

Genomic loci involved in sensing environmental cues and metabolism affect seasonal coat shedding in *Bos taurus* and *Bos indicus* cattle

Harly J. Durbin ^{1,2}, Helen Yampara-Iquise,³ Troy N. Rowan,^{1,4} Robert D. Schnabel,^{1,3,5} James E. Koltes,^{6,7} Jeremy G. Powell,⁷ Jared E. Decker ^{1,3,5,*}

¹Genetics Area Program, University of Missouri, Columbia, MO 65211, USA

²Syngenta, Research Triangle Park, NC 27709, USA

³Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA

⁴University of Tennessee Institute of Agriculture, Knoxville, TN 37996, USA

⁵Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65211, USA

⁶Department of Animal Science, Iowa State University, Ames, IA 50010, USA

⁷Department of Animal Science, University of Arkansas, Fayetteville, AR 72701, USA

*Corresponding author: Division of Animal Sciences, University of Missouri, 920 East Campus Drive, S134a Animal Sciences Research Center, Columbia, Missouri, 65211, USA.
Email: DeckerJE@missouri.edu

Seasonal shedding of winter hair at the start of summer is well studied in wild and domesticated populations. However, the genetic influences on this trait and their interactions are poorly understood. We use data from 13,364 cattle with 36,899 repeated phenotypes to investigate the relationship between hair shedding and environmental variables, single nucleotide polymorphisms, and their interactions to understand quantitative differences in seasonal shedding. Using deregressed estimated breeding values from a repeated records model in a genome-wide association analysis (GWAA) and meta-analysis of year-specific GWAA gave remarkably similar results. These GWAA identified hundreds of variants associated with seasonal hair shedding. There were especially strong associations between chromosomes 5 and 23. Genotype-by-environment interaction GWAA identified 1,040 day length-by-genotype interaction associations and 17 apparent temperature-by-genotype interaction associations with hair shedding, highlighting the importance of day length on hair shedding. Accurate genomic predictions of hair shedding were created for the entire dataset, Angus, Hereford, Brangus, and multibreed datasets. Loci related to metabolism and light-sensing have a large influence on seasonal hair shedding. This is one of the largest genetic analyses of a phenological trait and provides insight into both agriculture production and basic science.

Keywords: genotype-phenotype; phenology; hair shedding; molt; genetic interactions

Introduction

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 (Beltran et al. 2018). In cattle, molting occurs annually in the late spring and early summer when thick winter coats are exchanged for shorter 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 hours of sunlight per day (day length) and changes in temperature (Helm et al. 2013). Among ungulates and other mammals, the effects of temperature and day length interact to induce seasonal molting (Gebbie et al. 1999; Zimova et al. 2014, 2018; Schmidt et al. 2017). This interaction has never been explicitly demonstrated in cattle, although Yeates (1955) showed that artificial manipulation of day length can be used to perturb the timing of hair coat shedding regardless of temperature, while Murray (1965) found a moderate effect of temperature on hair coat shedding among cattle at similar latitudes.

The timing and completeness of molting are also influenced by variables intrinsic to the individual, including the plane of nutrition, life stage, and social status (Cowan et al. 1972; Heydon et al. 1995; Déry et al. 2019). In some species, inaccurate molt timing has a high fitness cost, and therefore phenotypic plasticity is limited (Zimova et al. 2014). In other species (including cattle), variation in molting has been documented within groups of contemporary individuals (Turner and Schleger 1960; Turner 1964), suggesting genetic variation influences an animal's 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 Ferreira et al. 2017, 2020) and to our knowledge only our group has associated genetic variants with phenotypic variation in seasonal hair shedding (Durbin et al. 2020). Though the cost of a “mismatched” seasonal phenotype may be lower in domestic species like cattle compared to many wild populations, they still have large impacts on productivity (St-Pierre et al. 2003; Baumgard et al. 2015).

Previous work has demonstrated the impact of poor early summer hair shedding upon economically relevant traits such as growth and milk production in cattle (Gray et al. 2011; Durbin et al. 2020). Here, we explore the genetic, environmental, and interaction effects that influence early summer hair shedding. We use a multibreed, repeated records dataset of early summer hair shedding scores collected across a range of latitudes, environmental conditions, and production systems to investigate how light, temperature, and metabolism interact with genomic loci to affect the timing of summer hair shedding in cattle.

Materials and methods

All analyses were reproducibly pipelined using Snakemake workflows (Köster and Rahmann 2018) and R (R Core Team 2020). Unless otherwise specified, the {tidyverse} (Wickham et al. 2019) suite of packages was used for data processing. The code is available at https://github.com/harjldurbin/mizzou_hairshed_public.

Phenotypes

Hair shedding scores were collected over 9 years by 77 beef cattle producers and university groups, with most scores collected between 2016 and 2019 (Supplementary Fig. 1a in Supplementary File 1). Hair shedding was classified on an integer 1–5 scale based on the systems developed by Gray et al. (2011) and Turner and Schleger (1960) as described in (Durbin et al. 2020), where a score of 5 indicated 100% winter coat remaining and a score of 1 indicated 0% winter coat remaining. Participants were asked to hair shedding score cattle when they observed the greatest amount of variation in shedding between contemporary individuals. 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 2 separate years (8,839 or 66.11% of all individuals; Supplementary Fig. 1b in Supplementary File 1).

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 $(n \times 365d) - 90d$ to $(n + 1) \times 365d - 90d$, where n is the age classification and d is days. This means that animals at least 9 months of age 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 (Cundiff et al. 2018) as in Durbin et al. (2020), and hair shedding scores recorded on animals fewer than 275 (i.e. 365–390) 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 (Déry et al. 2019). After filtering, 36,899 phenotypes from 13,364 cattle were retained for analysis.

Genotypes and imputation

Genotypes from SNP arrays 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 Bovine GeneSeek Genomic Profiler™ F250 (GGP-F250), a research assay enriched for low-frequency and putatively

functional SNPs (Rowan et al. 2019). On an assay-by-assay basis, markers with >10% missing data and markers significantly deviating from Hardy-Weinberg equilibrium (P -value < 10^{-50}) were set to missing. After marker-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 1 assay. The merged genotypes were then imputed to the union of the GGP-F250 and Illumina BovineHD assays using a multibreed reference panel and the two-step approach described in Rowan et al. (2019). Finally, SNPs with a minor allele frequency below 1% were removed, resulting in genotypes at 747,009 markers for 11,560 individuals. All genotype quality control was conducted using PLINK (Purcell et al. 2007).

Generation of the pedigree and genomic relatedness matrices

Using records provided by various participating breed associations, a 3-generation pedigree was constructed for registered animals with at least 1 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 1 breed association. Parentage was validated for genotyped animals with at least one genotyped parent using the SeekParentF90 program (Hayes 2011; Misztal et al. 2014). 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 with reported progeny were set to missing in the pedigree. In total, 106 sires and 130 dams were excluded for 236 individuals. The final 3-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 inverse ($\mathbf{H}^{-1}$) (Aguilar et al. 2011). $\mathbf{H}^{-1}$ is calculated as:

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

where $\mathbf{A}^{-1}$ represents the inverse of the numerator relationship matrix, $\mathbf{A}_{22}^{-1}$ represents $\mathbf{A}^{-1}$ subset to genotyped individuals, and $\mathbf{G}^{-1}$ represents the inverse of the genomic relatedness matrix (GRM) calculated using the VanRaden method (VanRaden 2008). Genomic and pedigree-relatedness matrices were blended with a weight of 0.05. 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

Full dataset

Genetic parameters and estimated breeding values (EBVs) of hair shedding were first estimated using records from all available animals in the following repeated records animal model, implemented with AIREMLF90 (Misztal et al. 2014):

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

Table 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 in 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 IGS multibreed evaluation: American Simmental Association, Red Angus Association of America, American Shorthorn Association, American Gelbvieh Association	4,722	10,996

In this model, $\mathbf{y}$ represents a vector of hair-shedding score phenotypes. Contemporary group effects are represented in the vector $\mathbf{c}$, where $\mathbf{X}_1$ is a matrix relating the elements of $\mathbf{c}$ to $\mathbf{y}$. 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 (1) 1, (2) 2–3, (3) 4–9, or (4) 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 (Gray et al. 2011; Durbin et al. 2020). Toxic fescue grazing status, or whether cattle grazed endophyte-infected fescue in spring of the recording year, was reported by participants as yes or no. A 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-d sliding window that maximized group size. Records from contemporary groups with fewer than 5 records were discarded. Additional models with various effects fit separately from contemporary group or removed from the model were fit to test the effects of environment and management on hair shedding.

Breed-specific datasets

When performing single-step genomic best linear unbiased prediction (ssGBLUP) in crossbred populations, including data from both purebred and crossbred animals yields the highest prediction accuracy (acc_{LR}), assuming individuals have sufficiently similar genetic structure (Alvarenga et al. 2020). 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 multibreed and cross-bred evaluations) can result in inflated estimates of the additive genetic variance (Dong et al. 1988). Thus, we chose to replicate analyses completed in the full dataset in 4 breed-specific subsets with sufficient sample size for independent genetic evaluation. Calculation of within-breed EBVs also allowed us to search for differences in additive genetic variation 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 fourth dataset consisted of cattle registered with partner breed associations participating in the International Genetics Solutions (IGS) multibreed evaluation (Bozeman, MT; <https://www.internationalgeneticsolutions.com/>). Further descriptions of these datasets can be found in Table 1. The 3-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).

Evaluation of breeding values

Prediction models were evaluated in the full and breed-specific datasets using several metrics proposed by Legarra and Reverter (2018) and explored further by Macedo et al. (2020). First, breeding values were estimated in 10 separate iterations within each dataset, excluding phenotypes from a randomly selected 25% of individuals. Within each iteration, “partial” EBVs (μ_p) were regressed on their corresponding “whole” EBVs (μ_w) from the model including all animals. In the absence of dispersion, the resulting slope is expected to be 1. Next, we estimated bias (d_p) by subtracting the absolute value of μ_p from the absolute value of μ_w . Finally, prediction accuracies were calculated as $acc_{LR} = \frac{\sqrt{cov(\mu_p, \mu_w)}}{\sqrt{(1-F^v)\sigma_a^2}}$, where $cov(\mu_p, \mu_w)$ represents the covariance between partial and whole EBVs, F^v represents the average inbreeding coefficient among animals with phenotypes randomly excluded, and σ_a^2 represents the additive genetic variance of hair shedding score estimated using all individuals.

In order to evaluate the effect of explicitly accounting for breed structure on acc_{LR} and model fit, we tested a model including principal components from a principal component analysis of all genotyped animals as fixed effects. The goal of this analysis was to determine if downstream analyses should use results from the previous (“full”) model discussed above. Principal component analysis was conducted using all 11,560 individuals and 747,009 SNPs with EIGENSOFT smartPCA v.7.2.1 (Patterson et al. 2006).

$$\mathbf{y} = \mathbf{X}_1\mathbf{c} + pc1\beta_1 + pc2\beta_2 + \mathbf{Z}_1\mathbf{u} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

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

The effects of temperature and photoperiod

Latitude and longitude coordinates were determined for each herd location using producer-provided addresses and the R package {tidygeocoder} (Cambon 2021). Based on these coordinates, the daily apparent high temperature, the sunrise time, and the sunset time were retrieved for the 30 d before each hair-shedding scoring date at each unique geographic location using the {darksky} R package (Rudis 2017). The {darksky} package interfaces with the Apple Dark Sky API (Application Programming Interface) to query National Oceanic and Atmospheric Administration historical weather records. For each score date-geographic coordinate combination, the resulting 30-d 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-d range to act as a proxy for light exposure before hair scoring. Next, we fit the following repeated records animal model using AIREMLF90 (Misztal et al. 2014).

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

In this model, $\mathbf{y}$ is a vector of hair shedding score phenotypes; $\mathbf{s}$, $\mathbf{f}$, $\mathbf{a}$, and $\mathbf{r}$, represent vectors of calving season, toxic fescue grazing status, age group, and year effects with matrices $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, and $\mathbf{X}_4$ relating observations to effects; β_1 represents the regression of $\mathbf{y}$ on mean apparent high temperature ($\mathbf{t}$) and β_2 represents the regression of $\mathbf{y}$ on mean day length ($\mathbf{l}$). 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 2 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 and then taking their product. All 3 of these models were compared to the base model using Akaike information criterion (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 (Cundiff et al. 2018)). 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 group effects. To test this, we compared the base model and 4 repeated records animal models similar to those explored in the previous section. The first of these can be described as:

$$y = X_1c + t\beta_1 + l\beta_2 + Z_1u + Z_2p + e$$

In this model, $\mathbf{y}$ represents a vector of hair shedding score phenotypes, and $\mathbf{c}$ 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 $\mathbf{y}$ on mean apparent high temperature ($\mathbf{t}$) and β_2 represents the regression of $\mathbf{y}$ on mean day length ($\mathbf{l}$). 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 (Ekine et al. 2014). 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 $1 - \frac{PEV}{(1+F)\sigma_e^2}$ (Aguilar et al. 2020), 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} (Wellmann 2019). Next, EBVs for genotyped animals were deregressed using the method proposed by Garrick et al. (2009), implemented in the {DRP} R package (Lopes 2016). The resulting 11,560 deregressed estimated breeding values (DEBVs) were used as pseudo-phenotypes in SNP1101 single-SNP regression (Sargolazei 2014). 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) method. These weights were used to construct the $\mathbf{R}^{-1}$ matrix of the mixed-model equations. Covariance between records due to relatedness was accounted for with a GRM constructed using the VanRaden method (VanRaden 2008). Postanalysis, P-values for single-SNP associations were adjusted by the estimated inflation factor (1.40) using the genomic control method (Devlin and Roeder 1999). After genomic control, P-values were converted to false discovery rate (FDR) adjusted q-values using the R package {qvalue} (Storey et al. 2017).

GWAA meta-analysis

Meta-analysis of summary statistics from multiple independent GWAA increases power and decreases FDR compared with single-study results. In addition to GWAA of DEBVs, we tested meta-analysis of year-specific GWAA.

First, we fit 4 separate mixed linear model association analyses in Genome-wide Complex Trait Analysis (GCTA-MLMA) software (Yang et al. 2011, 2014) using phenotypes recorded in each of the years between 2016 and 2019. Phenotypes were adjusted for environment and management using contemporary group designations as implemented in the *Estimation of breeding values and genetic parameters* section. In cases of animals with multiple phenotypes available within a given year, the phenotype recorded closest to May 1 was used. Records from contemporary groups with fewer than 5 records were excluded, resulting in 5,146, 4,989, 6,896, and 5,292 phenotypes included in the analyses associated with 2016, 2017, 2018, and 2019, respectively. Next, meta-analysis of the 4 year-by-year GWAA was performed using the first random effects (RE1) model implemented in METASOFT (Han and Eskin 2011). The RE1 method was used as it reported effect sizes and standard errors used in downstream analyses. The resulting meta-analysis P-values were adjusted by an inflation factor of 1.52 using the genomic control method (Devlin and Roeder 1999) and transformed into q-values (Storey et al. 2017).

Conditional and joint association analysis

Conditional and joint association analysis (COJO) is a stepwise model selection procedure used to identify independently associated SNPs and previously undetected secondary associations. Thus, it can be used to refine genome-wide association signals. COJO uses summary statistics (P-values as well as SNP effect betas and their associated standard errors) derived directly from a single GWAA or from meta-analysis of multiple GWAA. Betas and standard errors were not available from the analyses detailed in *Deregression of breeding values and single-SNP regression*, so we

performed COJO using P-values obtained in the GWAA meta-analysis section. Summary statistics from the METASOFT RE1 model were then used for COJO in GCTA (Yang et al. 2012) with linkage disequilibrium (LD) calculated within animals with a phenotype included in at least one of the year-by-year GWAA.

Genotype-by-environment interaction GWAA

Main effects GWAA is used to identify variants statistically associated with the phenotype of interest agnostic to the environment in which the phenotype was expressed, assuming appropriate model specification. By contrast, genotype-by-environment interaction (GxE) GWAA can be used to identify variants whose association with the phenotype of interest is dependent upon the environment in which the phenotype was expressed. Using the environmental data gathered in *The effects of temperature and photoperiod*, we fit GWAA modeling the interaction of genetics and either apparent high temperature or day length in Genome-wide Efficient Mixed Model Association (GEMMA) software (Zhou and Stephens 2012). Phenotypes were pre-adjusted by subtracting the contemporary group fixed effect estimate (in this analysis defined as farm, year, calving season, age group, and fescue status) from the raw hair shedding score. As such, records were excluded from contemporary groups containing fewer than 5 animals. One GWAA was fit per environmental variable (mean apparent high temperature or mean day length) per year (2016–2019) and in cases of animals with multiple phenotypes per year, 1 phenotype was randomly chosen to increase variation in environmental data. Meta-analysis of the resulting summary statistics was then performed in METASOFT using the second RE2 model (Han and Eskin 2011). The λ_{mean} (1.10 for temperature and 1.40 for day length) and λ_{hetero} (0.47 for temperature and 0.33 for day length) calculated in this first run of METASOFT were used to adjust RE2 P-values in a subsequent run. Finally, adjusted RE2 P-values were transformed to q-values to correct for multiple testing (Storey et al. 2017).

Annotation and enrichment

Gene set and QTL enrichment analyses were performed separately for (1) the results of the main effect GWAA using DEBVs, (2) the main effect GWAA meta-analysis, (3) COJO selected SNPs, and (4) the GxE interaction effect GWAA meta-analysis of mean day length.

First, genes within 50 kb of variants significant at FDR < 5% were identified using the {GALLO} package (Fonseca et al. 2020) and ARS-UCD1.2 bovine genome coordinates (Rosen et al. 2020). In the case of COJO selected SNPs with no genes within 50 kb ($n=13$), the nearest gene regardless of distance was included in the COJO gene set. Next, gene set enrichment analysis was performed for each gene set using the 'gost' function provided in the {gprofiler2} package (Kolberg and Raudvere 2020), which interfaces with the g:Profiler toolkit to query publicly available functional annotation databases. Considered data sources included gene ontology terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa and Goto 2000), Reactome pathways (Jassal et al. 2020), Comprehensive Resource of Mammalian (CORUM) protein complexes (Giurgiu et al. 2019), and human disease phenotypes from the Human Phenotype Ontology database (Köhler et al. 2021). Significance values for gene set enrichment results were corrected for multiple testing using the g:SCS algorithm, which is designed for hierarchically related, nonindependent tests.

We also performed QTL enrichment analysis using Animal QTLdb QTL annotations (Hu et al. 2019). The {GALLO} R package was used to identify QTL annotations within 50 kb of variants

significant at FDR < 5% and to calculate enrichments. Significant enrichments were determined using a Benjamini-Hochberg adjusted P-value threshold of 0.05.

Results

Estimation of breeding values and genetic parameters

Hair shedding scores were collected for cattle from 10 different breed groups across the Southeast and Midwest United States (Fig. 1) in late spring and early summer (Supplementary Fig. 2 in Supplementary File 1).

Models with and without principal components included as covariates were fit to assess the effect of population structure on breeding values. In principal component analysis, eigenvector 1 accounted for 35% of genetic variation, with Hereford animals at the negative end and Angus individuals at the other end. Eigenvector 2 accounted for differences in *Bos taurus* vs *Bos indicus* ancestry and explained 25% of the variance (Supplementary Fig. 3 in Supplementary File 1). The remaining principal components explained much less variation (Supplementary Fig. 4 in Supplementary File 1), and therefore we chose to include eigenvectors for only the first 2 principal components as covariates in the model evaluating the effect of explicitly accounting for breed structure. A likelihood ratio test comparing the models with and without principal components fit as fixed effects indicated that inclusion provided a moderately better fit ($-\log_{10}(P\text{-value}) = 3.04$). However, this model required the exclusion of phenotypes from un-genotyped animals. Further, including principal components resulted in lower acc_{LR} for hair shedding across 10 iterations (mean $acc_{LR} = 0.62$ with principal components vs mean $acc_{LR} = 0.68$ without principal components). Therefore, we chose to consider results from the "full" model excluding principal components and including all available individuals 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 our research (Durbin et al. 2020) and by Gray et al. (2011) (Table 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_A^2) estimate in the full dataset fell within the range of estimates of σ_A^2 in the breed-specific datasets, suggesting that this value was not inflated by the inclusion of crossbred animals. The permanent environmental variance (σ_{PE}^2) accounted for 5–7% of total variance with the exception of the Brangus dataset, for which σ_{PE}^2 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 1 or 2 years of data. This likely explains 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" where fuller hair coats are desirable (Shorthorn and Maine-Anjou) tended to have higher (i.e. less desirable) EBVs (Supplementary Fig. 5 in Supplementary File 1). Further, breeds with known *Bos indicus* ancestry (Brangus and Charolais; (Decker et al. 2014)) tended to have lower EBVs.

Estimated acc_{LR} ranged from 0.657 to 0.674 across 10 iterations using all available data. Among the breed-specific datasets, acc_{LR} tended to be lowest in the Hereford dataset and highest in the IGS dataset (Supplementary Table 1 in Supplementary File 1). The mean $b_{w,p}^v$ was between 1.00 and 1.04 in all datasets, suggesting minimal dispersion.

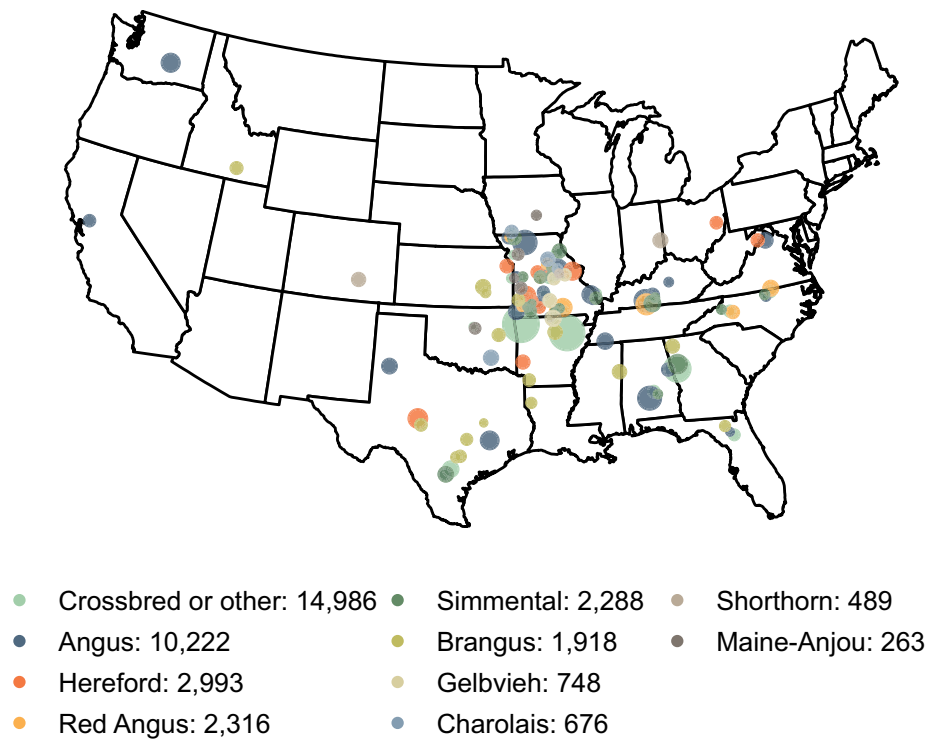


Fig. 1. Counts of hair shedding score records by reported breed. Most phenotypes came from 3 breeds and were recorded in the Midwest or South. For the purposes of this map, Angus, Hereford, Red Angus, Simmental, and Gelbvieh animals with at least 5% ancestry assigned to the given breed based on pedigree estimates were included in that breed. Animals with unknown ancestry, less than 5% ancestry assigned to 1 breed, or of a breed not listed above were called “Crossbred or other”. Map was plotted using public domain data from the US Department of Commerce, Census Bureau via the R package maps (<https://cran.r-project.org/web/packages/maps/>).

Table 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 4 breed-specific datasets. Approximated standard errors for heritability and repeatability estimates are in parentheses.

Dataset	σ^2_A	σ^2_{PE}	σ^2_E	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)

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 (Supplementary Fig. 6a in Supplementary File 1). The mean apparent high temperature ranged from 4.23 to 39.33°C with a mean of 25.87 and standard deviation of 5.37 (Supplementary Fig. 6b in Supplementary File 1). The base model including apparent temperature and day length provided a better fit over both the model with only apparent temperature ($-\log_{10}(P\text{-value}) = 92.35$) or day length ($-\log_{10}(P\text{-value}) = 156.52$), while the model including the interaction effect provided a better fit than the base model ($-\log_{10}(P\text{-value}) = 3.42$). The interaction model also had a lower AIC value than the base model and both of the reduced models (Table 3). The day length BLUE 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 d before 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 d before scoring.

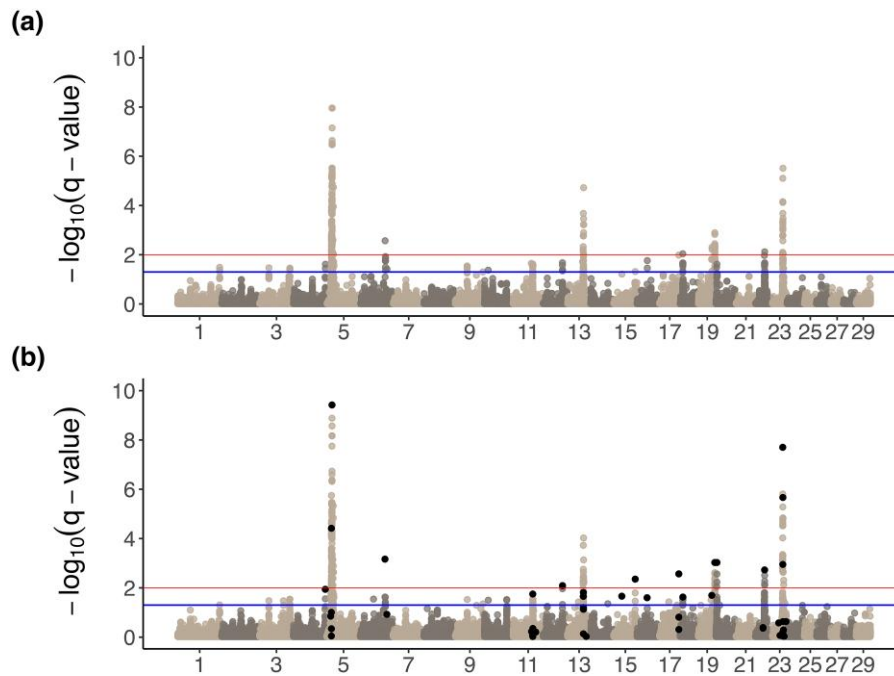
Calving season, toxic fescue grazing status, and age group BLUEs from all 4 models are in Supplementary Table 2 in Supplementary File 1. In general, BLUEs for grazing toxic fescue tended to be higher than BLUEs for not grazing toxic fescue and BLUEs for spring calving tended to be higher than BLUEs for fall calving, congruent with trends we previously reported (Durbin et al. 2020). The magnitude of the effect of grazing toxic fescue was estimated to be ~ 4 times larger in the day length-only model than in the temperature-only model, and ~ 2 times larger than in the models including both day length and temperature.

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.

Table 3. BLUES and AIC values from 4 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

**Fig. 2.** Manhattan plot of main effect $-\log_{10}(q\text{-values})$ a) using DEBVs and b) using meta-analysis. The lower (blue) lines at 1.3 represent FDR = 5% and the upper (red) lines at 2 represent FDR = 1%. Black points represent variants identified by COJO selection. Please see [Supplementary File 2](#), which lists genes within 50 kbp of the associated SNPs.

Genome-wide association

In the main effects GWAA using DEBVs, 377 variants were genome-wide significant at FDR = 5% (Fig. 2a). Of the 413 variants passing a FDR = 5% level in the meta-analysis of year-specific main effects GWAA, 20 were selected by COJO, plus an additional 37 variants with FDR < 5% (Fig. 2b). The 57 SNPs selected by COJO resided in 30 independent associations, with independent associations delimited by pairs of adjacent SNPs greater than 1 Mb apart as in (Yang et al. 2012).

Over half of all significant SNPs were found on chromosome 5 as in (Durbin et al. 2020) (Fig. 3a). COJO selection suggests that these results might represent 3 separate causal loci. The lead SNP in both main effects GWAA (BTA5:18,767,155) was also the most significant SNP identified in COJO selection. The nearest gene to this variant is a lncRNA approximately 300 kb downstream (ENSEMBL ID ENSBTAG00000053947), followed by KITLG approximately 400 kb upstream.

Gene set enrichment results were largely consistent between the 2 main effects GWAA strategies, both returning terms associated with regulation of apoptosis (Supplementary Table 3 in Supplementary File 1). QTL enrichment analysis of the main effects results returned 6 significant terms (Supplementary Table 2 in Supplementary File 1). Of these, “white spotting” was

the most significantly enriched QTL term ($P\text{-value} = 1.15 \times 10^{-16}$). Upon further examination, this signal appeared to be driven by SNPs within and near the *MITF* (*microphthalmia-associated transcription factor*) gene on chromosome 22, a master regulator of melanocyte production that is highly conserved across vertebrates (Steingrimsdottir et al. 2004; Levy et al. 2006; Hou and Pavan 2008).

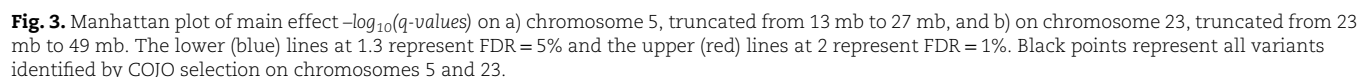
Meta-analysis of mean apparent high temperature and mean day-length GxE interaction effect GWAA resulted in 17 and 1,040 variants passing at FDR = 5%, respectively (Fig. 4). Unlike main effect gene set enrichment results, most terms enriched in the analysis of day length GxE interaction results were associated with keratinization and cytoskeleton formation (Supplementary Table 3 in Supplementary File 1).

Genes within 50 kbp of associated SNPs for all GWAA are reported in Supplementary File 2.

Discussion

In cattle, hair shedding is affected by nutrition and metabolism, temperature, and seasonal changes in the amount of daylight.

Calving season, toxic fescue grazing status, and age group BLUES shed light on the relationship between seasonal hair shedding and nutrition/management effects. For example, younger cows have



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. (Déry *et al.* 2019)). Since beef cows are typically culled from the herd or die between ages 11 and 12 (Azzam *et al.*

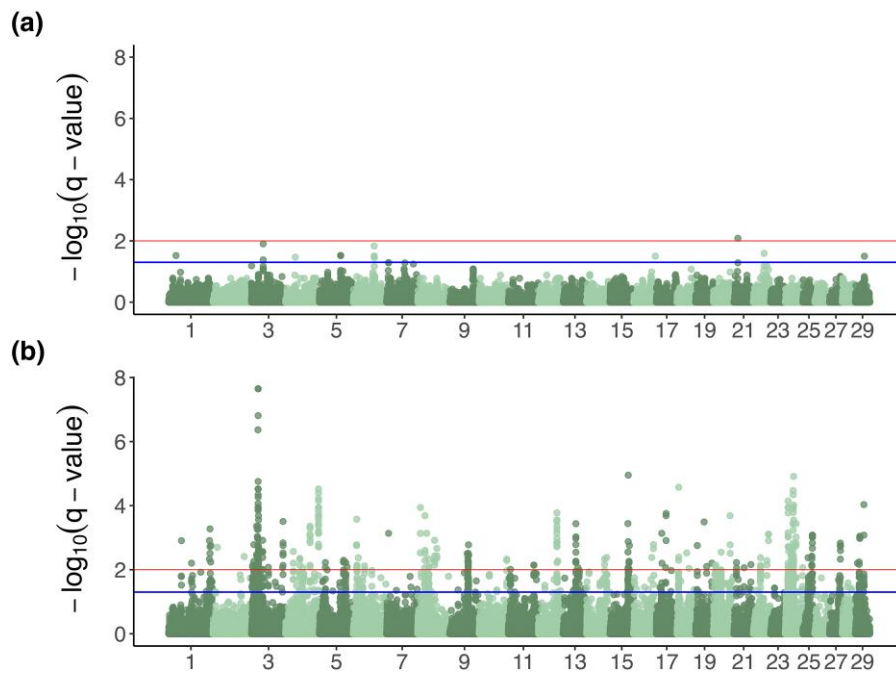


Fig. 4. Manhattan plots of a) mean apparent temperature and b) mean day length interaction effect meta-analysis $-\log_{10}(q\text{-values})$. Lower (blue) lines at 1.3 represent FDR = 5% and upper (red) lines at 2 represent FDR = 1%. Please see [Supplementary File 2](#), which lists genes within 50 kbp of the associated SNPs.

1990), it's likely that the advantage predicted here for very old cows is reflective of selection (i.e. culling) allowing well-adapted individuals to remain productive later into their lives.

The estimated effect of grazing toxic fescue was largest in the day length only model. This could reflect the relationship between temperature and tall fescue ergot alkaloid levels (see [Klotz 2015](#)) and the references therein), as major symptoms of fescue toxicosis is vasoconstriction increasing heat stress and decreased feed intake. Fescue toxins are also known to decrease serum prolactin ([Klotz 2015](#)), thus highlighting the relationship between fescue toxins, prolactin, and hair shedding. Selection for more rapid hair shedding may not only increase heat tolerance, but may also identify cattle more resilient to fescue toxicosis.

The length of daylight hours has a tremendous impact on hair shedding. The hormonal cascades underlying seasonal coat change, largely driven by prolactin, are highly conserved between species. In response to changes in day length, photoperiodic cues are forwarded from the eye to the pineal gland, which is responsible for the production of melatonin. Melatonin inhibits prolactin, and decreased melatonin production from increased daylight allows for elevated serum prolactin, thereby triggering the spring molt ([Zimova et al. 2018](#)). Temperature changes have also been implicated via interaction with day length ([Gebbie et al. 1999](#); [Zimova et al. 2014](#); [Schmidt et al. 2017](#)), though the mechanisms by which temperature influences seasonal biology (and by which mammals sense temperature in general) are unclear ([Caro et al. 2013](#)). The day length-temperature interaction has never been explicitly demonstrated in cattle, though Peters and Tucker (1978) showed the effect of photoperiod on serum prolactin is magnified by ambient temperature. Results from the models here quantifying the effects of external environmental variables on hair shedding score clearly point to roles for day length and temperature in regulating seasonal coat change. However, GxE GWAA of the 2 variables suggests a larger independent role for day length than temperature. Even though the 2 GxE GWAA had the exact same sample sizes, we observed substantially more associations in the daylength GxE GWAA. The lack of temperature

GxE GWAA results found here is consistent with previous work demonstrating the inability of temperature alone to initiate seasonal coat change ([Zimova et al. 2018](#)). Our results show that in cattle, there is little to no genetic variation for temperature sensing. However, there appears to be a large number of genetic interactions with day length that are associated with hair shedding variation.

Using 2 very different GWAA approaches (DEBVs as pseudo-phenotypes and meta-analysis of year-specific GWAA), we found nearly identical results ([Fig. 2](#); [Supplementary Table 3](#) and [Supplementary Table 4](#) in [Supplementary File 1](#)). In both main effects GWAA, over half of all significant variants were located on chromosome 5 ([Fig. 2](#)). We previously ([Durbin et al. 2020](#)) found a similarly large association on chromosome 5 in American Angus cattle but were unable to narrow down candidate genes. We previously theorized that the strength of the association, likely caused by extensive long-range LD, could contain multiple causal mutations affecting multiple genes ([Cannon and Mohlke 2018](#)). Using COJO selection, we find evidence for 3 separate associations between 14 and 24 Mb on chromosome 5 ([Fig. 3a](#)). The third association contains the lead COJO SNP and the lead SNP in both main effects GWAA, located at BTA5:18,767,155. The closest genes to this SNP are an unannotated lncRNA and KITLG, ~300 kb downstream and ~400 kb upstream, respectively. The gene CEP290 is between the 2nd and 3rd COJO peaks, and its action is involved in the biogenesis of the photoreceptor sensory cilia ([Rachel et al. 2012](#)). Finally, NTS (neurotensin) is located at 15.5 Mb on chromosome 5; NTS is known to affect prolactin levels ([McCann et al. 1982](#)). This connection to prolactin is interesting, based on the effect of prolactin on seasonal turnover of the haircoat.

The second largest association in the main effects GWAA was located on BTA23 in proximity to PRL (BTA23:35,332,705–35,341,607; [Fig. 3b](#)), the gene encoding prolactin. As previously mentioned, prolactin is the main hormonal regulator of seasonal response, making it an obvious candidate gene. Further, mutations in PRL and its receptor were previously linked to an abnormal hair coat and thermoregulatory phenotype ([Littlejohn et al. 2014](#)).

However, this association also contained *CDKAL1* (*CDK5 regulatory subunit associated protein 1 like 1*; BTA23:36,745,642–37,444,814), with many significant variants (including 5 COJO selected SNPs) within the gene itself (Fig. 3b). In multiple human populations, *CDKAL1* variants are associated with risk for type 2 diabetes mellitus and decreased insulin secretion (Steinthorsdottir et al. 2007). These results could point toward a relationship between metabolic regulation and seasonal hair shedding in cattle. In support of this idea, mature body weight and 18-month body weight QTL associations were enriched in the main effects results (Supplementary Table 4 in Supplementary File 1). Further, genes associated with “metabolism” (Supplementary Table 3 in Supplementary File 1) and QTL associated with “metabolic body weight” and “ketosis” (Supplementary Table 4 in Supplementary File 1) were enriched in the day length GxE GWAA results. Photoperiodic response is associated with patterns of seasonal growth and metabolic change across taxa in addition to coat change. 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 (Petitclerc et al. 1983; Tucker et al. 1984; Barenton et al. 1988; Dahl et al. 2000; Small et al. 2003).

The hair follicle cycle is composed of 3 main phases highly conserved across mammalian species (Stenn and Paus 2001): anagen, catagen, and telogen. During catagen, the hair follicle involutes via apoptosis (Botchkareva et al. 2006), or the process by which cells are eliminated after fulfilling a biological purpose. Our gene set enrichment analyses of the main effects GWAA results returned multiple terms associated with regulation of apoptosis. Additionally, terms associated with keratin formation were enriched in the day length GxE gene set enrichment analysis. In a transcriptomic study of mountain hare seasonal coat color, Ferreira et al. (2020) found similar enrichment of cell death terms and keratinization. This might suggest shared mechanisms between multiple forms of seasonal coat change.

QTL associated with “white spotting” were significantly enriched in the main effects results, driven by SNPs near and within *MITF* on BTA22. An association with *KITLG*, which is responsible for white and roan coats in Shorthorn cattle, is also tagged by SNPs associated with hair shedding. There are multiple potential explanations for this result, including confounding due to population structure, differences in hair color affecting hair shedding, effects on eye function, and crosstalk between melanocytes and neighboring cells.

Confounding due to population structure is unlikely. First, our *P*-values are well-calibrated. Second, none of the other genes with large frequency differences across breeds such as other coat color genes, the locus controlling horned vs polled, or other large effect genes related to breed differences are associated with hair shedding. Finally, the evaluation of EBVs shows that dispersion is not different from one and bias is not different from zero. Dispersion and bias are not well calibrated when breed effects unduly influence the predictions. Thus, we reject the interpretation that this result is due to confounding from population structure and breed effects.

We do not find strong evidence to support the conclusion that variation in coat color, including the amount of white on an animal, influences differences in hair shedding through heat absorption or other mechanisms. Most of the other genes related to cattle coat color differences, such as *MC1R*, *ASIP*, *TYRP1*, *TYR*, *PMEL*, and *KIT* (Schmutz 2012) are not associated with hair shedding variation. Further, we find the chromosome 5 association in Angus cattle (Durbin et al. 2020), which are predominately solid black cattle.

The effects of *MITF* and *KITLG* on retinal function cannot be ruled out. Perturbations to *MITF* are responsible for several audio-visual disorders across taxa, including Waardenburg syndrome (Tassabehji et al. 1994) and Tietz syndrome (Amiel et al. 1998) in humans. 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 (Wen et al. 2016; García-Llorca et al. 2019; Han et al. 2020). 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 (Reiter 1991). *KITLG* has 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 (Li et al. 2020).

The explanation we find most intriguing is the crosstalk between melanocytes and nearby cells, including keratinocytes. Not only do we find associations with *MITF* and *KITLG*, we find multiple genes related to melanocyte and melanogenesis gene pathways, including *SOX9*, *PBX1*, *PBX2*, *EDN3*, *FZD9*, *NOTCH2*, *CREB3L2*, and *GNAS* [(Baxter et al. 2010; D'Mello et al. 2016; Kanehisa et al. 2023) KEGG pathways hsa04916 and map04916] associated with variation in hair shedding. Further, various genes from the same families as melanogenesis genes, such as *SOX4*, *SOX21*, *NOTCH4*, *TRPM2*, *RAB22A*, *GRP84*, *GRP179*, *GRP89A*, and 12 genes from the solute carrier family are identified in our results. Crosstalk between melanocyte stem cells and hair follicle stem cells is believed to occur via TGF β and WNT signaling (Mort et al. 2015), but we only find 4 genes within these pathways in our results (*DCN*, *DKK4*, *LRP6*, *APCDD1L*). Although we find genes related to melanogenesis, we do not find many of the genes typically included in hair follicle gene networks, such as those in BMP (bone morphogenetic protein), FGF (fibroblast growth factors), and WNT (Wingless and Int-1) signaling pathways (Daszczuk et al. 2020). However, *SOX9*, *SOX21*, *GNAS*, *BCL2*, *TSPEAR*, *KRT74*, and *TFAP2C*, which have gene ontology annotations of “hair cycle” (GO:0042633), are identified in our analyses (Ashburner et al. 2000; The Gene Ontology Consortium 2023). Thus, based on a database and literature search of gene networks of melanogenesis and hair follicles, it appears to us that melanocytes affect hair molting, and there may not be a direct relationship between hair shedding and hair color.

In the future, imputation of genotypes to sequence level would enable fine-mapping of significantly associated regions and their functional annotations. At the genotype density used here, causal variants are unlikely to have been directly assayed. Sequence-level data in combination with our multibreed dataset could enable the refinement of causal variants to near the base-pair level.

Data collected and maintained by nonprofessionals are often under-utilized, as it can introduce certain errors and biases. However, it can afford researchers an increased analytical power via vastly increased sample size when statistically modeled correctly. In a similar effort, Nowak et al. (2020) quantified the effects of temperature and day length on molting in mountain goats using data and photographs collected by nonprofessionals. Similarly, we were able to explore the functional biology of a complex trait using farmer-sourced data (Decker 2015). This work reinforces the utility of “citizen-science” data collected by nonprofessionals 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 multibreed dataset, we were able to show that a previously published association found on chromosome 5 in American Angus cattle is likely driven by multiple causal variants at this locus. Together, these results point toward 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 sensing day length. Our results also point to crosstalk between melanocytes and keratinocytes affecting molting. 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, are largely shared across species (see Helm *et al.* (2013) and references therein). Therefore, this work also provides an important stepping-off point for research in other species.

Data availability

Genotype, phenotype, and metadata are available from the Dryad Digital Repository (<https://doi.org/10.5061/dryad.ngf1vhhz4>).

Acknowledgments

We thank the following participating breed associations for data, pedigree, and genotype sharing:

- The American Simmental Association, particularly Jackie Atkins, Rachel Endecott, and Jordan Bouwman
- The Red Angus Association of America, particularly Ryan Boldt
- The American Shorthorn Association, particularly Matt Woolfolk
- The American Gelbvieh Association, particularly Will Fiske and Taylor Buckley
- The American International Charolais Association, particularly Sally Northcutt
- The American Hereford Association, particularly Stacy Sanders
- The American Angus Association, particularly Steve Miller, Duc Lu, Kelli Retallick, and Dan Moser
- The International Brangus Breeders Association, particularly Macee Praise

Thank you to all undergraduate student workers and interns who helped with data collection, organization, and accession of DNA samples from 2016 to 2019:

- Will Shaffer
- Natalie Barr
- Sarah Van Ausdal
- Brian Arisman
- Aaron Mott

Thank you to the farm staff at the Savoy and Batesville Research Stations with the Division of Agriculture at the University of Arkansas System, in particular Don Hubbell and John Tucker at the Livestock and Forestry Research Station who coordinated extensive hair score phenotyping for this study.

Thank you, especially to the farmers and ranchers who contributed data, without whom these analyses would not have been possible.

Supplemental material available at G3 online.

Funding

This project was supported by Agriculture and Food Research Initiative Competitive Grant nos. 2016–68004-24827 and 2021-67021-33448 from the USDA National Institute of Food and Agriculture.

Conflicts of interest

Jared Decker and James Koltes were previously on the scientific advisory board of Vytelle, LLC.

Literature cited

- Aguilar I, Fernandez EN, Blasco A, Ravagnolo O, Legarra A. 2020. Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP. *J Anim Breed Genet.* 137(4):356–364. doi:10.1111/jbg.12470.
- Aguilar I, Misztal I, Legarra A, Tsuruta S. 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *J Anim Breed Genet.* 128(6):422–428. doi:10.1111/j.1439-0388.2010.00912.x.
- Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP, *et al.* 2023. The Gene Ontology knowledgebase in 2023. *Genetics.* 224(1). <https://doi.org/10.1093/genetics/iyad031>.
- Alvarenga AB, Veroneze R, Oliveira HR, Marques DBD, Lopes PS, Silva FF, Brito LF. 2020. Comparing alternative single-step GBLUP approaches and training population designs for genomic evaluation of crossbred animals. *Front Genet.* 11:263. doi:10.3389/fgene.2020.00263.
- Amiel J, Watkin PM, Tassabehji M, Read AP, Winter RM. 1998. Mutation of the MITF gene in albinism-deafness syndrome (tietz syndrome). *Clin Dysmorphol.* 7(1):17–20. doi:10.1097/00019605-199801000-00003.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* 2000. Gene Ontology: tool for the unification of biology. *Nat Genetics.* 25(1):25–29. doi:10.1038/75556.
- Azzam SM, Azzam AM, Nielsen MK, Kinder JE. 1990. Markov chains as a shortcut method to estimate age distributions in herds of beef cattle under different culling strategies. *J Anim Sci.* 68(1):5–14. doi:10.2527/1990.6815.
- Barenton B, Ravault J-P, Chabanet C, Daveau A, Pelletier J, Ortavant R. 1988. Photoperiodic control of growth hormone secretion and body weight in rams. *Domest Anim Endocrinol.* 5(3):247–255. doi:10.1016/0739-7240(88)90036-7.
- Baumgard LH, Keating A, Ross JW, Rhoads RP. 2015. Effects of heat stress on the immune system metabolism and nutrient partitioning: implications on reproductive success. *Revista Brasileira de Reproducao Animal.* 39(1):173–183.
- Baxter LL, Moreland RT, Nguyen AD, Wolfsberg TG, Pavan WJ. 2010. A curated online resource for SOX10 and pigment cell molecular genetic pathways. Database (Oxford). 2010:baq025. doi:10.1093/database/baq025.
- Beltran RS, Burns JM, Breed GA. 2018. Convergence of biannual moulting strategies across birds and mammals. *Proceed Roy Soc B Biol Sci.* 285(1878):20180318. doi:10.1098/rspb.2018.0318.
- Botchkareva NV, Ahluwalia G, Shander D. 2006. Apoptosis in the hair follicle. *J Invest Dermatol.* 126(2):258–264. doi:10.1038/sj.jid.5700007.
- Cambon J, Hernangómez D, Belanger C, Possenriede D. 2021. tidygeo-coder: An R package for geocoding. *J Open Sour Softw.* 6(65):3544. doi:10.21105/joss.03544. (R package version 1.0.5).

- Cannon ME, Mohlke KL. 2018. Deciphering the emerging complexities of molecular mechanisms at GWAS loci. *Am J Hum Genet.* 103(5):637–653. doi:[10.1016/j.ajhg.2018.10.001](https://doi.org/10.1016/j.ajhg.2018.10.001).
- Caro SP, Schaper SV, Hut RA, Ball GF, Visser ME. 2013. The case of the missing mechanism: how does temperature influence seasonal timing in endotherms? *PLoS Biol.* 11(4):e1001517. doi:[10.1371/journal.pbio.1001517](https://doi.org/10.1371/journal.pbio.1001517).
- Core Team R. 2020. R: A Language and Environment for Statistical Computing. Vienna (Austria): R Foundation for Statistical Computing.
- Cowan IM, Cowan IM, Raddi AG. 1972. Pelage and molt in the black-tailed deer (*Odocoileus hemionus* (rafinesque)). *Can J Zool.* 50(5): 639–647. doi:[10.1139/z72-088](https://doi.org/10.1139/z72-088).
- Cundiff LV, Van Vleck LD, Hohenboken WD, Cundiff LV, Van Vleck LD, Hohenboken WD, Cundiff LV, Van Vleck LD, Hohenboken WD. 2018. Animal evaluation. In: Cundiff LV, Van Vleck LD, Hohenboken WD, editors. *Guidelines for Uniform Beef Improvement Programs*. 9th ed. Prairie (MS): Beef Improvement Federation. p. 16–21. <https://beefimprovement.org/resource-center/bif-guidelines/>.
- Dahl GE, Buchanan BA, Tucker HA. 2000. Photoperiodic effects on dairy cattle: a review. *J Dairy Sci.* 83(4):885–893. doi:[10.3168/jds.S0022-0302\(00\)74952-6](https://doi.org/10.3168/jds.S0022-0302(00)74952-6).
- Daszczuk P, Mazurek P, Pieczonka TD, Olczak A, Boryń LM, Kobiela K. 2020. An intrinsic oscillation of gene networks inside hair follicle stem cells: an additional layer that can modulate hair stem cell activities. *Front Cell Dev Biol.* 8:595178. doi:[10.3389/fcell.2020.595178](https://doi.org/10.3389/fcell.2020.595178).
- Decker JE. 2015. Agricultural genomics: commercial applications bring increased basic research power (G. Gibson, ed.). *PLoS Genet.* 11(11):e1005621. doi:[10.1371/journal.pgen.1005621](https://doi.org/10.1371/journal.pgen.1005621).
- Decker JE, McKay SD, Rolf MM, Kim J, Molina Alcalá A, Sonstegard TS, Hanotte O, Götherström A, Seabury CM, Praharani L, et al. 2014. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet.* 10(3):e1004254. doi:[10.1371/journal.pgen.1004254](https://doi.org/10.1371/journal.pgen.1004254).
- Déry F, Hamel S, Côté SD. 2019. Getting ready for the winter: timing and determinants of molt in an alpine ungulate. *Ecol Evol.* 9(5): 2920–2932. doi:[10.1002/ece3.4970](https://doi.org/10.1002/ece3.4970).
- Devlin B, Roeder K. 1999. Genomic control for association studies. *Biometrics.* 55(4):997–1004. doi:[10.1111/j.0006-341x.1999.00997.x](https://doi.org/10.1111/j.0006-341x.1999.00997.x).
- D'Mello SAN, Finlay GJ, Baguley BC, Askarian-Amiri ME. 2016. Signaling pathways in melanogenesis. *Int J Mol Sci.* 17(7):1144. doi:[10.3390/ijms17071144](https://doi.org/10.3390/ijms17071144).
- Dong MC, Van Vleck LD, Wiggans GR. 1988. Effect of relationships on estimation of variance components with an animal model and restricted Maximum likelihood. *J Dairy Sci.* 71(11):3047–3052. doi:[10.3168/jds.S0022-0302\(88\)79904-X](https://doi.org/10.3168/jds.S0022-0302(88)79904-X).
- Durbin HJ, Lu D, Yampara-Iquise H, Miller SP, Decker JE. 2020. Development of a genetic evaluation for hair shedding in American angus cattle to improve thermotolerance. *Genet Sel Evol.* 52(1):63. doi:[10.1186/s12711-020-00584-0](https://doi.org/10.1186/s12711-020-00584-0).
- Ekine CC, Rowe SJ, Bishop SC, de Koning D-J. 2014. Why breeding values estimated using familial data should not be used for genome-wide association studies. *G3.* 4(4):341–347. doi:[10.1534/g3.113.008706](https://doi.org/10.1534/g3.113.008706).
- Ferreira MS, Alves PC, Callahan CM, Giska I, Farello L, Jenny H, Mills LS, Hackländer K, Good JM, Melo-Ferreira J. 2020. Transcriptional regulation of seasonal coat color change in hares. *Ecol Evol.* 10(3):1180–1192. doi:[10.1002/ece3.5956](https://doi.org/10.1002/ece3.5956).
- Ferreira MS, Alves PC, Callahan CM, Marques JP, Mills LS, Good JM, Melo-Ferreira J. 2017. The transcriptional landscape of seasonal coat colour moult in the snowshoe hare. *Mol Ecol.* 26(16): 4173–4185. doi:[10.1111/mec.14177](https://doi.org/10.1111/mec.14177).
- Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á. 2020. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *Gigascience.* 9(12):giaa149. doi:[10.1093/gigascience/giaa149](https://doi.org/10.1093/gigascience/giaa149).
- García-Llorca A, Aspelund SG, Ogmundsdóttir MH, Steingrímsson E, Eysteinnsson T. 2019. The microphthalmia-associated transcription factor (mitf) gene and its role in regulating eye function. *Sci Rep.* 9(1):15386. doi:[10.1038/s41598-019-51819-0](https://doi.org/10.1038/s41598-019-51819-0).
- Garrick DJ, Taylor JF, Fernando RL. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet Sel Evol.* 41(1):55. doi:[10.1186/1297-9686-41-55](https://doi.org/10.1186/1297-9686-41-55).
- Gebbie FE, Forsyth IA, Arendt J. 1999. Effects of maintaining solstice light and temperature on reproductive activity, coat growth, plasma prolactin and melatonin in goats. *J Reprod Fertil.* 116(1):25–33. doi:[10.1530/jrf.0.1160025](https://doi.org/10.1530/jrf.0.1160025).
- Giurgiu M, Reinhard J, Brauner B, Dunger-Kaltenbach I, Fobo G, Frishman G, Montrone C, Ruepp A. 2019. CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.* 47(D1):D559–D563. doi:[10.1093/nar/gky973](https://doi.org/10.1093/nar/gky973).
- Gray KA, Smith T, Maltecca C, Overton P, Cassady JP. 2011. Differences in hair coat shedding, and effects on calf weaning weight and BCS among angus dams. *Livest Sci.* 140(1-3):68–71. doi:[10.1016/j.livsci.2011.02.009](https://doi.org/10.1016/j.livsci.2011.02.009).
- Han S, Chen J, Hua J, Hu X, Jian S, Zheng G, Wang J, Li H, Yang J, Hejtmancik JF, et al. 2020. MITF protects against oxidative damage-induced retinal degeneration by regulating the NRF2 pathway in the retinal pigment epithelium. *Redox Biol.* 34: 101537. doi:[10.1016/j.redox.2020.101537](https://doi.org/10.1016/j.redox.2020.101537).
- Han B, Eskin E. 2011. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet.* 88(5):586–598. doi:[10.1016/j.ajhg.2011.04.014](https://doi.org/10.1016/j.ajhg.2011.04.014).
- Hayes BJ. 2011. Efficient parentage assignment and pedigree reconstruction with dense single nucleotide polymorphism data. *J Dairy Sci.* 94(4):2114–2117. doi:[10.3168/jds.2010-3896](https://doi.org/10.3168/jds.2010-3896).
- Helm B, Ben-Shlomo R, Sheriff MJ, Hut RA, Foster R, Barnes BM, Dominoni D. 2013. Annual rhythms that underlie phenology: biological time-keeping meets environmental change. *Proc Biol Sci.* 280(1765):20130016. doi:[10.1098/rspb.2013.0016](https://doi.org/10.1098/rspb.2013.0016).
- Heydon MJ, Milne JA, Brinklow BR, Loudon AS. 1995. 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.* 273(1):12–20. doi:[10.1002/jez.1402730103](https://doi.org/10.1002/jez.1402730103).
- Hou L, Pavan WJ. 2008. Transcriptional and signaling regulation in neural crest stem cell-derived melanocyte development: do all roads lead to mitf? *Cell Res.* 18(12):1163–1176. doi:[10.1038/cr.2008.303](https://doi.org/10.1038/cr.2008.303).
- Hu Z-L, Park CA, Reecy JM. 2019. Building a livestock genetic and genomic information knowledgebase through integrative developments of animal QTLdb and CorrDB. *Nucleic Acids Res.* 47(D1): D701–D710. doi:[10.1093/nar/gky1084](https://doi.org/10.1093/nar/gky1084).
- Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, et al. 2020. The reactome pathway knowledgebase. *Nucleic Acids Res.* 48(D1): D498–D503. doi:[10.1093/nar/gkz1031](https://doi.org/10.1093/nar/gkz1031).
- Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51(D1): D587–D592. doi:[10.1093/nar/gkac963](https://doi.org/10.1093/nar/gkac963).
- Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28(1):27–30. doi:[10.1093/nar/28.1.27](https://doi.org/10.1093/nar/28.1.27).

- Klotz JL. 2015. Activities and effects of ergot alkaloids on livestock physiology and production. *Toxins (Basel)*. 7(8):2801–2821. doi:[10.3390/toxins7082801](https://doi.org/10.3390/toxins7082801).
- Köhler S, Gargano M, Matentzoglou N, Carmody LC, Lewis-Smith D, Vasilevsky NA, Danis D, Balagura G, Baynam G, Brower AM, et al. 2021. The human phenotype ontology in 2021. *Nucleic Acids Res.* 49(D1):D1207–D1217. doi:[10.1093/nar/gkaa1043](https://doi.org/10.1093/nar/gkaa1043).
- Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. 2020. gprofiler2 – an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research*. 9(ELIXIR):709. doi:[10.12688/f1000research.24956.2](https://doi.org/10.12688/f1000research.24956.2).
- Köster J, Rahmann S. 2018. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics*. 34(20):3600. doi:[10.1093/bioinformatics/bty350](https://doi.org/10.1093/bioinformatics/bty350).
- Legarra A, Reverter A. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet Sel Evol.* 50(1):53. doi:[10.1186/s12711-018-0426-6](https://doi.org/10.1186/s12711-018-0426-6).
- Levy C, Khaled M, Fisher DE. 2006. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med.* 12(9):406–414. doi:[10.1016/j.molmed.2006.07.008](https://doi.org/10.1016/j.molmed.2006.07.008).
- Li H, Lian L, Liu B, Chen Y, Yang J, Jian S, Zhou J, Xu Y, Ma X, Qu J, et al. 2020. KIT ligand protects against both light-induced and genetic photoreceptor degeneration. *eLife*. 9:e51698. doi:[10.7554/eLife.51698](https://doi.org/10.7554/eLife.51698).
- Littlejohn MD, Henty KM, Tiplady K, Johnson T, Harland C, Lopdell T, Sherlock RG, Li W, Lukefahr SD, Shanks BC, et al. 2014. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. *Nat Commun.* 5(1):5861. doi:[10.1038/ncomms5861](https://doi.org/10.1038/ncomms5861).
- Lopes FB. 2016. DRP: R package to calculate deregressed proofs, and its reliabilities and weights. *github*. <https://github.com/camult/DRP>.
- Macedo FL, Reverter A, Legarra A. 2020. Behavior of the linear regression method to estimate bias and accuracies with correct and incorrect genetic evaluation models. *J Dairy Sci.* 103(1):529–544. doi:[10.3168/jds.2019-16603](https://doi.org/10.3168/jds.2019-16603).
- McCann SM, Vijayan E, Koenig J, Krulich L. 1982. The effects of neurotensin on anterior pituitary hormone secretion. *Ann N Y Acad Sci.* 400(1):160–171. doi:[10.1111/j.1749-6632.1982.tb31567.x](https://doi.org/10.1111/j.1749-6632.1982.tb31567.x).
- Misztal I, Tsuruta S, Lourenco D, Masuda Y, Aguilar I. 2014. BLUPF90 Family of Programs. Athens (GA): University of Georgia. <http://nce.ads.uga.edu/wiki/doku.php>.
- Mort RL, Jackson IJ, Elizabeth Patton E. 2015. The melanocyte lineage in development and disease. *Development (Cambridge)*. 142(4):620–632. doi:[10.1242/dev.106567](https://doi.org/10.1242/dev.106567).
- Murray DM. 1965. A field study of coat shedding in cattle under conditions of equal day-length but different temperatures. *J Agric Sci.* 65(3):295–300. doi:[10.1017/S0021859600048863](https://doi.org/10.1017/S0021859600048863).
- Nowak K, Berger J, Panikowski A, Reid DG, Jacob AL, Newman G, Young NE, Beckmann JP, Richards SA. 2020. Using community photography to investigate phenology: a case study of coat molt in the mountain goat (*Oreamnos americanus*) with missing data. *Ecol Evol.* 10:13488–13499. doi:[10.1002/ece3.6954](https://doi.org/10.1002/ece3.6954).
- Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet.* 2(12):e190. doi:[10.1371/journal.pgen.0020190](https://doi.org/10.1371/journal.pgen.0020190).
- Peters RR, Tucker HA. 1978. Prolactin and growth hormone responses to photoperiod in heifers. *Endocrinology*. 103(1):229–234. doi:[10.1210/endo-103-1-229](https://doi.org/10.1210/endo-103-1-229).
- Petitclerc D, Chapin LT, Emery RS, Tucker HA. 1983. Body growth, growth hormone, prolactin and puberty response to photoperiod and plane of nutrition in Holstein heifers. *J Anim Sci.* 57(4):892–898. doi:[10.2527/jas1983.574892x](https://doi.org/10.2527/jas1983.574892x).
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 81(3):559–575. doi:[10.1086/519795](https://doi.org/10.1086/519795).
- Rachel RA, Li T, Swaroop A. 2012. Photoreceptor sensory cilia and ciliopathies: focus on CEP290, RPGR and their interacting proteins. *Cilia*. 1(1):22. doi:[10.1186/2046-2530-1-22](https://doi.org/10.1186/2046-2530-1-22).
- Reiter RJ. 1991. Pineal melatonin: cell biology of its synthesis and of its physiological interactions. *Endocr Rev.* 12(2):151–180. doi:[10.1210/edrv-12-2-151](https://doi.org/10.1210/edrv-12-2-151).
- Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, Rowan TN, Low WY, Zimin A, Couldrey C, et al. 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 9(3):giaa021. doi:[10.1093/gigascience/giaa021](https://doi.org/10.1093/gigascience/giaa021).
- Rowan TN, Hoff JL, Crum TE, Taylor JF, Schnabel RD, Decker JE. 2019. A multi-breed reference panel and additional rare variants maximize imputation accuracy in cattle. *Genet Sel Evol.* 51(1):77. doi:[10.1186/s12711-019-0519-x](https://doi.org/10.1186/s12711-019-0519-x).
- Rudis B, Vilhelmsen L, Severski D. 2017. darksky: Tools to Work with the 'Dark Sky' 'API'. CRAN. <https://CRAN.R-project.org/package=darksky>.
- Sargolzaei M. 2014. SNP1101 User's Guide Version 1.0. Guelph (ON): HiggsGene Solutions Inc. p. 1–33.
- Schmidt K, Deichsel K, de Oliveira RA, Aurich J, Ille N, Aurich C. 2017. Effects of environmental temperature and season on hair coat characteristics, physiologic and reproductive parameters in shetland pony stallions. *Theriogenology*. 97:170–178. doi:[10.1016/j.theriogenology.2017.04.035](https://doi.org/10.1016/j.theriogenology.2017.04.035).
- Schmutz SM, Womack JE. 2012. Genetics of coat color in cattle. In: Womack JE, editor. *Bovine Genomics*. Ames (IA): John Wiley & Sons Inc. p. 20–33. <https://doi.org/10.1002/9781118301739.ch3>.
- Small JA, Glover ND, Kennedy AD, McCaughey WP, Ward DR. 2003. Photoperiod effects on the development of beef heifers. *Can J Anim Sci.* 83(4):721–730. doi:[10.4141/A03-005](https://doi.org/10.4141/A03-005).
- St-Pierre NR, Cobanov B, Schnitkey G. 2003. Economic losses from heat stress by US livestock industries. *J Dairy Sci.* 86:E52–E77. doi:[10.3168/jds.S0022-0302\(03\)74040-5](https://doi.org/10.3168/jds.S0022-0302(03)74040-5).
- Steingrímsson E, Copeland NG, Jenkins NA. 2004. Melanocytes and the microphthalmia transcription factor network. *Annu Rev Genet.* 38(1):365–411. doi:[10.1146/annurev.genet.38.072902.092717](https://doi.org/10.1146/annurev.genet.38.072902.092717).
- Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, et al. 2007. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet.* 39(6):770–775. doi:[10.1038/ng2043](https://doi.org/10.1038/ng2043).
- Stenn KS, Paus R. 2001. Controls of hair follicle cycling. *Physiol Rev.* 81(1):449–494. doi:[10.1152/physrev.2001.81.1.449](https://doi.org/10.1152/physrev.2001.81.1.449).
- Storey JD, Bass AJ, Dabney A, Robinson D. 2017. qvalue: Q-value estimation for false discovery rate control. *bioconductor*. doi:[10.18129/B9.bioc.qvalue](https://doi.org/10.18129/B9.bioc.qvalue).
- Tassabehji M, Newton VE, Read AP. 1994. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nat Genet.* 8(3):251–255. doi:[10.1038/ng1194-251](https://doi.org/10.1038/ng1194-251).
- Tucker HA, Petitclerc D, Zinn SA. 1984. The influence of photoperiod on body weight gain, body composition, nutrient intake and hormone secretion. *J Anim Sci.* 59(6):1610–1620. doi:[10.2527/jas1984.5961610x](https://doi.org/10.2527/jas1984.5961610x).
- Turner HG. 1964. Coat characters of cattle in relation to adaptation. *Proceed Austral Soc Animal Prod.* 5:181–187. <http://www.livestocklibrary.com.au/handle/1234/6335>.
- Turner HG, Schleger AV. 1960. The significance of coat type in cattle. *Aust J Agric Res.* 11(4):645–663. doi:[10.1071/ar9600645](https://doi.org/10.1071/ar9600645).
- VanRaden PM. 2008. Efficient methods to compute genomic predictions. *J Dairy Sci.* 91(11):4414–4423. doi:[10.3168/jds.2007-0980](https://doi.org/10.3168/jds.2007-0980).

- Wellmann R. 2019. Optimum contribution selection for animal breeding and conservation: the R package optiSel. *BMC Bioinformatics*. 20:25. doi:[10.1186/s12859-018-2450-5](https://doi.org/10.1186/s12859-018-2450-5).
- Wen B, Li S, Li H, Chen Y, Ma X, Wang J, Lu F, Qu J, Hou L. 2016. Microphthalmia-associated transcription factor regulates the visual cycle genes *rlbp1* and *rdh5* in the retinal pigment epithelium. *Sci Rep*. 6(1):21208. doi:[10.1038/srep21208](https://doi.org/10.1038/srep21208).
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. 2019. Welcome to the tidyverse. *J Open Source Softw*. 4(43):1686. doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Yang J, Ferreira T, Morris AP, Medland SE, Madden PAF, Heath AC, Martin NG, Montgomery GW, Weedon MN, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, et al. 2012. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 44(4):369–375. S1–S3. doi:[10.1038/ng.2213](https://doi.org/10.1038/ng.2213).
- Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 88(1):76–82. doi:[10.1016/j.ajhg.2010.11.011](https://doi.org/10.1016/j.ajhg.2010.11.011).
- Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. 2014. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet*. 46(2):100–106. doi:[10.1038/ng.2876](https://doi.org/10.1038/ng.2876).
- Yeates NTM. 1955. Photoperiodicity in cattle. I. Seasonal changes in coat character and their importance in heat regulation. *Aust J Agric Res*. 6(6):891–902. doi:[10.1071/ar9550891](https://doi.org/10.1071/ar9550891).
- Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 44(7):821–824. doi:[10.1038/ng.2310](https://doi.org/10.1038/ng.2310).
- Zimova M, Hackländer K, Good JM, Melo-Ferreira J, Alves PC, Mills LS. 2018. Function and underlying mechanisms of seasonal colour moulting in mammals and birds: what keeps them changing in a warming world? *Biol Rev Camb Philos Soc*. 93(3):1478–1498. doi:[10.1111/brev.12405](https://doi.org/10.1111/brev.12405).
- Zimova M, Mills LS, Lukacs PM, Mitchell MS. 2014. Snowshoe hares display limited phenotypic plasticity to mismatch in seasonal camouflage. *Proc Biol Sci*. 281(1782):20140029. doi:[10.1098/rspb.2014.0029](https://doi.org/10.1098/rspb.2014.0029).

Editor: A. Whitehead