



Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits

K. L. Parker Gaddis,*¹ D. J. Null,† and J. B. Cole†

*Department of Animal Sciences, University of Florida, Gainesville 32611

†Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

The objective of this study was to identify single nucleotide polymorphisms and gene networks associated with 3 fertility traits in dairy cattle—daughter pregnancy rate, heifer conception rate, and cow conception rate—using different approaches. Deregressed predicted transmitting abilities were available for approximately 24,000 Holstein bulls and 36,000 Holstein cows sampled from the National Dairy Database with high-density genotypes. Of those, 1,732 bulls and 375 cows had been genotyped with the Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA). The remaining animals were genotyped with various chips of lower density that were imputed to high density. Univariate and trivariate genome-wide association studies (GWAS) with both medium- (60,671 markers) and high-density (312,614 markers) panels were performed for daughter pregnancy rate, heifer conception rate, and cow conception rate using GEMMA (version 0.94; <http://www.xzlab.org/software.html>). Analyses were conducted using bulls only, cows only, and a sample of both bulls and cows. The partial correlation and information theory algorithm was used to develop gene interaction networks. The most significant markers were further investigated to identify putatively associated genes. Little overlap in associated genes could be found between GWAS using different reference populations of bulls only, cows only, and combined bulls and cows. The partial correlation and information theory algorithm was able to identify several genes that were not identified by ordinary GWAS. The results obtained herein will aid in further dissecting the complex biology underlying fertility traits in dairy cattle, while also providing insight into the nuances of GWAS.

Key words: fertility, genome-wide association, high-density genotypes, network analysis

INTRODUCTION

Selection for increased production has been very successful in the dairy industry. Simultaneously, however, cow fertility has undergone a significant decline. As a result, many of today's dairy cows experience fertility problems, resulting in a national 21-d pregnancy rate average of only approximately 15% (Norman et al., 2009). Despite a recent upward trend in traits such as daughter pregnancy rate (**DPR**), conception rate, days to last breeding after calving, and calving interval (Norman et al., 2009), fertility traits remain an area of high concern for dairy producers. Fertility problems are one of the most frequent reasons for culling (Bascom and Young, 1998; Liang, 2013). They are also one of the most costly problems to manage, with each lost pregnancy costing an average of approximately \$500 in 2006 (De Vries, 2006).

With the increased availability of dense SNP marker panels, genomic selection methods have been widely investigated and implemented in livestock species. Genomic selection may prove to be especially beneficial for traits such as fertility that can be difficult or expensive to measure (Calus et al., 2013). The decreasing cost of marker panels has resulted in more bulls and cows being genotyped. Improved prediction accuracy is achieved when both bull and cow populations are included in evaluations (Calus et al., 2013). Marker panels are now also available at a high-density level of approximately 800,000 SNP markers, compared with the initial average of approximately 50,000 markers. This increased density may provide more power when identifying significant associations (e.g., Khatkar et al., 2008; Meredith et al., 2013).

Fertility-associated phenotypes are considered complex traits with low heritabilities. Heritability of DPR has been estimated to be approximately 0.04 and heritabilities of cow and heifer conception rate (**CCR** and **HCR**, respectively) are approximately 0.01 (VanRaden and Cole, 2014). The antagonistic genetic relationship between cow fertility and production has been previously documented (e.g., VanRaden et al., 2004; Pritchard et al., 2013). Direct selection for female fertility was

Received September 24, 2015.

Accepted April 15, 2016.

¹Corresponding author: klpgaddis@ufl.edu

initiated in the United States in 2003 with the introduction of genetic evaluations for DPR (VanRaden et al., 2003). Since then, DPR has been incorporated into all major selection indices utilized by US dairy farmers, with a relative weight of approximately 7% of the total economic value. The Animal Improvement Programs Laboratory (now Animal Genomics and Improvement Laboratory, Beltsville, MD) began evaluations for HCR and CCR in 2010.

Genetic correlations between traits may be indicative of QTL having pleiotropic effects. Linkage experiments utilizing multiple trait analysis have shown increased power to detect QTL (Knott and Haley, 2000; Korol et al., 2001). Bolormaa et al. (2010) found that the statistical power to detect associations was as good or better when using multiple-trait rather than single-trait models to perform genome-wide association studies (GWAS). Additional associations have been identified in multiple-trait analyses compared with single-trait analyses without increasing the false discovery rate (Bolormaa et al., 2010). Using a multiple-trait model incorporating traits related to fertility may allow for putative fertility QTL to be identified.

Complex traits such as fertility are likely influenced by a large number of genes, each with a small absolute effect. In typical GWAS procedures, stringent significance thresholds are needed to avoid false positives, but this may consequently prevent significant genes with small effects from being identified (McCarthy et al., 2008). Systems biology approaches have been proposed to more thoroughly explore the genetic architecture of complex traits. Correlation networks are being used for analysis of differential gene expression data (Hudson et al., 2009, 2012) as well as genotype data (Fortes et al., 2010, 2013). In particular, the partial correlation and information theory (PCIT) algorithm has been shown to have higher sensitivity for identifying effects of smaller magnitude by exploring gene-to-gene associations (Reverter and Chan, 2008). The objectives of the following research were to identify genes and biological networks putatively associated with fertility in dairy cattle using several approaches. We expect that by using different reference populations, different genomic regions associated with fertility may be identified. We also anticipate that using the PCIT algorithm will allow additional associations to be identified that may not have previously met genome-wide significance.

MATERIALS AND METHODS

Phenotypic and Genotypic Data

Three traits of reproductive performance were analyzed herein: DPR, HCR, and CCR. All traits

were defined as described by the Council on Dairy Cattle Breeding (<https://www.cdcb.us/reference.htm>). Daughter pregnancy rate represents the lactating cow's interval of calving to conception. It is defined as the percentage of nonpregnant cows that become pregnant during each 21-d period. Heifer conception rate is the maiden heifer's ability to conceive and is defined as the percentage of inseminated heifers that become pregnant at each service. Cow conception rate is the lactating cow's ability to conceive, defined as percentage of inseminated cows that become pregnant at each service. Three population subsets were examined: bulls only, cows only, and a combination of bulls and cows. Traits were corrected for management group, permanent environment, and herd-sire interaction (VanRaden and Wiggans, 1991). For all analyses, deregressed PTA from the National Dairy Database (Council on Dairy Cattle Breeding, Bowie, MD) were used as the dependent variable by weighting each PTA by the squared reliability. There were 24,041 bulls with records and 36,210 cows with records that also had genotypes available after imposing the restriction that PTA reliability was greater than parent average reliability. This was done to ensure that animals had information beyond only their parent average. Minimum reliability resulting from this restriction was 10 in the cow population and 36 in the bull population, both for HCR. Few animals had these low reliabilities, however, as can be seen from the mean reliabilities (standard deviations) for each trait by population included in Table 1. A random sample with equal representation of bulls and cows was taken to create a combined population subset with comparable size to the bull-only and cow-only data sets with 24,880 records.

Analyses were performed using 2 different marker densities. The set of markers used in computing US genomic predictions (Wiggans et al., 2013) was defined as moderate density (MD). The BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) was used as a high-density marker panel (HD). Marker editing was performed to remove SNP with call rates less than 90%, allele frequencies that departed from Hardy-Weinberg equilibrium >0.15 , and those with more than 2% parent-progeny conflicts. In total, 1,732 bulls and 375 cows were genotyped with the HD chip. The remaining animals had genotypes from various chip densities that were imputed to the HD level using Findhap version 3 (VanRaden et al., 2011). Imputation from lower to higher density has been shown to reach imputation accuracies greater than 99% (VanRaden et al., 2013). After editing, 60,671 and 312,614 markers remained in the MD and HD analyses, respectively (Wiggans et al., 2010; VanRaden et al., 2013). All markers included in the MD analyses were also included in the HD analyses.

Table 1. Summary statistics of daughter pregnancy rate (DPR), cow conception rate (CCR), and heifer conception rate (HCR) including number of records, mean deregressed PTA and standard deviation (SD)

Item	Bulls	Cows	Bulls and cows
No. of records	24,041	36,210	24,880
DPR mean (SD)	-0.422 (2.12)	0.348 (2.87)	-0.024 (2.53)
DPR mean reliability (SD)	80.89 (6.31)	68.96 (4.59)	74.88 (8.27)
CCR mean (SD)	-1.604 (4.16)	-0.418 (5.13)	-0.976 (4.65)
CCR mean reliability (SD)	73.82 (8.77)	65.40 (4.82)	69.55 (8.37)
HCR mean (SD)	-0.200 (3.65)	0.696 (4.29)	0.249 (4.04)
HCR mean reliability (SD)	65.09 (8.77)	58.55 (5.05)	61.76 (8.06)

Statistical Analysis

Univariate analyses were carried out for each trait in each population at both marker densities. A trivariate analysis also was carried out incorporating all 3 traits for each predictor population at both marker densities. The general univariate model fit was as follows:

$$\mathbf{Y} = \mu + \mathbf{x}\beta + \mathbf{u} + \mathbf{e},$$

where \mathbf{Y} is an n -vector of traits for n individuals, μ represents the intercept term, \mathbf{x} is an n -vector of marker genotypes, β is the effect size of the marker, \mathbf{u} is an n -vector of random effects, and \mathbf{e} is an n -vector of errors. The vector of random effects was assumed to follow an n -dimensional multivariate normal (MVN) distribution $\mathbf{u} \sim MVN_n(0, \lambda\tau^{-1}\mathbf{K})$, where τ^{-1} is the variance of residual errors, λ is the ratio between the 2 variance components, and \mathbf{K} is a known $n \times n$ genomic relationship matrix. The vector of errors was assumed to follow $\mathbf{e} \sim MVN_n(0, \tau^{-1}\mathbf{I}_n)$, where \mathbf{I}_n is an $n \times n$ identity matrix. No fixed effects (with the exception of an intercept term) were included in the models because the phenotypes were precorrected. The trivariate model was similar with the exception of \mathbf{Y} being an $n \times d$ matrix of d phenotypes from n individuals incorporating the 3 traits ($d = 3$). The trivariate model is as follows:

$$\mathbf{Y} = \mu + \mathbf{x}\beta^T + \mathbf{U} + \mathbf{E},$$

where \mathbf{Y} is an $n \times d$ matrix of d phenotypes for n individuals, μ represents an intercept term, \mathbf{x} is an n -vector of marker genotypes, β is a d vector of marker effect sizes for the d phenotypes, T indicates the transpose of this vector, \mathbf{U} is an $n \times d$ matrix of random effects, and \mathbf{E} is an $n \times d$ matrix of errors. In the following description of variance assumptions, $MN_{n \times d}(0, \mathbf{V}_1, \mathbf{V}_2)$ represents an $n \times d$ matrix normal (MN) distribution with mean 0, row covariance matrix \mathbf{V}_1 with dimensions $n \times n$, and column covariance matrix \mathbf{V}_2 with dimensions $d \times d$. The random effects matrix, \mathbf{U} , was assumed to follow $MN_{n \times d}(0, \mathbf{K}, \mathbf{V}_g)$, where \mathbf{K} is a known $n \times n$ relat-

edness matrix and \mathbf{V}_g denotes a $d \times d$ symmetric matrix of genetic variance components. The error matrix, \mathbf{E} , was assumed to follow $MN_{n \times d}(0, \mathbf{I}_{n \times n}, \mathbf{V}_e)$, where $\mathbf{I}_{n \times n}$ is an $n \times n$ identity matrix and \mathbf{V}_e represents a $d \times d$ symmetric matrix of residual variance components (Zhou, 2014).

All analyses were performed with GEMMA version 0.94 software (<http://www.xzlab.org/software.html>); GEMMA allows a multivariate linear mixed model to be fitted for testing marker associations with multiple phenotypes simultaneously (Zhou and Stephens, 2014). Here, GEMMA was used to estimate a relatedness matrix by providing a mean genotype file and a phenotype file, both formatted as for BIMBAM (Servin and Stephens, 2007). The centered relatedness matrix (\mathbf{G}_c) for all analyses was computed as

$$\mathbf{G}_c = \frac{1}{p} \sum_{i=1}^p (x_i - 1_n \bar{x}_i)(x_i - 1_n \bar{x}_i)^T,$$

where \mathbf{X} is a matrix of n individuals by p markers, x_i is the i th column of \mathbf{X} representing the genotypes for the i th SNP, and \bar{x}_i is the sample mean (Zhou et al., 2013; Zhou, 2014).

A GWAS involves performing numerous tests of significance. The GEMMA software calculates a Wald test statistic and P -value for each SNP based on a null hypothesis of no SNP effect versus the alternative hypothesis of the SNP effect being different from zero. A method of control is needed to account for false positives. False discovery rate (FDR) is a more powerful and flexible method than controlling for family-wise error rate (Storey, 2002). The FDR was calculated as described in Bolormaa et al. (2010):

$$FDR = \frac{p(1-s)}{s(1-p)},$$

where p represents the significance threshold or P -value and s represents the proportion of significant SNP (number of significant SNP divided by the total number of SNP tested).

Correlation Network Analyses

To avoid imposing very stringent significance thresholds due to multiple testing, the PCIT algorithm was used to examine putative significant pathways. This algorithm allows the construction of gene co-association networks by combining concepts of partial correlation coefficients and information theory, allowing identification of significant gene-to-gene associations (Reverter and Chan, 2008). Selection for the network favors genes containing SNP that have significant association with the investigated phenotypes (Reverter and Fortes, 2013). Relevant SNP were identified from the association analysis by first selecting the top 0.2 or 0.1% most significant SNP from MD or HD chips, respectively, from each of the 3 fertility traits. These SNP were ranked based on P -value. Second, significant ($P < 0.05$) SNP were then selected for inclusion based on their distance from a gene determined using BEDTools version 2.21.0 (Quinlan and Hall, 2010). Significant ($P < 0.05$) markers that were close (i.e., within 2,500 bp) to a gene, as well as markers that were very far from a gene (i.e., more than 1.5 Mb), were selected to be included in the PCIT analysis. This was performed following Fortes et al. (2010) based on expected linkage disequilibrium, size of promoter region, and likelihood of *cis*-acting windows. Marker selection was restricted to autosomal regions due to poorer annotation of the bovine sex chromosomes. Gene information was gathered from the bovine UMD3.1 genome assembly (Zimin et al., 2009) utilizing bovine Ensembl gene IDs. After merging with the annotated gene data, approximately 30,501 markers remained in the MD analysis and 156,000 markers remained in the HD analysis to be used in the correlation network analysis.

From these selected SNP, an association weight matrix (AWM) was constructed. In the AWM, each column corresponds to a trait and each row corresponds to a SNP. Each cell in the AWM corresponds to the z -score normalized effect size for that particular SNP and trait. In constructing the AWM, a “1 SNP to 1 gene” rule was implemented such that if multiple SNP mapped to the same gene, only the most significant SNP was retained for that gene. Row-wise partial correlations were computed on the AWM using the PCIT algorithm as described by Reverter and Chan (2008). The PCIT package (Watson-Haigh et al., 2010) was utilized in R version 3.2.1 (R Core Team, 2014) to implement the algorithm. Significant correlations were identified by the PCIT algorithm and represented as edges between genes (nodes) in the network. Correlation networks were visualized using Cytoscape version 3.2.1 (Shannon et al., 2003). Gene ontology (GO) enrichment was conducted with the PCIT results using DAVID (version

Table 2. Heritability (on diagonal in bold), genetic correlation (above diagonal), and phenotypic correlation (below diagonal) of daughter pregnancy rate (DPR), cow conception rate (CCR), and heifer conception rate (HCR) as provided by VanRaden and Cole (2014)

	DPR	CCR	HCR
DPR	0.04	0.87	0.41
CCR	0.70	0.02	0.54
HCR	0.10	0.45	0.01

6.7; Huang et al., 2009a,b). Genes selected by the PCIT algorithm were compared with a background list of *Bos taurus* genes.

RESULTS AND DISCUSSION

Summary statistics for each of the bull, cow, and combined bull and cow data sets are provided in Table 1. Table 2 provides heritability estimates, as well as genetic and phenotypic correlations, for DPR, CCR, and HCR, as given in VanRaden and Cole (2014). These were calculated from PTA correlations of Holstein bulls with high reliabilities, also from the US Holstein population. Supplemental Tables S1 to S24 list markers evaluated in the study (<http://dx.doi.org/10.3168/jds.2015-10444>).

Genome-Wide Association Analyses

Bulls. The first analyses performed were classical GWAS to identify markers putatively associated with fertility. Several thresholds were investigated before determining a threshold of $-\log(P\text{-value}) > 3.0$ for MD data and $-\log(P\text{-value}) > 4.0$ for HD data that resulted in a reasonable number of markers to investigate further. There were 64 (0.94), 68 (0.88), and 81 (0.74) markers (FDR) exceeding the threshold for the CCR, DPR, and HCR analyses, respectively, in analyses using MD marker data. For the HD marker data, 40 (0.76), 58 (0.53), and 67 (0.46) markers exceeded this level in the CCR, DPR, and HCR analyses, respectively. A trivariate analysis including all 3 traits also was performed. There were 94 (0.64) markers (FDR) in the MD trivariate analysis and 93 (0.33) markers (FDR) in the HD trivariate analysis that exceeded the threshold after merging with the gene annotations. Results with HD data are depicted in Manhattan plots in Figure 1 for CCR, DPR, and HCR, and in Figure 2 for the trivariate analysis.

Similar significant regions were identified in previous studies related to fertility. Peters et al. (2013) identified a region on BTA8 (between 0.35 and 0.97 Mb) related to heifer pregnancy rate. Significant associations were identified herein for CCR using MD markers on BTA8 at approximately 0.6 and 0.8 Mb.

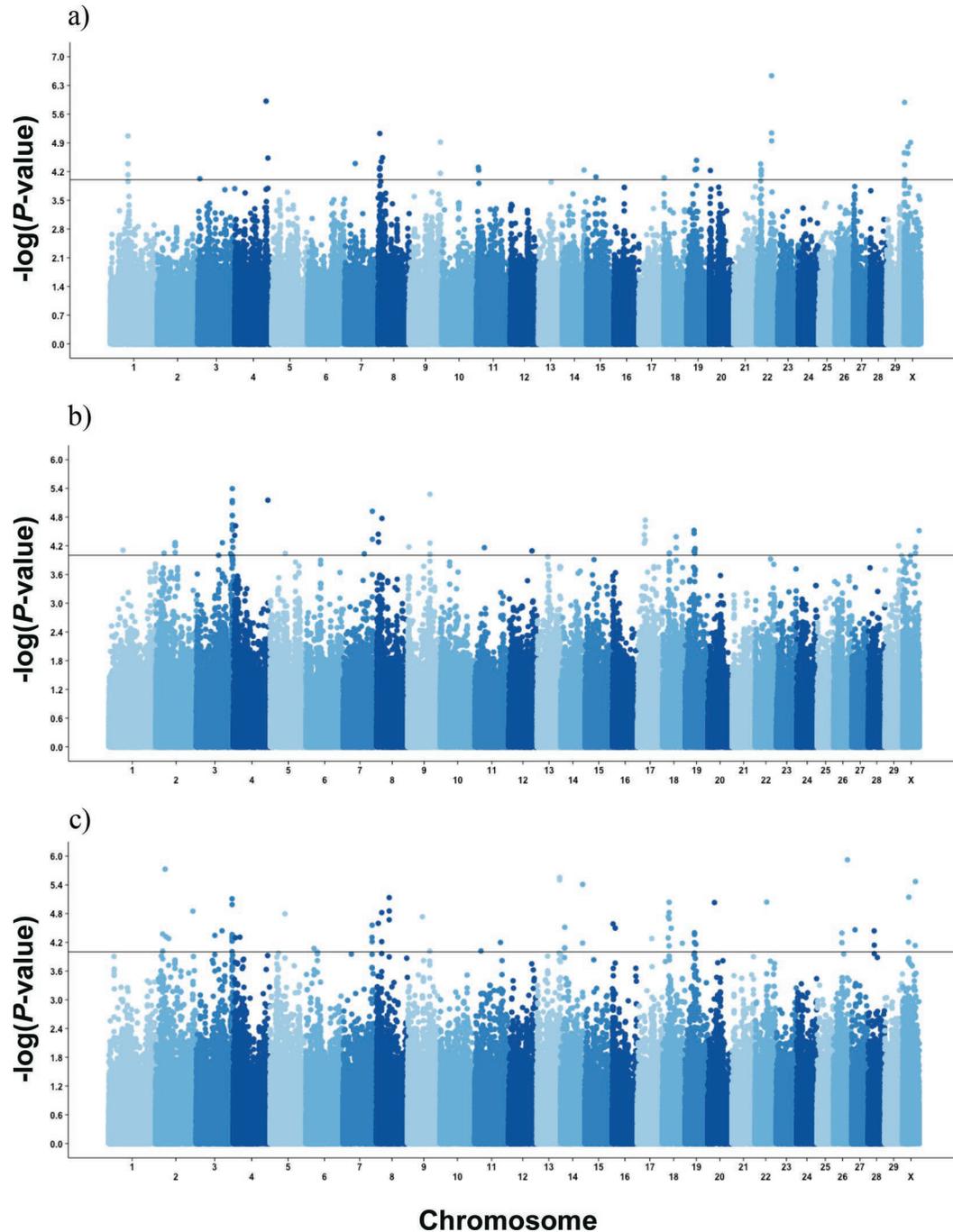


Figure 1. Manhattan plots of genome-wide association analysis results of cow conception rate (a), daughter pregnancy rate (b), and heifer conception rate (c) using high-density bull genotypes. Color version available online.

Regions on BTA10 (at approximately 25 and 48 Mb) have previously been associated with fertilization rate (Huang et al., 2010). Similar regions were identified in CCR, DPR, and trivariate analyses using MD markers (Supplemental Tables S1, S2, and S4; <http://dx.doi.org/10.3168/jds.2015-10444>). Kühn et al. (2003) also found a region on BTA10 between 34 and 57 Mb as-

sociated with paternal effect of nonreturn rate at 90 d. A region on BTA10 associated with CCR between 49 and 50 Mb was identified herein using MD markers. A second region on BTA10 was identified between 84 and 92 Mb associated with days from first to last insemination in cows (Höglund et al., 2009), which corresponded to regions found in analyses with MD mark-

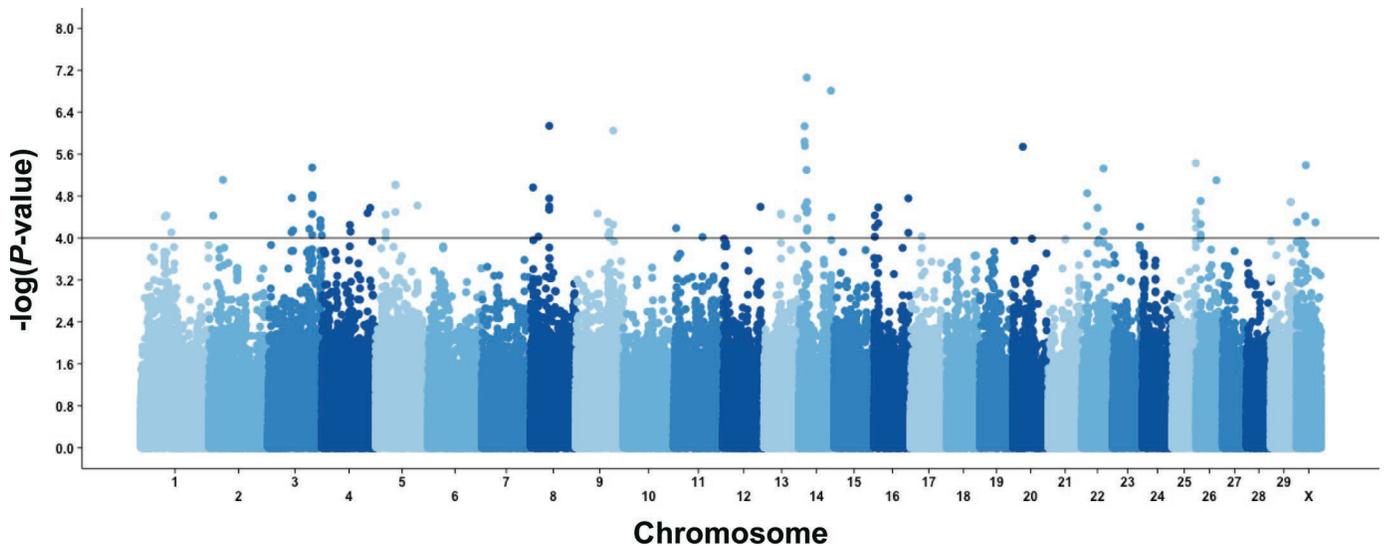


Figure 2. Manhattan plot of genome-wide association analysis results of cow conception rate, daughter pregnancy rate, and heifer conception rate in a trivariate analysis using high-density bull genotypes. Color version available online.

ers for HCR (at 86 Mb) and the trivariate analysis (at 84 Mb). Regions identified in the DPR, HCR, and trivariate analyses (with both MD and HD markers) on BTA3 corresponded to regions identified by Hawken and Zhang (2012) on BTA3 at 112.3 Mb associated with occurrence of first postpartum ovulation before weaning in the first rebreeding period. This region has also been shown to be associated with first-service conception (Peters et al., 2013). Cole et al. (2011) also identified a region of BTA3 at approximately 90 Mb associated with DPR. The study by Cole et al. (2011) analyzed a smaller number of cows, 1,555 of which overlapped with this study. A similar region was identified in the CCR and DPR analyses at MD, as well as the trivariate analyses at both density levels. Sahana et al. (2010) identified a region at approximately 87 Mb on BTA3 associated with the interval from calving to first insemination. This region was also identified in the analysis of HCR with HD markers. A region on BTA13 at approximately 68 Mb was also found to be associated with a fertility index (Sahana et al., 2010). A similar region at approximately 67 Mb was found in both the DPR and HCR analyses using MD markers. Hawken and Zhang (2012) identified a significant region on BTA15 between 31 and 38 Mb associated with age at first observed corpus luteum. An association in this region was also identified herein for DPR and HCR with MD markers, as well as for CCR with HD markers. Last, on BTA18, Ashwell et al. (2004) identified a region between 8 and 16 Mb associated with pregnancy rate. This corresponds to a region at approximately 14 Mb associated with HCR using HD markers.

Several markers were found at both levels of chip density, as expected. The number of significant markers found in common for each trait between chip densities was 6, 8, 8, and 7 markers for CCR, DPR, HCR, and the trivariate analysis, respectively. Among genes found in common between the chip densities, several have been previously documented to affect reproductive traits. For DPR, both chips contained a peak near *TRPV1* on chromosome 19. The protein that this gene encodes has been found predominantly in sperm tails, the apical region of the acrosome, and the postacrosomal region in the sperm head. It has also been identified in bull spermatozoa (Gervasi et al., 2011). Several genes in the HCR analysis were associated with peaks at both chip densities including *TRAF3IP1*, *SMARCA2*, and *CHFR*. Mutations in *TRAF3IP1* have been shown previously to result in defects in ciliogenesis and embryonic development (Berbari et al., 2011). The *SMARCA2* gene has been shown to have a role in early embryogenesis (Peddinti et al., 2010), as well as in chromatin remodeling in bovine and human oocytes (Adjaye et al., 2007; Peddinti et al., 2010). Finally, *CHFR* has been shown to be involved in the *MSX1* pathway, which delays pre-implantation development in bovine embryos in vitro (Tesfaye et al., 2010).

An additional way to identify genes that may have a role in fertility is to find those that overlap between univariate and trivariate analyses for a given population and chip density. In the bull population, 3 genes (*SLC35A5*, *GALNTL6*, *APPL1*) were identified in both the univariate analysis of CCR and the trivariate analysis at HD density. The *GALNTL6* gene has been

implicated as a candidate gene associated with cull cow carcass weights in Holstein Friesian cattle (Doran et al., 2014), but it has also been identified in humans as a potential gene involved in miscarriages (Koopers et al., 2014). Expression of *APPL1* has been previously found in the bovine ovary, large and small follicles, corpus luteum, cumulus cells, granulosa cells, follicular fluid, and the oocyte (Maillard et al., 2010). Analyses of DPR in the bull population with HD genotypes also resulted in a peak associated with *GALNTL6*. Additional genes that overlapped between DPR and trivariate analyses included *ESPNL*, *FAM132B*, *ILKAP*, *LINGO2*, and an uncharacterized protein on BTA9. Of particular interest to fertility, *ILKAP* is a C1-angiogenesis protein. Angiogenesis is typically limited to specific normal adult cells including the placenta, ovary, and endometrium (Redmer et al., 1988). Last, comparing overlap between the trivariate analysis with HCR, we found several genes that were in common with those found in the DPR analysis, including *ESPNL*, *ILKAP*, *GALNTL6*, and *LINGO2*. Another gene that overlapped with HCR with particular reproductive interest was *SMARCA2*, which was described above. Additional overlapping genes with HCR included *LRRN2*, *ERCC8*, and *LOC506670*.

Cows

Association analyses were performed for the cow population. Thresholds for MD data were increased to a $-\log(P\text{-value})$ of 4.0 in the cow population to identify a reasonable number of markers to investigate further. The threshold for the HD data set was also increased to $-\log(P\text{-value})$ of 5.0 for the cow population. There were 75 (0.08), 33 (0.18), 37 (0.16), and 83 (0.07) markers (FDR) from the MD marker panel exceeding the threshold for the CCR, DPR, HCR, and trivariate analyses, respectively. In the HD analysis, there were 37 (0.08), 48 (0.06), 68 (0.05), and 116 (0.03) markers (FDR) for each trait (DPR, CCR, HCR, and trivariate analysis, respectively). Results from HD GWAS in the cow populations are shown in Manhattan plots in Figures 3 and 4. Tables listing the markers exceeding the threshold and their closest gene are included as supplementary materials (Supplemental Tables S5–S8 and S17–S20; <http://dx.doi.org/10.3168/jds.2015-10444>).

As in the bull analyses, significant associations were found within the cow analyses that corresponded to previous reports in literature. Berry et al. (2012) identified an association with postpartum interval to commencement of luteal activity on BTA2 at approximately 134 Mb. The analysis herein with MD markers for HCR also identified this region of BTA2. A region on BTA3 was identified in the CCR analysis using MD markers,

corresponding to a region identified by Hawken and Zhang (2012) associated with occurrence of first postpartum ovulation before weaning in the first rebreeding period. Peters et al. (2013) identified an association on BTA3 at approximately 96 Mb with first service conception. This is near an association with HCR identified herein with MD markers (Supplemental Table S7; <http://dx.doi.org/10.3168/jds.2015-10444>). Trivariate analyses and DPR analyses at both marker densities identified a region at approximately 88 to 90 Mb on BTA6 corresponding to a region at approximately 89 Mb previously found to be associated with interval from calving to first insemination (Sahana et al., 2010). In CCR analyses with MD markers, regions were identified on BTA8 and BTA13 at approximately 25 and 80 Mb, respectively. This region on BTA8 is close to regions associated with first service conception (Peters et al., 2013) and fertilization rate (Huang et al., 2010). The region on BTA13 is close to regions associated with fertilization rate (Huang et al., 2010) as well as heifer pregnancy (Peters et al., 2013). Previous research by Cochran et al. (2013) identified an association between DPR and HCR with the *CACNA1D* gene. In our analysis of DPR using HD markers in the cow population, we identified the same gene but with a different subunit (1E). On BTA7, a region at approximately 15.4 Mb has been previously associated with productive life, somatic cell score, and DPR (Cole et al., 2011). Analyses with MD markers of CCR and HCR, as well as the trivariate analysis, identified a similar location on BTA7. Ashwell et al. (2004) identified a region at approximately 70 Mb on BTA16 associated with pregnancy rate. This region was also identified in the trivariate analysis, as well as CCR and DPR, using HD markers. Kühn et al. (2003) identified putative QTL on BTA18 in the region of 62 to 65 Mb associated with both the maternal and paternal effects of nonreturn rate at 90 d. A large block on BTA18 ranging from approximately 57 to 62 Mb was found to be associated in analyses using HD markers of DPR and HCR, as well as the trivariate analysis. Additional research has identified associations with fertility traits on BTA18 (Sahana et al., 2010). Ashwell et al. (2004) also identified a region associated with fertility traits on BTA18; however, it was in the region spanning 38 to 43 Mb. Analyses of DPR, HCR, and trivariate herein identified regions slightly beyond this at 44 to 46 Mb.

The number of significant markers in common between the MD and HD analyses ranged from 1 for CCR, 9 for HCR, and 7 in the trivariate analyses. Analyses at both chip densities for CCR identified a peak near the *AFF1* gene located on chromosome 6. In mice, *Aff1* is expressed in the kidney, brain, lung, liver, spleen,

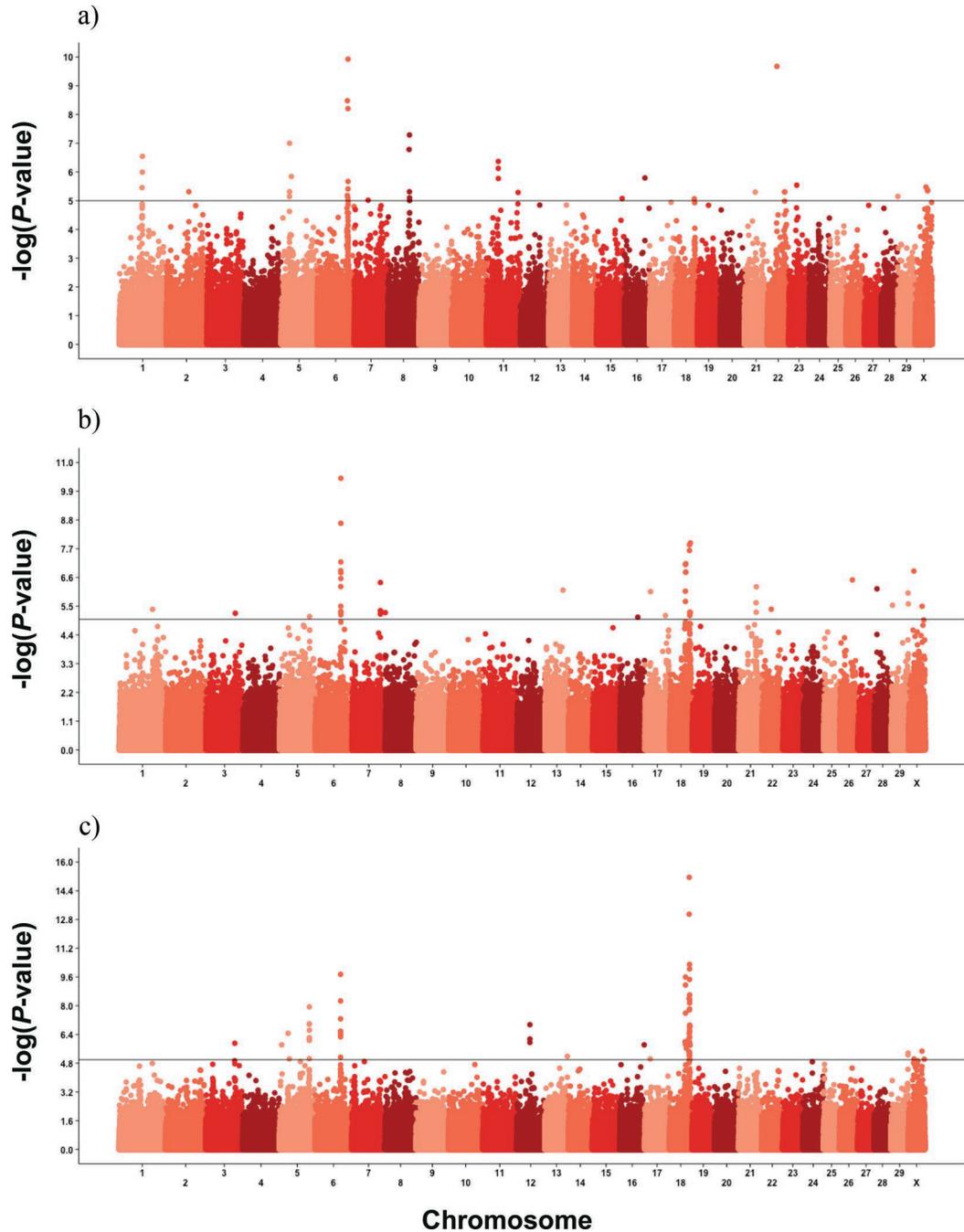


Figure 3. Manhattan plots of genome-wide association analysis results of cow conception rate (a), daughter pregnancy rate (b), and heifer conception rate (c) using high-density cow genotypes. Color version available online.

skeletal muscle, and testis. Specific gonadal studies identified *Aff1* expression in the ovary, epididymis, and testis of mice (Alves et al., 2011). In cattle, Fortes et al. (2011) identified *AFF1* as a significant transcription factor in regulatory gene networks underlying puberty in beef cattle. The DPR analyses identified a peak near the *TP53BP1* gene on chromosome 21. This

gene, which encodes a key p53 binding protein, was also included in the association matrix for puberty of beef cattle (Fortes et al., 2010). The p53 family of proteins have roles not only in cancer and development but also in maternal reproduction (Levine et al., 2011; Hu et al., 2007). A gene of interest identified in the HCR analyses was *SIGLEC12* (*LOC618463*), which has previously

been identified as an interesting target gene for calving traits (Cole et al., 2009).

More genes overlapped when comparing HD results in the cow population between the trivariate analysis and each univariate analysis. For CCR, overlapping genes with implications in fertility included *AFF1*, *CCL19*, *KIF27*, *LAMC3*, *SMYD2*, and *LRRC4B*. Both *AFF1* and *LRRC4B* were also identified in the cow population using MD in CCR and HCR, respectively. Guerin et al. (2011) has suggested that *CCL19* plays a role in controlling lymphocyte populations that are associated with embryo implantation in humans. They found that seminal fluid increases expression of *CCL19* in uterine epithelial cells, which in turn increases circulating regulatory T-cells (Guerin et al., 2011).

Bulls and Cows

When the bull and cow populations were combined, a threshold of $-\log(P\text{-value})$ equal to 3.0 was used for MD GWAS and $-\log(P\text{-value})$ equal to 4.0 for HD GWAS. At this level, 61 (0.98), 57 (0.99), 61 (0.98), and 41 (0.99) markers (FDR) were identified for CCR, DPR, HCR, and trivariate analysis, respectively. In the HD GWAS, 154 (0.20), 130 (0.24), 173 (0.18), and 27 (0.99) markers (FDR) for CCR, DPR, HCR, and trivariate analyses, respectively. Results from HD GWAS in the combined bull and cow populations are shown in Manhattan plots in Figures 5 and 6. Markers exceeding the threshold and their closest gene are included in Supplemental Tables S9 to S12 and S21 to S24 ([\[dx.doi.org/10.3168/jds.2015-10444\]\(http://dx.doi.org/10.3168/jds.2015-10444\)\) for both MD and HD analyses.](http://</p>
</div>
<div data-bbox=)

Overlap was also identified when comparing the results from using a combined bull and cow data set to results previously reported in the literature. Trivariate analysis using HD markers identified a region at approximately 129 to 130 Mb on BTA2, which corresponded to an association with postpartum interval to commencement of luteal activity in the same region identified by Berry et al. (2012). The aforementioned study by Cole et al. (2011) identified a region on BTA7 at 15.4 Mb associated with DPR, productive life, and SCS. This corresponded to regions identified in the trivariate analysis with MD markers and HCR analysis with HD markers in the combined bull and cow data set. Additionally, a region associated with DPR on BTA3 aligned with a region identified in the HCR analysis with MD markers. A region identified using MD markers on BTA8 at approximately 71 Mb associated with DPR corresponded to a region identified by Berry et al. (2012) associated with days from calving to first observed heat. Using HD markers, trivariate and CCR analyses identified a region on BTA10 at approximately 8 Mb. A similar region spanning 8 to 15 Mb was identified previously as being associated with fertility treatments in third parity. Fertility treatments included hormonal reproductive disorders, ovarian cyst treatments, and infective reproductive disorders (Höglund et al., 2009). Corresponding regions were also identified on BTA22 with fertility treatment in first-parity animals in the region of 57 Mb (Höglund et al.,

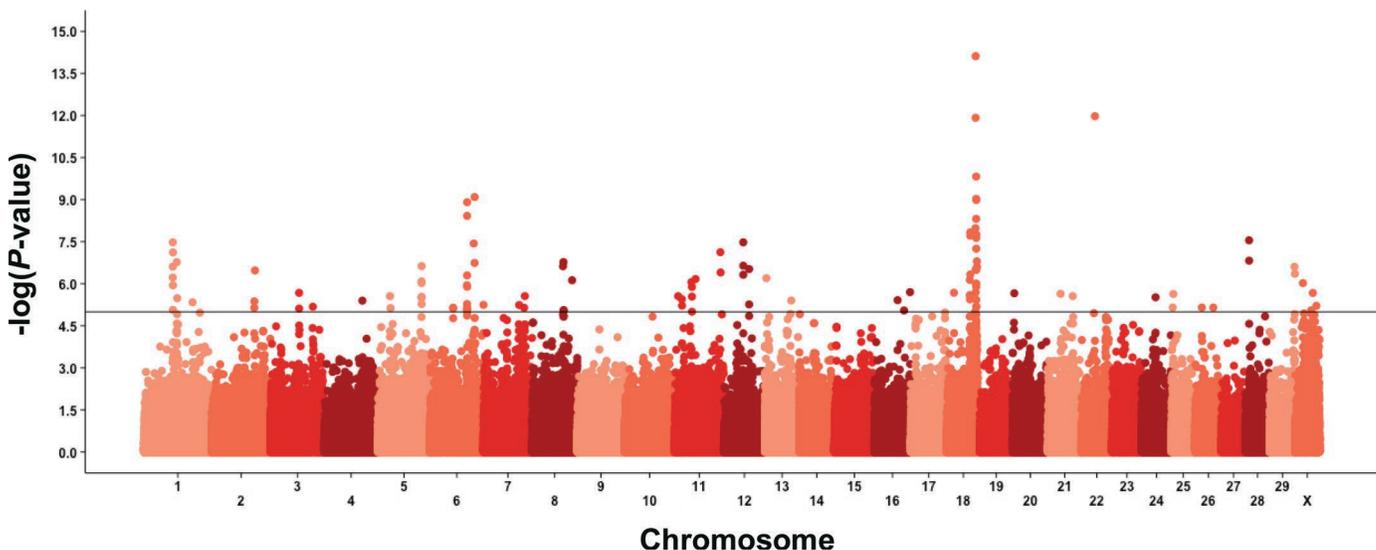


Figure 4. Manhattan plot of genome-wide association analysis results of cow conception rate, daughter pregnancy rate, and heifer conception rate in a trivariate analysis using high-density cow genotypes. Color version available online.

2009). All analyses herein with HD markers identified associations in a similar region of BTA22 spanning 54 to 57 Mb. Huang et al. (2010) reported associations with fertilization rate on BTA10 (at approximately 25 Mb) and BTA29 (at approximately 45 Mb). Similar regions were identified herein using MD markers for DPR (see Supplemental Table S10; <http://dx.doi.org/10.3168/jds.2015-10444>).

Using HD markers in the trivariate analysis, an association identified on BTA2 corresponded to a region identified by Huang et al. (2010) associated with fertilization rate. Sahana et al. (2010) cited significant fertility-related associations on BTA10 at 40.7, 52.7, and 93 Mb. Trivariate and DPR analyses using MD markers identified corresponding

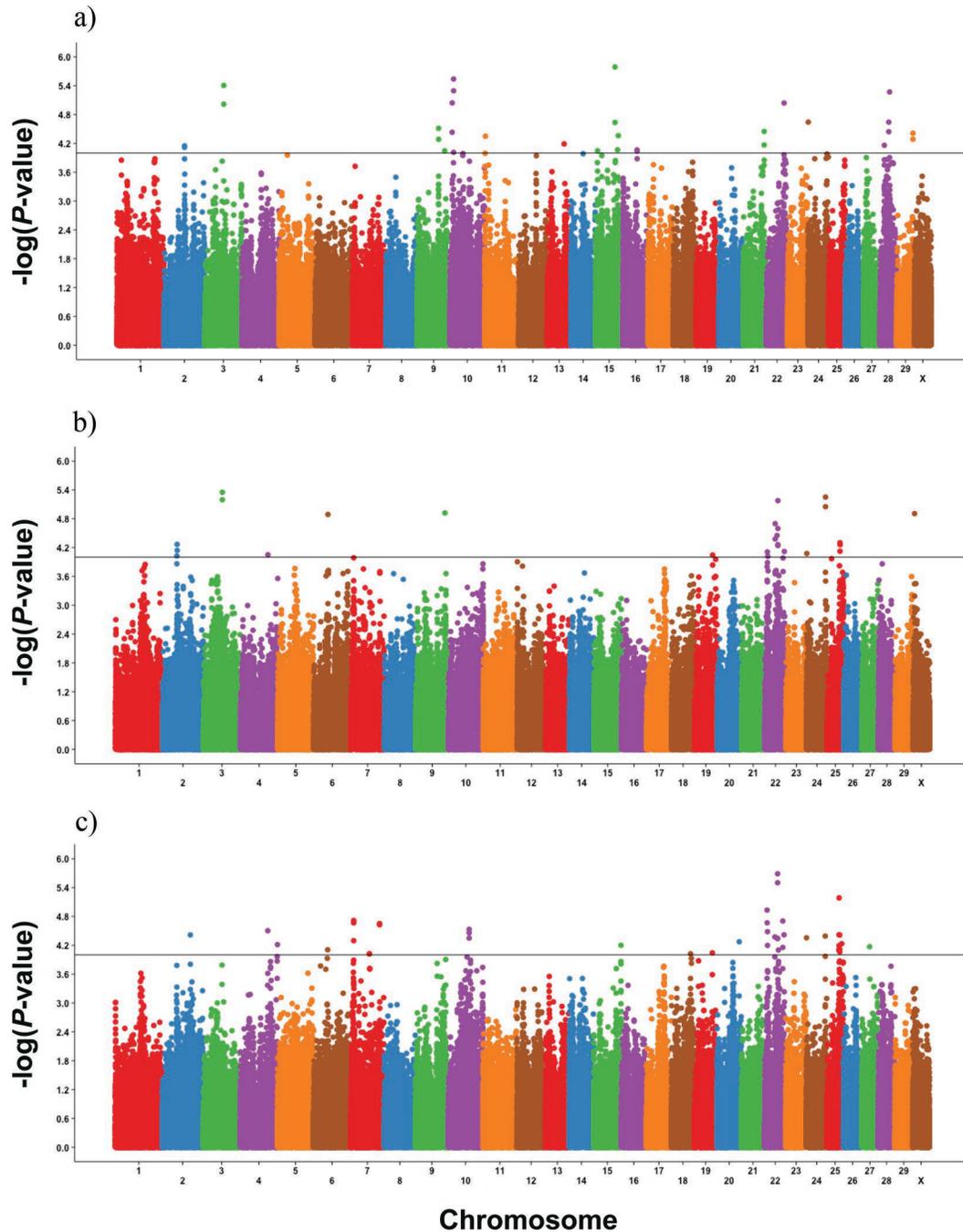


Figure 5. Manhattan plots of genome-wide association analysis results of cow conception rate (a), daughter pregnancy rate (b), and heifer conception rate (c) using high-density bull and cow genotypes. Color version available online.

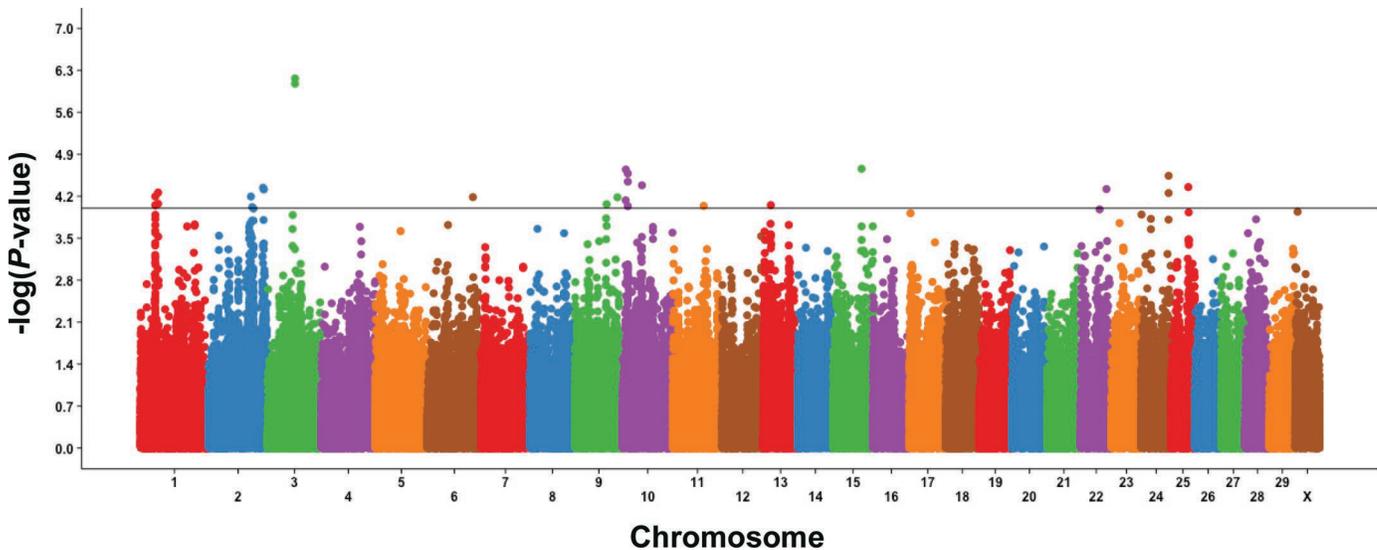


Figure 6. Manhattan plot of genome-wide association analysis results of cow conception rate, daughter pregnancy rate, and heifer conception rate in a trivariate analysis using high-density bull and cow genotypes. Color version available online.

associations in the region of 52 Mb on BTA10. Kühn et al. (2003) also identified a similar region on BTA10 spanning 34 to 57 Mb associated with paternal effect of nonreturn rate at 90 d. In addition to associations with the trivariate and DPR analyses mentioned above, this region also includes an association with HCR using MD markers at approximately 37 to 38 Mb. A region of BTA18 spanning 62 to 65 Mb has been previously associated with both the maternal and paternal effects of nonreturn rate at 90 d (Kühn et al., 2003). Corresponding regions on BTA18 were found to be associated with DPR (at 62 Mb) and HCR (at 64 Mb) using MD markers. The *APBB1* gene has previously been identified as being associated with DPR and HCR (Cochran et al., 2013). Using MD markers for the analysis of DPR, we identified an association with *APBB2*. Similar to what was found in the cow analysis, we identified *CACNA1A* as associated with HCR, whereas Cochran et al. (2013) identified an association of *CACNA1D* with DPR and HCR. Last, *MRGPRF* was associated with CCR using HD markers and was also previously associated with DPR (Cochran et al., 2013).

There were no significant markers in common between the 2 marker densities in the combined population; however, several genes were identified in both the trivariate analysis and univariate analysis for CCR and DPR with HD data. No genes overlapped between HCR and the trivariate analysis. Genes associated with large peaks in both the trivariate analysis and CCR analysis that have previous implications in reproduction included *PDE8B* and *EIF3M*. The *PDE8B* gene is part of the phosphodiesterase family and is expressed in cumulus

cells, granulosa cells from small follicles, and cumulus-oocyte complexes. It has also been identified in rat brain, mouse ovary, and bovine testis (Sasseville et al., 2009). Zeng et al. (2013) showed that embryos lacking functional *EIF3M* died at the peri-implantation stage, implying that murine *EIF3M* is essential for embryonic development. Comparing the results from DPR with the trivariate analysis, the only gene that overlapped was *MAS1*. The *MAS1* gene is a proto-oncogene, G protein-coupled receptor that has been implicated to play a role in regulation of the ovulatory process in cattle (Tonello dos Santos et al., 2012).

Correlation Network Analyses

Results from merging significant markers with their closest gene are included as supplementary materials for each reference population, density, and trait (Supplemental Tables S1 to S24; <http://dx.doi.org/10.3168/jds.2015-10444>).

Bulls

The largest network identified with the PCIT algorithm using HD bull data incorporated 24 genes or markers and is shown in Figure 7. Among the genes in this network, only *GALNTL6* was also identified in the prior GWAS analyses. Genes included in the network that have had prior citations involving reproductive traits include *CUX1*, *EPSTI1*, *ESRRA*, *KIF5B*, *PDE5A*, and *WNT7A*. Forde and Lonergan (2012) found *EPSTI1* to be differentially expressed as part

of the early endometrial response to a conceptus. The *WNT7A* gene is involved in cell proliferation (Robinson et al., 2008), with reports of significant effects on reproductive traits. Complete deletion of *WNT7A* results in abnormal Müllerian duct patterning, specification, and cell fate in a developing fetus. Mice that carry mutated *WNT7A* lacked endometrial glands, but all other uterine cell types appeared normal. It also appears that mutations in *WNT7A* interfere with blastocyst implantation (Dunlap et al., 2011).

Using the genes selected by the PCIT algorithm, GO enrichment identified processes involved in steroid metabolism (GO:0019218, GO:0030301, GO:0015918, GO:0008202, GO:0045940, GO:0050810). Changes in reproduction of dairy cattle have previously been associated with elevated steroid metabolism (Wiltbank et al., 2006). Additional related processes that were identified included regulation of hormone levels (GO:0010817) and in utero embryo development (GO:0001701).

Cows

The largest network identified by the PCIT algorithm in the cow population using HD data involved 13 genes or markers, as shown in Figure 8. None of these genes were previously identified by the GWAS analysis discussed above; however, several have been implicated in reproductive processes. Two genes with the greatest amount of literature supporting their role in reproduction were *DAZL* and *BAX*. The *DAZL* gene has been implicated to have an important role in gametogenesis, with deletions or mutations resulting in sterility in vertebrates (Zhang et al., 2008). It is believed to be important for the transcriptional regulation of mRNA expression (Zhang et al., 2008). The *BAX* gene is interesting, as it seems to have different functions in males and females. In females, *BAX* has been identified as a pro-apoptotic gene (Lazzari et al., 2011) that, when deleted, results in increased oocyte and follicle numbers in mice (Greenfeld et al., 2007). In males, however, it has been shown that the absence of *BAX* results in infertility. It is hypothesized that the balance between apoptotic and anti-apoptotic factors plays an important role in normal spermatogenesis (Matzuk and Lamb, 2002).

From genes identified with the PCIT algorithm in the cow data set, GO enrichment identified several processes related to cell death and apoptosis (GO:0042981, GO:0043067, GO:0010941, GO:0043068, GO:0010942). Associations with cell death and apoptosis as related to fertility traits may be associated with embryo survival. Additionally, several processes involving ion binding were identified (GO:0043167, GO:0043169, GO:0046872).

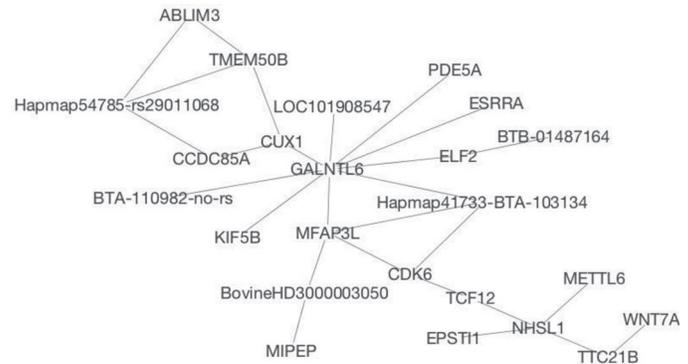


Figure 7. Largest network identified using partial correlation and information theory algorithm with high-density marker fertility results in the bull population. *ABLIM3* = actin binding LIM protein family, member 3; *TMEM50B* = transmembrane protein 50B; *CCDC85A* = coiled-coil domain containing 85A; *CUX1* = cut-like homeobox 1; *LOC101908547* = olfactory receptor 4D2-like; *PDE5A* = phosphodiesterase 5A; *ESRRA* = estrogen related receptor alpha; *ELF2* = E74-like factor 2; *GALNTL6* = polypeptide N-acetylgalactosaminyltransferase-like 6; *KIF5B* = kinensin family member 5B; *MFAP3L* = micro-fibrillar associated protein 3 like; *CDK6* = cyclin-dependent kinase 6; *TCF12* = transcription factor 12; *METTL6* = methyltransferase like 6; *NHSL1* = NHS like 1; *WNT7A* = wingless-type MMTV integration site family member 7A; *TTC21B* = tetratricopeptide repeat domain 21B; *EPSTI1* = epithelial stromal interaction 1; *MIPEP* = mitochondrial intermediate peptidase.

Bulls and Cows

The combined data set of bulls and cows produced a network including 13 genes and markers using HD data as shown in Figure 9. Several of the identified genes were not identified in the GWAS analysis; however, they did have previous citations in the literature related to reproductive function. For example, the *NOS2* gene is a nitric oxide synthase and has been identified in

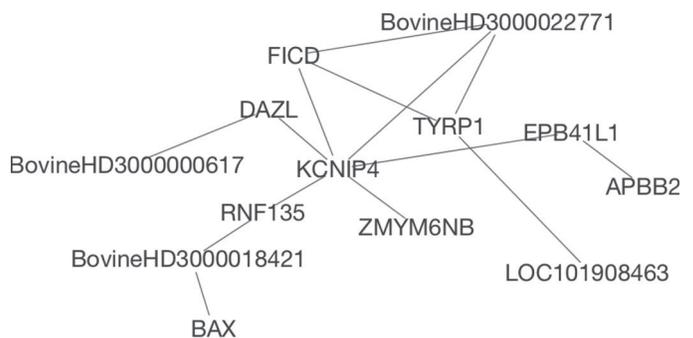


Figure 8. Largest network identified using partial correlation and information theory algorithm with high-density marker fertility results in the cow population. *FICD* = FIC domain containing; *DAZL* = deleted in azoospermia-like; *TYRP1* = tyrosinase-related protein 1; *EPB41L1* = erythrocyte membrane protein band 4.1-like 1; *APBB2* = amyloid beta precursor protein binding family B member 2; *KCNIP4* = Kv channel interacting protein 4; *ZMYM6NB* = ZMYM6 neighbor; *RNF135* = ring finger protein 135; *BAX* = BCL2-associated X protein.

human, bovine, and rat oviducts. It was hypothesized that regulation of *NOS* at estrus in the isthmus results in increased oviduct motility by acting on smooth muscle activity (Ulbrich et al., 2006). It has also been postulated that nitric oxide may contribute to survival of granulosa cells (Zamberlam et al., 2011).

Gene ontology analysis with the genes identified by the PCIT algorithm in the combined bull and cow population did not result in many processes that could clearly be associated with fertility. Several processes involved with wound healing were identified; however, their relationship with fertility is not clear (e.g., GO:0009611, GO:0042060). Two processes identified that were more clearly associated with fertility were those related to sexual reproduction (GO:0019953) and response to hormone stimulus (GO:0043434).

General Discussion

In addition to identifying genes putatively associated with fertility, the analyses performed herein provide insight into several nuances of GWAS. First, we were able to discern differences when using a medium-density versus high-density genotyping platform. A set of genes had large associations with the traits of interest, regardless of the chip density used. This implies that some association signals are identified irrespective of marker density. This also lends support that the analysis is identifying regions of the genome that are truly associated with the traits. Conversely, we observed additional genomic regions with the HD panel compared with the MD panel. Some analyses with MD markers had very high FDR (e.g., bull population, combined bull and cow population). Using HD markers resulted in lower FDR in all analyses. High FDR in analyses with MD markers may be partly the result of association strength with these complex traits. It has previously been seen that increasing the density of marker panels leads to increased power and resolution to detect significant loci (Khatkar et al., 2008; Spencer et al., 2009).

Most GWAS results are reported on a univariate basis, but multivariate analysis can potentially identify additional significant associations. Multivariate GWAS studies can achieve higher statistical power compared with univariate GWAS (Korol et al., 2001; Bolormaa et al., 2010). Multivariate GWAS may also aid in identifying genes that are involved in overarching biological processes affecting fertility. Overlap among identified genes observed between univariate and trivariate analyses herein may indicate genes that are involved in fertility processes in general. Differences between traits are also to be expected, however, as different aspects of fertility are represented (e.g., heifer conception versus cow conception). A better understanding of the biological

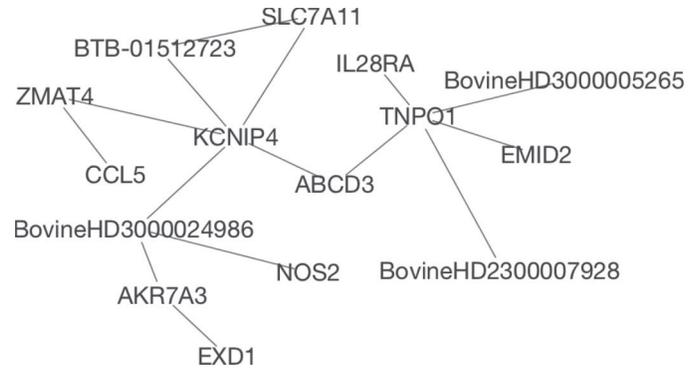


Figure 9. Largest network identified using partial correlation and information theory algorithms with high-density marker fertility results in the combined bull and cow population. *SLC7A11* = solute carrier family 7 (anionic amino acid transporter light chain, xc-system); *IL28RA* = interferon, lambda receptor 1; *TNPO1* = transportin 1; *EMID2* = collagen type XXVI alpha 1; *ABCD3* = ATP binding cassette subfamily D member 3; *KCNIP4* = Kv channel interacting protein 4; *CCL5* = chemokine (C-C motif) ligand 5; *ZMAT4* = zinc finger, matrin-type 4; *AKR7A3* = aflatoxin B1 aldehyde reductase member 3; *EXD1* = exonuclease 3'-5' domain containing 1.

mechanisms controlling fertility will allow more precise and rapid improvement of cattle fertility to be made.

In contrast, little overlap of genes could be identified when comparing results across reference populations. Genes identified in analyses using data only from bulls were very rarely also identified as significant in analyses using only cow data. This may be the result of different power levels due to the sample size difference between bull and cow populations. It may also be the result of differences in reliability between the bull and cow populations. Despite having the same reliability constraints in place, the cow population had a larger spread of reliabilities, which may have introduced additional noise.

Finally, network analyses using the PCIT algorithm bring additional GWAS implications to the forefront. The PCIT algorithm allows the genetic dissection of complex traits and the development of networks derived based solely on the data itself. It also allows us to be more lenient in deeming markers as significant. This should prove especially beneficial for low-heritability traits such as fertility, where the prevailing assumption is that many genes, each with a small effect, affect the trait. In support of this, the PCIT network analyses performed herein identified several genes that were not identified as significant in the ordinary GWAS analysis. Many of these genes have been implicated in reproductive functions. The gene networks presented include fewer genes than those previously presented in literature (e.g., Fortes et al., 2010). It should be noted that the number of phenotypes used herein for PCIT analysis is fewer than used in previous research, possi-

bly resulting in smaller networks. The smaller resulting networks may also be the result of using too stringent a selection threshold for these traits. In developing a PCIT network, the significance threshold here is only one method to determine if genes should be included for further analysis. The number of genes included in a network can also depend on chip density and quality of the annotation of SNP to the genome (Reverter and Fortes, 2013).

CONCLUSIONS

The most significant markers from the present GWAS were investigated to identify genes putatively associated with fertility traits (specifically, CCR, DPR, and HCR). We also explored differences in results based on chip density, reference population, and dependent variable(s). The PCIT algorithm was used to develop gene interaction networks that identified several genes not previously identified with the typical GWAS analysis. The results obtained will aid in further dissecting the complex biology underlying fertility traits in dairy cattle.

ACKNOWLEDGMENTS

The Council on Dairy Cattle Breeding (Reynoldsburg, OH) and the Cooperative Dairy DNA Repository (Columbia, MO) are acknowledged for providing data used in this study. Financial support was provided by Agriculture and Food Research Initiative Grant 2013-68004-20365, "Improving Fertility of Dairy Cattle Using Translational Genomics," from the National Institute of Food and Agriculture (Washington, DC). J. B. Cole and D. J. Null also were supported by appropriated project 1265-31000-096-00, "Improving Genetic Predictions in Dairy Animals Using Phenotypic and Genomic Information," of the Agricultural Research Service of the United States Department of Agriculture (Washington, DC). Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

REFERENCES

- Adjaye, J., R. Herwig, T. C. Brink, D. Herrmann, B. Greber, S. Sudheer, D. Groth, J. W. Carnwath, H. Lehrach, and H. Niemann. 2007. Conserved molecular portraits of bovine and human blastocysts as a consequence of the transition from maternal to embryonic control of gene expression. *Physiol. Genomics* 31:315–327.
- Alves, B. C. A., P. R. Tobo, R. Rodrigues, J. C. Ruiz, V. F. M. Hospician de Lima, and C. A. Moreira-Filho. 2011. Characterization of bovine transcripts preferentially expressed in testis and with a putative role in spermatogenesis. *Theriogenology* 76:991–998. <http://dx.doi.org/10.1016/j.theriogenology.2011.04.027>.
- Ashwell, M. S., D. W. Heyen, T. S. Sonstegard, C. P. Van Tassell, Y. Da, P. M. VanRaden, M. Ron, J. I. Weller, and H. A. Lewin. 2004. Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. *J. Dairy Sci.* 87:468–475. [http://dx.doi.org/10.3168/jds.S0022-0302\(04\)73186-0](http://dx.doi.org/10.3168/jds.S0022-0302(04)73186-0).
- Bascom, S. S., and A. J. Young. 1998. A summary of the reasons why farmers cull cows. *J. Dairy Sci.* 81:2299–2305. [http://dx.doi.org/10.3168/jds.S0022-0302\(98\)75810-2](http://dx.doi.org/10.3168/jds.S0022-0302(98)75810-2).
- Berbari, N. F., N. W. Kin, N. Sharma, E. J. Michaud, R. A. Kesterson, and B. K. Yoder. 2011. Mutations in *Traf3ip1* reveal defects in cilogenesis, embryonic development, and altered cell size regulation. *Dev. Biol.* 360:66–76.
- Berry, D. P., J. W. M. Bastiaansen, R. F. Veerkamp, S. Wijga, E. Wall, B. Berglund, and M. P. L. Calus. 2012. Genome-wide association for fertility traits in Holstein-Friesian dairy cows using data from experimental research herds in four European countries. *Animal* 6:1206–1215.
- Bolormaa, S., J. E. Pryce, B. J. Hayes, and M. E. Goddard. 2010. Multivariate analysis of a genome-wide association study in dairy cattle. *J. Dairy Sci.* 93:3818–3833. <http://dx.doi.org/10.3168/jds.2009-2980>.
- Calus, M. P. L., Y. de Haas, and R. F. Veerkamp. 2013. Combining cow and bull reference populations to increase accuracy of genomic prediction and genome-wide association studies. *J. Dairy Sci.* 96:6703–6715. <http://dx.doi.org/10.3168/jds.2012-6013>.
- Cochran, S. D., J. B. Cole, D. J. Null, and P. J. Hansen. 2013. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genet.* 14:49 <http://dx.doi.org/10.1186/1471-2156-14-49>.
- Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. *J. Dairy Sci.* 92:2931–2946. <http://dx.doi.org/10.3168/jds.2008-1762>.
- Cole, J. B., G. R. Wiggans, L. Ma, T. S. Sonstegard, T. J. Lawlor, B. A. Crooker, C. P. Van Tassell, J. Yang, S. Wang, L. K. Matukumalli, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics* 12:408 <http://dx.doi.org/10.1186/1471-2164-12-408>.
- De Vries, A. 2006. Economic value of pregnancy in dairy cattle. *J. Dairy Sci.* 89:3876–3885.
- Doran, A. G., D. P. Berry, and C. J. Creevey. 2014. Whole genome association study identifies regions of the bovine genome and biological pathways involved in carcass trait performance in Holstein-Friesian cattle. *BMC Genomics* 15:837 <http://dx.doi.org/10.1186/1471-2164-15-837>.
- Dunlap, K. A., J. Filant, K. Hayashi, E. B. Rucker, G. Song, J. M. Deng, R. R. Behringer, F. J. DeMayo, J. Lydon, J.-W. Jeong, and T. E. Spencer. 2011. Postnatal deletion of *Wnt7a* inhibits uterine gland morphogenesis and compromises adult fertility in mice. *Biol. Reprod.* 85:386–396. <http://dx.doi.org/10.1095/biolreprod.111.091769>.
- Forde, N., and P. Lonergan. 2012. Transcriptomic analysis of the bovine endometrium: What is required to establish uterine receptivity to implantation in cattle? *J. Reprod. Dev.* 58:189–195. <http://dx.doi.org/10.1262/jrd.2011-021>.
- Fortes, M. R. S., K. L. Deatley, S. A. Lehnert, B. M. Burns, A. Reverter, R. J. Hawken, G. Boe-Hansen, S. S. Moore, and M. G. Thomas. 2013. Genomic regions associated with fertility traits in male and female cattle: advances from microsatellites to high-density chips and beyond. *Anim. Reprod. Sci.* 141:1–19. <http://dx.doi.org/10.1016/j.anireprosci.2013.07.002>.
- Fortes, M. R. S., A. Reverter, S. H. Nagaraj, Y. Zhang, N. N. Jonsson, W. Barris, S. Lehnert, G. B. Boe-Hansen, and R. J. Hawken. 2011. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J. Anim. Sci.* 89:1669–1683. <http://dx.doi.org/10.2527/jas.2010-3681>.
- Fortes, M. R. S., A. Reverter, Y. Zhang, E. Collis, S. H. Nagaraj, and N. N. Jonsson. 2010. Association weight matrix for the genetic

- dissection of puberty in beef cattle. *Proc. Natl. Acad. Sci. USA* 107:13642–13647. <http://dx.doi.org/10.1073/pnas.1002044107>.
- Gervasi, M. G., C. Osycka-Salut, J. Caballero, M. Vazquez-Levin, E. Pereyra, S. Billi, A. Franchi, and S. Perez-Martinez. 2011. Anandamide capacitates bull spermatozoa through CB1 and TRPV1 activation. *PLoS ONE* 6:e16993.
- Greenfield, C. R., M. E. Pepling, J. K. Babus, P. A. Furth, and J. A. Flaws. 2007. BAX regulates follicular endowment in mice. *Reproduction* 133:865–876.
- Guerin, L. R., L. M. Moldenhauer, J. R. Prins, J. J. Bromfield, J. D. Hayball, and S. A. Robertson. 2011. Seminal fluid regulates accumulation of FOXP3+ regulatory T cells in the preimplantation mouse uterus through expanding the FOXP3+ cell pool and CCL19-mediated recruitment. *Biol. Reprod.* 85:397–408. <http://dx.doi.org/10.1095/biolreprod.110.088591>.
- Hawken, R., and Y. Zhang. 2012. Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J. Anim. Sci.* 90:1398–1410. <http://dx.doi.org/10.2527/jas.2011-4410>.
- Höglund, J. K., B. Guldbbrandtsen, G. Su, B. Thomsen, and M. S. Lund. 2009. Genome scan detects quantitative trait loci affecting female fertility traits in Danish and Swedish Holstein cattle. *J. Dairy Sci.* 92:2136–2143. <http://dx.doi.org/10.3168/jds.2008-1104>.
- Hu, W., Z. Feng, A. K. Teresky, and A. J. Levine. 2007. p53 regulates maternal reproduction through LIF. *Nature* 450:721–724.
- Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44–57. <http://dx.doi.org/10.1038/nprot.2008.211>.
- Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37:1–13.
- Huang, W., B. W. Kirkpatrick, G. J. M. Rosa, and H. Khatib. 2010. A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. *Anim. Genet.* 41:570–578. <http://dx.doi.org/10.1111/j.1365-2052.2010.02046.x>.
- Hudson, N. J., B. P. Dalrymple, and A. Reverter. 2012. Beyond differential expression: The quest for causal mutations and effector molecules. *BMC Genomics* 13:356.
- Hudson, N. J., A. Reverter, and B. P. Dalrymple. 2009. A differential wiring analysis of expression data correctly identifies the gene containing the causal mutation. *PLOS Comput. Biol.* 5:e1000382 <http://dx.doi.org/10.1371/journal.pcbi.1000382>.
- Khatkar, M. S., F. W. Nicholas, A. R. Collins, K. R. Zenger, J. A. Cavanagh, W. Barris, R. D. Schnabel, J. F. Taylor, and H. W. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics* 9:187 <http://dx.doi.org/10.1186/1471-2164-9-187>.
- Knott, S. A., and C. S. Haley. 2000. Multitrait least squares for quantitative trait loci detection. *Genetics* 156:899–911.
- Kooper, A. J. A., B. H. W. Faas, I. Feenstra, N. de Leeuw, and D. F. C. M. Smeets. 2014. Best diagnostic approach for the genetic evaluation of fetuses after intrauterine death in first, second or third trimester: QF-PCR, karyotyping and/or genome wide SNP array analysis. *Mol. Cytogenet.* 7:6.
- Korol, A. B., Y. I. Ronin, A. M. Itskovich, J. Peng, and E. Nevo. 2001. Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits. *Genetics* 157:1789–1803.
- Kühn, C., J. Bennewitz, N. Reinsch, N. Xu, H. Thomsen, C. Loof, G. A. Brockmann, M. Schwerin, C. Weimann, S. Hiendleder, G. Erhardt, I. Medjugorac, M. Förster, B. Brenig, F. Reinhardt, R. Reents, I. Russ, G. Averdunk, J. Blümel, and E. Kalm. 2003. Quantitative trait loci mapping of functional traits in the German Holstein cattle population. *J. Dairy Sci.* 86:360–368.
- Lazzari, G., S. Colleoni, R. Duchi, A. Galli, F. D. Houghton, and C. Galli. 2011. Embryonic genotype and inbreeding affect preimplantation development in cattle. *Reproduction* 141:625–632.
- Levine, A. J., R. Tomasini, F. D. McKeon, T. W. Mak, and G. Melino. 2011. The p53 family: Guardians of maternal reproduction. *Nat. Rev. Mol. Cell Biol.* 12:259–265. <http://dx.doi.org/10.1038/nrm3086>.
- Liang, D. 2013. Estimating the economic losses from diseases and extended days open with a farm-level stochastic model. MS Thesis. University of Kentucky, Lexington.
- Maillard, V., S. Uzbekova, F. Guignot, C. Perreau, C. Ramé, S. Coyral-Castel, and J. Dupont. 2010. Effect of adiponectin on bovine granulosa cell steroidogenesis, oocyte maturation and embryo development. *Reprod. Biol. Endocrinol.* 8:23.
- Matzuk, M. M., and D. J. Lamb. 2002. Genetic dissection of mammalian fertility pathways. *Nat. Cell Biol.* 4:s41–s49.
- McCarthy, M. I., G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little, J. P. A. Ioannidis, and J. N. Hirschhorn. 2008. Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat. Rev. Genet.* 9:356–369.
- Meredith, B. K., D. P. Berry, F. Kearney, E. K. Finlay, A. G. Fahey, D. G. Bradley, and D. J. Lynn. 2013. A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front. Genet.* 4:229. <http://dx.doi.org/10.3389/fgene.2013.00229>.
- Norman, H. D., J. R. Wright, S. M. Hubbard, R. H. Miller, and J. L. Hutchison. 2009. Reproductive status of Holstein and Jersey cows in the United States. *J. Dairy Sci.* 92:3517–3528.
- Peddinti, D., E. Memili, and S. C. Burgess. 2010. Proteomics-based systems biology modeling of bovine germinal vesicle stage oocyte and cumulus cell interaction. *PLoS ONE* 5:e11240.
- Peters, S. O., K. Kizilkaya, D. J. Garrick, R. L. Fernando, J. M. Reecy, R. L. Weaver, G. A. Silver, and M. G. Thomas. 2013. Heritability and Bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. *J. Anim. Sci.* 91:605–612. <http://dx.doi.org/10.2527/jas.2012-5580>.
- Pritchard, T., M. Coffey, R. Mrode, and E. Wall. 2013. Genetic parameters for production, health, fertility and longevity traits in dairy cows. *Animal* 7:34–46.
- Quinlan, A. R., and I. M. Hall. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.
- R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Redmer, D. A., A. T. Grazul, J. D. Kirsch, and L. P. Reynolds. 1988. Angiogenic activity of bovine corpora lutea at several stages of luteal development. *J. Reprod. Fertil.* 82:627–634. <http://dx.doi.org/10.1530/jrf.0.0820627>.
- Reverter, A., and E. K. F. Chan. 2008. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics* 24:2491–2497. <http://dx.doi.org/10.1093/bioinformatics/btn482>.
- Reverter, A., and M. R. S. Fortes. 2013. Building single nucleotide polymorphism-derived gene regulatory networks: Towards functional genome-wide association studies. *J. Anim. Sci.* 91:530–536. <http://dx.doi.org/10.2527/jas.2012-5780>.
- Robinson, R. S., A. Hammond, D. Wathes, M. Hunter, and G. Mann. 2008. Corpus luteum–endometrium–embryo interactions in the dairy cow: Underlying mechanisms and clinical relevance. *Reprod. Domest. Anim.* 43:104–112. <http://dx.doi.org/10.1111/j.1439-0531.2008.01149.x>.
- Sahana, G., B. Guldbbrandtsen, C. Bendixen, and M. S. Lund. 2010. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. *Anim. Genet.* 41:579–588. <http://dx.doi.org/10.1111/j.1365-2052.2010.02064.x>.
- Sasseville, M., F. K. Albuz, N. Côté, C. Guillemette, R. B. Gilchrist, and F. J. Richard. 2009. Characterization of novel phosphodiesterases in the bovine ovarian follicle. *Biol. Reprod.* 81:415–425. <http://dx.doi.org/10.1095/biolreprod.108.074450>.
- Servin, B., and M. Stephens. 2007. Imputation-based analysis of association studies: Candidate regions and quantitative traits. *PLoS Genet.* 3:e114 <http://dx.doi.org/10.1371/journal.pgen.0030114>.
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape:

- A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498–2504. <http://dx.doi.org/10.1101/gr.1239303>.
- Spencer, C. C. A., Z. Su, P. Donnelly, and J. Marchini. 2009. Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet.* 5:e1000477. <http://dx.doi.org/10.1371/journal.pgen.1000477>.
- Storey, J. D. 2002. A direct approach to false discovery rates. *J. R. Stat. Soc., B* 64:479–498.
- Tesfaye, D., A. Regassa, F. Rings, N. Ghanem, C. Phatsara, E. Tholen, R. Herwig, C. Un, K. Schellander, and M. Hoelker. 2010. Suppression of the transcription factor MSX1 gene delays bovine preimplantation embryo development in vitro. *Reproduction* 139:857–870.
- Tonellotto dos Santos, J., R. Ferreira, B. G. Gasperin, L. C. Siqueira, J. F. de Oliveira, R. A. Santos, A. M. Reis, and P. B. Gonçalves. 2012. Molecular characterization and regulation of the angiotensin-converting enzyme type 2/angiotensin-(1–7)/MAS receptor axis during the ovulation process in cattle. *J. Renin Angiotensin Aldosterone Syst.* 13:91–98. <http://dx.doi.org/10.1177/1470320311417273>.
- Ulbrich, S. E., S. Rehfeld, S. Bauersachs, E. Wolf, R. Rottmayer, S. Hiendleder, M. Vermehren, F. Sinowatz, H. H. D. Meyer, and R. Einspanier. 2006. Region-specific expression of nitric oxide synthases in the bovine oviduct during the oestrous cycle and in vitro. *J. Endocrinol.* 188:205–213. <http://dx.doi.org/10.1677/joe.1.06526>.
- VanRaden, P. M., and J. B. Cole. 2014. Net merit as a measure of lifetime profit: 2014 revision. Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD. Accessed Aug. 12, 2015. <http://aipl.arsusda.gov/reference/nmcalc-2014.htm>.
- VanRaden, P. M., D. J. Null, M. Sargolzaei, G. R. Wiggans, M. E. Tooker, J. B. Cole, T. S. Sonstegard, E. E. Connor, M. Winters, J. B. C. H. M. van Kaam, A. Valentini, B. J. Van Doormaal, M. A. Faust, and G. A. Doak. 2013. Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* 96:668–678. <http://dx.doi.org/10.3168/jds.2012-5702>.
- VanRaden, P. M., J. R. O'Connell, G. R. Wiggans, and K. A. Weigel. 2011. Genomic evaluations with many more genotypes. *Genet. Sel. Evol.* 43:10.
- VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, and H. D. Norman. 2003. Daughter pregnancy rate evaluation of cow fertility. AIPL Research Report: DPR1 (11–02). Accessed Aug. 12, 2015. https://aipl.arsusda.gov/reference/fertility/DPR_rpt.htm.
- VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn, and G. R. Wiggans. 2004. Development of a national genetic evaluation for cow fertility. *J. Dairy Sci.* 87:2285–2292. [http://dx.doi.org/10.3168/jds.S0022-0302\(04\)70049-1](http://dx.doi.org/10.3168/jds.S0022-0302(04)70049-1).
- VanRaden, P. M., and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* 74:2737–2746. [http://dx.doi.org/10.3168/jds.S0022-0302\(91\)78453-1](http://dx.doi.org/10.3168/jds.S0022-0302(91)78453-1).
- Watson-Haigh, N. S., H. N. Kadarmideen, and A. Reverter. 2010. PCIT: An R package for weighted gene co-expression networks based on partial correlation and information theory approaches. *Bioinformatics* 26:411–413. <http://dx.doi.org/10.1093/bioinformatics/btp674>.
- Wiggans, G. R., T. A. Cooper, D. J. Null, and P. M. VanRaden. 2013. Genomic evaluations using 61,013 markers. Changes to Evaluation System (December 2013). Council on Dairy Cattle Breeding. Accessed Aug. 12, 2015. <https://www.cdcb.us/reference/changes/chglst.cfm?run=1312>.
- Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. *J. Dairy Sci.* 93:2287–2292. <http://dx.doi.org/10.3168/jds.2009-2773>.
- Wiltbank, M., H. Lopez, R. Sartori, S. Sangsritavong, and A. Gümen. 2006. Changes in reproductive physiology of lactating dairy cows due to elevated steroid metabolism. *Theriogenology* 65:17–29. <http://dx.doi.org/10.1016/j.theriogenology.2005.10.003>.
- Zamberlam, G., V. Portela, J. F. C. de Oliveira, P. B. D. Gonçalves, and C. A. Price. 2011. Regulation of inducible nitric oxide synthase expression in bovine ovarian granulosa cells. *Mol. Cell. Endocrinol.* 335:189–194. <http://dx.doi.org/10.1016/j.mce.2011.01.013>.
- Zeng, L., Y. Wan, D. Li, J. Wu, M. Shao, J. Chen, L. Hui, H. Ji, and X. Zhu. 2013. The m subunit of murine translation initiation factor eIF3 maintains the integrity of the eIF3 complex and is required for embryonic development, homeostasis, and organ size control. *J. Biol. Chem.* 288:30087–30093. <http://dx.doi.org/10.1074/jbc.M113.506147>.
- Zhang, Q., Q. Li, J. Li, X. Li, Z. Liu, D. Song, and Z. Xie. 2008. b-DAZL: A novel gene in bovine spermatogenesis. *Prog. Nat. Sci.* 18:1209–1218.
- Zhou, X. 2014. GEMMA User Manual. University of Chicago, Chicago, IL.
- Zhou, X., P. Carbonetto, and M. Stephens. 2013. Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genet.* 9:e1003264. <http://dx.doi.org/10.1371/journal.pgen.1003264>.
- Zhou, X., and M. Stephens. 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11:407–409. <http://dx.doi.org/10.1038/nmeth.2848>.
- Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassel, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yorke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10:R42.