Sustainable production: genomic predictions for selecting locally adapted beef cattle

Sara M. Nilson[1], Troy N. Rowan[1], Robert D. Schanbel[1,2], Jared E. Decker[1,2]

[1]Division of Animal Science, University of Missouri, Columbia, MO
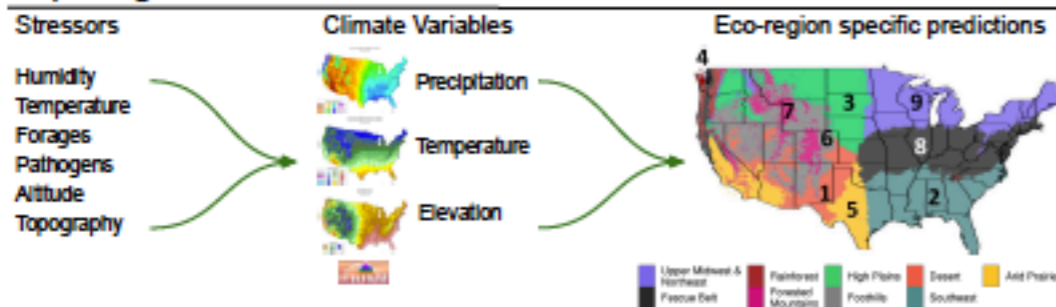[2]Institute for Data Science and Informatics, University of Missouri, Columbia, MO

Abstract

To feed ~10 billion people by 2050, beef production needs to increase. To sustainably increase production, environment-specific genomic predictions were formulated to improve selection of environmentally adapted beef cattle. To classify environments, K-means clustering was performed with 30-year normals of precipitation, temperature, and elevation resulting in 9 environmental regions for the United States. This multivariate approach additionally captures local effects including but not limited to pathogens, forages, and humidity. Simmental, Red Angus, and Gelbvieh breeds were utilized with over 17,000, 15,000, and 12,000, individuals respectively. Phenotypes of birth weight, weaning weight, and yearling weight were chosen due to the large number of available records. Genotypes were phased and imputed up to ~850,000 SNPs and SNPs with a minor allele frequency >0.01 were retained. In BLUPF90, bivariate models were analyzed for each region combination, phenotype, and breed. Region 3 and Region 8 for birth weight for Simmental. Regions with less than 1,000 animals were not included. Multivariate models of all regions were analyzed for each phenotype and breed. Genetic correlations indicate substantial GxE effects among regions, reflected by different allele substitution effects between regions. Environment-specific heritabilities were calculated. Breeding values were estimated (EBV) for every animal for each region and phenotype. Accuracy was defined as the correlation between EBVs of the full dataset and EBVs when the youngest 25% of animals were excluded. Animals significantly reranked between regions. These predictions will serve as valuable tools to identify exceptional animals that will perform in their environment before breeding or purchasing an animal.
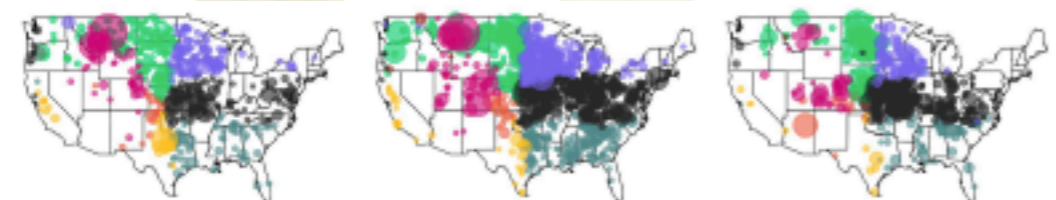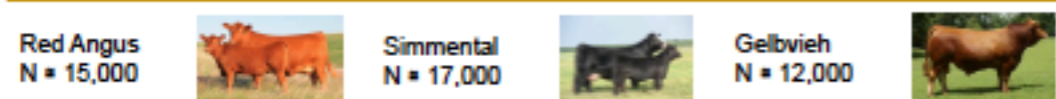
Methods

- Birth weight (BW), weaning weight (WW), and yearling weight (YW) were the phenotypes selected to identify genotype-by-environment interactions in life stages and due to the number of records available.
- Red Angus, Simmental, and Gelbvieh breed associations shared their genotypes, phenotypes and pedigree information.
  - Genotypes were imputed[1] to ~850,000 SNPs.
- K-means clustering was performed with the stats package[2] on 30-year normals of precipitation, temperature and elevation in R program[2] identifying 9 eco-regions.
- Bivariate models between regions with over 1,000 phenotypes for each BW, WW, and YW were run in the BLUPF90 suite[4] to estimate breeding values (EBV) and SNP effects.
  - Models accounted for contemporary group, sex, and maternal effect (BW, WW).
  - For each bivariate model: each animal had an EBV for each region, and each SNP had an effect estimate for each region.
- Pearson and Spearman correlations were calculated between the EBV for the bivariate models with the stats package[2] in R program[2].
- The difference in SNP effects for each bivariate model were calculated and plotted in R[2].

Capturing environmental effects

Stressors
Humidity
Temperature
Forages
Pathogens
Altitude
Topography

Climate Variables
Precipitation
Temperature
Elevation

Eco-region specific predictions

Upper Midwest & Northeast | Rainforest | High Plains | Desert | Arid Prairie
Fescue Belt | Forested Mountains | Foothills | Southeast

Large impact with large populations widely distributed

Red Angus N = 15,000

Simmental N = 17,000

Gelbvieh N = 12,000

Number of Simmental phenotypes per eco-region

	Southeast	High Plains	Forested Mountains	Fescue Belt	Upper Midwest & Northeast
Birth Weight	1305	3457	2538	4348	2528
Weaning Weight	1012	3459	2489	3435	2226
Yearling Weight	639	2626	2065	2240	1535

Citations and acknowledgements

1. Rowan, Troy N., Jesse L. Hoff, Tamar E. Crum, Jeremy F. Taylor, Robert D. Schnabel, and Jared E. Decker. 2019. "A Multi-Breed Reference Panel and Additional Rare Variants Maximize Imputation Accuracy in Cattle." *Genetics, Selection, Evolution: GSE* 51 (1): 77.
2. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
3. Misztal, I., S. Tsuruta, D. A. L. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, Z. Vitezica. 2018. Manual for BLUPF90 family programs. University of Georgia. http://nce.ads.uga.edu/ wiki/doku.php?id=documentation

Simmental: EBVs change across eco-regions

r = 0.965, ρ = 0.963 (High Plains vs Southeast)
r = 0.954, ρ = 0.951 (Forested Mountains vs Southeast)
r = 0.999, ρ = 0.999 (Fescue Belt vs Southeast)
r = 0.781, ρ = 0.768 (Upper Midwest & Northeast vs Southeast)
r = 0.876, ρ = 0.872 (Forested Mountains vs High Plains)
r = 0.967, ρ = 0.965 (Fescue Belt vs High Plains)
r = 0.987, ρ = 0.986 (Upper Midwest & Northeast vs High Plains)
r = 0.973, ρ = 0.973 (Fescue Belt vs Forested Mountains)
r = 0.891, ρ = 0.883 (Upper Midwest & Northeast vs Forested Mountains)
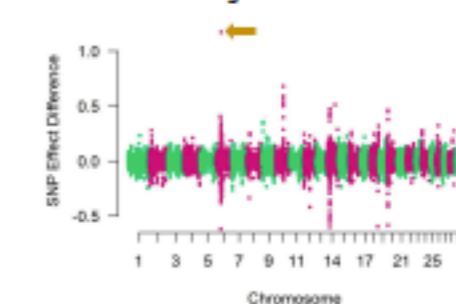r = 1, ρ = 1 (Upper Midwest & Northeast vs Fescue Belt)

SNP effect differences between eco-regions

Bivariate model of High Plains and Forested Mountains

High Plains | Forested Mountains
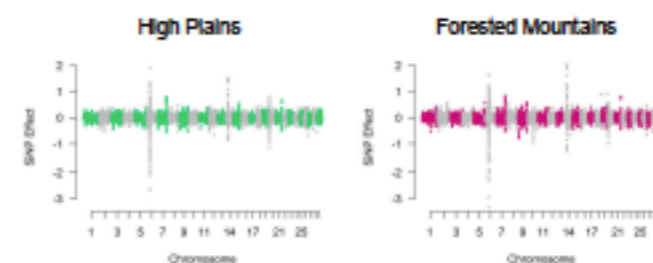
Difference between High Plains & Forested Mountains

Genomic regions of SNP differences

Chromosome 6
QTL:
- Milk yield
- Body weight (yearling)
- Average daily gain

Gene:
- HERC3
- NAP1L5

Chromosome 14
QTL:
- Residual feed intake
- Body weight (birth)
- Shear force
- Carcass weight
- Scrotal circumference

Gene:
- LYN

Chromosome 10
QTL:
- Fatty acid content
- Shear force

Gene:
- WDR72

Chromosome 18
QTL:
- Calving ease
- Calf size
- Dry matter intake
- Stillbirth

Chromosome 20
QTL:
- Body weight

Conclusions and future directions

- An animal's estimated breeding value and rank will change among eco-regions due to differences in genotype-by-environment interactions.
- SNP effect estimates are dependent on the regions included in the model.
  - Pearson correlation between SNP effect estimates for High Plains:
    - When paired separately with Southeast and Forested Mountains, r = 0.711
- Many genomic regions of interest due to differences in effect contain known QTLs related to growth and production traits.
- These eco-region specific genomic predictions will better able help producers decide where to select cattle from instead of the national rank which does not take into account genotype-by-environment interactions.
- By implementing these predictions into practice, the industry will be able to take advantage of underlying interactions to create locally adapted beef cattle in a sustainable manner.
- Multivariate models with all eco-regions would allow for the simultaneous estimate of estimated breeding values with genotype-by-environment interactions allowing for direct comparisons.
- To reduce noise and improve accuracy, the number of SNPs utilized will be reduced to those identified as significant by environmental GWAS.

Contact Us

Email for presenting author:
smnz7@mail.missouri.edu

Follow on Twitter: @DrSNilson