Neupane, M., Ben Zaabza, H., Toghiani, S., Norman, H., Cole, J., Van Tassell, C. (2026) Genomic regions associated with heterosis in U.S. Holstein-Jersey crossbreds, Proceedings 13th World Congress on Genetics Applied to Livestock Production 24306:
https://doi.org/10.31274/wcgalp.24306 © 2026 *Neupane, et al.* CC BY-NC-ND

Genomic regions associated with heterosis in U.S. Holstein-Jersey crossbreds

M. Neupane[1*], H. Ben Zaabza[2], S. Toghiani[1], H.D. Norman[3], J.B. Cole[3,4,5,6], and C.P. Van Tassell[1]

[1]Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA; [2]University of Georgia, Athens, Georgia, GA, USA; [3]Council on Dairy Cattle Breeding, Bowie, MD, USA; [4]University of Florida, Gainesville, FL, USA; [5]North Carolina State University, Raleigh, NC, USA; [6]University of Guelph, ON, Canada; [*]Corresponding author: mahesh.neupane@usda.gov

Introduction

Holstein (HO) × Jersey (JE) crossbreeding leverages the complementary strengths of two globally important dairy breeds, combining the high milk yield of HO with the superior milk component concentrations, fertility, and feed efficiency characteristics of JE (Weigel and Barlass, 2003). Crossbreeding also helps mitigate the accumulation of inbreeding by increasing genome-wide heterozygosity. Genomic breed composition provides a quantitative description of admixture, while runs of homozygosity (ROH) summarize genome-wide autozygosity (inbreeding) and can highlight candidate regions shaped by selection. Conversely, runs of heterozygosity (ROHet) can identify genomic intervals where heterozygosity is maintained in crossbreds, potentially reflecting balancing selection, or heterozygote advantage (Ghoreishifar *et al.*, 2014). The objective of this study was to characterize genomic breed composition, ROH, and ROHet patterns in HO-JE crossbred dairy cattle to better understand the genetic architecture underlying heterosis.

Materials & Methods

Animals and genotypes. This study analyzed 5,660 HO-JE crossbred males with <90% breed-base representation (BBR) from either breed, genotyped using 69,200 imputed markers. The BBR is an estimate of the proportion of an individual animal's genome that that most closely matches each purebred reference populations (Holstein, Jersey, Ayrshire, Brown Swiss, and Guernsey in U.S. dairy evaluations) (VanRaden *et al.*, 2020). After quality control (autosomal markers; call rate ≥0.90; minor allele frequency≥0.01), 5,652 animals and 65,937 markers remained. To assess ancestry-dependent patterns, crossbreds were further grouped into three JE-ancestry classes: 10-37.5% (n=333), 37.5-62.5% (n=4,007), and 62.5-90%, (n=1,312). Purebred references (35,210 HO and 15,529 JE) were also included and processed using the same quality control criteria.

Principal component and admixture analysis. Principal component analysis was conducted using the approximation algorithm implemented in PLINK 2.0. Admixture proportions were inferred using ADMIXTURE v1.3.0 (Alexander *et al*., 2009) in supervised mode, specifying pure HO and JE animals as reference populations for k=2.

Runs of homozygosity and heterozygosity. ROH and ROHet were identified using the detectRUNS v 0.9.6 package (Biscarini *et al.*, 2019) in R 4.5. ROH were defined as segments ≥500 kb with ≥20 SNPs, allowing up to one heterozygous and one missing genotype, and inter-marker gaps <500 kb, following parameters from previous cattle studies (Meyermans *et al.*, 2020). ROHet

segments were defined using identical density and length parameters but required the segments to have heterozygosity instead of homozygosity. For each animal, FROH was calculated as the proportion of the autosomal genome contained within ROH.

Mapping and enrichment analysis. Genes within 1 Mb of ROH and ROHet islands (contiguous genomic regions with high frequency of ROH and ROHet) were annotated using the ARS-UCD1.2 assembly. QTLs for these regions were retrieved with the GALLO 1.5 R package (Fonseca *et al*., 2020) using animal QTLdb, and trait enrichment analysis was conducted to identify overrepresented traits using an FDR-adjusted threshold of $P < 0.05$.

Results & Discussion

PCA showed clear clustering by ancestry (Figure 1a), and the first two PCs explained 85% of the variation (Figure 1b). The supervised admixture plot, using HO and JE as reference groups, revealed a broad gradient of intermediate ancestry proportions in the crossbreds, consistent with varying HO-JE admixture levels (Figure 1c). Admixture-derived ancestry proportions were highly correlated with independent BBR breed-composition estimates for 5,652 crossbred animals (Figure 1d).

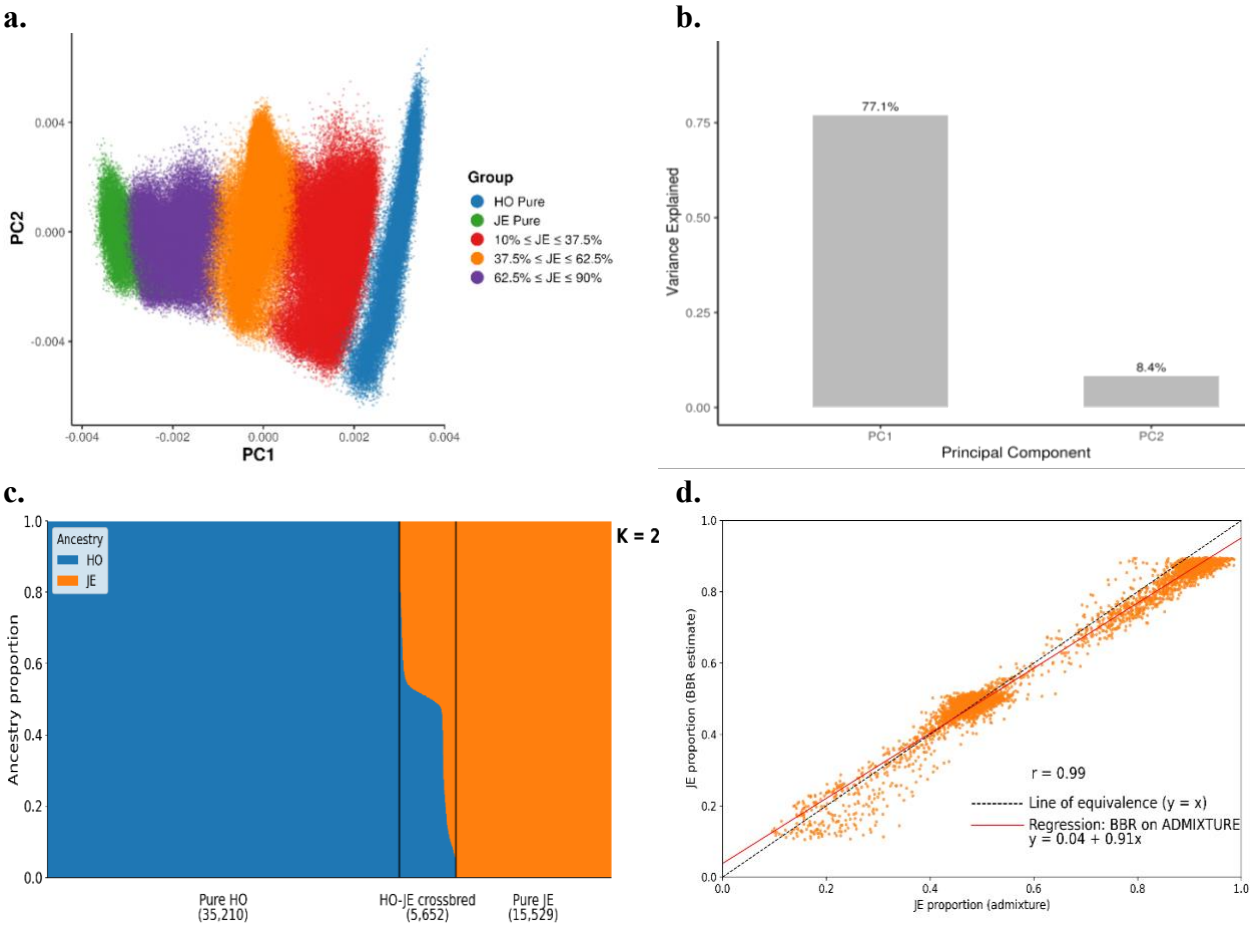

Figure1. a. Principal component analysis of HO-JE crossbreds and pure breeds b. Variance explained by the first two principal components. c. Supervised admixture analysis d. Concordance between admixture estimates and breed base representation.

ROH and FROH revealed strong ancestry-dependent autozygosity (Table 1). The lowest ROH burden in the intermediate JE 37.5-62.5% group is consistent with increased heterozygosity expected in crossbreds closest to an F1-like background. Elevated ROH/FROH at low and high JE ancestry likely reflects increased contributions of within-breed haplotypes. ROHet provided complementary insight, with total lengths remaining small relative to ROH and reached their highest levels in the intermediate JE ancestry group.

Table1. Summary of ROH, FROH and ROHet across JE ancestry groups.

JE ancestry percentage (n)	Total ROH (Mb) mean	Total ROH (Mb) median	FROH mean	FROH median	Total ROHet (Mb) median
10-37.5% (333)	260.77	244.90	0.11	0.10	2.33
37.5-62.5% (4,007)	105.54	81.37	0.04	0.03	2.62
62.5-90% (1,312)	465.25	466.35	0.19	0.19	2.04

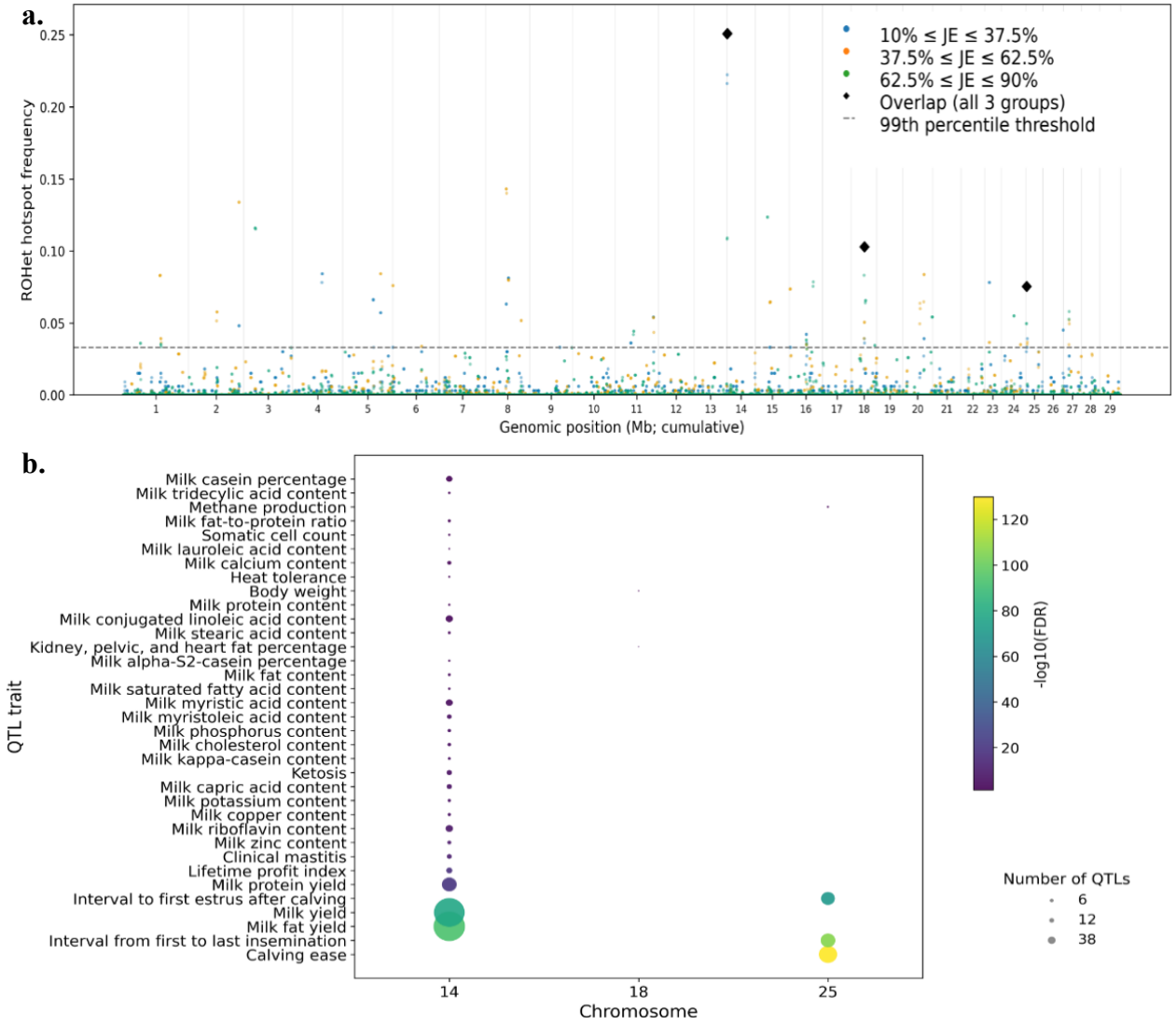

Figure 2. a. Manhattan-like plot of runs of heterozygosity hotspot across three Jersey ancestry groups b. QTL enrichment analysis (FDR <0.05).

Hotspot (regions where runs of ROH or ROHet occur much more frequently than elsewhere) mapping identified genomic regions that overlapped ROH or ROHet segments across different JE ancestry groups. A prominent ROHet island shared across all three JE groups was located at Chr14:0-1 Mb (Figure 2a). This gene dense interval (58 genes) contains *DGAT1*, along with multiple additional loci associated with milk yield, milk fat yield, milk fat content, and related milk-composition traits in QTL analysis. The high heterozygosity at this Chromosome 14 locus across all ancestry groups suggests a potential heterozygote advantage, where maintenance of both HO and JE haplotypes may be favored for optimal milk composition. Given its biological relevance and the concentration of functionally important genes, this chromosome 14 region represents a high-priority for follow-up analyses, including haplotype characterization and association testing for milk, fertility, and health traits in HO-JE crossbred population. Additional shared regions of runs of heterozygosity were detected on chromosome 18, containing QTL associated with body weight and body fat percentage, and on chromosome 25, which harbored loci linked to fertility and methane production (Figure 2b). These multi-chromosome findings suggest that heterosis in HO-JE crosses likely involves a complex interplay of traits, ranging from production efficiency to improved environmental and reproductive outcomes.

Conclusions

Integrating breed composition provides an interpretable framework to monitor retained heterozygosity in HO-JE crossbreds and prioritizes candidate genomic regions for downstream trait association analysis of heterosis.

Acknowledgements

The authors thank Cooperative Dairy DNA Repository (CDDR) partners, Council on Dairy Cattle Breeding (CDCB, Bowie, MD) and dairy farmers for providing data used in this study. This research used resources provided by the SCINet project and/or the AI Center of Excellence of the USDA Agricultural Research Service, ARS project numbers 0201-88888-003-000D and 0201-88888-002-000D.

References

Alexander, D.H., Novembre, J., and Lange, K. (2009) Genome Res., 19(9):1655–64. https://doi.org/10.1101/gr.094052.109

Biscarini F., Cozzi P., Gaspa G., Marras G. (2019) R package version 0.9.6. https://cran.r-project.org/package=detectRUNS

Fonseca P.A.S., Suárez-Vega A., Marras G., *et al.* (2020) Gigascience 9(12):giaa149. https://doi.org/10.1093/gigascience/giaa149

Ghoreishifar, M., Vahedi, S. M., Salek Ardestani, *et al.* (2023). BMC Genomics, 24(1):230. https://doi.org/10.1186/s12864-023-09298-1

Meyermans R., Gorssen W., Buys N., and Janssens S. (2020) BMC Genomics 21:1-14. https://doi.org/10.1186/s12864-020-6463-x

VanRaden, P.M., Tooker M.E., Chud, T.C.S., *et al.* (2020) J. Dairy Sci., 103(2) 1620-31. https://doi.org/10.3168/jds.2019-16634

Weigel, K. A. & Barlass, K. A. (2003). J. Dairy Sci., 86(12), 4148-54. https://doi.org/10.3168/jds.S0022-0302(03)74029-6