee C. On the Negative Estimates of Direct and Maternal Genetic Correlation - A Review. Asian-australas J Anim Sci [Internet]. Asian-Australasian Association of Animal Production Societies; 2002 [cited 2020 Apr 2];15:1222–6. Available from: http://www.ajas.info/journal/view.php?number=20251

38. Falconer DS, Mackay TFC. Introduction to quantitative genetics [Internet]. 4th ed. Harlow: Prentice Hall; 1996. Available from: https://openlibrary.org/books/OL22815221M.opds

39. Meyer K. Estimates of genetic parameters for mature weight of Australian beef cows and its relationship to early growth and skeletal measures. Livestock Production Science [Internet]. 1995;44:125–37. Available from: http://www.sciencedirect.com/science/article/pii/0301622695000674

40. Freetly HC, Nienaber JA, Brown-Brandl T. Partitioning of energy during

lactation of primiparous beef cows. J Anim Sci [Internet]. 2006;84:2157–62. Available from: http://dx.doi.org/10.2527/jas.2005-534

41. Schons D, Hohenboken WD, Hall JD. Population Analysis of a Commercial Beef Cattle Herd. J Anim Sci [Internet]. Oxford Academic; 1985 [cited 2020 Apr 2];61:44–54. Available from: https://academic.oup.com/jas/article-pdf/61/1/44/22985636/jan0610010044.pdf

42. Azzam SM, Azzam AM, Nielsen MK, Kinder JE. Markov chains as a shortcut method to estimate age distributions in herds of beef cattle under different culling strategies. J Anim Sci [Internet]. 1990;68:5–14. Available from: http://dx.doi.org/10.2527/1990.6815

43. Strickland JR, Bailey EM, Abney LK, Oliver JW. Assessment of the mitogenic potential of the alkaloids produced by endophyte (Acremonium coenophialum)-infected tall fescue (Festuca arundinacea) on bovine vascular smooth muscle in vitro. J Anim Sci [Internet]. 1996;74:1664–71. Available from: http://dx.doi.org/10.2527/1996.7471664x

44. Klotz JL, Kirch BH, Aiken GE, Bush LP, Strickland JR. Bioaccumulation of ergovaline in bovine lateral saphenous veins in vitro. J Anim Sci [Internet]. 2009;87:2437–47. Available from: http://dx.doi.org/10.2527/jas.2008-1692

45. Falconer DS. The Problem of Environment and Selection. Am Nat [Internet]. [University of Chicago Press, American Society of Naturalists]; 1952;86:293–8. Available from: http://www.jstor.org/stable/2457811

46. Guerre P. Ergot alkaloids produced by endophytic fungi of the genus Epichloë. Toxins [Internet]. 2015;7:773–90. Available from: http://dx.doi.org/10.3390/toxins7030773

47. Fribourg HA, Chestnut AB, Thompson RW, McLaren JB, Carlisle RJ, Gwinn KD, et al. Steer Performance in Fescue-Clover Pastures with Different Levels of Endophyte Infestation. Agron J [Internet]. 1991;83:777. Available from: https://www.agronomy.org/publications/aj/abstracts/83/5/AJ0830050777

48. Littlejohn MD, Henty KM, Tiplady K, Johnson T, Harland C, Lopdell T, et al. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nat Commun [Internet]. 2014;5:5861. Available from: http://dx.doi.org/10.1038/ncomms6861

49. Strickland JR, Oliver JW, Cross DL. Fescue toxicosis and its impact on animal agriculture. Vet Hum Toxicol [Internet]. 1993;35:454–64. Available from: https://www.ncbi.nlm.nih.gov/pubmed/8249272

50. Neibergs HL, Seabury CM, Wojtowicz AJ, Wang Z, Scraggs E, Kiser JN, et al. Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. BMC Genomics [Internet]. 2014;15:1164. Available from: http://dx.doi.org/10.1186/1471-2164-15-1164

51. Santana ML, Eler JP, Cardoso FF, Albuquerque LG, Ferraz JBS. Phenotypic plasticity of composite beef cattle performance using reaction norms model with unknown covariate. Animal [Internet]. 2013;7:202–10. Available from:

http://dx.doi.org/10.1017/S1751731112001711

52. Carvalheiro R, Costilla R, Neves HHR, Albuquerque LG, Moore S, Hayes BJ. Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. Genet Sel Evol [Internet]. 2019;51:29. Available from: http://dx.doi.org/10.1186/s12711-019-0470-x

53. Lee C, Pollak EJ. Genetic antagonism between body weight and milk production in beef cattle. J Anim Sci [Internet]. 2002;80:316–21. Available from: http://dx.doi.org/10.2527/2002.802316x

54. Lee C. An analytical approach to sire-by-year interactions in direct and maternal genetic evaluation. Asian-australas J Anim Sci [Internet]. Asian-Australasian Association of Animal Production Societies; 1998 [cited 2020 Apr 28];11:441–4. Available from: https://www.ajas.info/journal/view.php?number=19307

55. Bauman DE, Currie WB. Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. J Dairy Sci [Internet]. 1980;63:1514–29. Available from: http://dx.doi.org/10.3168/jds.s0022-0302(80)83111-0

56. Baumgard LH, Collier RJ, Bauman DE. A 100-Year Review: Regulation of nutrient partitioning to support lactation. J Dairy Sci [Internet]. 2017;100:10353–66. Available from: http://dx.doi.org/10.3168/jds.2017-13242

57. Rhoads ML, Rhoads RP, VanBaale MJ, Collier RJ, Sanders SR, Weber WJ,

et al. Effects of heat stress and plane of nutrition on lactating Holstein cows: I. Production, metabolism, and aspects of circulating somatotropin. J Dairy Sci [Internet]. 2009;92:1986–97. Available from: http://dx.doi.org/10.3168/jds.2008-1641

58. Burnight ER, Wiley LA, Drack AV, Braun TA, Anfinson KR, Kaalberg EE, et al. CEP290 gene transfer rescues Leber congenital amaurosis cellular phenotype. Gene Ther [Internet]. 2014;21:662–72. Available from: http://dx.doi.org/10.1038/gt.2014.39

59. Coppieters F, Lefever S, Leroy BP, De Baere E. CEP290, a gene with many faces: mutation overview and presentation of CEP290base. Hum Mutat [Internet]. 2010;31:1097–108. Available from: http://dx.doi.org/10.1002/humu.21337

60. Peirson SN, Brown LA, Pothecary CA, Benson LA, Fisk AS. Light and the laboratory mouse. J Neurosci Methods [Internet]. 2018;300:26–36. Available from: http://dx.doi.org/10.1016/j.jneumeth.2017.04.007

61. Zhang Y, Seo S, Bhattarai S, Bugge K, Searby CC, Zhang Q, et al. BBS mutations modify phenotypic expression of CEP290-related ciliopathies. Hum Mol Genet [Internet]. 2014;23:40–51. Available from: http://dx.doi.org/10.1093/hmg/ddt394

62. Cannon ME, Mohlke KL. Deciphering the Emerging Complexities of Molecular Mechanisms at GWAS Loci. Am J Hum Genet [Internet]. 2018;103:637–53. Available from: http://dx.doi.org/10.1016/j.ajhg.2018.10.001

63. Carabaño MJ, Ramón M, Menéndez-Buxadera A, Molina A, Díaz C. Selecting for heat tolerance. Anim Front [Internet]. 2019;9:62–8. Available from: http://dx.doi.org/10.1093/af/vfy033

64. Robinson TP, Wint GRW, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, et al. Mapping the global distribution of livestock. PLoS One [Internet]. 2014;9:e96084. Available from: http://dx.doi.org/10.1371/journal.pone.0096084

65. Helm B, Ben-Shlomo R, Sheriff MJ, Hut RA, Foster R, Barnes BM, et al. Annual rhythms that underlie phenology: biological time-keeping meets environmental change. Proc Biol Sci [Internet]. 2013;280:20130016. Available from: http://dx.doi.org/10.1098/rspb.2013.0016

66. Murray DM. A field study of coat shedding in cattle under conditions of equal day-length but different temperatures. J Agric Sci [Internet]. Cambridge University Press; 1965 [cited 2020 Nov 12];65:295–300. Available from: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/EE1C48F19A860EB5A5DAA5DADC0013E5/S002185960004 8863a.pdf/div-class-title-a-field-study-of-coat-shedding-in-cattle-under-conditions-of-equal-day-length-but-different-temperatures-div.pdf

67. Cowan IM, McT. Cowan I, Raddi AG. Pelage and molt in the black-tailed deer (Odocoileus hemionus (Rafinesque)). Can J Zool [Internet]. 1972;50:639–47. Available from: http://dx.doi.org/10.1139/z72-088

68. Zimova M, Mills LS, Lukacs PM, Mitchell MS. Snowshoe hares display limited phenotypic plasticity to mismatch in seasonal camouflage. Proc Biol Sci [Internet]. royalsocietypublishing.org; 2014;281:20140029. Available from: http://dx.doi.org/10.1098/rspb.2014.0029

69. Turner HG. Coat characters of cattle in relation to adaptation. Proceedings of the Australian Society of Animal Production [Internet]. 1964. p. 181–7. Available from: https://pdfs.semanticscholar.org/afaf/90c8979ed17e1c79714801b5a96ebf628675 .pdf

70. Ferreira MS, Alves PC, Callahan CM, Giska I, Farelo L, Jenny H, et al. Transcriptomic regulation of seasonal coat color change in hares. Ecol Evol [Internet]. 2020;10:1180–92. Available from: http://dx.doi.org/10.1002/ece3.5956

71. Ferreira MS, Alves PC, Callahan CM, Marques JP, Mills LS, Good JM, et al. The transcriptional landscape of seasonal coat colour moult in the snowshoe hare. Mol Ecol [Internet]. 2017;26:4173–85. Available from: http://dx.doi.org/10.1111/mec.14177

72. Durbin HJ, Lu D, Yampara-Iquise H, Miller SP, Decker JE. Development of a genetic evaluation for hair shedding in American Angus cattle to improve thermotolerance. Genet Sel Evol [Internet]. 2020;52:63. Available from: https://doi.org/10.1186/s12711-020-00584-0

73. Hayes BJ. Efficient parentage assignment and pedigree reconstruction with

dense single nucleotide polymorphism data. J Dairy Sci [Internet]. 2011;94:2114–7. Available from: http://dx.doi.org/10.3168/jds.2010-3896

74. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet [Internet]. 2006;2:e190. Available from: http://dx.doi.org/10.1371/journal.pgen.0020190

75. Alvarenga AB, Veroneze R, Oliveira HR, Marques DBD, Lopes PS, Silva FF, et al. Comparing Alternative Single-Step GBLUP Approaches and Training Population Designs for Genomic Evaluation of Crossbred Animals. Front Genet [Internet]. 2020;11:263. Available from: http://dx.doi.org/10.3389/fgene.2020.00263

76. Dong MC, Van Vleck LD, Wiggans GR. Effect of Relationships on Estimation of Variance Components with an Animal Model and Restricted Maximum Likelihood. J Dairy Sci [Internet]. Elsevier; 1988;71:3047–52. Available from: https://doi.org/10.3168/jds.S0022-0302(88)79904-X

77. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/

78. Cambon J. tidygeocoder: Geocoding Made Easy [Internet]. 2020. Available from: https://CRAN.R-project.org/package=tidygeocoder

79. Rudis B. darksky: Tools to Work with the "Dark Sky" "API" [Internet]. 2017. Available from: https://CRAN.R-project.org/package=darksky

80. Ekine CC, Rowe SJ, Bishop SC, de Koning D-J. Why breeding values estimated using familial data should not be used for genome-wide association studies. G3 [Internet]. 2014;4:341–7. Available from: http://dx.doi.org/10.1534/g3.113.008706

81. Aguilar I, Fernandez EN, Blasco A, Ravagnolo O, Legarra A. Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP. J Anim Breed Genet [Internet]. 2020; Available from: http://dx.doi.org/10.1111/jbg.12470

82. Wellmann R. optiSel: Optimum Contribution Selection and Population Genetics [Internet]. 2020. Available from: https://CRAN.R-project.org/package=optiSel

83. Garrick DJ, Taylor JF, Fernando RL. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genet Sel Evol [Internet]. 2009;41:55. Available from: http://dx.doi.org/10.1186/1297-9686-41-55

84. Lopes FB. DRP: Deregressing estimated breeding values [Internet]. 2016. Available from: https://github.com/camult/DRP

85. Storey JD, Bass AJ, Dabney A, Robinson D. qvalue: Q-value estimation for false discovery rate control [Internet]. 2017. Available from: http://github.com/StoreyLab/qvalue

86. Fonseca P, Suarez-Vega A, Marras G, Cánovas A. GALLO: Genomic Annotation in Livestock for Positional Candidate LOci [Internet]. 2020. Available from: https://CRAN.R-project.org/package=GALLO

87. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience [Internet]. 2020;9. Available from: http://dx.doi.org/10.1093/gigascience/giaa021

88. Hu Z-L, Park CA, Reecy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Res [Internet]. 2019;47:D701–10. Available from: http://dx.doi.org/10.1093/nar/gky1084

89. Kolberg L, Raudvere U. gprofiler2: Interface to the "g:Profiler" Toolset [Internet]. 2020. Available from: https://CRAN.R-project.org/package=gprofiler2

90. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment Map: A Network-Based Method for Gene-Set Enrichment Visualization and Interpretation. PLoS One [Internet]. Public Library of Science; 2010 [cited 2020 Nov 9];5:e13984. Available from: https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0013984&type=printable

91. Decker JE, McKay SD, Rolf MM, Kim J, Molina Alcalá A, Sonstegard TS, et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. PLoS Genet [Internet]. 2014;10:e1004254. Available from: http://dx.doi.org/10.1371/journal.pgen.1004254

92. Steingrímsson E, Copeland NG, Jenkins NA. Melanocytes and the

microphthalmia transcription factor network. Annu Rev Genet [Internet].

2004;38:365–411. Available from:

http://dx.doi.org/10.1146/annurev.genet.38.072902.092717

93. Levy C, Khaled M, Fisher DE. MITF: master regulator of melanocyte

development and melanoma oncogene. Trends Mol Med [Internet]. 2006;12:406–

14. Available from: http://dx.doi.org/10.1016/j.molmed.2006.07.008

94. Hou L, Pavan WJ. Transcriptional and signaling regulation in neural crest

stem cell-derived melanocyte development: do all roads lead to Mitf? Cell Res

[Internet]. 2008;18:1163–76. Available from:

http://dx.doi.org/10.1038/cr.2008.303

95. Maurel D, Coutant C, Boissin J. Effects of photoperiod, melatonin implants

and castration on molting and on plasma thyroxine, testosterone and prolactin

levels in the European badger (Meles meles). Comp Biochem Physiol A Comp

Physiol [Internet]. 1989;93:791–7. Available from: http://dx.doi.org/10.1016/0300-

9629(89)90503-3

96. Lincoln GA, Ebling FJ. Effect of constant-release implants of melatonin on

seasonal cycles in reproduction, prolactin secretion and moulting in rams. J

Reprod Fertil [Internet]. rep.bioscientifica.com; 1985;73:241–53. Available from:

http://dx.doi.org/10.1530/jrf.0.0730241

97. Boissin-Agasse L, Maurel D, Boissin J. Seasonal variations in thyroxine and

testosterone levels in relation to the moult in the adult male mink (Mustela vison

Peale and Beauvois). Can J Zool [Internet]. NRC Research Press; 1981;59:1062–6. Available from: https://doi.org/10.1139/z81-147

98. Tassabehji M, Newton VE, Read AP. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. Nat Genet [Internet]. 1994;8:251–5. Available from: https://doi.org/10.1038/ng1194-251

99. Amiel J, Watkin PM, Tassabehji M, Read AP, Winter RM. Mutation of the MITF gene in albinism-deafness syndrome (Tietz syndrome). Clin Dysmorphol [Internet]. 1998;7:17–20. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9546825

100. Philipp U, Lupp B, Mömke S, Stein V, Tipold A, Eule JC, et al. A MITF mutation associated with a dominant white phenotype and bilateral deafness in German Fleckvieh cattle. PLoS One [Internet]. 2011;6:e28857. Available from: http://dx.doi.org/10.1371/journal.pone.0028857

101. Wen B, Li S, Li H, Chen Y, Ma X, Wang J, et al. Microphthalmia-associated transcription factor regulates the visual cycle genes Rlbp1 and Rdh5 in the retinal pigment epithelium. Sci Rep [Internet]. 2016;6:21208. Available from: http://dx.doi.org/10.1038/srep21208

102. García-Llorca A, Aspelund SG, Ogmundsdottir MH, Steingrimsson E, Eysteinsson T. The microphthalmia-associated transcription factor (Mitf) gene and its role in regulating eye function. Sci Rep [Internet]. 2019;9:15386. Available from: http://dx.doi.org/10.1038/s41598-019-51819-0

103. Han S, Chen J, Hua J, Hu X, Jian S, Zheng G, et al. MITF protects against oxidative damage-induced retinal degeneration by regulating the NRF2 pathway in the retinal pigment epithelium. Redox Biol [Internet]. 2020;34:101537. Available from: http://dx.doi.org/10.1016/j.redox.2020.101537

104. Reiter RJ. Pineal melatonin: cell biology of its synthesis and of its physiological interactions. Endocr Rev [Internet]. 1991;12:151–80. Available from: http://dx.doi.org/10.1210/edrv-12-2-151

105. Uz E, Alanay Y, Aktas D, Vargel I, Gucer S, Tuncbilek G, et al. Disruption of ALX1 causes extreme microphthalmia and severe facial clefting: expanding the spectrum of autosomal-recessive ALX-related frontonasal dysplasia. Am J Hum Genet [Internet]. 2010;86:789–96. Available from: http://dx.doi.org/10.1016/j.ajhg.2010.04.002

106. Mészáros G, Petautschnig E, Schwarzenbacher H, Sölkner J. Genomic regions influencing coat color saturation and facial markings in Fleckvieh cattle. Anim Genet [Internet]. 2015;46:65–8. Available from: http://dx.doi.org/10.1111/age.12249

107. Tucker HA, Petitclerc D, Zinn SA. The influence of photoperiod on body weight gain, body composition, nutrient intake and hormone secretion. J Anim Sci [Internet]. 1984;59:1610–20. Available from: http://dx.doi.org/10.2527/jas1984.5961610x

108. Peters RR, Tucker HA. Prolactin and growth hormone responses to

photoperiod in heifers. Endocrinology [Internet]. 1978;103:229–34. Available

from: http://dx.doi.org/10.1210/endo-103-1-229

109. Petitclerc D, Chapin LT, Emery RS, Tucker HA. Body Growth, Growth

Hormone, Prolactin and Puberty Response to Photoperiod and Plane of Nutrition

in Holstein Heifers. J Anim Sci [Internet]. Oxford Academic; 1983 [cited 2020 Nov

18];57:892–8. Available from: https://academic.oup.com/jas/article-

abstract/57/4/892/4665271

110. Dahl GE, Buchanan BA, Tucker HA. Photoperiodic Effects on Dairy Cattle:

A Review. J Dairy Sci [Internet]. 2000;83:885–93. Available from:

http://www.sciencedirect.com/science/article/pii/S0022030200749526

111. Small JA, Glover ND, Kennedy AD, McCaughey WP, Ward DR. Photoperiod

effects on the development of beef heifers. Can J Anim Sci [Internet]. NRC

Research Press; 2003;83:721–30. Available from: https://doi.org/10.4141/A03-

005

112. Barenton B, Ravault J-P, Chabanet C, Daveau A, Pelletier J, Ortavant R.

Photoperiodic control of growth hormone secretion and body weight in rams.

Domest Anim Endocrinol [Internet]. 1988;5:247–55. Available from:

http://www.sciencedirect.com/science/article/pii/0739724088900367

113. Hayes BJ, Bowman PJ, Chamberlain AJ, Savin K, van Tassell CP,

Sonstegard TS, et al. A validated genome wide association study to breed cattle

adapted to an environment altered by climate change. PLoS One [Internet].

2009;4:e6676. Available from: http://dx.doi.org/10.1371/journal.pone.0006676

114. Rachel RA, Li T, Swaroop A. Photoreceptor sensory cilia and ciliopathies: focus on CEP290, RPGR and their interacting proteins. Cilia [Internet]. 2012;1:22. Available from: http://dx.doi.org/10.1186/2046-2530-1-22

115. Li H, Lian L, Liu B, Chen Y, Yang J, Jian S, et al. KIT ligand protects against both light-induced and genetic photoreceptor degeneration. Elife [Internet]. 2020;9. Available from: http://dx.doi.org/10.7554/eLife.51698

116. Yap CX, Sidorenko J, Wu Y, Kemper KE, Yang J, Wray NR, et al. Dissection of genetic variation and evidence for pleiotropy in male pattern baldness. Nat Commun [Internet]. 2018;9:5407. Available from: http://dx.doi.org/10.1038/s41467-018-07862-y

117. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nat Genet [Internet]. 2018;50:362–7. Available from: http://dx.doi.org/10.1038/s41588-018-0056-5

118. Saatchi M, Beever JE, Decker JE, Faulkner DB, Freetly HC, Hansen SL, et al. QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. BMC Genomics [Internet]. 2014;15:1004. Available from: http://dx.doi.org/10.1186/1471-2164-15-1004

119. Bovine HapMap Consortium, Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science [Internet]. 2009;324:528–32. Available from: http://dx.doi.org/10.1126/science.1167936

120. Nowak K, Berger J, Panikowski A, Reid DG, Jacob AL, Newman G, et al. Using community photography to investigate phenology: A case study of coat molt in the mountain goat (Oreamnos americanus) with missing data. Ecol Evol [Internet]. 2020;16:80. Available from: https://onlinelibrary.wiley.com/doi/10.1002/ece3.6954

121. Bradford HL, Fragomeni BO, Bertrand JK, Lourenco DAL, Misztal I. Regional and seasonal analyses of weights in growing Angus cattle. J Anim Sci [Internet]. 2016;94:4369–75. Available from: http://dx.doi.org/10.2527/jas.2016-0683

122. Mousseau TA, Fox CW. The adaptive significance of maternal effects. Trends Ecol Evol [Internet]. 1998;13:403–7. Available from: http://dx.doi.org/10.1016/s0169-5347(98)01472-4

123. Oliveira DP, Lourenco DAL, Tsuruta S, Misztal I, Santos DJA, de Araújo Neto FR, et al. Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. J Anim Sci [Internet]. 2018;96:27–34. Available from: http://dx.doi.org/10.1093/jas/skx006

124. Ravagnolo O, Misztal I. Genetic component of heat stress in dairy cattle,

parameter estimation. J Dairy Sci [Internet]. 2000;83:2126–30. Available from: http://dx.doi.org/10.3168/jds.S0022-0302(00)75095-8

125. Robertson A. The Sampling Variance of the Genetic Correlation Coefficient. Biometrics [Internet]. [Wiley, International Biometric Society]; 1959;15:469–85. Available from: http://www.jstor.org/stable/2527750

126. PRISM Climate Group, Oregon State University. PRISM Gridded Climate Data [Internet]. 2004. Available from: http://prism.oregonstate.edu

127. Mattar M, Silva LOC, Alencar MM de, Cardoso FF. Genotype× environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. J Anim Sci [Internet]. Oxford University Press; 2011;89:2349–55. Available from: https://academic.oup.com/jas/article-abstract/89/8/2349/4764417

128. Aguilar I, Legarra A, Cardoso F, Masuda Y, Lourenco D, Misztal I. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. Genet Sel Evol [Internet]. 2019;51:28. Available from: http://dx.doi.org/10.1186/s12711-019-0469-3

129. Tiensuu H, Haapalainen AM, Karjalainen MK, Pasanen A, Huusko JM, Marttila R, et al. Risk of spontaneous preterm birth and fetal growth associates with fetal SLIT2. PLoS Genet [Internet]. 2019;15:e1008107. Available from: http://dx.doi.org/10.1371/journal.pgen.1008107

130. Wang A, Zhang Y, Li M, Lan X, Wang J, Chen H. SNP identification in FBXO32 gene and their associations with growth traits in cattle. Gene [Internet].

Elsevier; 2013;515:181–6. Available from:

http://dx.doi.org/10.1016/j.gene.2012.11.054

131. Brunes LC, Baldi F, Lopes FB, Lôbo RB, Espigolan R, Costa MF, et al.

Weighted single-step genome-wide association study and pathway analyses for

feed efficiency traits in Nellore cattle. J Anim Breed Genet [Internet]. 2020;

Available from: http://dx.doi.org/10.1111/jbg.12496

132. Strillacci MG, Vega-Murillo VE, Román-Ponce SI, López FJR, Cozzi MC,

Gorla E, et al. Looking at genetic structure and selection signatures of the

Mexican chicken population using single nucleotide polymorphism markers. Poult

Sci [Internet]. 2018;97:791–802. Available from:

http://dx.doi.org/10.3382/ps/pex374

133. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or

closely linked QTL segregate within and across ten US cattle breeds. BMC

Genomics [Internet]. 2014;15:442. Available from: http://dx.doi.org/10.1186/1471-

2164-15-442

134. Seabury CM, Oldeschulte DL, Saatchi M, Beever JE, Decker JE, Halley YA,

et al. Genome-wide association study for feed efficiency and growth traits in U.S.

beef cattle. BMC Genomics [Internet]. 2017;18:386. Available from:

http://dx.doi.org/10.1186/s12864-017-3754-y

135. Wei Q, Zhong L, Zhang S, Mu H, Xiang J, Yue L, et al. Bovine lineage

specification revealed by single-cell gene expression analysis from zygote to

blastocyst. Biol Reprod [Internet]. 2017;97:5–17. Available from:

http://dx.doi.org/10.1093/biolre/iox071

136. Flori L, Gonzatti MI, Thevenon S, Chantal I, Pinto J, Berthier D, et al. A

quasi-exclusive European ancestry in the Senepol tropical cattle breed highlights

the importance of the slick locus in tropical adaptation. PLoS One [Internet].

2012;7:e36133. Available from: http://dx.doi.org/10.1371/journal.pone.0036133

137. Noble DW, Radersma R, Uller T. Plastic responses to novel environments

are biased towards phenotype dimensions with high additive genetic variation.

Proc Natl Acad Sci U S A [Internet]. 2019;116:13452–61. Available from:

http://dx.doi.org/10.1073/pnas.1821066116

138. Gibson G, Dworkin I. Uncovering cryptic genetic variation. Nat Rev Genet

[Internet]. 2004;5:681–90. Available from: http://dx.doi.org/10.1038/nrg1426

139. Hoffmann AA, Merilä J. Heritable variation and evolution under favourable

and unfavourable conditions. Trends Ecol Evol [Internet]. 1999;14:96–101.

Available from: http://dx.doi.org/10.1016/s0169-5347(99)01595-5

140. Aldridge MN, Vandenplas J, Bergsma R, Calus MPL. Variance estimates

are similar using pedigree or genomic relationships with or without the use of

metafounders or the algorithm for proven and young animals. J Anim Sci

[Internet]. 2020;98. Available from: http://dx.doi.org/10.1093/jas/skaa019

141. Kirkpatrick M, Lande R. The Evolution of Maternal Characters. Evolution

[Internet]. [Society for the Study of Evolution, Wiley]; 1989;43:485–503. Available

from: http://www.jstor.org/stable/2409054

142. Willham RL. The Covariance between Relatives for Characters Composed of Components Contributed by Related Individuals. Biometrics [Internet]. [Wiley, International Biometric Society]; 1963;19:18–27. Available from: http://www.jstor.org/stable/2527570

143. Willham RL. The role of maternal effects in animal breeding. 3. Biometrical aspects of maternal effects in animals. J Anim Sci [Internet]. 1972;35:1288–93. Available from: http://dx.doi.org/10.2527/jas1972.3561288x

144. Koch RM. The role of maternal effects in animal breeding. VI. Maternal effects in beef cattle. J Anim Sci [Internet]. 1972;35:1316–23. Available from: http://dx.doi.org/10.2527/jas1972.3561316x

145. Via S, Lande R. Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. Evolution [Internet]. [Society for the Study of Evolution, Wiley]; 1985;39:505–22. Available from: http://www.jstor.org/stable/2408649

VITA

Harly Jane Durbin was born on August 10, 1994 near Houston, Texas and grew up in Fort Worth, Texas. Part of a multi-generational ranching family, her passion for cattle production was fostered from a young age.

Harly attended Texas A&M University in College Station, Texas, graduating in 2016 with a B.Sc. in Animal Science. During her time at A&M, Harly had the opportunity to participate in a reciprocal exchange program with University College Dublin in Ireland, where she took a course in equine genomics that sparked her interest in animal breeding and genetics. Harly also worked for three years as a genetics student worker in the Texas A&M College of Veterinary Medicine Gastrointestinal Lab. Together, these experiences led her to pursue a graduate degree.

In 2016, Harly began a Ph.D. program at the University of Missouri under the mentorship of Dr. Jared Decker. During her graduate career, Harly had the opportunity to participate in numerous industry-facing activities, including extension field days, publication of multiple popular-press articles, and collaboration with breed associations. Most notably, her work managing the Mizzou Hair Shedding Project led to the collaborative development and deployment of a research breeding value with the American Angus Association. Harly also had the opportunity to attend and presented at several international conferences.

After graduation, Harly hopes to find an industry position where she can take

advantage of her passions for data science, quantitative genetics, and breeding.