

RESEARCH

Open Access



Validation of loci and genes associated with fertility in Holstein cows using gene-set enrichment analysis-SNP and genotype-by-sequencing

Jennifer N. Kiser¹, Christopher M. Seabury², Mahesh Neupane³, Joao G. N. Moraes⁴, Allison L. Herrick¹, Joseph Dalton⁵, Gregory W. Burns⁶, Thomas E. Spencer⁷ and Holly L. Neibergs^{1*}

Abstract

Background The financial strain fertility issues cause the dairy cattle industry is substantial, with over \$7 billion in lost revenue accrued annually due to a relatively low cow conception rate (CCR; 30–43%) for US dairy cows. While CCR has been improving through genomic selection, identification of causal mutations would help improve the rate of genetic progress with genomic selection and provide a better understanding of infertility. The objectives of this study were to: (1) identify genes and gene-sets associated with CCR to the first breeding (CCR1) and the number of breedings required to conceive (TBRD) in Holstein cows and (2) identify putative functional variants associated with CCR1 and TBRD through a custom genotype-by-sequencing array. The study consisted of 1,032 cows (494 pregnant to first breeding, 472 pregnant to subsequent [2–20] services, and 66 that never conceived). Cows were artificially inseminated, and pregnancy was determined 35d later by rectal palpation of uterine contents. Gene-set enrichment analyses with SNP data (GSEA-SNP) were conducted for CCR1 and TBRD with a normalized enrichment score (NES) ≥ 3.0 required for significance. Leading edge genes (LEG) and positional candidate genes from this and 26 additional studies were used to validate 100 loci associated ($P < 1 \times 10^{-5}$) with cow fertility using a custom sequencing genotyping array of putative functional variants (exons, promoters, splice sites, and conserved regions).

Results GSEA-SNP identified 95 gene-sets (1,473 LEG) enriched for CCR1 and 67 gene sets enriched (1,438 LEG) for TBRD (NES ≥ 3). Thirty-four gene-sets were shared between CCR1 and TBRD along with 788 LEG. The association analysis for TBRD identified three loci: BTA1 at 83 Mb, BTA1 at 145 Mb, and BTA 20 at 46 Mb ($P < 1 \times 10^{-5}$). The loci associated with TBRD contained candidate genes with functions relating to implantation and uterine receptivity. No loci were associated with CCR1, however a single locus on BTA1 at 146 Mb trended toward significance with an FDR of 0.04.

Conclusions The validation of three loci associated with CCR and TBRD in Holsteins can be used to improve fertility through genomic selection and provide insight into understanding infertility.

*Correspondence:
Holly L. Neibergs
neibergs@wsu.edu

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Keywords Fertility, GBS, GSEA-SNP, Holstein cattle, Uterine receptivity

Background

Revenues from dairy cattle products account for roughly 21.3% (\$41.8 billion) of all animal cash receipts within the United States [1]. Worldwide, global revenue attributed to milk is expected to amount to \$350 billion in 2024 and continue to grow in subsequent years [2]. Prior to lactation and the generation of dairy products, cattle must first conceive and maintain the pregnancy through the normal gestation period, making fertility critical to dairy profitability.

The dairy industry has struggled with fertility for decades, with the average cow conception rate (CCR) ranging between 30–43% [3–4]. Previous studies have estimated the average economic cost of pregnancy loss from breeding to day 60 of gestation in dairy cows to be \$2,333 [5]. This cost is dependent on several factors including length of pregnancy, increased chance of culling, and the cost of rebreeding. To put this on a national scale, if 35% of the 9.38 million U.S. cows [6] lost a single pregnancy it would result in a cost of greater than \$7.6 billion dollars to the dairy industry. This illustrates the importance of selecting cattle that can conceive and maintain pregnancies.

Prior to the early 2000s, genetic selection in the dairy industry primarily focused on milk yield or milk components [7–8]. Selection for female fertility began with the inclusion of daughter pregnancy rate in 2003 [9], and expanded to include overall heifer conception rate and CCR as well as conception rates to first breeding in 2009 [10]. Additional fertility measures such as early first calving (the age at which a heifer calves) were added to selection indexes in 2019 [11].

Since its inclusion in the Council for Dairy Cattle Breeding genomic evaluations, cow conception rate to the first breeding (CCR1) has improved by 7% points reaching 44%, while CCR to any service has improved by 6% points reaching 42% [4]. This improvement in CCR has been achieved through genomic selection using single nucleotide polymorphisms (SNPs) that are in linkage disequilibrium (LD) with the mutation responsible for the phenotype (causal mutation). To date, few causal variants are known or used in genomic selection for female fertility traits as finding the causal mutation can be both time consuming and challenging. The advantages to identifying the causal mutations is that the accuracy of prediction of fertility will be higher because the prediction is not reliant on the lack of changing LD with successive generations of meiotic events or recombination differences between breeds. Identifying the causal mutation also offers insights into the molecular mechanism of fertility [12].

Given the importance of fertility to the dairy industry and the lack of characterization of casual mutations associated with CCR, the goals of the current study were to: (1) identify leading edge genes and gene sets enriched for CCR and the number of times a cow was bred to achieve a pregnancy (TBRD) using gene set enrichment analysis using SNPs (GSEA-SNP) and (2) perform an association analysis of loci validated for an association with fertility (from at least two independent studies) using a custom genotype by sequencing (GBS) panel to identify putative causal variants associated with the CCR and TBRD in Holstein cows.

Materials and methods

Study population

This study was conducted with the approval of Washington State University's Institutional Animal Care and Use Committee (#4295) and the authors confirm that this study is reported in accordance with ARRIVE guidelines. The Holstein study population consisted of 2,015 Holstein cows sampled from six dairies located in central Washington [13]. All dairies provided written agreement to participate in the study. No animals were euthanized for this study. Whole blood (~16 ml) was collected into EDTA tubes from each cow via venipuncture of the tail vein. Blood was collected from cows without being anesthetized, as approved by the Washington State University's Institutional Animal Care and Use Committee (#4295). Cows ($n=951$) were removed from the study if they experienced an inflammatory event or disease that could potentially have impacted pregnancy retention such as: lameness, fever, mastitis, uterine disease, spontaneous abortions, dystocia, foot disease, pink eye, respiratory disease, and metabolic issues. All cows were artificially inseminated (AI) at either observed estrus or using timed AI. Pregnancy status was determined at 35 days post insemination by rectal palpation of uterine contents. Pregnancy status after d35 was not evaluated for the current study, as this study aims to focus on early embryonic loss not fetal losses. Animals remaining in the study ($n=1064$) were genotyped using the Illumina BovineHD Beadchip (San Diego, CA) at Neogen Laboratories (Lincoln, NE). Prior to any analysis, samples were quality control filtered for duplicates and individual call rate (<90%) with 32 cattle removed. This left a total of 1,032 cows for the study: 494 cows that conceived to the first service, 472 cows that were bred 2–20 times before conceiving, and 66 infertile cows that never conceived after a minimum of 6 AI services (range 6–20 AI attempts) [13]. The phenotypes of CCR1 and TBRD were used for the GSEA-SNP and the association analysis.

Phenotypes were coded as follows: CCR1 using 1 (successful conception to first service) and 0 (no conception at first service); and TBRD which used a scale from 1 to 20 depending on the number of inseminations a cow received.

For this study, only first lactation (first parity) cow records were used to ensure the entire study population were the same age for comparison. This data was collected retrospectively through farm management records. The average 305-day milk yields did not differ ($P=0.14$) across the different phenotypic groups of cattle. The average 305-day milk yield for the entire population was 28,831 kg, additional milk yield information can be found in Supplemental Table 1. Conception rate did not differ between AI technicians ($P>0.05$) or between service sires ($P=0.99$) [13]. A total of 435 sires (433 Holstein and 2 Angus bulls) were used across all dairies with an average conception rate of 26.8% [13]. Most cows that failed to conceive to one sire were subsequently rebred to a different service sire reducing the risk for a cow to be categorized as infertile in the TBRD phenotype due solely to a service sire effect. Dairy of origin did have an effect on CCR1 and TBRD ($P<1 \times 10^{-10}$) and was included as a covariate in all analyses.

Gene-set enrichment analysis-SNP

The genotypes from the Illumina BovineHD Beadchip (San Diego, CA) at Neogen Laboratories (Lincoln, NE), were used to conduct an association analysis to identify the SNPs that would serve as representatives for the genes in the gene sets [13]. A total of 625,093 SNP were analyzed and mapped to 21,039 protein-coding genes within the ARS-UCD 1.2 genome assembly (https://www.animalgenome.org/repository/cattle/UMC_bovine_coordinates/). SNPs with the greatest evidence of an association with CCR1 or TBRD were used as the gene proxies for the GSEA-SNP. One SNP represented each gene within a 17 kb region up- and down-stream of the gene. This 17 kb region was based on the average haplotype block size of a large ($n=4,800$) Holstein cattle population [13–15] also genotyped with the BovineHD BeadChip calculated using the method described by Gabriel et al., in 2002 [16]. A single SNP could represent more than one gene if it fell within the parameters outlines above. Genes were ranked by their significance (P -value) for their association with CCR1 or TBRD from the genome wide association analysis.

The GSEA-SNP was performed following the methods of Wang et al. (2007) [17] and five gene set databases were used: Biocarta (217 gene sets; <http://www.genecart.com/>), Gene Ontology or GO (3,147 gene sets; <http://www.geneontology.com>), Kyoto Encyclopedia of Genes and Genomes or KEGG (186 gene sets; <http://www.genome.jp.kegg>), Protein Analysis Through Evolutionar

y Relationships or PANTHER (165 gene sets; <http://www.pantherdb.org>), and Reactome (674 gene sets; <http://www.reactome.org>). An enrichment score for each gene set was computed using the highest value from the running sum statistics, similar to a weighted Kolmogorov-Smirnov-like statistic [18]. Each gene set received a permuted P -value calculated using 10,000 phenotype-based permutations in GenABEL in R [19]. To account for the varying number of genes within each gene set, a normalized enrichment score (NES) was computed, and enriched gene sets were those that had an $NES \geq 3.0$. For enriched gene sets, a list of genes that contributed to the peak NES were identified as leading edge genes (LEG).

Genotype by sequencing

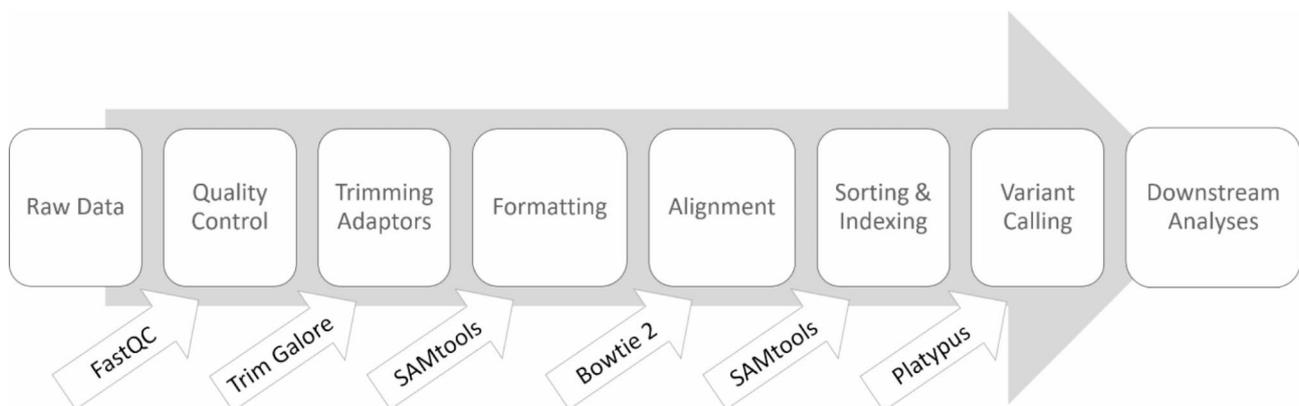
Fertility validated loci were identified by comparing results from twenty-six studies [13, 15, 20–43] (Table 1), as well as the GSEA-SNP from the current study, that utilized a range of fertility phenotypes and cattle breeds. A locus was defined as SNP associated with fertility that were in linkage disequilibrium (LD) with each other ($D' > 0.75$). This D' threshold was used as it falls within ranges previously used to define SNPs within a locus [42]. To identify putative causal mutations in the 202 loci validated by two or more fertility studies, a custom genotyping array was designed after whole genome sequencing was conducted on 24 Holsteins at 10X coverage using the Illumina HiSeq X Ten at the University of Missouri sequencing core. Sequences of the 202 loci were mapped to the ARS-UCD 1.2 reference assembly. There were 5,102 variants identified within the 202 loci. Variants were screened to determine which ones were located in regulatory regions ($n=317$), exons ($n=633$), splice sites ($n=22$), and evolutionarily conserved regions ($n=392$). After screening, 1,358 variants within 152 loci were sent for probe design at Neogen Laboratories (Lincoln, NE). Of the variants identified, 284 variants (21%) failed the probe design due to the presence of secondary structures, high or low GC content, or repetitive motifs of the targeted sequence. Probes were successfully designed for 713 variants covering both directions and 361 variants where probes were successfully designed for a single direction. This resulted in the creation of a genotyping array with 1,787 probes which covered 100 of the 152 loci identified across studies. Paired-end 150 bp libraries for 1,064 cows were prepared and analyzed at Neogen laboratories (Lincoln, NE) following Tecan's targeted genotyping V2 Allegro protocols (Männedorf, Switzerland). Sequencing was performed using an Illumina NextSeq2000 system (San Diego, CA).

Raw sequencing data was processed to create variant call format (VCF) files for the association analysis (Fig. 1). Raw sequencing reads were assessed for quality using the FastQC software [44] and adaptors were

Table 1 List of cattle fertility studies used for loci validation

Study ¹	Phenotype(s) ²	Breeds ³
Akanno et al. 2015 [31]	AFC, PR	Crossbred beef ^{Ca}
Blaschek et al. 2011 [24]	Non-compensatory fertility	Holstein ^{Ur}
Cochran et al. 2013 [27]	DPR	Holstein ^{Ur}
Cole et al. 2011 [25]	DPR	Holstein ^U
Fonseca et al. 2018 [41]	Reproduction	Brangus ^U , Brahman ^A , Tropical Composite ^A
Galliou et al. 2020 [42]	HCR1, TBRD	Holstein ^U
Höglund et al. 2009 [20]	AISC, AISH, ICF, IFLC, IFLH, NRRC, NRRH	Holstein ^N
Höglund et al. 2014 [30]	AISC, AISH, FTI, ICF, IFLC, IFLH, NRRC, NRRH	Holstein ^N , Jersey ^N , Nordic Red ^N
Höglund et al. 2015 [32]	AISC, AISH, FTI, ICF, IFLC, IFLH, NRRC, NRRH	Nordic Red ^N
Huang et al. 2010 [21]	FR, BR	Holstein ^U
Iso-Touru et al. 2016 [33]	Milk production, AISC, AISH, FTI, ICF, IFLC, IFLH, NRRC, NRRH	Nordic Red ^N
Kiser et al. 2019a [13]	CCR1, TBRD	Holstein ^U
Kiser et al. 2019b [15]	HCR1, TBRD	Holstein ^U
Liu et al. 2017	AFS, AISC, CCR, DO, ICF, IFLC, NRRC	Holstein ^{C, N}
Minozzi et al. 2013 [28]	CI, DFS, FTI, NRR	Holstein ^{It}
Minten et al. 2013 [29]	High vs. Low Fertile	Crossbred Beef ^U
Moore et al. 2016 [34]	CI	Holstein ^{A, I} , Jersey ^A
Müller et al. 2017 [38]	CTFS, DO, FSTCC, FSTCH, NRRC, NRRH,	Holstein ^G
Nayeri et al. 2016 [35]	CTFS, DO, FSTCH	Holstein ^{Ca}
Neupane et al. 2017 [39]	P28	Angus crosses ^U
Oliver et al. 2020 [43]	HCR1, TBRD	Crossbred Beef ^U
Olsen et al. 2011 [26]	NRRC, NRRH	Nordic Red ^N
Ortega et al. 2016 [36]	DPR	Holstein ^U
Ortega et al. 2017 [40]	DPR	Holstein ^U
Pryce et al. 2010 [22]	Fertility	Holstein ^A , Jersey ^A
Sahana et al. 2010 [23]	AISC, AISH, FTI, IFLC, IFLH, ICF, NRRC, NRRH	Holstein ^N

¹The citation number for each study is listed in superscript brackets. ²Traits abbreviated as follows: AFC - age at first calving; AISC - number of inseminations to conception in cows; AISH - number of inseminations to conception in heifers; AFS - age at first insemination; BR - blastocyst rate; CCR1 - conception rate to first insemination in cows; CI - calving interval; CTFS - days from calving to first insemination; DFS - days to first service; DO - days open; DPR - daughter pregnancy rate; FSTCC - days from first service to conception in cows; FSTCH - days from first service to conception in heifers; FR - fertility rate; FTI - fertility index; HCR - heifer conception rate; HCR1 - conception rate to first insemination in heifers; ICF - interval (in days) from calving to first insemination; IFLC - days from first to last insemination in cows; IFLH - days from first to last insemination in heifers; NRR - 56 day non return rate; NRRC - 56 day non return rate in cows; NRRH - 56 day non return rate in heifers; P28 - pregnancy success at day 28 post embryo transfer; P42 - pregnancy success within first 42 days of mating; PR - pregnancy rate; TBRD - number of times bred to conception. ³Cattle breeds loci were previously identified are listed with the country or region the population was from indicated in superscript as follows: Australia - A; Canada - Ca; Chinese - C; Germany - G; Ireland - I; Nordic - N; United States - U; Unreported - Ur

**Fig. 1** Processing pipeline for GBS data

trimmed off with Trim Galore and Cutadapt [45, 46]. Files were aligned to the ARS-UCD 1.2 reference assembly using bowtie2 [47], and sorted and indexed using SAMtools [48]. Variant calling was successful for 1,222 or

the 1,787 probes using Platypus [49]. The Platypus VCF files were then uploaded into SNP and Variation Suite (SVS) version 8.9.1 (Golden Helix, Bozeman, MT) for analysis. Quality control filtering of SNPs was performed

on the 1,222 called variants. SNPs were then removed for call rate < 85% ($n=127$) or minor allele frequency < 1% ($n=80$), leaving 1,015 SNPs spanning 100 loci that were further analyzed.

Association analysis

Three association analyses were conducted for each phenotype (CCR1 and TBRD) using the GBS variants in conjunction with the Illumina BovineHD BeadChip (San Diego, CA) genotypes the cattle were originally genotyped on. Given that the model of inheritance for cow fertility is unknown, three modes of inheritance were analyzed for each phenotype: additive, dominant, and recessive. SNPs were considered significant if their uncorrected P -value met the Wellcome Trust significance threshold of $P < 1 \times 10^{-5}$ [50]. The association analyses were performed using SVS software and an efficient mixed-model association expedited (EMMAX) model [51]. This model was described in detail previously [13, 15, 51]. Briefly, the model is described by $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$, \mathbf{y} is the $n \times 1$ vector of observed phenotypes, \mathbf{X} is an $n \times f$ matrix of fixed effects (f), $\boldsymbol{\beta}$ is an $f \times 1$ vector containing the fixed effect coefficients, \mathbf{Z} is an $n \times t$ matrix relating the random effects (t) to the phenotype, and \mathbf{u} is the random effect of the mixed model. The model assumes residuals to be independent with an identical distribution such that $Var(\mathbf{u}) = \sigma_g^2 \mathbf{K}$ and $Var(\boldsymbol{\epsilon}) = \sigma_e^2 \mathbf{I}$, and $Var(\mathbf{y}) = \sigma_g^2 \mathbf{Z}\mathbf{K}\mathbf{Z}' + \sigma_e^2 \mathbf{I}$. For this study, \mathbf{K} is a matrix of pairwise genomic relationships and \mathbf{Z} is the identity matrix, \mathbf{I} [52].

Observed heritability was estimated using a genomic best linear unbiased predictor (GBLUP) [53] analysis using AI-REML (average information algorithm) [54, 55]. While computationally intensive, this method to estimate heritability provides a more accurate estimation of heritability when sample size is limited.

Results

Gene-set enrichment analysis-SNP

The GSEA-SNP identified 95 gene sets with $NES \geq 3$ (Supp. Table 2) that were enriched for CCR1. There were 1,473 unique LEG across the 95 gene sets, with 2 LEG, scinderin (*SCIN*) and villin 1 (*VIL1*), found in 40 of the 95 gene sets. Gene sets with greatest evidence for enrichment with CCR1 ($NES > 4.0$) are shown in Table 2. For TBRD, 67 gene sets were enriched ($NES \geq 3$; Supp. Table 3) and 4 gene sets had an $NES \geq 4$ (Table 3). Two gene sets with $NES > 4.0$ for TBRD (recruitment of mitotic centrosome proteins and complexes; loss of NLP from mitotic centrosomes) also were enriched for CCR1 with $NES > 4.0$. There were 1,438 unique LEG for TBRD. A single LEG, Rac family small GTPase 1 (*RAC1*), was shared in 28 gene sets.

Thirty-four gene sets were enriched in both CCR1 and TBRD (Table 4), which included 788 shared LEG. The three most commonly shared LEG were twinfilin actin binding protein 1 (*TWFI*) identified in 58 gene sets, *RAC1* in 59 gene sets and *VIL1* in 61 gene sets that contribute to the function of the cytoskeleton.

Many of the shared gene sets can be grouped into categories based on common functions. These groups include gene sets related to cellular component organization and structure ($n=17$), mitotic cell cycle/mitosis ($n=6$), RNA and mRNA metabolic processes ($n=3$), and signaling pathways ($n=3$). The remaining five gene sets have functions that do not fall in any of the above-mentioned categories (Table 4).

Association analyses

There were no SNP associated with CCR1 in any model ($P < 1 \times 10^{-5}$). However, a single locus on BTA1 at 146 Mb in the additive model had a false discovery rate (FDR) of 0.04, though the P -value did not meet the Wellcome Trust significance threshold. The observed heritability for this model was 0.19 ± 0.05 .

The TBRD association analysis identified 4 SNP (3 loci), two on BTA1 and one on BTA20 in the additive model ($P < 1 \times 10^{-5}$; Table 5). The locus on BTA1 at 146 Mb that was associated with TBRD was the same locus in the additive model of CCR1 that trended toward significance (FDR=0.04). This locus contained a synonymous variant within disco interacting protein 2 homolog A (*DIP2A*). The other locus on BTA1 was located at 83 Mb. The two SNP within this locus were a missense and a synonymous variant within the first exon of *ENSBTAG00000032217*. This missense mutation (rs110205198), which results in an amino acid change from cysteine to serine, was further examined to determine if there were any predicted detrimental effects caused by the amino acid substitution. This was done using the sequencing homology-based tool SIFT (sorts intolerant from tolerant) which predicts if substitutions of amino acids could result in a phenotypic effect of the protein through a change in its structure with values ranging from 0 (intolerable change) to 1 (tolerated change) [56]. The predicted damage score for the amino acid substitution caused by this mutation was only 0.9 (with low confidence) indicating the substitution is likely tolerated. However, the low confidence warning indicates that there was not enough diversity between the sequences to accurately predict altered protein function [56]. The locus on BTA20 contained a variant within a CNV (nsv810323). No SNP were associated with TBRD in the dominant or recessive models and observed heritability for TBRD was 0.18 ± 0.05 .

The significant GBS SNP were then evaluated for LD with the original SNP that were used to create the custom assay. When compared to the original nine SNP used in

Table 2 Gene set enrichment analysis– single nucleotide polymorphism results for conception rate to first service (CCR1) in Holstein cows

Gene Sets ^a	# Genes (# LEG) ^b	NES ^c	Leading Edge Genes ^d
Cell Cycle, Mitotic (R-HSA-69278)	300 (93)	5.08	<i>RAD21, PAFAH1B1, CEP290, CLIP1, RBL2, TUBB4B, CENPH, CCNB1, CDK6, CCNE2, CEP250, CENPJ, CASCS, PRIM2, KNTC1, CDK1, YWHAE, CLASP1, BTRC, CDC14A, SDCCAG8, CDK7, CDKN2A, CDKN2B, CEP41, CCNA2, NUF2, KIF23, PSMA4, PSMB8, PSMB9, MCM7, MCM3, MCM2, UBE2C, OFD1, RPA3, KIF2B, NUP133, AHCTF1, PRKAR2B, NUDC, PRIM1, TYMS, ALMS1, CENPT, ANAPC5, CEP72, CEP70, CUL1, DHFR, ANAPC7, SKA1, AKAP9, NINL, PRKACA, STAG2, CDC45, CCNE1, PLK1, PSMD11, DYRK1A, CDT1, PSMB7, ITGB3BP, MAD1L1, NUMA1, INCENP, CEP135, RBL1, FBXO5, PPP2R5E, ANAPC10, DCTN1, Sect. 13, PSMB10, CDC7, PTTG1, BUB1B, CENPP, BUB1, PSMA2, LINS2, PSME1, PSME2, TUBGCP3, FGFR1OP, PSMB2, PPP2R2A, PSMC1, RB1, YWHAG, CCND2</i>
Cell Cycle (R-HSA-1640170)	364 (126)	4.87	<i>RAD21, PAFAH1B1, CEP290, CLIP1, RBL2, TUBB4B, CENPH, CCNB1, CDK6, RFWD2, SYNE2, H2AFX, CCNE2, CEP250, CENPJ, CASCS, PRIM2, KNTC1, CDK1, YWHAE, CLASP1, BTRC, CDC14A, SDCCAG8, CDK7, CDKN2A, CDKN2B, CEP41, CCNA2, NUF2, RUVBL1, KIF23, PSMA4, POT1, PSMB8, PSMB9, MCM7, MCM3, ATR, MCM2, UBE2C, OFD1, HIST1H4D, RPA3, KIF2B, NUP133, AHCTF1, PRKAR2B, SYNE1, NUDC, PRIM1, TYMS, ALMS1, CENPT, ANAPC5, CEP72, CEP70, CUL1, FKBP6, DHFR, ANAPC7, SKA1, HIST1H2BN, SMC1B, AKAP9, NINL, PRKACA, STAG2, CDC45, CCNE1, PLK1, PSMD11, DYRK1A, MIS18A, RAD1, CDT1, PSMB7, ITGB3BP, MAD1L1, NUMA1, INCENP, HUS1, CEP135, RBL1, FBXO5, PPP2R5E, ANAPC10, DCTN1, Sect. 13, PSMB10, CDC7, PTTG1, BUB1B, CENPP, BUB1, PSMA2, LINS2, PSME1, PSME2, TUBGCP3, FGFR1OP, PSMB2, PPP2R2A, PSMC1, RB1, YWHAG, CCND2, HIST1H2BI, REC8, CEP76, HJURP, CHEK2, HIST1H2BJ, PSMC6, HIST3H2BB, RSF1, GINS4, MCM8, ORC2, PCNT, LINS4, ANAPC1, PSMC2, HIST1H2BB, ANAPC2, SSNA1</i>
Protein Polymerization (GO:0051258)	79 (21)	4.81	<i>TWF1, TUBB4B, ARFIP2, TUBA8, CAPZA2, F2, SCIN, WASF1, SLAIN2, ANG, VIL1, CAPZB, TMSB4, RASA1, SNCA, ARPIN, ARPC1A, UBE2C, RAC1, WASL, ARPC5</i>
Recruitment of Mitotic Centrosome Proteins and Complexes (R-HSA-380270)	62 (22)	4.77	<i>PAFAH1B1, CEP290, TUBB4B, CCNB1, CEP250, CENPJ, CDK1, YWHAE, CLASP1, SDCCAG8, CEP41, OFD1, PRKAR2B, ALMS1, CEP72, CEP70, AKAP9, NINL, PRKACA, PLK1, NUMA1, CEP135</i>
Mitotic G2-G2/M phases (R-HSA-453274)	76 (24)	4.74	<i>PAFAH1B1, CEP290, TUBB4B, CCNB1, CEP250, CENPJ, CDK1, YWHAE, CLASP1, SDCCAG8, CDK7, CEP41, CCNA2, OFD1, PRKAR2B, ALMS1, CEP72, CEP70, AKAP9, NINL, PRKACA, PLK1, NUMA1, CEP135</i>
Loss of NLP From Mitotic Centrosomes (R-HSA-380259)	55 (20)	4.68	<i>PAFAH1B1, CEP290, TUBB4B, CEP250, CENPJ, CDK1, YWHAE, CLASP1, SDCCAG8, CEP41, OFD1, PRKAR2B, ALMS1, CEP72, CEP70, AKAP9, NINL, PRKACA, PLK1, CEP135</i>
Cellular Component Disassembly (GO:0022411)	81 (20)	4.64	<i>TWF1, DPP4, MICAL3, CCNB1, GABARAPL1, MMP13, CAPZA2, SCIN, VIL1, CAPZB, ENDOG, FOXL2, BAX, KIF2B, FIS1, NAPB, STMN1, DSTN, SBDS, MTRF1L</i>
Macromolecular Complex Assembly (GO:0065003)	372 (84)	4.47	<i>TWF1, SRSF1, TUBB4B, CNOT7, HES1, SF3A1, H1FOO, CENPH, CHMP4A, DPAGT1, CAMK2D, KCTD1, SAMHD1, VAMP4, ARFIP2, RAD51, SRR, TUBA8, SLC6A4, CAPZA2, TBCA, OAT, F2, FECH, SCIN, WASF1, SLAIN2, ANG, VIL1, PEX5, VWF, STRAP, EIF6, COX7A2L, CAPZB, SNAP29, MGST1, TMSB4, RASA1, IGF1R, GEMIN8, LRRC6, NAP1L1, CBR4, SNCA, NUDT21, ARPIN, COX19, ARPC1A, ASF1A, UBE2C, NLRC4, NDUFAF6, MICU1, CLU, BAX, LMO4, RAC1, CAT, WASL, ARPC5, FIS1, NAPB, PRPF19, CYBA, SBDS, DRC1, LONP1, ARPC3, TARBP2, PDCL, IL5, H4, HIST1H1A, ATP6V0A2, PSMD11, RDX, ARL6, MIS18A, FAS, LUC7L3, KCTD5, SAR1A, MIF</i>
Signaling by NGF (R-HAS-9031628)	207(80)	4.42	<i>NTRK2, SOS2, APH1B, RAP1A, PRKCE, ADCY2, ADCY5, ATF1, VAV3, MAPK8, AKT3, MYD88, SORCS3, CDK1, AKAP13, ADAM17, YWHAE, MCF2, TIAM2, CAMK4, TIAM1, PRKCI, FOXO3, ADCY1, DUSP6, FGD2, ADCY9, SQSTM1, FGD4, ABR, PIK3R1, NR4A1, ITSN1, PDPK1, OBSCN, RALB, MEF2A, PDE1A, AP2A1, PRKAR2A, RAC1, RPS6KA3, CHUK, KIDINS220, AP2B1, PRKAR2B, FURIN, ADCYAP1R1, PIK3CB, RASGRF2, ARHGEF3, SRC, ARHGAP4, MAPKAP1, PHLPP1, SHC3, CASP9, FGD3, LINGO1, ARHGEF6, PRKACA, MAP2K5, ITPR3, ADCY8, ADCY7, ITGB3BP, MEF2C, TRIO, FOXO1, PLCG1, AKT1S1, MTOR, PIK3CA, CLTC, CLTA, RIPK2, ADCY3, TRAF6, PRKAR1B, APM1A</i>
Protein Complex Assembly (GO:0065003)	303 (73)	4.38	<i>TWF1, TUBB4B, HES1, H1FOO, CENPH, CHMP4A, DPAGT1, CAMK2D, KCTD1, SAMHD1, VAMP4, ARFIP2, RAD51, SRR, TUBA8, SLC6A4, CAPZA2, TBCA, OAT, F2, SCIN, WASF1, SLAIN2, ANG, VIL1, PEX5, VWF, COX7A2L, CAPZB, SNAP29, MGST1, TMSB4, RASA1, IGF1R, LRRC6, NAP1L1, CBR4, SNCA, NUDT21, ARPIN, COX19, ARPC1A, ASF1A, UBE2C, NLRC4, NDUFAF6, MICU1, CLU, BAX, LMO4, RAC1, CAT, WASL, ARPC5, FIS1, NAPB, CYBA, DRC1, LONP1, ARPC3, PDCL, IL5, H4, HIST1H1A, ATP6V0A2, PSMD11, RDX, ARL6, MIS18A, FAS, KCTD5, SAR1A, MIF</i>
Regulation of RNA Splicing (GO:0043484)	33 (13)	4.31	<i>SRSF1, SF3A1, RBFOX1, CLK3, SNW1, HNRNP, SNRNP70, PIK3R1, RBM22, CELF3, RBFOX2, PRPF19, NSRP1</i>
Actin Filament Polymerization (GO:0030041)	48 (15)	4.29	<i>TWF1, ARFIP2, CAPZA2, SCIN, WASF1, ANG, VIL1, CAPZB, TMSB4, RASA1, ARPIN, ARPC1A, RAC1, WASL, ARPC5</i>

Table 2 (continued)

Gene Sets ^a	# Genes (# LEG) ^b	NES ^c	Leading Edge Genes ^d
Actin Polymerization or Depolymerization (GO:0008154)	58 (17)	4.27	<i>TWF1, MICAL3, ARFIP2, CAPZA2, SCIN, WASF1, ANG, VIL1, CAPZB, TMSB4, RASA1, ARPIN, ARPC1A, RAC1, WASL, ARPCS, DSTN</i>
Cilium Morphogenesis (GO:0060271)	46 (8)	4.24	<i>CEP290, ASAP1, IFT46, IFT43, TLL1, C7H5ORF30, CEP41, LRRC6</i>
Protein Complex Biogenesis (GO:0070271)	306 (73)	4.23	<i>TWF1, TUBB4B, HES1, H1FOO, CENPH, CHMP4A, DPAGT1, CAMK2D, KCTD1, SAMHD1, VAMP4, ARFIP2, RAD51, SRR, TUBA8, SLC6A4, CAPZA2, TBCA, OAT, F2, SCIN, WASF1, SLAIN2, ANG, VIL1, PEX5, VWF, COX7A2L, CAPZB, SNAP29, MGST1, TMSB4, RASA1, IGF1R, LRRC6, NAP1L1, CBR4, SNCA, NUDT21, ARPIN, COX19, ARPC1A, ASF1A, UBE2C, NLRC4, NDUFAF6, MICU1, CLU, BAX, LMO4, RAC1, CAT, WASL, ARPCS, FIS1, NAPB, CYBA, DRC1, LONP1, ARPC3, PDCL, ILS, H4, HIST1H1A, ATP6VOA2, PSMD11, RDX, ARL6, MIS18A, FAS, KCTD5, SAR1A, MIF</i>
Macromolecular Complex Binding (GO:0044877)	297 (97)	4.13	<i>PAFAH1B1, RAP1A, H1FOO, PPARGC1A, CENPH, GABARAPL1, MMP13, MUM1, SLC6A4, YWHAE, PLS1, MYO10, F2, STRN3, COLEC12, SCIN, WASF1, SEC61A2, MYB, HMGB4, TDRD3, GNAO1, VIL1, VWF, EIF6, PRKCB, CAPZB, GNAI1, TFAM, POLR3A, IGF1R, SEC61A1, DNAJC2, HMGNA4, PIK3R1, USH1C, PDGFA, MEF2A, ARPC1A, ASF1A, PRIMPOL, LMO2, MYO1D, NANOG, MTA3, CCNT1, TNNC1, MTM1, KIF2B, MSR1, MEN1, KDM8, ZNHIT1, PPIH, SCARB1, SBDS, NR5A1, CDC5L, SPARC, PLAC8, ARPC3, MSH2, SKA1, RNF169, FBLN5, RNF20, HIST1H1A, THBS4, GNAT3, ITGB6, PLK1, BAP1, RBPJ, ANKRD54, CETN1, FOXP1, KCTD5, MEF2C, FSCN2, UHRF1, MCMBP, HMGNA3, GNB1, ERMN, GNAT1, REEP4, MAP1LC3B, ASPN, CITED2, AAK1, TMOD1, ITGB1, ARPC1B, URI1, PPARG, SPDL1, MYL12A</i>
Cellular Protein Complex Disassembly (GO:0032984)	40 (13)	4.06	<i>TWF1, MICAL3, CCNB1, CAPZA2, SCIN, VIL1, CAPZB, KIF2B, NAPB, STMN1, DSTN, SBDS, MTRF1L</i>
Cellular Macromolecular Complex Assembly (GO:0006461)	245 (55)	4.06	<i>TWF1, SRSF1, TUBB4B, CNOT7, SF3A1, H1FOO, CENPH, VAMP4, ARFIP2, TUBA8, CAPZA2, TBCA, F2, SCIN, WASF1, SLAIN2, ANG, VIL1, STRAP, EIF6, COX7A2L, CAPZB, SNAP29, TMSB4, RASA1, GEMIN8, LRRC6, NAP1L1, SNCA, ARPIN, COX19, ARPC1A, ASF1A, UBE2C, NDUFAF6, RAC1, WASL, ARPCS, NAPB, PRPF19, CYBA, SBDS, DRC1, LONP1, ARPC3, TARBP2, H4, HIST1H1A, ATP6VOA2, PSMD11, RDX, ARL6, MIS18A, LUC7L3, SAR1A</i>
Regulation of Actin Filament Polymerization (GO:0030833)	44 (6)	4.03	<i>TWF1, CAPZA2, SCIN, VIL1, CAPZB, TMSB4</i>

^a Accession code for each gene set is in parentheses: Reactome – R; Gene Ontology – GO. ^b Total number of genes included in a gene set, with the number of leading edge genes (LEG) listed in parentheses. ^c Normalized enrichment score (NES) for each gene set, calculated by Kolmogorov-Smirnov-like statistics. ^d Leading edge gene identified by each gene set, listed in order of significance and identified by gene symbols as listed in the National Center for Biotechnology Information gene database (<https://www.ncbi.nlm.nih.gov/gene/>; accessed: 9 January, 2022)

to create the custom assay, the two GBS SNP in the locus on BTA1 at 84 Mb were in LD with only two of the original SNP ($D' > 0.75$). The other significant GBS SNP on BTA1 at 146 Mb was in LD with one of the two original SNPs ($D' = 0.80$), while the GBS SNP on BTA20 at 46 Mb was in high LD with both of the original SNP used to create the custom assay. The high level of LD with the original markers for the loci on BTA1 @146 Mb and BTA20 at 46 Mb suggests that progress can be made using the SNP on the currently available BeadChip. However, it should be noted that the two GBS SNP from the locus on BTA1 @84 Mb are only in LD with two of the nine original SNP indicating that the new GBS markers would be more useful for selection of fertility in cattle.

Loci associated with fertility in this study were compared to loci reported by Kiser et al. (2019a) [13] using only SNPs on the Illumina BovineHD BeadChip to determine if the GBS markers were helpful in further defining the validated CCR1 and TBRD loci. None of the three loci from the current studies were associated with fertility in Kiser et al. 2019a [13], meaning they were associated with cow fertility for the first time within the current

analyses and could provide functional targets for further analysis to identify putative causal mutations. Originally, these loci were identified in two or more of the external fertility analyses investigated [15, 20, 30, 42] and associated with heifer fertility instead of cow fertility.

Discussion

Most of the enriched gene sets have functions that fit within the organization of cellular components. These gene sets contain genes that are expressed within macromolecular complexes of the cell [57]. Gene Ontology's cellular component designation characterizes genes whose molecular function occurs within a cellular compartment of the cell such as the plasma membrane or the cytoskeleton. The three most identified leading edge genes for CCR1 and TBRD (villin 1 (*VIL1*), *RAC1*, and *TWF1*) were present within the gene sets in the organization of cellular components as they control different aspects of the cytoskeleton [58–60]. *VIL1* is often expressed on cell borders to help maintain their shape and their adhesion with other cells [61]. and its expression is also used as a marker for cells that have undergone

Table 3 Gene set enrichment analysis– single nucleotide polymorphism results for number of breeding services required to conceive (TBRD) in Holstein cows

Gene Sets ^a	# Genes (# LEG) ^b	NES ^c	Leading Edge Genes ^d
Recruitment of mitotic centrosome proteins and complexes (R-HSA-380270)	62 (13)	4.327	<i>TUBB4B, CEP290, PAFAH1B1, YWHAE, CEP250, CCNB1, CDK1, CENPJ, PLK1, CSNK1D, SDCCAG8, CLASP1, PRKAR2B</i>
Loss of NLP from mitotic centrosomes (R-HSA-380259)	55 (12)	4.293	<i>TUBB4B, CEP290, PAFAH1B1, YWHAE, CEP250, CDK1, CENPJ, PLK1, CSNK1D, SDCCAG8, CLASP1, PRKAR2B</i>
Cellular component assembly (GO:0022607)	515 (173)	4.08	<i>NDOR1, TUBB4B, TWF1, SF3A1, SRSF1, RAP1A, HES1, SAMHD1, IFT46, H1FOO, CAMK2D, KCTD1, SNAP29, CNOT7, CENPH, SERPINF2, SLC6A4, CHMP4A, EIF6, DPAGT1, CAPZA2, TMSB4, COX7A2L, WASF1, TUBA8, NRXN3, UBE2C, ASF1A, GABARAPL1, ANG, PRPF19, MYO10, NUDT21, TMEM231, FECH, MEF2A, SLAIN2, VAMP4, RASA1, ATPAF2, CSRP3, NDUFAF6, GTPBP8, MGST1, RAC2, TMEM216, CSNK1D, RAD51, SRR, CYBA, MAP1LC3A, PEX5, CRTC2, ATP6V0D1, ARPC3, OGFOD1, ITGB1BP1, ARPC5L, TLL1, LSM14A, KCTD19, FBLN5, COX19, VIL1, TBCA, IL5, LUC7L3, SNAP25, KCTD5, ATL1, PDCL, STRAP, PSMD11, BIRC5, PPP1R16B, VWF, ARFIP2, SRPX2, GAP43, LMO4, NAPB, KCNB2, RAP1B, CAV1, CDC42EP2, SHMT1, SLU7, MPP7, CAPZB, ATP6V0A2, GPM6A, CEP41, MEF2C, F2, MIS18A, IGF1R, UPK1A, MAPT, LCMT1, PRKACA, PSMG2, PSMD9, RAPGEF2, ARPIN, IFT20, CAT, EIF3G, CAPN3, EMP2, CLSTN3, FOXP1, NDUFS4, DRC1, KCTD18, SPICE1, ACTG1, ATG3, UBE2S, PICK1, APP, TBCD, GRB7, WASL, CCDC151, TBC1D7, ATP6V1D, FIS1, ARPC1A, SNRPE, KIT, CAND1, S100A10, PAK1, TMEM138, DDX39B, CLGN, TCAP, RAC1, APOA1, NAP1L1, RDX, HAUS1, NAP1L4, COX15, SAR1A, COG4, SCO2, MCIDAS, F2RL1, MMS19, DBNL, USO1, ATG5, TUBB2B, CHAF1A, TARBP2, SFRP1, DCXR, COL17A1, LONP1, NASP, KCNJ2, SNCA, SLC39A12, BRX1, CFL2, NDUFAF4, PRKAR1A, MFAP4, SNRPF, VDACC2, SLC2A1, GEMIN8</i>
Regulation of RNA splicing (GO:0043484)	33 (9)	4.062	<i>SF3A1, SRSF1, RBFOX1, HNRNPF, PRPF19, PIK3R1, CLK3, NSRP1, SNW1</i>

^a Accession code for each gene set is in parentheses: Reactome– R; Gene Ontology– GO. ^b Total number of genes included in a gene set, with the number of leading edge genes (LEG) listed in parentheses. ^c Normalized enrichment score (NES) for each gene set, calculated by Kolmogorov-Smirnov-like statistics. ^d Leading edge gene identified by each gene set, listed in order of significance and identified by gene symbols as listed in the National Center for Biotechnology Information gene database (<https://www.ncbi.nlm.nih.gov/gene/>; accessed: 9 January, 2022)

the epithelial-to-mesenchymal transition, which is important for normal female reproductive organ function [62]. Similarly, *RAC1* is involved with uterine cell remodeling. A deficiency of *RAC1* expression in mice during the peri-implantation period impairs the function of the luminal epithelia reducing uterine receptivity [63]. Additional, several cattle studies have identified associations between several p21 activated kinases genes, which are regulated by *RAC1*, and fertility traits such as conception rate and pregnancy rate in Holsteins [25, 64]. Twinfilin Actin Binding Protein 1 is a highly conserved actin binding protein that regulates the assembly of the cytoskeleton through the binding of actin monomers [65]. While its function in cattle fertility is not characterized, in mice *TWF1* was the most common isoform of twinfilin in embryos as well as in non-muscle cell types within adults [65]. In general, the cytoskeleton plays an important role in all cells and is considered to have three overarching functions: (1) organizing the contents within a cell, (2) enabling the cell to move and change shape, and (3) connecting cells to their external environments [66].

Prior to blastocyst implantation, changes to the actin cytoskeleton of uterine epithelial cells can result in a loss of cell polarity [67, 68]. The loss of cell polarity can result in actin cytoskeleton disorganization leading to abnormal endometrial receptivity and implantation failure.

In women, this failure influences the plasma membrane transformation period during the mid-secretory phase of the menstrual cycle [69]. In mice, cytoskeleton disorganization negatively impacts embryo development prior to implantation [70].

The second largest functional category of enriched gene sets is related to mitosis and the cell cycle. Mitotic cell division is an essential process for sustaining life. In the uterus, successful regulation of mitosis is important for fertility for the developing uterus, placenta, and embryo. Dysregulation of mitotic cell division leads to defects in uterine receptivity through alterations to stromal proliferation [71, 72] which could play a role in recurrent implantation failure in women [73]. During mitosis, centrosomes play a crucial role in the formation of the spindle assembly [74]. During reproduction, centrosomes are susceptible to rearrangements and changes to their complex structure can result in aneuploidy, mis-segregation of chromosomes, and destabilization of chromosomal integrity [75, 76]. Issues with centrosomes and their components can also result in oocyte abnormalities. Wang et al. (2001) [76] found that a main indicator of oocyte quality and thus chances of successful fertilization was centrosome integrity and its impact on spindle integrity.

Three gene sets (Neurotrophin Signaling Pathway, Notch Signaling Pathway, and Signaling by Neurotrophin

Table 4 List of gene sets enriched for both cow conception rate at first service (CCR1) and the number of breeding services required for conception (TBRD)

Functional Category	Gene set ^a	NES ^b		
		CCR1	TBRD	
<i>Cellular Component Organization</i>	Actin cytoskeleton organization (GO:0030036)	3.055	3.32	
	Actin filament based process (GO:0030029)	3.101	3.245	
	Cellular component assembly (GO:0022607)	3.34	4.08	
	Cellular component biogenesis (GO:0044085)	3.261	3.626	
	Cellular localization (GO:0051641)	3.765	3.083	
	Cellular macromolecular complex assembly (GO:0065003)	4.055	3.434	
	Cellular macromolecule localization (GO:0070727)	3.031	3.146	
	Cellular protein complex assembly (GO:0065003)	3.932	3.35	
	Cytoskeleton organization (GO:0007010)	3.331	3.183	
	Gap junction (K - hsa04540)	3.966	3.5	
	Macromolecular complex assembly (GO:0065003)	4.472	3.967	
	Macromolecular complex binding (GO:0044877)	4.131	3.348	
	Macromolecular complex subunit organization (GO:0043933)	3.261	3.785	
	Macromolecule localization (GO:0033036)	3.051	3.057	
	Organelle localization (GO:0051640)	3.509	3.15	
	Protein complex assembly (GO:0065003)	4.384	3.633	
	Protein complex biogenesis (GO:0070271)	4.234	3.488	
	<i>Mitotic Cell Cycle</i>	Cell cycle (R-HSA-1640170)	4.869	3.414
		Cell Cycle, Mitotic (R-HSA-69278)	5.077	3.831
Cell division (GO:0051301)		3.092	3.649	
Loss of NLP from mitotic centrosomes (R-HSA-380259)		4.676	4.293	
Mitotic G2-G2/M phases (R-HSA-453274)		4.744	3.523	
Recruitment of mitotic centrosome proteins and complexes (R-HSA-380270)		4.769	4.327	
<i>RNA & mRNA Metabolic Processes</i>	Regulation of mRNA metabolic process (GO:1903311)	3.423	3.318	
	Regulation of mRNA processing (GO:0050684)	3.411	3.222	
	Regulation of RNA splicing (GO:0043484)	4.311	4.062	
<i>Signaling Pathways</i>	Neurotrophin signaling pathway (K - hsa04722)	3.749	3.293	
	Notch signaling pathway (P00045)	3.139	3.629	
	Signaling by NGF (R-HSA-187037)	4.423	3.208	
<i>Other</i>	Cilium morphogenesis (GO:0060271)	4.236	3.237	
	Positive regulation of lipid biosynthetic process (GO:0046889)	3.281	3.112	
	Protein polymerization (GO:0051258)	4.806	3.422	
	Retinol metabolism (K - map00830)	3.503	3.537	
	Transmembrane transport of small molecules (R-HSA-382551)	3.43	3.23	

^a Accession code for each gene set is in parentheses: Reactome– R; Gene Ontology– GO; Protein Analysis Through Evolutionary Relationships– P; Kyoto Encyclopedia for Genes and Genomes - K. ^b Normalized enrichment score (NES) for each gene set for CCR1 and TBRD

Table 5 Single nucleotide polymorphisms associated with cow fertility using GBS variants

BTA (Mb) ^a	# SNP in Locus	Phenotype ^b	Model ^c	GBS P-Value	Region Feature(s) ^d
1 (83.7)	2	CCR1	Add	0.04*	ENSBTAG00000032217
		TBRD	Add	2.10×10^{-11}	
1 (146.1)	1	TBRD	Add	2.23×10^{-08}	DIP2A
20 (46.0)	1	TBRD	Add	1.82×10^{-06}	nsv810323

^a Chromosome location of the locus followed by the location of the locus in megabases (Mb) as denoted in the ARS-UCD 1.2 reference genome assembly (https://www.animalgenome.org/repository/cattle/UMC_bovine_coordinates/; accessed 5 September 2023). ^b Cow conception rate to first service (CCR1); Number of breeding services required to achieve conception (TBRD). ^c Inheritance model: additive (Add), dominant (Dom), recessive (Rec). ^d Genomic feature that the locus/SNP are located within: gene, regulatory regions, copy number variants (CNVs), conserved regions. *FDR < 0.05

Growth Factors [NGF]) in the signaling pathway category are involved in cellular proliferation and differentiation. Previous studies have found that NGF can increase cell proliferation through Notch signaling [77], and that the inhibition of Notch Signaling can result in the downregulation of NGF during acute inflammatory responses [78]. In addition to their nervous system roles, NGF and Notch Signaling function in ovulation [79] and pregnancy establishment [80], respectively. After systemic injection with NGF, luteinizing hormone and progesterone increase to aid in ovulation as well as establish and maintain pregnancy in llamas [79] and dairy cattle [81]. This increased expression of fertility hormones is suggested to increase uterine receptivity and reproductive efficiency. Notch Signaling has a broad role in female fertility from prenatal development, the estrus cycle, implantation and pregnancy [80]. Within the female reproductive tract, Notch Signaling activation is influenced by progesterone levels and has a role in pre-implantation angiogenesis by influencing the vasculature of the endothelial cells [80, 82]. Additionally, inhibition of Notch Signaling within the endometrium has been linked to repeated implantation failure, endometriosis, and polycystic ovary syndrome which cause infertility in women [83].

Gene sets enriched for CCR1 and TBRD that are related to RNA and mRNA (Table 4) could have several different influences on uterine receptivity and pregnancy maintenance. For example, some RNA and mRNAs are known to influence uterine receptivity during implantation through the regulation of Wnt signaling [84]. In humans, a meta-analysis on genes that were differentially expressed during embryo implantation and/or the endometrium identified 39 mRNA genes as uterine receptivity markers for uterine receptivity [85]. These genes have functions related to the immune response, exosomes, and the complement cascade pathway during successful pregnancies [85]. Alternative splicing of RNA and mRNA can also influence uterine receptivity. In mice, alternative splicing of epithelial splicing regulatory protein 1 (*Esrp1*) is associated with fertility, whereas *Esrp1* knockout female mice are infertile. Female *Esrp1*^{-/-} mice present with smaller ovaries and have impaired ovulation [86]. In humans, alternative splicing helps regulate gene expression within the myometrium to maintain quiescence throughout gestation [87]. Proteins involved in alternative splicing regulation like SF2/ASF and hnRNPA1 upregulate the production of certain GTP-binding protein isoforms during pregnancy leading to elevated cAMP levels and myometrial quiescence [88–89]. Alterations to RNA and mRNA processing and splicing have the potential to alter uterine receptivity and ultimately pregnancy success.

The association analyses identified 3 loci associated with CCR ($P < 1 \times 10^{-5}$) from the 100 loci GBS panel.

Two of the associated loci contained genes. Disco Interacting Protein 2 Homolog A (*DIP2A*) has potential functions in uterine receptivity while the function of *ENSBTAG00000032217* is currently uncharacterized. However, *ENSBTAG00000032217* is an ortholog to L-lactate dehydrogenase which is an enzyme that catalyzes the conversion of pyruvate to lactate as it converts NAD⁺ to NADH to produce energy in nearly all living cells. Previous work in mice has shown that *DIP2A* is expressed in the endometrium [90], and in bovine conceptuses during pregnancy recognition [91]. The *DIP2A* protein is also a receptor for follistatin which is essential for uterine receptivity [92, 93]. A successful pregnancy hinges on the ability of the embryo to implant into the endometrium. Further functional analyses are needed to determine how and if the variants identified in this study might influence uterine receptivity, implantation, and ultimately conception rate in dairy cows.

The observed heritability estimates from this study were moderate (CCR1 0.19 ± 0.05 ; TBRD 0.18 ± 0.05). When compared to the original heritability estimates calculated by Kiser et al. (2019a) [13] based on genome wide association analysis using 625,093 SNPs (0.56 ± 0.06 for CCR1 and 0.42 ± 0.07 for TBRD), the estimates from the current study were lower but similar to the heritability reported by others [94–97]. However, the current study heritability estimates confirm that significant improvement in fertility can be achieved by selecting for fertility traits. As one might expect, when the genetic correlation between CCR1 and TBRD was calculated it results in a genetic correlation of -1. This indicates that as the number of breedings needed to result in pregnancy increases, the conception rate to the first service decreases. Identifying variants and genes that are associated with both traits will help improve overall fertility within the industry.

The gene-sets, leading edge genes from the GSEA-SNP, the positional candidate genes, and the putative functional variants from the association study have potential roles in uterine receptivity and embryo implantation. Of the two loci that harbor positional candidate genes, one missense variant was located within the currently uncharacterized *ENSBTAG00000032217*, while two synonymous mutations were found within *DIP2A* and *ENSBTAG00000032217*. While the one missense variant is not thought to impact the function of the *ENSBTAG00000032217* protein, additional investigation is needed to confirm this due to the low confidence warning in the SIFT prediction.

Exploring putative causal mutations by genotyping through sequencing aided in defining the loci associated with fertility in dairy cattle. These results have practical implications in genomic selection as they could be incorporated into current commercial genotyping panels to

facilitate genomic predicted transmitting abilities. This improvement in fertility would help alleviate financial losses dairy producers incur from failed pregnancies and increase their efficiency and profitability.

Abbreviations

AI	Artificial insemination
CCR	Cow conception rate
CCR1	Cow conception rate to first service
EMMAX	Efficient mixed-model association expedited
FDR	False discovery rate
GBLUP	Genomic best linear unbiased predictor
GBS	Genotype by sequencing
GSEA-SNP	Gene-set enrichment analysis with SNP data
LEG	Leading edge gene
NES	Normalized enrichment score
NGF	Neurotrophin growth factors
SNP	Single nucleotide polymorphism
SVS	SNP variation suite
TBRD	Number of times bred to conception
VCF	Variant call format

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11364-9>.

Supplementary Material 1: Supplemental_Tables.xlsx: Additional file 1 (.xlsx): **Table S1.** Milk yield data. **Table S2.** Gene sets associated with cow conception rate to the first breeding in Holstein cows. **Table S3.** Gene sets associated with number of breedings to conception in Holstein cows.

Acknowledgements

The authors would like to acknowledge the central Washington dairies for their assistance in collecting phenotypes and genotypes, as well as Neogene for their help designing the custom assay.

Author contributions

HN, TS, and CS were responsible for the conception and design of the study. HN, JK, MN, JGNM, GB and TS were involved in the acquisition of data. JK, AH, and CS were responsible for the analysis of the data. JK and HN were involved in the interpretation of the data. JK and HN were involved in the drafting the manuscript. JK, HN, JD, MN, JGNM, and AH were involved in revision of the manuscript. All authors read and approved the final manuscript.

Funding

This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2013-68004-20365 from the USDA National Institute of Food and Agriculture. The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Data availability

Data generated and/or analysed during the current study are available in the CattleQTLdb repository (https://www.animalgenome.org/cgi-bin/QTLdb/BT/pubtails?PUBMED_ID=31718557) and the GBS data is pending review and acceptance by USDA Ag Data Commons (<https://figshare.com/s/a748d0d75fbaf4303648>).

Declarations

Ethics approval and consent to participate

This study's procedures were conducted with the approval of Washington State University's Institutional Animal Care and Use Committee (#4295) and each dairy owner provided written agreement to participate in the study. This study is reported in accordance with ARRIVE guidelines (<https://arriveguidelines.org>).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Animal Sciences, Washington State University, Pullman, WA, USA

²Department of Veterinary Pathobiology, College of Veterinary Medicine, Texas A&M University, College Station, TX, USA

³Animal Genomics and Improvement Laboratory, United States Department of Agriculture, Beltsville, MD, USA

⁴Department of Animal and Food Sciences, Oklahoma State University, Stillwater, OK, USA

⁵Department of Animal, Veterinary and Food Sciences, University of Idaho, Caldwell, ID, USA

⁶College of Agriculture and Natural Resources, Michigan State University, East Lansing, MI, USA

⁷Division of Animal Sciences, University of Missouri, Columbia, MO, USA

Received: 28 October 2024 / Accepted: 14 February 2025

Published online: 21 February 2025

References

- USDA, Economic Research Service. Farm income and wealth statistics. 2023. <https://data.ers.usda.gov/reports.aspx?ID=17845>
- Statista. Milk - Worldwide. Retrieved October 15, 2024, from <https://www.statista.com/outlook/cmo/food/dairy-products-eggs/milk/worldwide>
- Fernandez-Novo A, Fargas O, Loste JM, Sebastian F, Perez-Villalobos N, Pesantez-Pacheco JL, et al. Pregnancy loss (28–110 days of pregnancy) in Holstein cows: a retrospective study. *Animals*. 2020;10(6):925.
- Council on Dairy Cattle Breeding. 2022. CDCB Conception rate in 2022 for all breeds of cows. <https://webconnect.usdcb.com/#/national-performance-metrics>. Accessed 13 September 2024.
- Lee J, Kim I. Pregnancy loss in dairy cows: the contributing factors, the effects on reproductive performance and the economic impact. *J Vet Sci*. 2007;8(3):283–8.
- Quick Stats. United States Department of Agriculture National Agriculture Statistics Service. <https://quickstats.nass.usda.gov/>. Accessed 27 July 2023.
- Crowe MA. Fertility in dairy cows– the conference in perspective. In *Fertility in Dairy Cows bridging the gaps*. Eds Royal MD, Friggens NC and Smith RF. UK: British Society of Animal Science / Cambridge University Press. 2007;156–60.
- Berry DP, Wall E, Pryce JE. Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal*. 2014;8(S1):105–21.
- VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD. Daughter pregnancy rate evaluation of cow fertility. 2002. https://aipl.arsusda.gov/reference/fertility/DPR_rpt.htm#announce. Accessed 3 July 2023.
- Norman HD, VanRaden RM, Fok G, Hutchison J. Conception rate evaluations for males and females. In: *Changes to evaluations system*. 2010. <https://aipl.arsusda.gov/reference/changes/eval1008.html>. Accessed 3 July 2023.
- Norman HD, Hutchison J. New trait: Early first calving. 2019. https://queries.usdcb.com/News/CDCB_Connection_Early_First_Calving_03_2019.pdf. Accessed 10 July 2023.
- Misztal I, Aguilar I, Lourenco D, Ma L, Steibel JP, Toro M. Emerging issues in genomic selection. *J Anim Sci*. 2021;99(6):skab092.
- Kiser JN, Clancey E, Moraes JGN, Dalton J, Burns GW, Spencer TE, Neiberghs HL. Identification of loci associated with conception rate in primiparous holstein cows. *BMC Genomics*. 2019a;20:840.
- Neiberghs HL, Seabury CM, Wojtowicz AJ, Wang Z, Scraggs E, Kiser JN, et al. Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. *BMC Genomics*. 2014;15:1164.
- Kiser JN, Keuter EM, Seabury CM, Neupane M, Moraes JGN, Dalton J, et al. Validation of 46 loci associated with female fertility traits in cattle. *BMC Genomics*. 2019b;20:576.
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. *Science*. 2002;296:2225–9.

17. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genome-wide association studies. *Am J Hum Genet.* 2007;81:1278–83. <https://doi.org/10.1086/522374>
18. Holden M, Deng S, Wojnowski L, Kulle B. GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics.* 2008;24:2784–5. <https://doi.org/10.1093/bioinformatics/btn516>
19. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics.* 2007;23:1294–6. <https://doi.org/10.1093/bioinformatics/btm108>
20. Höglund JK, Guldbbrandsten B, Su G, Thomsen B, Lund MS. Genome scan detects quantitative trait loci affecting female fertility traits in Danish and Swedish holstein cattle. *J Dairy Sci.* 2009;92(5):2136–43.
21. Huang W, Kirkpatrick BW, Rosa GJM, Khatib H. A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. *Anim Genet.* 2010;41(6):570–8.
22. Pryce JE, Bolormaa S, Chamberlain AJ, Bowman PJ, Savin K, Goddard ME, Hayes BJ. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. *J Dairy Sci.* 2010;93:3331–45.
23. Sahana G, Guldbbrandtsen B, Bendixen C, Lund MS. Genome-wide association mapping for female fertility traits in Danish and Swedish holstein cattle. *Anim Genet.* 2010;41(6):579–88.
24. Blaschek M, Kaya A, Zwald N, Memili E, Kirkpatrick BW. A whole-genome association analysis of noncompensatory fertility in Holstein bulls. *J Dairy Sci.* 2011;94(9):4695–9.
25. Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ Jr, Crooker BA, Van Tassel CP, Yang J, Wang S, Matukumalli L, Da Y. Genome-wide association analysis of thirty one production, health, and reproduction and body conformation traits in contemporary US Holstein cows. *BMC Genomics.* 2011;12:408.
26. Olsen HG, Hayes BJ, Kent MP, Nome T, Svendsen M, Larsgard AG, Lien S. Genome-wide association mapping in Norwegian red cattle identifies quantitative trait loci for fertility and milk production on BTA12. *Anim Genet.* 2011;42(5):466–74.
27. Cochran SD, Cole JB, Null DJ, Hansen PJ. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genet.* 2013;14:49.
28. Minozzi G, Nicolazzi EL, Stella A, Biffani S, Negrini R, Lazzari B, Ajmone-Marsan P, Williams JL. Genome wide analysis of fertility and production traits in Italian holstein cattle. *PLoS ONE.* 2013;8:e80219.
29. Minten MA, Bilby TR, Bruno RGS, Allen CC, Madsen CA, Wang Z, Sawyer JE, Tibary A, Neiberghs HL, Geary TW, Bauersachs S, Spencer TE. Effects of fertility on gene expression and function of the bovine endometrium. *PLoS ONE.* 2013;8:e69444.
30. Höglund JK, Sahana G, Guldbbrandtsen B, Lund MS. Validation of associations for female fertility traits in Nordic Holstein, Nordic Red, and Jersey dairy cattle. *BMC Genet.* 2014;15:8.
31. Akanno EC, Plastow G, Fitzsimmons C, Miller SP, Baron V, Ominski K, et al. Genome-wide association for heifer reproduction and calf performance traits in beef cattle. *Genome.* 2015;58(12):549–57.
32. Höglund JK, Buitenhuis B, Guldbbrandtsen B, Lund MS, Sahana G. Genome-wide association study for female fertility in Nordic Red cattle. *BMC Genet.* 2015;16:110.
33. Iso-Touru T, Sahana G, Guldbbrandtsen B, Lund MS, Vikki J. Genome-wide association analysis of milk yield traits in Nordic Red cattle using imputed whole genome sequence variants. *BMC Genet.* 2016;17:55.
34. Moore SG, Pryce JE, Hayes BJ, Chamberlain AJ, Kemper KE, Berry DP, McCabe M, Cormican P, Lonergan P, Fair T, Butler ST. Differentially expressed genes in endometrium and corpus luteum of Holstein cows selected for high and low fertility are enriched for sequence variants associated with fertility. *Biol Reprod.* 2016;94:19.
35. Nayeri S, Sargolzaei M, Abo-Ismael MK, May N, Miller SP, Schenkel F, Moore SS, Stothard P. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet.* 2016;17:75.
36. Ortega MS, Denicol AC, Cole JB, Null DJ, Hansen PJ. Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. *Anim Genet.* 2016;47(3):288–97.
37. Liu A, Wang Y, Sahana G, Zhang Q, Liu L, Lund MS, Su G. Genome-wide association studies for female fertility traits in Chinese and nordic holsteins. *Sci Rep.* 2017;7:8487.
38. Müller MP, Rothammer S, Seichter D, Russ I, Hinrichs D, Tetens J, Thaller G, Medugorac I. Genome-wide mapping of 10 calving and fertility traits in Holstein dairy cattle with special regard to chromosome 18. *J Dairy Sci.* 2017;100(3):1987–2006.
39. Neupane M, Geary TW, Kiser JN, Burns GW, Hansen PJ, Spencer TE, Neiberghs HL. Loci and pathways associated with uterine capacity for pregnancy and fertility in beef cattle. *PLoS ONE.* 2017;12(12):e0188997.
40. Ortega MS, Denicol AC, Cole JB, Null DJ, Taylor JF, Schnabel RD, Hansen PJ. Association of single nucleotide polymorphisms in candidate genes previously related to genetic variation in fertility with phenotypic measurements of reproductive function in Holstein cows. *J Dairy Sci.* 2017;100(5):3725–34.
41. Fonseca PAS, Id-Lahoucine S, Reverter A, Medrano JF, Fortes MS, Casellas J, et al. Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS ONE.* 2018;13:e0205295.
42. Galliou JM, Kiser JN, Oliver KF, Seabury CM, Moraes JGN, Burns GW, et al. Identification of loci and pathways associated with heifer conception rate in U.S. Holsteins. *Genes.* 2020;11(7):767.
43. Oliver KF, Geary TW, Kiser JN, Galliou JM, Van Emon ML, Seabury CM, et al. Loci associated with conception rate in crossbred beef heifers. *PLoS ONE.* 2020;15(4):e0230422.
44. Andrews S, FastQC. A quality control tool for high throughput sequence data. 2010. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
45. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal.* 2011;17:10–2. <https://doi.org/10.14806/ej.17.1.200>
46. Babraham Bioinformatics - Trim Galore. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 3 January 2023. <https://doi.org/10.5281/zenodo.7598955>
47. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357–9.
48. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10(2):giab008.
49. Rimmer A, Phan H, Mathieson I, Iqbal Z, Tqigg SRF, WGS500 Consortium. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* 2014;46(8):912–8.
50. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447:661–78.
51. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet.* 2010;42(4):348–54.
52. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E. Efficient control of population structure in model organism association mapping. *Genetics.* 2008;178(3):1709–23.
53. Taylor J. Implementation and accuracy of genomic selection. *Aquaculture.* 2014;420–421:58–14.
54. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;88:76–82.
55. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. *Bioinformatics.* 2012;28:2540–2.
56. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812–4.
57. Thomas PD. The Gene Ontology and the meaning of biological function. *Methods Mol Biol.* 2017;1446:15–24.
58. Bosco EE, Mulloy JC, Zheng Y. Rac1 GTPase: a rac of all trades. *Cell Mol Life Sci.* 2009;66(3):370–4.
59. Nakamura E, Satoh T, Iwakawa M, Nakawatari M, Oki A, Matsumoto K, et al. Villin 1, a diagnostic marker for endometrial adenocarcinoma with high grade nuclear atypia. *Cancer Biol Ther.* 2011;12(3):181–90.
60. Huo G, Wang Y, Chen J, Song Y, Zhang C, Guo H, et al. A pan-cancer analysis of the oncogenic role of twinfilin actin binding protein in human tumors. *Front Oncol.* 2021;11:692136.
61. Khurana S, George SP. Regulation of cell structure and function by actin-binding proteins: Villin's perspective. *FEBS Lett.* 2008;582(14):2128–39.
62. Bilyk O, Coathman M, Jewer M, Postovit L. Epithelial-to-mesenchymal transition in the female reproductive tract: from normal functioning to disease pathology. *Front Oncol.* 2017;7:145.
63. Tu Z, Wang Q, Cui T, Wang J, Ran H, Bao H, et al. Uterine RAC1 via Pak1-ERM signaling directs normal luminal epithelial integrity conducive to on-time embryo implantation in mice. *Cell Death Differ.* 2016;23:169–81.

64. Jiang J, Ma L, Prakapenka D, VanRaden PM, Cole JB, Da Y. A large-scale genome-wide association study in U.S. Holstein cattle. *Front Genet.* 2019;14:10.
65. Vartiainen MK, Sarkkinen EM, Matilainen T, Salminen M. Mammals have two twinfilin isoforms whose subcellular localizations and tissue distributions are differentially regulated. *J Biol Chem.* 2003;278(36):34347–55.
66. Fletcher DA, Mullins RD. Cell mechanics and the cytoskeleton. *Nature.* 2010;463(7280):485–92.
67. Luxford KA, Murphy CR. Cytoskeletal alterations in the microvilli of uterine epithelial cells during early pregnancy. *Acta Histochem.* 1989;87(2):131–6.
68. Jalali BM, Likso P, Andronowska A, Skarzynski DJ. Alterations in the distribution of actin and its binding proteins in the porcine endometrium during early pregnancy: possible role in epithelial remodeling and embryo adhesion. *Theriogenology.* 2018;116:17–27.
69. Whitby S, Zhou W, Dimitriadis E. Alterations in epithelial cell polarity during endometrial receptivity a systematic review. *Front Endocrinol.* 2020;11:596324.
70. Tan K, An L, Wang S, Wang Z, Zhang Z, Miao K, et al. Actin disorganization plays a vital role in impaired embryonic development of in vitro-produced mouse preimplantation embryos. *PLoS ONE.* 2015;10(6):e0130382.
71. Tsai C, Chou C, Yang C, Lai Y, Liang C, Chen C, Tsai T. Hurler deficiency in mice leads to embryo infertility caused by an implantation defect. *J Biol Chem.* 2008;283(39):26302–6.
72. Winuthayanon W, Hewitt SC, Korach KS. Uterine epithelial cell estrogen receptor alpha-dependent and -independent genomic profiles that underlie estrogen responses in mice. *Biol Reprod.* 2014;91(5):110.
73. Zhou X, Cao Y, Zhou M, Han M, Liu M, Hu Y, et al. Decreased CD44v3 expression impairs endometrial stromal cell proliferation and decidualization in women with recurrent implantation failure. *Reprod Biol Endocrinol.* 2022;20:170.
74. Barra V, Fachinetti D. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat Commun.* 2018;9:4340.
75. McKinley KL, Cheeseman IM. The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol.* 2016;17(1):16–29.
76. Wang WH, Meng L, Hackett RJ, Oldenbourg R, Keefe DL. The spindle observation and its relationship with fertilization after intracytoplasmic sperm injection in living human oocytes. *Fertil Steril.* 2001;75(2):348–53.
77. Park JC, Chang IB, Ahn JH, Kim JH, Song JH, Moon SM, Park Y. Nerve growth factor stimulates glioblastoma proliferation through Notch1 receptor signaling. *J Korean Neurosurg Soc.* 2018;61(4):441–9.
78. Yin J, Hu H, Li X, Xue M, Cheng W, Wang Y, et al. Inhibition of notch signaling pathway attenuates sympathetic hyperinnervation together with the augmentation of M2 macrophages in rats post-myocardial infarction. *Am J Physiol Cell Physiol.* 2016;310:C41–53.
79. Ratto MH, Huanca W, Singh J, Adams GP. Comparison of the effect of ovulation-inducing factor (OIF) in the seminal plasma of llamas, alpacas, and bulls. *Theriogenology.* 2006;66(5):1102–6.
80. Moldovan GE, Miele L, Fazleabas AT. Notch signaling in reproduction. *Trends Endocrinol Metab.* 2021;32(12):1044–57.
81. Gajardo G, Paiva L, Ulloa-Leal C, Valderrama X, López G, Carrasco A, et al. Administration of beta-nerve growth factor during preovulatory stage improves endocrine and luteal function in dairy heifers. *Animals.* 2023;13(6):1004.
82. Shawber CJ, Lin L, Gnarra M, Sauer MV, Papaioannou, Kitajewski JK, Douglas NC. Vascular notch proteins and notch signaling in the peri-implantation mouse uterus. *Vasc Cell.* 2015;7:9.
83. Amjadi F, Salehi E, Zandieh Z, Rashidi M, Taleahmad S, Javedani masrouf M, et al. Comparative evaluation of NOTCH signaling molecules in the endometrium of women with various gynecological diseases during the window of implantation. *Iran J Basic Med Sci.* 2019;22(4):426–31.
84. Shekibi M, Heng S, Nie G. MicroRNAs in the regulation of endometrial receptivity for embryo implantation. *Int J Mol Sci.* 2022;23(11):6210.
85. Altmäe S, Koel M, Vösa U, Adler P, Suhorutšenko M, Laisk-Podar T, et al. Meta-signature of human endometrial receptivity: a meta-analysis and validation study of transcriptomic biomarkers. *Nature.* 2017;7:10077.
86. Yu L, Zhang H, Guan Z, Qin D, Zhou J, Wu X. Alternative splicing in the germline: ESRP1 essential for female fertility. *Development.* 2021;148(2):e148e0202.
87. Tyson-Capper AJ. Alternative splicing: an important mechanism for myometrial gene regulation that can be manipulated to target specific genes associated with preterm labour. *BMC Pregnancy Childbirth.* 2007;7(Suppl 1):S13.
88. Europe-Finner GN, Phaneuf S, Tolkovsky AM, Watson SP, López Bernal A. Down-regulation of G alpha s in human myometrium in term and preterm labor: a mechanism for parturition. *J Clin Endocrinol Metab.* 1994;79(6):1835–9.
89. Pollard AJ, Krainer AR, Robson SC, Europe-Finner GN. Alternative splicing of the adenylyl cyclase stimulatory G-protein G alpha(s) is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNP1) and involves the use of an unusual TG 3'-splice site. *J Biol Chem.* 2002;277(18):15241–51.
90. Zhang L, Mabwi HA, Palange NJ, Jia R, Ma J, Bah FB, et al. Expression patterns and potential biological roles of Dip2a. *PLoS ONE.* 2015;10(11):e0143284.
91. Estepa IM, Tinning H, Vasconcelos EJR, Fernandez-Fuertes B, Sánchez JM, Burns GW, et al. Protein synthesis by day 16 bovine ceptuses during the time of maternal recognition of pregnancy. *Int J Mol Sci.* 2020;21(8):2870.
92. Ouchi N, Asaumi Y, Ohashi K, Higuchi A, Sono-Romanelli, Oshima Y, Walsh K. DIP2A functions as a FSTL1 receptor. *J Biol Chem.* 2010;285(10):7127–34.
93. Fullerton PT Jr, Monsivais D, Kommagani R, Matzuk MM. Follistatin is critical for mouse uterine receptivity and decidualization. *Proc Natl Acad Sci USA.* 2017;114(24):E4772–81.
94. Kuhn MT, Hutchison JL, Wiggans GR. Characterization of Holstein heifer fertility in the United States. *J Dairy Sci.* 2006;89:4907–20.
95. Tiezzi F, Maltecca C, Cecchinato A, Penasa M, Bittante G. Genetic parameters for fertility in dairy heifers and cows at different parties and relationships with production traits in first lactation. *J Dairy Sci.* 2012;95:7355–62.
96. Bowley FE, Green RE, Amer PR, Meier S. Novel approaches to genetic analysis of fertility traits in New Zealand dairy cattle. *J Dairy Sci.* 2015;98(3):2005–12.
97. Seabury CM, Smith JL, Wilson ML, Bhattarai E, Santos JEP, Chebel RC, et al. Genome-wide association and genomic prediction for a reproductive index summarizing fertility outcomes in U.S. Holsteins. *G3.* 2023;13(9):jkad043.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.