Ladeira, G., Pinedo, P., Santos, J., Thatcher, W., Rezende, F. (2026) CNV-assisted genetic parameters and genome-wide associations for low-heritability traits, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 24249: https://doi.org/10.31274/wcgalp.24249
©2026 *Ladeira, et al.* CC BY-NC-ND

CNV-assisted genetic parameters and genome-wide associations for low-heritability traits

G.C. Ladeira[1*], P.J. Pinedo[2], J.E.P. Santos[1], W.W. Thatcher[1], and F.M. Rezende[1]

[1]University of Florida, Gainesville, FL, USA; [2]Colorado State University, Fort Collins, CO, USA; [*]Corresponding author: g.coelholadeira@ufl.edu

Introduction

Although single nucleotide polymorphisms (SNPs) have been invaluable in genomic selection, they may capture only part of the genetic architecture of complex traits. Structural genomic variants, such as copy number variations (CNVs), may complement SNP-based models by capturing variation in gene and regulatory element dosage that is not accounted for by SNPs. In fact, the low SNP-CNV linkage disequilibrium suggests that these markers may capture largely distinct genomic information (Lee *et al*., 2020). Evidence that CNVs are associated with fertility and disease phenotypes in dairy cattle (Oliveira *et al*., 2024) further suggests that these variants may explain additional genetic effects relevant to low-heritability traits that are overlooked by SNPs. Despite this potential, the additive genetic contribution of CNVs to low-heritability traits remains largely uncharacterized, and CNV-based GWAS are still scarce. Therefore, the objectives of this study were (1) to quantify the additive genetic variance explained by CNVs beyond that captured by SNPs for uterine health and reproductive traits in Holstein cattle and (2) to identify CNV-based trait-associated genomic regions.

Materials & Methods

Population and genotyping. The study population consisted of 3,601 genotyped Holsteins, including 3,387 cows from 16 herds located in seven U.S. states and 214 commercially available artificial insemination (AI) sires. The cohort of genotyped cows calved between 2012 and 2014. All animals were genotyped using the Illumina BovineHD BeadChip array (n = 777,962 SNPs), and SNP coordinates were updated to the ARS-UCD1.2 genome version. Next, CNVs were detected using PennCNV software v1.0.5 (Wang *et al*., 2007), as described in Ladeira *et al*. (2025). The CNV calling procedure yielded 4,113 non-redundant, high-confidence CNVs distributed across all bovine autosomal chromosomes (BTA) in 3,463 animals.

Phenotypes. Retained fetal membrane (RFM; n = 3,305) was defined as retention of fetal membranes for more than 24 hours after calving. Metritis (MET; n = 3,305) was evaluated by vaginal discharge at 7 ± 3 days in milk (DIM) using a 0–5 scoring system (0 = no mucus, 1 = crystalline, 2 = flecks of pus, 3 = mucopurulent < 50% pus, 4 = purulent > 50% pus, and 5 = watery, reddish/brownish, fetid discharge), where a score 5 indicated the presence of metritis. Clinical endometritis (CE; n = 3,303) was assessed at 28 ± 7 DIM using the same 0–5 vaginal discharge scoring system, with scores > 2 indicating CE. Resumption of ovarian cyclicity (ROC; n = 3,303) was determined via transrectal ultrasonography or serum progesterone levels at 40 ± 3 and 54 ± 3 DIM. Ovulation was defined as a visible corpus luteum on at least one ultrasound scan or progesterone concentration > 1 ng/mL in at least one serum sample. Pregnancy on day 60 after AI (P60; n = 3,207) was diagnosed via ultrasonography at 60 ± 3 days after insemination. All traits

were recorded as binary outcomes, where 0 and 1 represented the absence and presence of the event, respectively.

Variance components estimation. The SNP-derived genomic relationship matrix ($\mathbf{G}_{\mathrm{SNP}}$) was created using 578,677 SNPs with minor allelic frequency (MAF) ≥ 0.01 and call rate ≥ 90%, while the CNV-derived genomic relationship matrix ($\mathbf{G}_{\mathrm{CNV}}$) was created using all the 4,154 CNVs. Here, to model copy number at mixed CNV loci (i.e., loci where both deletions and duplications were observed across individuals), the 41 mixed loci were split into deletion- and duplication-specific loci, increasing the number of analyzed CNVs from 4,113 to 4,154. $\mathbf{G}_{\mathrm{SNP}}$ and $\mathbf{G}_{\mathrm{CNV}}$ were independently constructed following the first method described in VanRaden (2008), with SNP and CNV loci coded as 0,1,2. For each low-heritability uterine health and reproductive trait, variance components were estimated on the liability scale by fitting two independent single-trait threshold animal models (SNP-based and SNP+CNV-based) implemented in ASReml-R (Butler *et al*., 2023). Both models included parity category and farm-year-season as fixed effects and assumed a binomial distribution for the observed responses. The SNP-based model estimated additive genetic variance using SNP information only, whereas the SNP+CNV-based model included two polygenic effects, one based on SNPs and the other on CNVs, to quantify the additive genetic variance attributable to each marker type.

Genome-wide association study. A total of 2,654 CNVs located within CNV regions with MAF ≥ 0.01 and 578,677 SNPs were individually tested as fixed effects for each binary trait (RFM, MET, ROC, and P60) using single-trait generalized linear mixed models with a logit link function in the GCTA software (Yang *et al*., 2011). Each model included the fixed effects of a single marker genotype (CNV or SNP), farm, year, season, and parity, and a random polygenic effect to account for relatedness using a SNP-derived genomic relationship matrix constructed from 50,000 SNPs randomly sampled proportional to chromosome length. Finally, gene annotation was performed by retrieving all genes located within a ±500 kb window around CNVs that were significantly ($P < 0.05$) or suggestively ($P < 0.10$) associated with the analyzed traits.

Results

The SNP+CNV-based model captured more additive genetic variance for all traits than the SNP-based model, resulting in higher heritability estimates (Table 1). As expected, MET, CE, ROC, and P60 exhibited low heritability in both SNP-based and SNP+CNV-based models. However, the contribution of CNVs to additive genetic variance was trait-specific, with some low-heritability traits showing more than a twofold increase when CNVs were included in the model. Although SNP-based and SNP+CNV-based models exhibited similar log-likelihoods across traits, the log-likelihood ratio test showed that incorporating CNV significantly improved model fit for MET ($P < 0.01$). In contrast, the RFM models failed to converge due to a highly imbalanced case–control phenotypic ratio.

The CNV-GWAS revealed a duplication on BTA29 associated with CE and P60 and suggestively associated with MET. Remarkably, this region was not detected by SNPs. This duplication flagged candidate genes for uterine health, including the *GSTP1*, *PPP1CA*, and *KDM2A*. Markedly, cows carrying the BTA29 duplication showed approximately twice the prevalence of CE and MET compared with non-carriers and had about a fivefold reduction in confirmed pregnancy. Additionally, a deletion on BTA22 was suggestively associated with RFM. Interestingly, this genomic region, identified exclusively by the CNV-GWAS, harbors the *DBNL*, *LANCL2*, and

UBE2D4 candidate genes for RFM. Furthermore, carriers of the BTA22 deletion exhibited roughly seven times greater RFM prevalence than non-carriers. No CNVs were associated with ROC. Overall, CNVs revealed trait-associated genomic regions that were not captured by the SNP-GWAS.

Table 1. Additive genetic variance (σ^2_a) and heritability (h^2) estimates with standard errors (SE), on the liability scale, from SNP-based and SNP+CNV-based models for uterine health and reproductive traits in Holstein cattle.

Trait	Model	$\sigma^2_a \pm$ SE	σ^2_a change (%)	h^2	h^2 change (%)
MET	SNP-based	0.076 ± 0.029	+27.6%	0.071 ± 0.025	+23.9%
	SNP+CNV-based	0.097 ± 0.031		0.088 ± 0.026	
CE	SNP-based	0.041 ± 0.027	+121.9%	0.039 ± 0.025	+115.4%
	SNP+CNV-based	0.091 ± 0.031		0.084 ± 0.026	
ROC	SNP-based	0.088 ± 0.030	+26.1%	0.081 ± 0.025	+23.5%
	SNP+CNV-based	0.111 ± 0.032		0.100 ± 0.026	
P60	SNP-based	0.099 ± 0.030	+117.2%	0.090 ± 0.025	+96.7%
	SNP+CNV-based	0.215 ± 0.039		0.177 ± 0.026	

MET = metritis; CE = clinical endometritis; ROC = resumption of ovarian cyclicity; P60 = pregnancy on day 60 after artificial insemination. All estimates are on the liability scale. Models for retained fetal membranes (RFM) failed to converge due to a case–control imbalance and are therefore not shown.

Discussion

The SNP+CNV model revealed that CNVs contribute substantially to additive genetic variance and heritability estimates for the low-heritability traits studied here. This may be explained by three key mechanisms. First, CNVs can alter gene dosage, delete or duplicate regulatory elements, or create breakpoint fusions (Feuk *et al*., 2006), providing an important source of genetic variation and potentially modifying gene expression patterns. Second, CNV regions (CNVRs) mapped in our Holstein population overlapped 1,630 reproduction- and 458 health-related QTLs reported in the cattle genome (Ladeira *et al*., 2025), suggesting that CNVs underlie genetic variation in low-heritability traits. In the same study, a CNVR-based QTL enrichment analysis further revealed non-random overlap between reproduction- and health-related QTLs, implying that CNVs and these QTLs segregate in a coordinate manner. Third, this same population showed low SNP-CNV linkage disequilibrium, consistent with reports in other dairy cattle populations (Lee *et al*., 2020), indicating that CNV-driven variations are not fully captured by SNPs and may be overlooked in traditional SNP-based models. By capturing additional additive genetic variance beyond SNPs, CNV-inclusive models could improve prediction accuracy of breeding values for uterine health and reproductive traits, which should be tested in future studies.

The CNV-GWAS also uncovered candidate genes for uterine health and reproductive traits that were not flagged by traditional SNP-GWAS. Indeed, the *GSTP1* gene on BTA29, previously associated with endometriosis and endometrial hyperplasia in women (Ertunc *et al*., 2005; Elsaid *et al*., 2015), is likely involved in uterine detoxification and tissue remodeling, potentially affecting susceptibility to CE and MET. The *PPP1CA* gene, whose overexpression reduces senescence of endometrial epithelial cells in sheep (Jiao *et al*., 2024), may influence endometrial renewal and uterine recovery postpartum, supporting its association with uterine health traits. Moreover, the *KDM2A* gene, which is involved in innate immunity and cell proliferation in the uterus of cows

exposed to *Escherichia coli* (Najafi *et al*., 2022), overlaps by about 70% with the BTA29 duplication, further supporting its potential role in uterine inflammatory responses.
The BTA22 deletion revealed candidate genes for RFM, including *DBNL*, which is implicated in cytoskeletal remodeling (Martinez-Garay, 2020), *LANCL2*, which is involved in epithelial detachment (Landlinger *et al*., 2006), and *UBE2D4*, which contributes to placental invasion through ubiquitination (Kalko *et al*., 2014). Hence, these genes participate in processes critical for normal placental separation and expulsion, making them candidate genes for RFM. Overall, the CNV-GWAS unveiled candidate genes for uterine health and reproductive traits not detected by traditional SNP-based analyses.

Conclusions

Incorporating CNVs into genomic models captured additional additive genetic variance for low-heritability traits beyond that explained by SNPs, supported by CNV-QTL overlap and low CNV-SNP linkage disequilibrium. Similarly, the CNV-GWAS revealed candidate genes not flagged by SNPs. Therefore, integrating SNP- and CNV-based GWAS provides a more comprehensive understanding of the genetic architecture underlying low-heritability traits.

References

Butler D.G., Cullis B.R., Gilmour A.R., Gogel B.J., and Thompson R. (2023) https://asreml.kb.vsni.co.uk/knowledge-base/asreml_r_documentation/

Elsaid A., Al-Kholy W., Ramadan R., and Elshazli R. (2015) Egypt. J. Med. Hum. Genet. 16:361–365. https://doi.org/10.1016/j.ejmhg.2015.04.005

Ertunc D., Aban M., Tok E.C., Tamer L., Arslan M., et al. (2005) Hum. Reprod. 20:2157–2161. https://doi.org/10.1093/humrep/dei053

Feuk L., Carson A.R., and Scherer S.W. (2006) Nat. Rev. Genet. 7:85–97. https://doi.org/10.1038/nrg1767

Jiao X., Jiao Y., Cui J., Zhang H., Li X., et al. (2024) Front. Vet. Sci. 11:1466482. https://doi.org/10.3389/fvets.2024.1466482

Kalko S.G., Paco S., Jou C., Rodríguez M.A., Meznaric M., et al. (2014) BMC Genomics 15:91. https://doi.org/10.1186/1471-2164-15-91

Ladeira G.C., Pinedo P.J., Santos J.E.P., Thatcher W.W., and Rezende F.M. (2025) BMC Genomics 26:381 https://doi.org/10.1186/s12864-025-11536-7

Landlinger C., Salzer U., and Prohaska R. (2006) Biochim. Biophys. Acta Biomembr. 1758(11):1759–1767. https://doi.org/10.1016/j.bbamem.2006.07.018

Lee Y.L., Bosse M., Mullaart E., Groenen M.A.M., Veerkamp R.F., et al (2020) BMC Genomics 21:89. https://doi.org/10.1186/s12864-020-6496-1

Martinez-Garay I. (2020) Front. Cell Dev. Biol. 8:588152. https://doi.org/10.3389/fcell.2020.588152

Najafi M., Guo Y., Andersson G., Humblot P., and Bongcam-Rudloff E. (2022) Genes 13(12):2342 https://doi.org/10.3390/genes13122342

Oliveira H.R., Chud T.C.S., Oliveira G.A., Hermisdorff I.C., Narayana S.G., et al. (2024) J. Dairy Sci. 107(9): 7052-7063.https://doi.org/10.3168/jds.2023-24295

VanRaden P.M. (2008) J. Dairy Sci. 91(11):4414–4423.https://doi.org/10.3168/jds.2007-0980

Wang K., Li M., Hadley D., Liu R., Glessner J., et al. (2007) Genome Res. 17:1665–1674. https://doi.org/10.1101/gr.6861907

Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) Am J Hum Genet. 88(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011