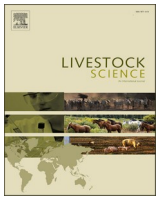
Livestock Science 253 (2021) 104731

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci



Single-step genome-wide association study uncovers known and novel candidate genomic regions for endocrine and classical fertility traits in Swedish Red and Holstein dairy cows

G.M. Tarekegn [a,b,*], E. Strandberg [a], S. Andonov [a], R. Båge [c,d], P. Ask-Gullstrand [a], E. Rius-Vilarrasa [d], J.M. Christensen [e], B. Berglund [a,*]

[a] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, P. O. Box 7023, SE-750 07 Uppsala, Sweden
[b] Department of Animal Production and Technology, School of Animal Sciences and Veterinary Medicine, Bahir Dar University, P. O. Box 79, Bahir Dar, Ethiopia
[c] Department of Clinical Sciences, Swedish University of Agricultural Sciences, P. O. Box 7054, SE-750 07 Uppsala, Sweden
[d] Växa Sverige, P.O. Box 30204, SE-104 25 Stockholm, Sweden
[e] Lattec, Slangerupsgade 69, 3400 Hillerod, Denmark

HIGHLIGHTS

- In this study, the novel endocrine and classical fertility traits of Swedish dairy herds were evaluated using in-line milk progesterone profiles in ssGWAS approach.
- The genomic regions associated with the endocrine and classical traits were identified.
- A total of 20 QTLs, of which 18 of them are novel, were detected. The novel QTL regions embedded candidate genes (*ATG7, GPI, GJA1, PPP3CA, ECT2, ARHGAP20, PHLDB1, CACNA1D, CCNE1, CDH13, PLD1, FBN2, KIF3A, FGF12, KCNMB2, MAN1A1, KCNN2, SMAD6, MAPK8IP1, PHF21A, LPXN, MMRN1, KCNIP4, NID2, PCDHGA8, GRIA1, PCDHGB4, PHLDB2, STXBP5L, PTPRR, SRGAP1, SNX27, SPTA1, S100A10, TBC1D20* and *ITCH*) that are associated with endocrine and classical fertility traits.
- The candidate QTL regions provide insights into the genetic basis of endocrine and classical fertility in SR and Holstein dairy breeds.

ARTICLE INFO

Keywords:
Dairy cow
Endocrine trait
Fertility
Folliculogenesis
Progesterone

ABSTRACT

In a study aiming to identify candidate genomic regions associated with endocrine and classical fertility traits in Swedish Red (SR) and Holstein cows, data on 3955 lactations in 1164 SR and 1672 Holstein cows were examined. The dataset comprised milk progesterone (P4) levels (n = 341,212) in 14 Swedish herds, automatically collected and analyzed in-line using the DeLaval Herd Navigator™. Endocrine traits studied were: days from calving to commencement of luteal activity (C-LA), first luteal phase length (LPL), length of first inter-luteal interval, length of first inter-ovulatory interval (IOI), luteal activity during the first 60 DIM, and proportion of samples with luteal activity during the first 60 DIM. Classical fertility traits based on insemination data were also investigated, such as days from calving to last insemination and calving interval. A total of 180 SR and 312 Holstein cows were genotyped with a low-density SNP chip and imputed to 50 K. Single-step genome-wide association (ssGWAS) was used to explore candidate genomic regions associated with fertility traits. A mixed linear single-trait animal model was fitted, considering season and parity as fixed effects and animal and permanent environment as random effects. The results revealed 990 and 415 SNPs above the threshold (-log (p-value) $\geq$4) for SR and Holstein cows, respectively. The breeds shared only eight SNPs significantly associated with fertility traits. Annotation analysis revealed 281 SNPs located in 241 genes. Functional enrichment analysis using DAVID tools reduced the number to 80 genes, which were mediated in various biological processes and KEGG pathways in multiple functions, including folliculogenesis, embryogenesis, uterine growth and development, immune response, and ovarian cysts. Of the 80 genes, 67 were associated with fertility traits in SR cows and 13 in Holstein. Most genes were associated with LPL and IOI in SR cows, but in Holstein the only association with an endocrine trait was with C-LA. Twenty QTL regions that embedded 40 genes were associated with fertility traits

* Corresponding authors at: Animal Breeding and Genetics, Swedish University of Agricultural Sciences: Sveriges lantbruksuniversitet, August Sodermans Vag 4 lgh 1102, 75649 Uppsala, Sweden.
E-mail addresses: getinet.mekuriaw@aau.edu.et (G.M. Tarekegn), britt.berglund@slu.se (B. Berglund).

https://doi.org/10.1016/j.livsci.2021.104731
Received 7 March 2021; Received in revised form 2 October 2021; Accepted 5 October 2021
Available online 8 October 2021

in both breeds. All the QTLs detected, except at BTA2 and BTA19 are novel QTL regions that were not reported previously. These novel QTL regions embedded the candidate genes that include ARHGAP20, PHLDB1, CACNA1D, ATG7, CCNE1, GPI, CDH13, ECT2, PLD1, FBN2, KIF3A, FGF12, KCNMB2, GJA1, MAN1A1, KCNN2, SMAD6, MAPK8IP1, PHF21A, LPXN, MMRN1, KCNIP4, NID2, PCDHGA8, GRIA1, PCDHGB4, PHLDB2, STXBP5L, PPP3CA, PTPRR, SRGAP1, SNX27, SPTA1, S100A10, TBC1D20 and ITCH. The candidate regions may help to improve genetic progress in female fertility if used in selection decisions. A challenge for future research is to determine why different regions seem relevant for different traits and breeds, and the practical implications for genomic selection.

1. Introduction

Genome-wide association studies (GWAS) is mainly determined by the heritability of the trait evaluated, the availability of phenotypic data linked to genotyped animals (Daetwyler et al., 2008), the number of genes influencing the trait (Berry et al., 2012), and the distribution of the allelic substitution effects. Classical fertility phenotypes (e.g. calving interval and number of inseminations per series) in dairy cows are partly influenced by management decisions and have low heritabilities, which has resulted in slow genetic progress (Berry et al., 2012). Some traits also require a long time to generate phenotypes, thus delaying selection. Moreover, the antagonistic relationship between production and fertility traits complicates genetic improvement of both traits (e.g., Roxström et al., 2001; Nyman et al., 2014). Thus, identifying novel fertility traits with higher heritability and utilizing genomic information that explains a large proportion of the genetic variation in fertility is crucial (Berry et al., 2012). Recently, there has been increasing interest in milk progesterone-based evaluation of cow fertility. Endocrine fertility traits are novel traits that directly reflect the physiology of the cow. These traits are not influenced by management decisions like voluntary waiting period, estrus detection, and efficiency of insemination, which may facilitate identification of genes that play a major role in controlling and regulating dairy cow fertility. There have been very few GWAS on endocrine fertility traits. To our knowledge, only Berry et al. (2012), Tenghe et al. (2016), and Nyman et al. (2019) have applied classical GWAS methods to find genomic regions related to endocrine fertility traits in Holstein-Friesian cows. Those studies revealed various genomic regions related to several endocrine fertility traits.

Single-step genome-wide association study (ssGWAS) is increasingly becoming the method of choice for studying associations between genotype and phenotype in livestock breeding, especially when there are many non-genotyped animals with phenotypes and known pedigree information. Whereas the classical GWAS has lower power for detection of quantitative trait loci (QTL) and precision for fine mapping (Wang et al., 2012). Similarly, its multiple-step procedure could cause loss of information, and heterogeneity through the different amounts of information in the original dataset could lead to bias (VanRaden et al., 2009). Moreover, the multi-step method is suboptimal despite its capacity to consider ungenotyped animals (Wang et al., 2012). As an alternative, ssGWAS increases both power and precision without increasing genotyping costs. This is because it takes advantage of phenotypes from other related animals using the traditional pedigree relationships with those genotyped, and converts the genomic estimated breeding values (GEBVs) to marker effects and weights (Wang et al., 2012). The flexibility of ssGWAS to apply mixed models allows for both simple and complex analyses involving multiple traits and confounding factors that include environmental, epigenetic, or maternal environmental effects (Wang et al., 2012). It integrates genomically derived relationships (G) with pedigree-based relationships (A) into a combined relationship matrix (H) and allows for genomic evaluation in a single step (Misztal et al., 2013). It is also suggested that the single-step genomic BLUP (ssGBLUP) method (Misztal et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010) can yield more accurate GEBVs for estimating single-nucleotide polymorphism (SNP) effects in populations with smaller numbers of genotyped animals compared to the classical approach (Oliveira et al., 2019). Again, this is because ssGWAS methods utilize GEBVs based on combined pedigree, genomic, and phenotypic information (Wang et al., 2014).

Previous GWAS studies on endocrine fertility traits (Berry et al., 2012; Tenghe et al., 2016; Nyman et al., 2019) used the classical GWAS approach, including information from genotyped animals only. Moreover, there is no GWAS report on Swedish Red cows. Thus the aim in the present study was to identify candidate genomic regions associated with endocrine and classical fertility traits in Swedish Red and Holstein dairy cattle breeds using ssGWAS.

2. Materials and methods

2.1. Animals and phenotypes

Data on a total of 3955 lactations in Swedish Red (SR) ($n = 1164$) and Holstein ($n = 1672$) cows were included in the study. The data comprised milk progesterone (P4) levels automatically collected and analyzed in-line with the Herd Navigator™ system (DeLaval International, Tumba, Sweden) ($n = 341,212$), and were collected from October 2015 to February 2019 in 14 commercial herds in Sweden. Endocrine fertility traits were defined from the P4 profile data. Pedigree, calving, and insemination data were obtained from Växa Sverige (Stockholm, Sweden), a consulting and service provider company mainly in dairy and beef in Sweden (https://svenskkooperation.se/goda-affarer/vaxa-sverige/). The pedigree information was traced back 25 generations for 14,715 SR animals and 23 generations for 13,323 Holstein animals (Table 1). To define luteal activity, a P4 concentration of 5 ng/mL was used as a threshold. Data editing was performed as described in Tarekegn et al. (2019).

The endocrine traits studies include commencement of luteal activity (C-LA) that refers to interval from calving to commencement of luteal activity (a threshold $\geq$ 5 ng/mL was used as indication for luteal activity); first luteal phase length (LPL): the interval from the first day of elevated P4 level ($\geq$5 ng/mL) to the last consecutive day of elevated P4 level ($\geq$5 ng/mL); length of inter-luteal interval (ILI): that refers to the period of an-estrus between the demise of one corpus luteum and the rise of the next; length of first inter-ovulatory interval (IOI): the interval between P4 level rise from the corpus luteum of 1 estrus cycle and the P4 level rise from corpus luteum of the next estrus cycle; luteal activity during the first 60 DIM (LA60; categorical 0 or 1): defined as the presence (LA60 = 1) or absence (LA60 = 0) of luteal activity between 20 and

Table 1
Summary of number of lactations, animals with or without phenotypes and genotypes, and parents for Swedish Red and Holstein cows.

Item	Swedish Red	Holstein	Total
Lactations	1593	2362	3955
Animals with records	1164	1672	2836
Animals with genotypes and records	180	312	492
Animals with records or genotypes	1226	1682	2908
Number of animals without genotypes and with records	4	5	9
Parents without records or genotypes	14,715	13,323	28,038
Total number of animals	15,945	15,010	30,955

60 days in milk (DIM); and proportion of samples with luteal activity during the first 60 DIM (PLA). These definitions of the endocrine traits can be found in previous studies (Tenghe et al., 2015; Tarekegn et al., 2019). The classical traits included were: calving interval (CI), interval from calving to first insemination (CFI), interval from calving to last insemination (CLI), interval from first to last insemination (FLI), number of inseminations per series (NINS), and the mixed (endocrine and classical) trait interval from commencement of luteal activity to first service (CLAFS): refers to the interval from the initiation of luteal activity to first service. Data on all traits except LA60, PLA, and NINS were (natural) log-transformed before statistical analysis.

2.2. Genotyping and quality control

A total of 492 cows (180 SR and 312 Holstein) were genotyped according to the routines of the Nordic Cattle Genetic Evaluation (NAV) (NAV, 2021). Genotyping was done by Eurofins GenoScan (Aarhus, Denmark) with low density (7 K) panel (6910 markers) (Illumina, San Diego, CA, USA), and imputed to medium density (50 K SNP chip) (Illumina, San Diego, CA, USA), separately by breed The imputation was made using the software FImpute (Sargolzaei et al., 2014). The reference populations for SR and Holstein consist of proven bulls and females from Denmark, Sweden and Finland. Additionally, the reference groups also include foreign bulls that have Interbull proofs and where genotypes are exchanged with NAV. The genotype data was obtained from 9 of the 14 commercial Herd Navigator™ herds included. A call rate of >95% was applied at both animal and SNP level, and MAF>5% was used as the first quality control (QC) criterion for imputed data. After the QC assessment, the imputed markers (50 K) were reduced to 42,446 SNP markers. These markers were further reduced to 38,975 SNPs after checking for Mendelian conflicts and monomorphic and missing markers using postGSf90 (Aguilar et al., 2014), in preparation for ssGWAS (Table 1).

To assess the differentiation and genetic background of the two dairy breeds, analysis of molecular variance (AMOVA) and population structure were applied using the software ARLEQUIN v3.5.2 (Excoffier and Lischer, 2010) and STRUCTURE v.2.3.4 (Pritchard et al., 2000), respectively. The population genetic structure assumed was hypothetical population clusters (K) ranging between 2 and 5. Five runs of 20,000 Markov Chain Monte Carlo (MCMC) iterations were performed for each K-value after a burn-in of 10,000 iterations. The STRUCTURE output was further analyzed in STRUCTURE HARVESTER (Earl, 2012) to determine the optimal number of genetic groups (ΔK) (Evanno et al., 2005) present in the dataset.

2.3. Variance components and GWAS

The variance components were estimated separately for each breed, using a mixed linear single-trait animal model. Animal and permanent environmental effects were considered as random effects, and the variance components were estimated using the remlf90 package in the blupf90 family (Misztal et al., 2002). The following single-trait general model was fitted:

$$\boldsymbol{Y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Zpe} + \boldsymbol{e}$$

where: Y = vector of phenotypic observations

X = incidence matrix linking phenotypes to fixed effects
b = vector of fixed effects (season: December-February, March-May, June-August, September-November; parity: 1, 2, 3, 4, 5+)
Z = incidence matrix relating animals and permanent environmental effects to phenotypes
a = vector of individual animal effects, $\boldsymbol{a} \sim N(0, \mathbf{H}\sigma_a^2)$
$\boldsymbol{pe}$ = vector of permanent environmental effects, $\boldsymbol{pe} \sim N(0, \boldsymbol{I}_{\boldsymbol{pe}}\sigma_{pe}^2)$
e = vector of residuals (residual variance) with $\boldsymbol{e} \sim N(0, \boldsymbol{I}\sigma_e^2)$,

where $\boldsymbol{a}$, $\boldsymbol{pe}$, and e are assumed to be uncorrelated; σ_a^2 and σ_e^2 are additive genetic and residual variance, respectively; H is a matrix combining pedigree and genomic relationships (Aguilar et al., 2010); and I_{pe} and $\boldsymbol{I}$ are identity matrices.

(Co)variance components for single-step genomic prediction were estimated using procedures implemented in REMLf90 package of blupf90 (Misztal et al., 2002, 2013). The pedigree of the animals was combined with their genomic information to estimate GEBV for all animals in the pedigree. The inverse of H combining the additive and genomic relationship matrices, A and G, respectively, was calculated as:

$$\boldsymbol{H}^{-1} = \boldsymbol{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{G}^{-1} - \boldsymbol{A}_{22}^{-1} \end{bmatrix}$$

where: $\boldsymbol{A}^{-1}$ = inverse of relationship matrix based on pedigree information

$\boldsymbol{G}^{-1}$ = inverse of genomic relationship matrix (VanRaden et al., 2008)
$\boldsymbol{A}_{22}^{-1}$ = inverse of pedigree-based relationship matrix for genotyped animals

The H matrix was built scaling G based on $\boldsymbol{A}_{22}$ considering that the average of the diagonal of G should be equal to the average of the diagonal of $\boldsymbol{A}_{22}$, and average off-diagonal values of G should be equal to average off-diagonal values of $\boldsymbol{A}_{22}$ (Vitezica et al., 2011; Medeiros de Oliveira Silva et al., 2017).

The breeding values of each animal were estimated by the ssGBLUP method using the blupf90 software (Aguilar et al., 2011). To obtain the SNP effects for all fertility traits, they were back-solved from the GEBVs as implemented by the program postGSf90 (version 1.46) (Wang et al., 2012; Aguilar et al., 2014). The SNP effects were derived by decomposing the animal effects into those for genotyped ($\boldsymbol{a}_g$) and ungenotyped ($\boldsymbol{a}_n$) animals (Wang et al., 2012). Hence, the animal effects of genotyped animals were taken as a function of SNP effects, defined as:

$$\boldsymbol{a}_g = \boldsymbol{Zu}$$

where: $Z = a$ matrix that relates genotypes of each locus
u = vector of marker effects
Hence, variance of animal effects ($\boldsymbol{a}_g$) is:

$$\text{var}\ (\boldsymbol{a}_g) = \text{var}(\boldsymbol{Zu}) = \boldsymbol{ZDZ}'\sigma_u^2 = \boldsymbol{G}^*\sigma_a^2$$

where: D = diagonal matrix of weights for variances of markers ($\boldsymbol{D} = \boldsymbol{I}$ for GBLUP)
$\boldsymbol{\sigma}_u^2$ = additive genetic variance captured by each SNP when no weights are present
$\boldsymbol{G}^*$ = the weighted genomic relationship matrix.
The joint (co)variance of animal effects ($\boldsymbol{a}_g$) and SNP effects ($\boldsymbol{a}_n$) was evaluated as described in (Wang et al., 2012):

$$\text{Var}\begin{bmatrix} \boldsymbol{a}_g \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{ZDZ}' & \boldsymbol{ZD}' \\ \boldsymbol{DZ}' & \boldsymbol{D} \end{bmatrix}\sigma_u^2$$

$$\boldsymbol{G}^* = \frac{\text{var}(\boldsymbol{a}_g)}{\sigma_a^2} = \frac{\text{var}(\boldsymbol{Zu})}{\sigma_a^2} = \boldsymbol{ZDZ}'\frac{\sigma_u^2}{\sigma_a^2} = \boldsymbol{ZDZ}'\lambda;\ \lambda = \frac{\sigma_u^2}{\sigma_a^2}$$
$$= \frac{1}{\sum_{i=1}^{N} 2p_i(1-p_i)}$$

where: λ = variance ratio or normalizing constant (VanRaden et al., 2009)
N = number of SNPs
p_i = frequency of the second allele in the i^{th} SNP
According to Stranden and Garrick (2009):

$$\hat{u} = \frac{\sigma_u^2}{\sigma_a^2} \boldsymbol{DZ}' \boldsymbol{G}^{*-1} \hat{a}_g$$

Hence to predict the SNP effects using weighted relationship matrix $\boldsymbol{G}^*$(Wang et al., 2012):

$\hat{u} = \lambda \boldsymbol{DZ}' \boldsymbol{G}^{*-1} \hat{a}_g = \boldsymbol{DZ}' [\boldsymbol{ZDZ}']^{-1} \hat{a}_g$, and all weights were assumed to be equal to each other as 1. Estimates of SNP effects were used to estimate individual variance of each SNP effect as in Zhang et al. (2010):

$$\hat{\boldsymbol{\delta}}^2_{u,i} = \hat{u}_i^2 2\boldsymbol{p}_i(1-\boldsymbol{p}_i)$$

The estimation of D was performed with the algorithm as employed in Wang et al. (2012). In the present study, three iterations were employed to obtain the marker effects and the iterative procedures considering D for SNP effect estimation were performed according to the procedures described in Wang et al. (2012). The SNP p-value method was used to evaluate the SNP effects, as suggested in Aguilar et al. (2019), using the 20 adjacent SNPs as windows. Manhattan plots were generated using postGSf90 software (Aguilar et al., 2014).

2.4. Declaring significant association

To test significant association of the variants, a more liberal threshold, i.e. 10^{-4}, was applied for all fertility traits, as suggested by Teyssèdre et al. (2012). For all traits, all markers with -log (p-value) ≥4 were considered candidate genomic regions associated with the traits. Less stringent thresholds could help to establish the association of markers for complex traits that mainly have low heritability estimates. Indeed, we are aware of the low number of genotyped animals, and by keeping a liberal significance threshold, we aimed not to discard any potential QTL. The confirmation of QTLs would be in research on a larger number of genotyped animals.

2.5. Functional enrichment analysis of the candidate regions and detection of QTLs

Gene annotation was performed to identify candidate genes, using the bovine genome assembly (ARS-UCD 1.2; Rosen et al., 2020) as a reference genome embedded in Ensembl Genome Browser (http://www.ensembl.org/index.html). To classify the genes according to their biological pathways and molecular functions, gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analyses were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v.6.7 (http://david.abcc.ncifcrf.gov/). STRING Genomics 11.0 (Szklarczyk et al., 2019) was applied to significantly enriched candidate genes in DAVID tools, to replicate the functional enrichment analysis using protein-protein interaction (PPI) networks and GO terms encoded by the candidate genes. The candidate genes were significantly enriched in DAVID tools, and thus a less stringent confidence interval (PPI enrichment score >0.3) was applied using *Bos taurus* as background species and constructing the global network. STRING stipulates known PPI from curated databases or experiments. In addition, it predicts interacting genes based on gene neighborhoods, fusions, and co-occurrences, text mining in literature, co-expression, or protein homology. To define the QTL regions, at least two significantly enriched SNPs found in the 10 Mb window in a chromosome were used, as defined in Tenghe et al. (2016), and the SNPs which revealed functional enriched genes were used for QTL detection.

3. Results

3.1. Descriptive statistics and differentiation of the study populations

Descriptive statistics for the endocrine and classical fertility traits are summarized in Table 2. The fertility traits that include LPL ILI, IOI and FLI had also zero minimum value. This is because occasionally the hormone cycle of a cow may go up and down on same day for the first three traits and the first insemination can be effective to the cow to conceive for the latter trait. The estimates in the descriptive statistics showed that the average intervals of all traits for Holstein were longer (and higher for PLA and NINS) than the estimates for SR cows, with the exception of C-LA. The results from differentiation and structure of the two breeds indicated that SR and Holstein cows are highly differentiated from each other (Fig. 1). AMOVA demonstrated 9.2% variation (data not shown) between breeds in the sample and indicated that both breeds have a pure genetic background (SR: 98.4%; Holstein: 87.9%) with the Holsteins having ~10% additional genetic background (Fig. 1a; the light blue genetic background at cluster 2). Similarly, PCA (PC1: 54.11%; PC2: 5.06%) clearly differentiated the two breeds (Fig. 1b). However, the variation within Holstein was larger compared to that within SR, which could be attributed to the ~10% additional genetic background in Holstein cows.

Table 2
Descriptive statistics of classical and endocrine fertility traits of Swedish Red (SR) and Holstein cows.

Trait[1]	SR			Holstein		
	N	Mean ± SD	Min/Max	N	Mean ± SD	Min/Max
C-LA (d)	1344	34.8 ± 16.6	20/208	2010	32.5 ± 16.4	20/234
LPL (d)	893	9.5 ± 12.8	0/216	1496	11.5 ± 14.7	0/252
ILI (d)	858	6.5 ± 6.0	0/79	1440	6.6 ± 6.1	0/82
IOI (d)	822	21.5 ± 12.2	0/219	1408	23.7 ± 14.6	1/212
CLAFS (d)	830	49.2 ± 29.1	4/257	1387	56.5 ± 32.8	5/403
LA60 (%)	1344	0.94±0.23	0/1	2010	0.94±0.23	0/1
PLA (%)	1504	0.59±0.19	0/1	2241	0.64±0.21	0/1
CI (d)	590	379.9 ± 48.1	297/734	882	385.9 ± 51.4	265/683
CFI (d)	1384	84.7 ± 30.0	36/375	2078	89.2 ± 32.2	15/438
CLI (d)	1384	138.9 ± 101.4	39/623	2078	150.3 ± 101.6	43/830
FLI (d)	1384	54.2 ± 98.9	0/518	2078	61.1 ± 97.5	0/731
NINS (N)	1397	1.9 ± 1.1	1/8	2140	2.2 ± 1.4	1/8

[1]C-LA = interval from calving to commencement of luteal activity; LPL = first luteal phase length; ILI = length of inter-luteal interval; IOI = length of first inter-ovulatory interval; CLAFS = interval from commencement of luteal activity to first service; SR = Swedish Red; LA60 = luteal activity during the first 60 DIM; PLA = proportion of samples with luteal activity during the first 60 DIM; CI = calving interval; CFI = interval from calving to first insemination; CLI = interval from calving to last insemination; FLI = interval from first to last insemination; NINS = number of inseminations per series; d = day.

3.2. SNPs significantly differentiated in ssGWAS

Estimates of additive genetic and residual variances are presented in Table 3, and significantly differentiated SNPs in Table 4. The ssGWAS revealed 1405 significantly associated SNPs at -log (p-value) ≥4 for across all 12 endocrine and classical fertility traits in SR and Holstein cows. These SNPs demonstrated genetic variances (GV) explained by each SNP ranging between 0.004 to 0.726 (27.9%: 0.004 < GV explained by the SNP < 0.05; 30.8%: 0.05 ≤ GV explained by the SNP < 0.10; 41.3%: 0.10 ≤ GV explained by the SNP ≤ 0.726; Supplementary Table S1). Of these significantly associated SNPs, only eight variants were shared by both breeds. The highest number of the SNPs was associated with LPL (409 SNPs: 406 SNPs in SR, 3 SNPs in Holstein) and IOI (278 SNPs: 259 SNPs in SR, 19 SNPs in Holstein) (Table 4). The top nine associated SNPs were for LPL (-log (p-value) = 10.42 to 24.35; GV explained by the SNP = 0.01 to 0.42) in SR cows (Table 5; Supplementary Table S1). Annotation of significantly differentiated variants was performed to identify candidate genomic regions. From the total of 1405 differentiated SNPs at -log (p-value) ≥4 for all traits (Fig. 2a-f), 281 SNPs were located in 241 genes, whereas the majority were found at

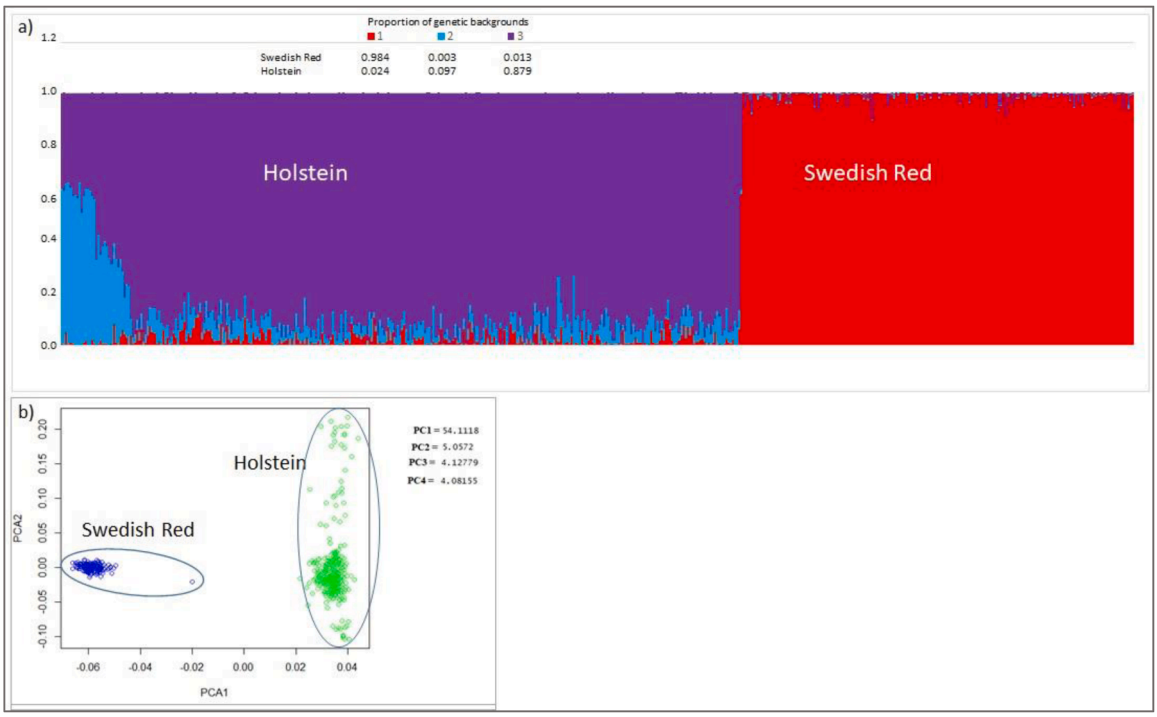


Fig. 1. Population structure (a) at $K = 3$ and principal component (b) analyses of Swedish Red and Holstein dairy cows.

Table 3
Estimates of additive genetic and residual variances of classical and endocrine fertility traits for Swedish Red (SR) and Holstein cows.

Trait[1]	SR		Holstein	
	Additive genetic variance	Residual variance	Additive genetic variance	Residual variance
C-LA (d)	0.01254	0.09106	0.02968	0.09701
LPL (d)	0.01194	0.3268	0.11430	0.3490
IOI (d)	0.007714	0.2027	0.0308	0.2359
LA60 (%)	0.00151	0.04033	0.005324	0.04575
PLA (%)	0.004771	0.02797	0.002622	0.03185
CLAFS (d)	0.02637	0.2732	0.05964	0.2214
CI (d)	0.001302	0.008631	0.001729	0.01175
CFI (d)	0.01136	0.07359	0.03047	0.1395
CLI (d)	0.006899	0.2653	0.0412	0.2229
FLI (d)	0.0032	0.6707	0.07079	0.5463
NINS (N)	0.03624	1.177	0.1297	1.687

1C-LA = interval from calving to commencement of luteal activity; LPL = first luteal phase length; IOI = length of first inter-ovulatory interval; LA60 = luteal activity during the first 60 DIM; PLA = proportion of samples with luteal activity during the first 60 DIM; CLAFS = interval from commencement of luteal activity to first service; CI = calving interval; CFI = interval from calving to first insemination; CLI = interval from calving to last insemination; FLI = interval from first to last insemination; NINS = number of inseminations per series; *d* = day.

Table 4
Total number of differentiated SNPs [-log (*p*-value) ≥ 4] and differentiated SNPs located in genes of classical and endocrine fertility traits for Swedish Red and Holstein cows.

Trait[1]	Number of differentiated SNPs [-log (*p*-value) ≥ 4]		Differentiated SNPs located in genes (N)	
	Swedish Red	Holstein	Swedish Red	Holstein
C-LA (d)	0	6	0	1
LPL (d)	406	3	92	0
ILI (d)	0	0	0	0
IOI (d)	259	19	67	2
LA60 (%)	12	5	5	1
PLA (%)	15	126	0	4
CLAFS (d)	21	29	4	1
CLI (d)	127	10	24	2
CFI (d)	6	1	1	0
CI (d)	76	70	20	18
FLI (d)	0	65	0	21
NINS (N)	68	81	16	2
Total	990	415	229	52

[1]C-LA = interval from calving to commencement of luteal activity; LPL = first luteal phase length; ILI = length of inter-luteal interval; IOI = length of first inter-ovulatory interval; LA60 = luteal activity during the first 60 DIM; PLA = proportion of samples with luteal activity during the first 60 DIM; CLAFS = interval from commencement of luteal activity to first service; CI = calving interval; CFI = interval from calving to first insemination; CLI = interval from calving to last insemination; FLI = interval from first to last insemination; NINS = number of inseminations per series; *d* = day.

intergenic spaces in the genome (Table 4; Supplementary Table S2). Compared with Holstein ($n = 46$), more genes ($n = 195$) were differentiated for SR, of which four genes (*SUFU, CREB5, IP6K1*, and *TRAIP*) were shared by both breeds (Fig. 3). For endocrine traits, a larger number of candidate genes was identified in SR than in Holstein ($n = 142$ and $n = 7$, respectively) (Fig. 3). The endocrine fertility traits found to be associated with the highest number of genes were LPL and IOI, and mainly in SR (Table 4). No significant *p*-values were obtained for the SNP effect for ILI in either of the breeds, and this trait was excluded from the downstream genomic evaluation. Manhattan plots for the remaining endocrine and classical fertility traits are shown in Supplementary Figure S1.

3.3. Gene annotation and functional enrichment analysis

The GO revealed 24 biological processes, nine molecular functions, 18 cellular components, and 20 KEGG pathways (Supplementary Table S3). *In utero* embryonic development, embryonic skeletal system development, cell morphogenesis, fibroblast activations, gonadotropin

Table 5
The most significantly associated SNPs (-log*P-value*) identified on BTA chromosomes from ssGWAS of classical and endocrine fertility traits in Swedish Red (SR) and Holstein (HOL) cows, and the genetic variances explained by the SNPs in parenthesis.

BTA	Trait[1]	-log*P-value (*)*	Position	Breed	BTA	Trait	-log*P-value (*)*	Position	Breed
1	LPL	10.4 (0.065)	14,438,946	SR	10	LPL	17.1 (0.027)	54,735,314	SR
2	LPL	12.5 (0.010)	80,096,393	SR	11	CLI	10.2 (0.123)	*104,182,498*	SR
2	LPL	12.0 (0.010)	79,388,083	SR	11	LPL	7.5 (0.025)	103,458,444	SR
2	LPL	11.9 (0.021)	79,158,290	SR	11	LPL	7.5 (0.030)	*103,478,873*	SR
2	LPL	9.0 (0.088)	77,420,390	SR	11	LPL	7.4 (0.013)	82,688,494	SR
2	LPL	9.0 (0.091)	77,444,377	SR	12	LPL	11.2 (0.060)	*63,103,894*	SR
2	LPL	9.0 (0.092)	77,583,642	SR	13	LPL	8.8 (0.034)	71,766,574	SR
2	LPL	9.0 (0.114)	77,620,108	SR	13	LPL	8.2 (0.005)	57,367,631	SR
2	LPL	8.6 (0.079)	55,566,541	SR	13	LPL	7.7 (0.013)	*53,091,922*	SR
2	LPL	7.7 (0.129)	78,691,609	SR	15	LPL	10.1 (0.020)	81,869,228	SR
2	LPL	7.2 (0.024)	50,632,840	SR	15	LPL	8.8 (0.078)	*21,732,355*	SR
3	IOI	9.5 (0.175)	18,769,778	SR	15	LPL	7.4 (0.160)	43,039,782	SR
3	IOI	8.9 (0.070)	22,921,514	SR	16	LPL	24.4 (0.059)	39,639,717	SR
3	LPL	8.9 (0.350)	118,963,574	SR	16	LPL	10.4 (0.427)	54,846,459	SR
3	LPL	7.5 (0.151)	114,911,444	SR	19	LPL	7.4 (0.102)	59,984,238	SR
3	IOI	7.3 (0.146)	*18,819,153*	SR	21	LPL	7.3 (0.047)	51,325,226	SR
4	LPL	7.6 (0.004)	9,726,535	SR	22	LPL	8.4 (0.138)	50,338,409	SR
4	LPL	7.2 (0.013)	67,701,182	SR	22	IOI	8.3 (0.028)	48,764,434	SR
5	NINS	7.8 (0.107)	120,264,251	SR	22	LPL	7.6 (0.160)	53,520,681	SR
5	CLI	7.1 (0.021)	25,800,581	SR	22	LPL	7.6 (0.127)	53,695,239	SR
6	LPL	11.4 (0.072)	*110,836,285*	SR	22	LPL	7.5 (0.075)	43,175,304	SR
6	LPL	7.8 (0.278)	118,481,842	SR	22	LPL	7.3 (0.033)	51,419,074	SR
6	LPL	7.4 (0.063)	89,572,513	SR	24	IOI	8.7 (0.026)	*12,899,222*	SR
7	IOI	8.9 (0.100)	73,344,236	SR	24	IOI	7.7 (0.062)	21,153,197	SR
9	LPL	9.2 (0.157)	54,690,915	SR	26	PLA	7.4 (0.149)	27,972,961	HOL
9	LPL	9.2 (0.222)	54,628,399	SR	26	PLA	7.2 (0.102)	*30,508,089*	HOL
9	LPL	8.2 (0.122)	77,832,949	SR	26	PLA	7.1 (0.065)	31,213,256	HOL
9	LPL	8.1 (0.087)	75,194,809	SR	28	LPL	9.6 (0.052)	*7,852,154*	SR
9	LPL	7.7 (0.563)	54,216,888	SR	28	LPL	7.9 (0.094)	*4,493,022*	SR
9	LPL	7.2 (0.119)	35,516,891	SR	29	CLI	7.6 (0.021)	*25,219,612*	SR

[1]LPL = first luteal phase length; IOI = length of first inter-ovulatory interval; PLA = proportion of samples with luteal activity during the first 60 DIM; CLI = interval from calving to last insemination; NINS = number of inseminations per series; BTA = *Bos taurus*; * = Genetic variance explained by the SNP. Twelve loci are imputed loci *italicized* in blue, and 48 of them are actual loci.

releasing hormone (GnRH) signaling pathway, and aldosterone synthesis and secretion were some of the most relevant biological processes and pathways identified in the GO and KEGG enrichment analyses (Table 6). Of the total of 241 genes, 80 were significantly enriched and spanned the whole genome except for three chromosomes (BTA17, BTA24, and BTA28). Functionally enriched genes were associated with four endocrine (C-LA, LPL, IOI and LA60), five classical (CI, CFI, CLI, FLI, and NINS) and one mixed (CLAFS) fertility traits. The *GNAQ, CACNA1D, ITPR2, CREB5*, and *ATF2* genes were found to be mediated in 8–12 pathways each in three traits (LPL, IOI, and CI). All chromosomes except BTA20, BTA21, BTA23, BTA25, and BTA29 that showed at least one gene each displayed 2–8 genes associated with the fertility traits. Table 6.

Protein-protein interaction (PPI) network analysis implemented using STRING (Szklarczyk et al., 2019) confirmed the biological pathways obtained using DAVID. In the STRING analysis, the network proteins encoded by the 80 candidate genes had more interactions among themselves than expected for a random set of proteins of similar size drawn from the genome (number of nodes = 76; observed number of edges = 73; expected number of edges = 47; Fig. 4), indicating that the candidate genes are biologically connected. STRING revealed three and 36 significantly enriched ($P<0.05$) molecular functions and KEGG pathways, respectively, that mediated 2–16 candidate genes each (Supplementary Table S4). The UniPort, PFAM, InterPRO, and SMART protein domains uncovered nine, four, 18, and nine processes, respectively, that contained 2–30 significantly enriched candidate genes (Supplementary Table S4).

Detection of QTL regions was performed using SNP variants/genes that were significantly enriched in various biological processes and pathways, considering 10 Mbp as a maximum QTL window. Of the 80 genes, 40 were found in 20 QTL regions in 12 chromosomes associated with four endocrine fertility traits (C-LA, IOI, LA60, LPL), four classical fertility traits (CI, CLI, FLI, NINS) and one mixed (classical and endocrine: CLAFS) fertility trait (Table 7). The most associated endocrine fertility trait was LPL (associated with 11 QTL regions), followed by IOI (7 QTL regions) and CI and CLI (4 QTL regions each). Swedish Red and Holstein cows shared only three QTL regions; on BTA2 (21.78–23.82 Mb) for CI and LPL, on BTA5 (42.68–49.73Mb) for CI, and on BTA22 (47.45–55.37 Mb) for IOI and CLA. The top associated SNP from all QTL regions was for IOI and LPL (-log*(p-value)* = 10.1) on BTA15 (75.94–81.87 Mb) for SR cows, and three genes (*MAPK8IP1, PHF21A*, and *LPXN*) that were identified based on four SNPs were embedded in this QTL region. All the QTLs detected that embedded the candidate genes, except the QTLs at BTA2 and BTA19, are novel QTL regions that were not reported before.

4. Discussion

4.1. Differentiation of the study populations

Swedish Red and Holstein cows were found to have almost completely different genetic backgrounds (Fig. 1). This finding was supported by the population differentiation revealed by AMOVA, which showed 9.2% variation between the two study populations. Edwards et al. (2011) found that the Nordic Red dairy breed carries a unique Y1–98–158 haplotype that reflects a founder effect associated with prehistoric development of dairy farming in the region (Beja-Pereira et al., 2003). Thus, the Red breed in the Nordic region might have developed unique adaptation over time compared with Holstein cows. The population structure showed that Holstein cows have an additional genetic background that constitutes approximately 10% (Fig. 1). This was supported by PCA, which demonstrated that some Holstein cows differentiated from the majority.

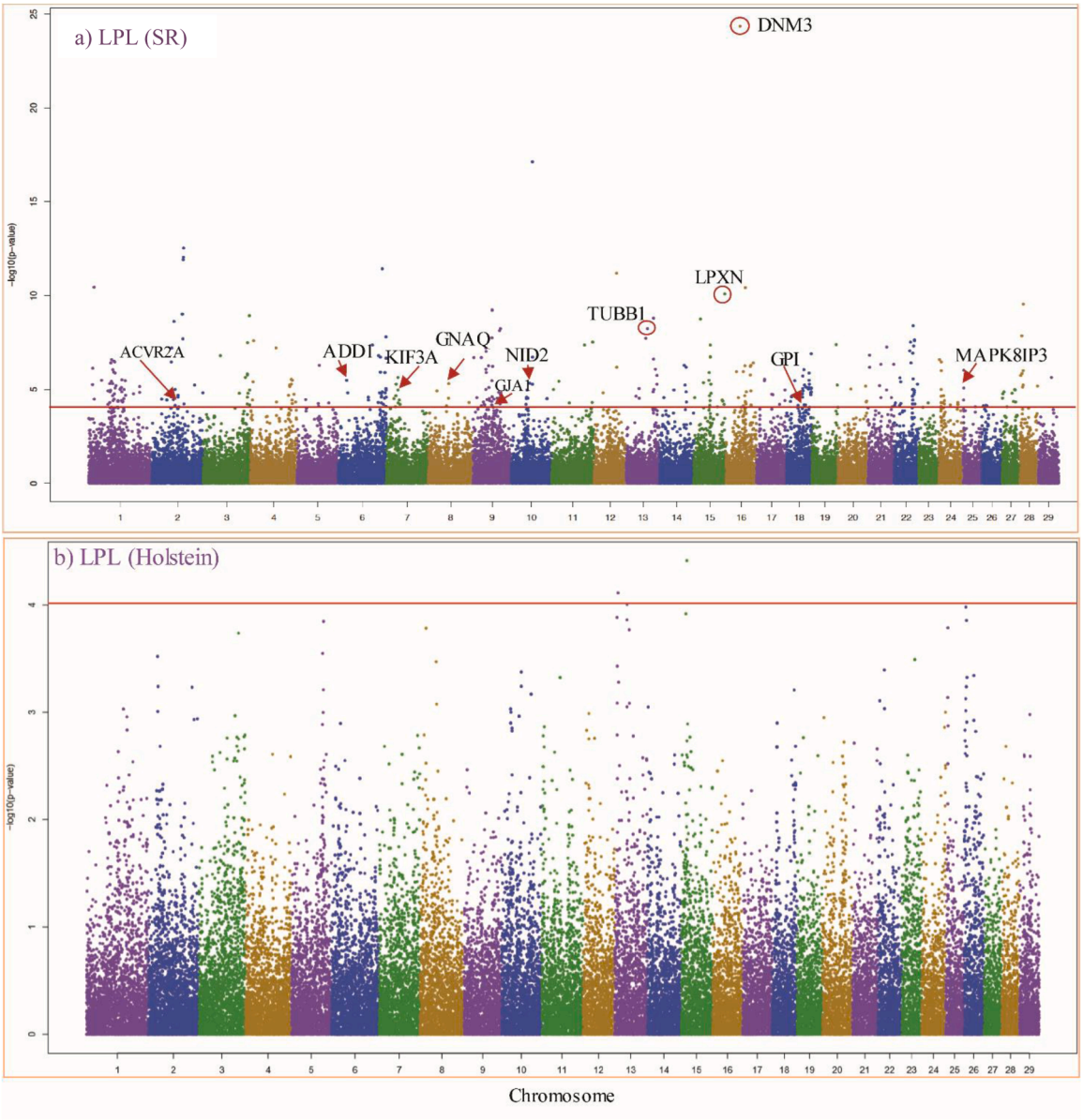


Fig. 2. a-f. Manhattan plot generated for Swedish Red (SR) and Holstein cows. LPL = First luteal phase length; IOI = First inter ovulatory interval; C-LA = Commencement of luteal activity.

4.2. Significant association and functional enrichment

The GO and KEGG enrichment functional analysis revealed that 33.2% of the genes analyzed ($n = 80$) were significantly enriched within various biological processes and KEGG pathways in a wide range of functions mainly associated with fertility. Of the significantly enriched genes, 67 were detected in SR and 13 in Holstein cows (Table 5), with only one gene associated with fertility traits, *CREB5*, found to be present in both breeds (Table 6). The different genetic background of the two dairy breeds (Fig. 1) may contribute to each breed having different allele frequencies at any given locus. In addition, the level of gene expression for those candidate genomic regions may vary between the breeds. Another reason could be that Holstein are inferior in some endocrine fertility traits, such as LA60, compared to SR cows (Tarekegn et al., 2019), which could be associated to a higher selection intensity for milk yield in Holstein. Most of the genes were associated with LPL ($n = 25$) and IOI ($n = 18$) in SR cows (Table 6). Moreover, in SR, four genes (*ECT2, MAPK8IP1, MAPK8IP3*, and *PPP3CA*) were associated with both LPL and IOI. In Holstein, we did not detect any gene associated with any endocrine fertility trait except for C-LA.

4.3. Candidate genes associated with endocrine and classical fertility traits

We identified several candidate genes associated with endocrine and classical fertility traits (Table 6). The traits found to have most associations with candidate genes are discussed below.

4.3.1. Candidate genes related to commencement of luteal phase length (C-LA) and first luteal phase length (LPL)

Amongst the endocrine fertility traits, the autophagy-related 7 gene (*ATG7*) with the genetic variance explained by the SNP of 0.075 (Supplementary Table S1) was the only gene found to be associated with C-LA, and then only in Holstein cows. The *ATG7* gene protects against ovarian follicle loss in mice (Song et al., 2015) and aids in the development of zygotes by refining excessive maternal factors during early

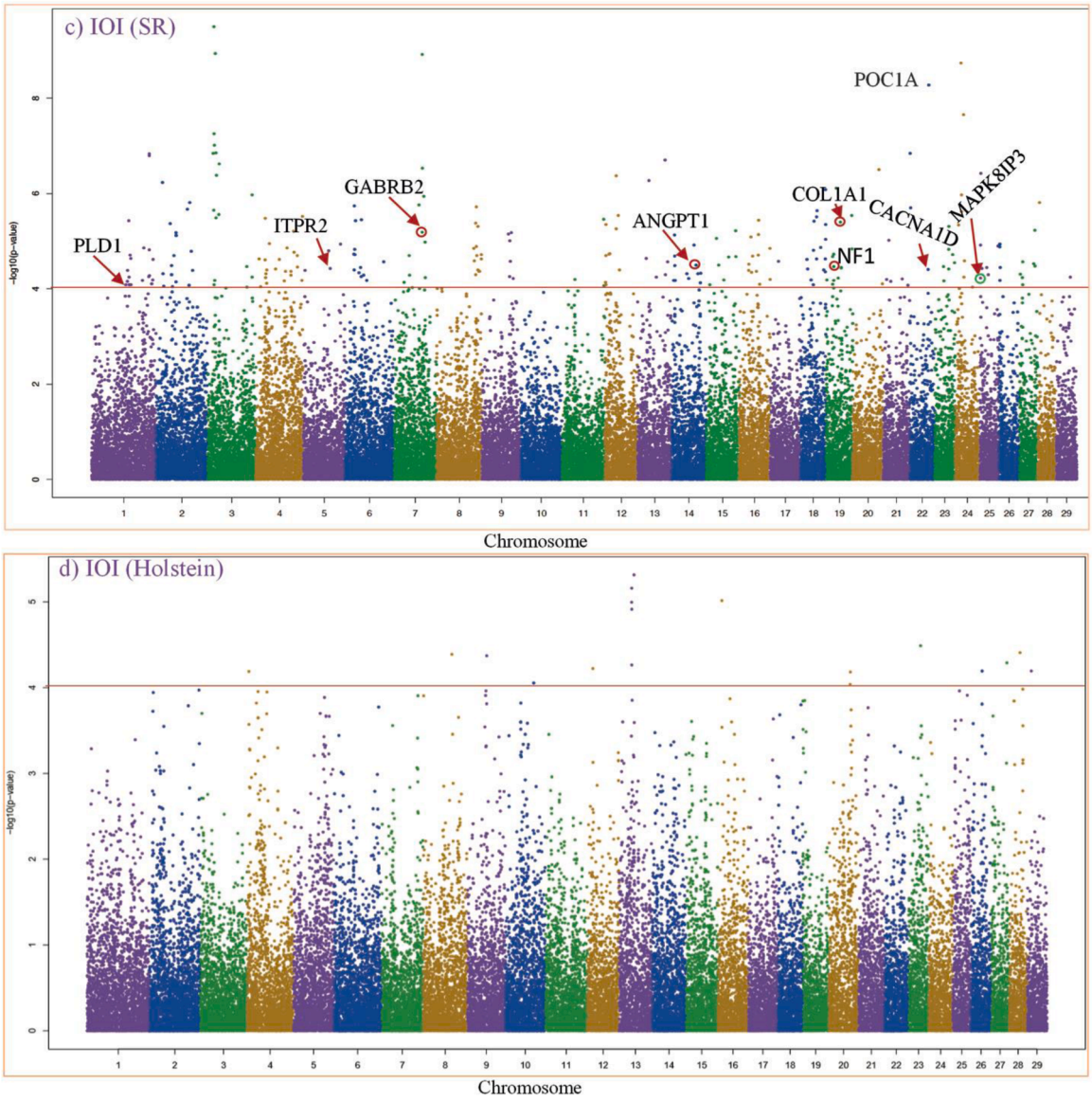


Fig. 2. (*continued*).

embryo development in bovines (Song et al., 2012).

The candidate genes *NID2, KIF3A, ADD1, GPI, GNAQ, ACVR2A*, and *PPP3CA* were among the genes found to be associated with LPL (Table 6). The Nidogen-2 (*NID2*) gene was associated with IOI. As LPL and IOI have a part-whole relationship, this finding was not unexpected. The *NID2* gene plays a direct role in conceptus elongation in cattle, and establishes uterine receptivity to implantation (Forde et al., 2012). In the study by Forde et al. (2012), expression of the *NID2* gene increased on day 13 compared with day 7 after synchronization ($P<0.05$), and the magnitude of increase was significantly diminished in heifers with low P4 compared with heifers grouped randomly as controls. Low P4 level, mainly caused by luteal insufficiency and increased metabolic clearance by the liver (Sangsritavong et al., 2002; Wiltbank et al., 2011), has been identified as a cause of infertility in lactating dairy cows (Diskin et al., 2008; Nyman et al., 2018).

Studies show that the *KIF3A* gene is essential for all assembly and function of primary cilia that have projecting roles in GnRH neurons (Marszalek et al., 1999; Shin et al., 2014) and favors meiotic division in Holstein oocytes (Ticianelli et al., 2017). Primary cilia drive pituitary regulation of gonadal function and fertility (Shin et al., 2014). Within the KIF family, the *KIF4A* gene, which plays important roles in cell proliferation associated with endometrial function during the estrus cycle and pregnancy, has been reported as a candidate gene for fertility in beef cattle (Neupane et al., 2017). Similarly, studies have shown that the Adducin 1 alpha gene (*ADD1*) is upregulated in the luteal phase in post-pubertal Brahman heifers (Tahir et al., 2019), and is essential for pregnancy success (Valcarce et al., 2013). Hence, the *KIFA3* and *ADD1* genes may play a role in luteal activity in SR cows. The glucose-6-phosphate isomerase gene (*GPI*) is expressed by the corpus luteum (Schulz and Bahr, 2003) during pregnancy and plays a role in endometrial remodeling in preparation for embryo implantation in cattle (Mamo et al., 2012). Forde et al. (2012) reported high expression of the *GPI* gene associated with the specific effect of P4 on day 13 of the estrus cycle, when conceptus elongation is initiated in cattle.

The guanine nucleotide-binding protein G(q) subunit alpha gene (*GNAQ*) was mediated in 12 biological processes and pathways (e.g., GnRH signaling pathway) (Table 6). The *GNAQ* gene is reported to be associated with ovarian follicle development in crossbred beef cattle (Zielak-Steciwko et al., 2014) and to act as a signal for the *KISS1* gene (León et al., 2016). Kisspeptin, encoded by the *KISS1* gene, is a major molecule that stimulates the hypothalamus region to release GnRH, which in turn triggers the pituitary glands to secrete follicle-stimulating

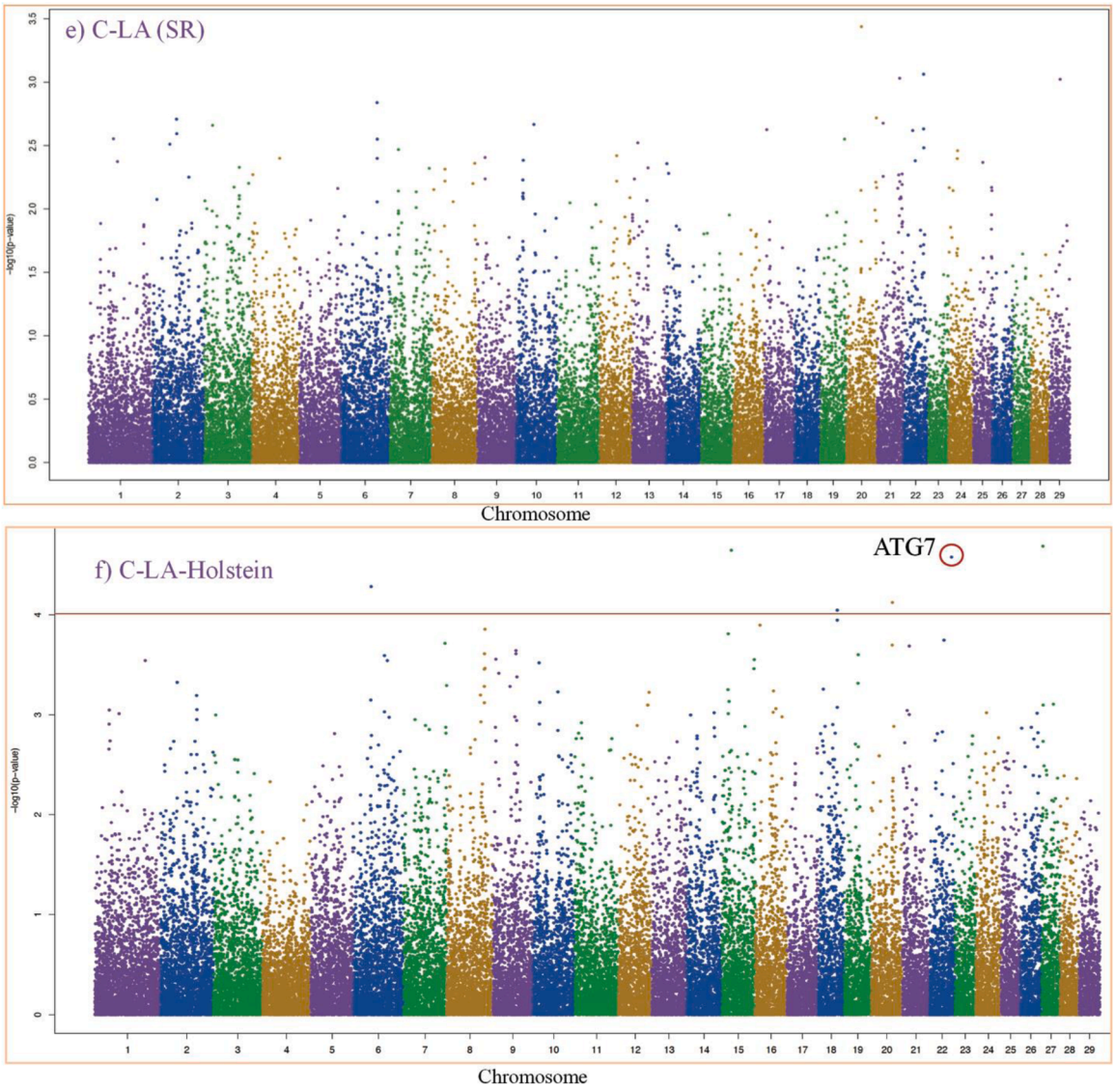


Fig. 2. (*continued*).

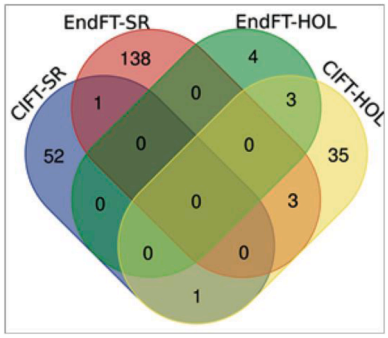


Fig. 3. Venn diagram of the candidate genes. ClFT-SR: classical fertility traits of Swedish Red cows; EndFT-SR: endocrine fertility traits of Swedish Red cows; ClFT-HOL: classical traits of Holstein cows; EndFT-HOL: endocrine traits of Holstein cows.

hormone (FSH) and luteinizing hormone (LH) in cattle (Ahmed et al., 2009), as in many other species. Kisspeptin-induced LH secretion has been shown to stimulate follicular development and ovulation in Japanese Black beef calves and cows (Ezzat Ahmed et al., 2009; Naniwa et al., 2013). Other studies show that Kisspeptin signaling regulates ovarian function (Dorfman et al., 2014; Gaytan et al., 2014; Babwah, 2015; Bhattacharya and Babwah, 2015). Overall, the candidate genes identified in association with LPL in this study may be more expressed and control the initial stages of luteolysis, conceptus elongation, and the luteal phase in the SR breed.

4.3.2. *Candidate genes related to inter-ovulatory interval (IOI)*

In this study, the Angiopoietin-1 gene (*Angpt1*), *CACNA1D, RGCC, GABRB2, PPP3CA, NID2*, and *SLC25A17* were among the candidate genes observed in association with IOI in the SR breed (Table 6; Supplementary Table 5). According to Parborell et al. (2008), the angiopoietin encoded by *Angpt1* is induced by gonadotropins, and is involved in regulation of follicular development and ovulation. Similarly, the calcium voltage-gated channel subunit alpha-1D gene (*CACNA1D*) is involved in the physiological mechanisms that are dependent on calcium, including the release of some hormones such as estrogens (Ji et al., 2016). In dairy cows, Hatzirodos et al. (2014a) reported that the *CACNA1D* gene is downregulated in small atretic follicles compared with in healthy follicles, indicating the importance of the gene for follicular development. By definition, IOI includes the follicular and luteal phases. However, other studies have shown the *CACNA1D* gene to be associated

Table 6
Significantly enriched GO and KEGG pathways of classical and endocrine fertility traits for Swedish Red and Holstein cows.

Category[1]	Term	Pathways	Gene(N)	*P* value	Gene symbol	Trait	Breed
BP	GO:0,070,507	Regulation of microtubule cytoskeleton organization	3	0.0178	PHLDB1	CI	HOL
					EFNA5	CLI	SR
					PHLDB2	LPL	SR
BP	GO:0,045,732	Positive regulation of protein catabolic process	4	0.020	STX5	FLI	HOL
					ATG7	CLA	HOL
					GJA1	LPL	SR
					ITCH	CLI	SR
BP	GO:0,001,701	In utero embryonic development	7	0.024	GPI	LPL	SR
					KIF3A	LPL	SR
					PTPRR	CI	SR
					MAPK8IP3	IOI, LPL	SR
					GJA1	LPL	SR
					ANGPT1	IOI	SR
					ADD1	LPL	SR
BP	GO:0,007,162	Negative regulation of cell adhesion	3	0.028	CDH13	LA60, LPL	SR
					LPXN	LPL	SR
					ANGPT1	IOI	SR
BP	GO:0,000,902	Cell morphogenesis	4	0.035	FRYL	CLI, NINS	SR
					TENM3	LA60	SR
					ECT2	IOI, LPL	SR
					ADD1	LPL	SR
BP	GO:0,007,605	Sensory perception of sound	5	0.038	GABRB2	IOI	SR
					ATP6V0A4	LPL	SR
					COL1A1	IOI	SR
					USH2A	CLI, FLI	HOL
					CACNA1D	IOI	SR
BP	GO:0,072,537	Fibroblast activation	2	0.041	IL17A	IOI	SR
					RGCC	IOI	SR
BP	GO:0,048,706	Embryonic skeletal system development	3	0.043	ACVR2A	LPL	SR
					PBX1	CI	HOL
					COL1A1	IOI	SR
MF	GO:0,035,091	Phosphatidylinositol binding	5	0.013	PLD1	IOI	SR
					SNX27	IOI	SR
					SNX7	FLI	HOL
					ITPR2	IOI	SR
					PLEKHA2	LPL	SR
MF	GO:0,005,096	GTPase activator activity	10	0.001	ARHGAP20	FLI	HOL
					RGS7, PREX2	IOI	SR
					TBC1D20	NINS	SR
					GNAQ, STXBP5L	LPL	SR
					ARHGAP15	CLI	SR
					ASAP3	CLI, FLI	HOL
					SRGAP1	CI	HOL
					ECT2	IOI, LPL	SR
MF	GO:0,005,509	Calcium ion binding	15	0.015	CAPSL, EFHD1	NINS	SR
					PCDHGA8, PCDHGB4	CI	SR
					S100A10	IOI	SR
					FBN2, NID2, MAN1A1, SLC24A2	LPL	SR
					MMRN1	CI, CLI	SR
					KCNIP4, ITPR2	IOI	SR
					CDH13	LA60,LPL	SR
					SPTA1	CLAFS	SR
					VCAN	CFI,CLI,CLAFS	SR
MF	GO:0,030,165	PDZ domain binding	3	0.029	ACVR2A, PLEKHA2, GJA1	LPL	SR
MF	GO:0,003,682	Chromatin binding	9	0.045	TDRD3, PHF21A	IOI	SR
					MEF2A, DDX1, TPR	LPL	SR
					BRD3	CLI	SR
					CBX5	CI, CLI, NINS	SR
					ATF2, SMAD6	CI	HOL
KEGG	bta04911	Insulin secretion	7	0.001	GNAQ, KCNMB2, RAPGEF4	LPL	SR
					CACNA1D	IOI	SR
					CREB5, ATF2, KCNN2	CI	HOL
KEGG	bta04010	MAPK signaling pathway	11	0.001	RIPK14	LA60	SR
					NF1, PPM1B	IOI	SR
					PTPRR	CI	SR
					MAPK8IP3, PPP3CA	IOI, LPL	SR
					CACNA1D	IOI	SR
					FLNB, FGF12	LPL	SR
					ATF2	CI	HOL
					MAPK8IP1	IOI,LPL	SR
KEGG	bta04724	Glutamatergic synapse	7	0.003	PLD1, ITPR2, CACNA1D, GRIA1	IOI	SR
					GNAQ	LPL	SR
					GRIK2	LPL	SR
					PPP3CA	IOI, LPL	SR

(*continued on next page*)

Table 6 (*continued*)

Category[1]	Term	Pathways	Gene(N)	*P* value	Gene symbol	Trait	Breed
KEGG	bta04022	cGMP-PKG signaling pathway	8	0.004	MEF2A, GNAQ, KCNMB2	LPL	SR
					CREB5, ATF2	CI	HOL
					PPP3CA	IOI, LPL	SR
					ITPR2, CACNA1D	IOI	SR
KEGG	bta04728	Dopaminergic synapse	7	0.005	GNAQ	LPL	SR
					CREB5, ATF2	CI	HOL
					PPP3CA	IOI, LPL	SR
					ITPR2, GRIA1, CACNA1D	IOI	SR
KEGG	bta05031	Amphetamine addiction	5	0.010	GRIA1, CACNA1D	IOI	SR
					CREB5, ATF2	CI	HOL
					PPP3CA	IOI, LPL	SR
KEGG	bta04918	Thyroid hormone synthesis	5	0.011	TG, GNAQ	LPL	SR
					ATF2, CREB5	CI	HOL
					ITPR2	IOI	SR
KEGG	bta04925	Aldosterone synthesis and secretion	5	0.017	GNAQ	LPL	SR
					CACNA1D, ITPR2	IOI	SR
					ATF2, CREB5	CI	HOL
KEGG	bta04151	PI3K-Akt signaling pathway	10	0.028	CCNE1, FGF12	LPL	SR
					FGF14	LA60	SR
					CREB5, ATF2	CI	HOL
					EFNA5	CLI	SR
					ANGPT1, COL1A1	IOI	SR
KEGG	bta04922	Glucagon signaling pathway	5	0.030	GNAQ, CREB5	LPL	SR
					PPP3CA	IOI, LPL	SR
					ATF2, CREB5	CI	HOL
					ITPR2	IOI	SR
KEGG	bta04723	Retrograde endocannabinoid signaling	5	0.040	GNAQ	LPL	SR
					ITPR2, GRIA1, CACNA1D, GABRB2	IOI	SR
KEGG	bta04912	GnRH signaling pathway	4	0.042	GNAQ	LPL	SR
					ITPR2, PLD1, CACNA1D	IOI	SR
KEGG	bta04720	Long-term potentiation	4	0.047	GNAQ	LPL	SR
					PPP3CA	IOI, LPL	SR
					ITPR2, GRIA1	IOI	SR
KEGG	bta04924	Renin secretion	4	0.049	GNAQ	LPL	SR
					PPP3CA	IOI, LPL	SR
					CACNA1D, ITPR2	IOI	SR

[1]KEGG = Kyoto encyclopedia of genes and genomes (KEGG); BP = Biological processes; MF = Molecular function; SR = Swedish Red; HOL = Holstein.

with classical fertility traits in dairy cows, such as daughter pregnancy rate, conception rate, and age at first calving (Parker-Gaddis et al., 2016; Ortega et al., 2017; Ortega, 2018; Fernández et al., 2019).

The regulator gene for cell cycle (*RGCC*) and the gamma-aminobutyric acid receptor subunit beta-2 gene (*GABRB2*) are highly differentiated in the granulosa cells of pre-antral follicles in mural granulosa and cumulus cells (Wigglesworth et al., 2015). These two cell types play crucial roles in oocyte maturation and ovulation in all animal species. The oocyte maturation and ovulation occurring during IOI suggests that the *RGCC* and *GABRB2* can potentially be candidate genes for IOI in SR cows. Polymorphisms (ARS-BFGL-NGS-103,398: -log *P*-value = 4.82, SNP variance = 0.105; HAPMAP49541-BTA-24,412: -log *P*-value = 4.71, SNP variance = 0.069; HAPMAP50782-BTA-78,177: -log *P*-value = 4.81, SNP variance = 0.076; Supplementary Tables S1 and S2) detected on the protein phosphatase 3 catalytic subunit alpha isozyme gene (*PPP3CA*), which showed co-expression with *GABRB2* and *CACNA1D* in this study (Fig. 4), facilitates early pregnancy in heifers (Dias et al., 2015; Barbero et al., 2017). In another study, Dias et al. (2017) detected SNPs associated with early puberty in the *PPP3CA* gene, and suggested that this gene could be considered a strong predictor of puberty in cattle. The *PPP3CA* gene may influence follicular development after calving. However, the influence may vary from breed to breed, e.g., in this study SR cows had on average 2 days shorter IOI and LPL than Holstein (Table 2). Although only the first luteal phase length and inter-ovulatory intervals were considered here, the shorter luteal phase and inter-ovulatory interval may indicate better fertility of the SR breed. According to Ahmad et al. (1997), cows that become pregnant have a shorter luteal phase than non-pregnant cows after insemination, but cows that have two follicle waves have lower pregnancy percentage than cows that have three follicle waves. In addition, in an analysis based on intensive sampling in Herd Navigator™, a significant negative linear relationship was observed between luteal phase length and probability of pregnancy (Bruinjé et al., 2019).

The SLC24A2 protein is a potassium-dependent sodium-calcium exchanger and can transport metal ions across cell membranes (Wang et al., 2017). The *SLC25A17* gene was associated with LPL in SR cows in the present study. Associations with genes in the SLC family have been reported previously, e.g., Tenghe et al. (2016) found *SLC40A1* to be associated with CLAFS in Holstein cows, and Liu et al. (2017) identified *SLC39A12* as a strong candidate gene for female fertility in Holstein-Frisian. Overall, the candidate genes found in association with IOI can be potential candidate genes for improving female fertility in SR cows.

4.3.3. Candidate genes related to calving to last insemination (CLI) and calving interval (CI)

Candidate genes associated with the classical fertility traits CLI and CI in SR and Holstein cows include *EFNA5, PBX1, SMAD6, ATF2*, and *CREB5*. The Ephrin-A5 gene (*EFNA5*), induced by FSH, was found to be associated with CLI in SR cows in this study. The *EFNA5* gene is expressed in the ovaries and is involved in follicular growth and ovulation (Worku et al., 2018). Hatzirodos et al. (2014b) found higher expression of the *EFNA5* gene in large antral follicles than in small follicles in cow ovaries. A possible explanation is that SR cows have alleles that result in higher *EFNA5* gene expression, which may induce earlier cellular changes for follicular growth and ovulation compared with Holstein cows. The Pre-B-cell leukemia transcription factor 1 gene (*PBX1*) was found to be associated with CI in Holstein cows in the present study. Studies on mice and humans indicate that the *PBX1* gene is involved in several biological processes during embryogenesis, including steroidogenesis and sexual development (Kim et al., 2002; Bijl et al., 2005; Duque-Afonso et al., 2015). This protein functions in the

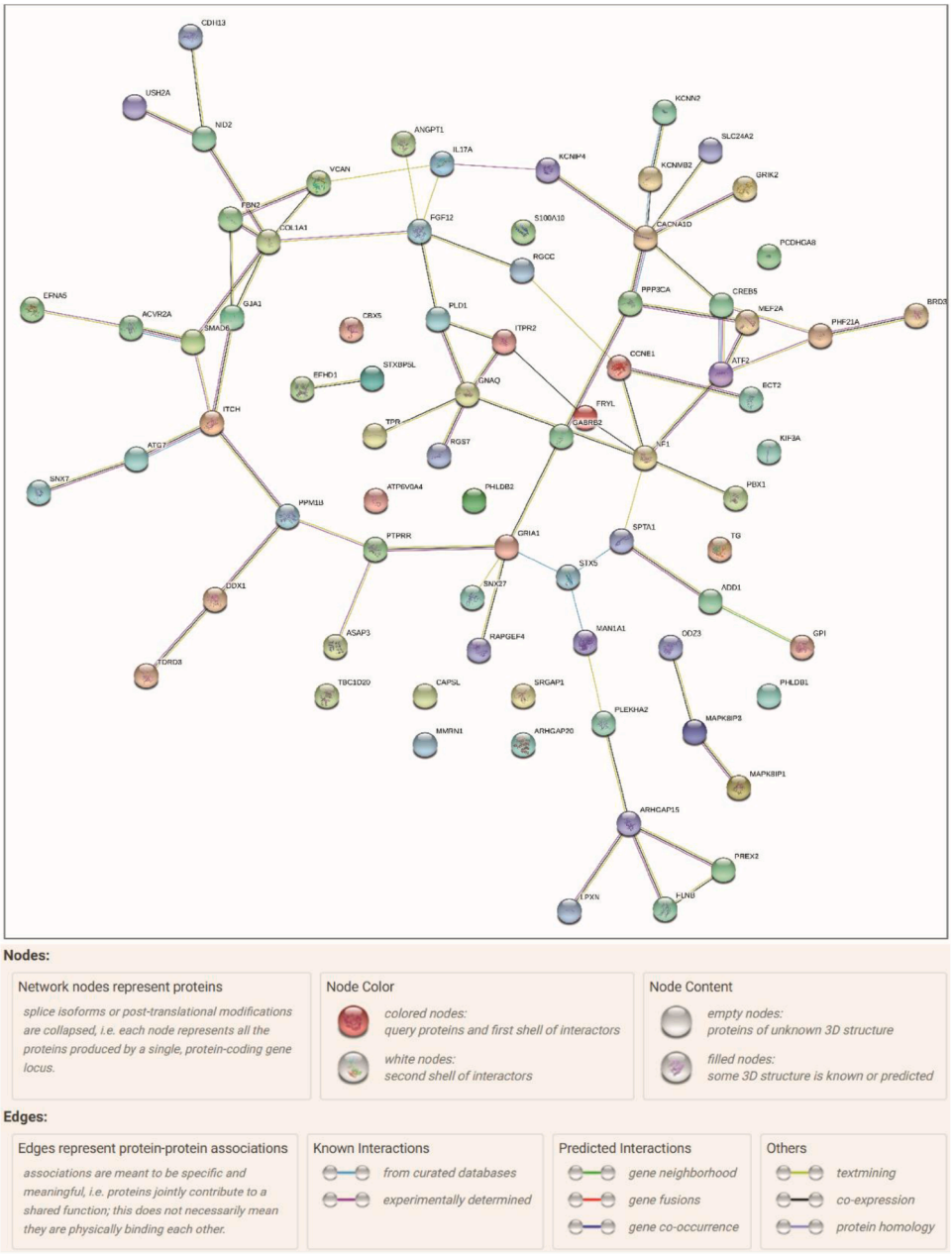


Fig. 4. Protein-protein interaction of functionally significantly enriched genes for classical and endocrine fertility traits in Swedish Red (SR) and Holstein cows revealed by STRING: number of node =76; observed number of edges = 73; expected number of edges 47; average node degree = 1.92; average local clustering coefficient = 0.371; PPI enrichment *p*-value = 0.000319.

development of several organ systems and plays a role in skeletal patterning and programming (Kim et al., 2002).

The *SMAD* (*SMAD6*) gene was found to be associated with CI in Holstein. This gene shows a negative correlation (−0.29) with the *ACVR2A* gene (Satheesh et al., 2016). The *ACVR2A* gene, which was associated with LPL in SR cows (Table 6), acts as a receptor for myostatin (MSTN) protein involved in inhibiting satellite cell proliferation and differentiation (Lee and McPherron, 2001). The *SMAD* gene regulates initiation of growth, development, and maintenance of primordial follicles in mice (Yang et al., 2013). Two SNPs in the *SMAD6* gene are reported to be associated with increased ovulation rate in synchronized daughters of bulls carrying the haplotype (Kirkpatrick and Morris, 2015). Thus, it can be hypothesized that the differences in allele frequencies because of the differentiated SNP in the *ACVR2A* and *SMAD6* genes, and their expression pattern, may vary between SR and Holstein dairy cows, which could cause the breeds to differ in follicular development and, later, in fetus growth. Gene expression analysis in both dairy breeds will provide better insights on this.

Activating transcription factor 2 gene (*ATF2*), a cAMP-dependent transcription factor, was mediated in chromatin-binding molecular function (GO:0,003,682) and in nine other KEGG pathways with the *CREB5* gene (Table 6). The ATF and CREB proteins can homodimerize or heterodimerize to form complexes that regulate stress responses, embryonic development, disease development, and cell death (Lopez-Bergami et al., 2010). Recent studies have shown that the *CREB5* gene is involved in the pathogenesis of recurrent pregnancy loss in women (Yu et al., 2018). In the present study, we detected *ATF2* and *CREB5* genes in the cGMP-PKG signaling pathway associated with CI in Holstein cows. Cyclic GMP (cGMP) is also involved in mediation of hormone-induced cell death during development and regulation of steroidogenesis

Table 7
QTL[1] regions on the chromosomes of *Bos taurus* (BTA) and genes embedded in the QTL regions associated to classical and endocrine fertility traits.

BTA	Start QTL in bp [-log(*p-value*) ≥4]	End QTL in bp [-log(*p-value*) ≥4]	Gene	No[2]	Traits[3]	Breed[4]
1	56,648,072 (6.4)	65,663,517 (4.8)	PHLDB2, STXBP5L	4	LPL	SR
1	75,146,323 (4.4)	88,327,865 (4.4)	FGF12, KCNMB2	4	LPL	SR
1	94,730,717 (5.5)	95,750,768 (4.1)	ECT2, PLD1	3	IOI, LPL	SR
2	21,778,683 (4.2)	23,824,758 (4.5)	ATF2, RAPGEF4	2	CI, LPL	HOL, SR
2	48,356,704 (4.02)	53,113,209 (4.5)	ACVR2A, ARHGAP15	2	LPL, CLI	SR
3	11,040,167 (4.8)	19,093,097 (4.1)	SPTA1, S100A10, SNX27	3	CLAFS, IOI	SR
5	42,680,103 (4.8)	49,727,902 (5.5)	PTPRR, SRGAP1	2	CI	SR, HOL
6	23,460,345 (4.8)	23,738,304 (4.7)	PPP3CA	3	LPL, IOI	SR
6	34,686,041 (4.8)	41,443,081 (5.4)	MMRN1, KCNIP4	4	CLI, IOI	SR
7	21,677,064 (4.5)	25,577,569 (4.1)	FBN2, KIF3A	2	LPL	SR
7	52,489,226 (6.2)	65,071,894 (5.8)	PCDHGB4, PCDHGA8, GRIA1	2	CI, IOI	SR
9	29,755,031 (4.3)	31,833,368 (4.2)	GJA1, MAN1A1	2	LPL	SR
10	3,274,229 (4.1)	13,567,343 (4.2)	KCNN2, SMAD6	2	CI	HOL
10	44,868,505 (4.5)	44,899,012 (4.02)	NID2	2	LPL	SR
13	60,593,663 (5.3)	63,875,561 (5.3)	TBC1D20, ITCH	2	NINS, CLI	SR
15	20,795,418 (4.1)	29,271,879 (5.5)	ARHGAP20, PHLDB1	2	FLI, CLI	HOL
15	75,937,148 (5.2)	81,869,228 (10.1)	MAPK8IP1, PHF21A, LPXN	4	IOI, LPL	SR
18	40,426,150 (4.6)	44,801,596 (4.2)	CCNE1, GPI	2	LPL	SR
18	9,515,675 (5.1)	9,582,369 (4.3)	CDH13	2	LA60	SR
22	47,445,142 (4.4)	55,365,371 (4.6)	CACNA1D, ATG7	3	IOI, CLA	SR, HOL

[1] QTL = Quantitative trait loci defined as a chromosomal region with two or more significant SNPs within 10 Mbp window. [2]No = Number of significant SNPs within the QTL. [3]C-LA = interval from calving to commencement of luteal activity; LPL = first luteal phase length; IOI = length of first inter-ovulatory interval; LA60 = luteal activity during the first 60 DIM; CI = calving interval; CLI = interval from calving to last insemination; NINS = number of inseminations per series; FLI = Interval from first to last insemination; CLAFS = interval from commencement of luteal activity to first service. [4]SR = Swedish Red; HOL = Holstein. All the QTLs detected except at BTA2 are novel QTL regions.

(Goy, 1991; Anderson et al., 1994; Grealy et al., 1997). In the PPI network shown in Fig. 4, the *ATF2* and *CREB5* genes showed co-expression, indicating that their gene products are correlated in expression.

4.4. Identification of quantitative trait loci regions

The QTL regions were identified based on the potential candidate genes enriched in the DAVID tool. A total of 20 QTLs were found in association with the fertility traits, of which 11 QTLs were specific to endocrine fertility traits, four QTLs to classical fertility traits, four QTLs to both endocrine and classical traits, and one to CLAFS and IOI (Table 7). In a recent study, no QTL region was detected for endocrine or classical fertility traits in Holstein-Friesian cows using the same SNP chip panel as in the present study (50 K) (Nyman et al., 2019). However, several significant associations with delayed cyclicity and C-LA were demonstrated when fine-mapping QTL regions on chromosomes BTA8, BTA17, and BTA23 using imputed sequence variants to the whole genome (Nyman et al., 2019). In the present study, the QTL regions detected on BTA1–2, BTA7, BTA9–10, BTA15, and BTA18 were associated with LPL; on BTA1, BTA3, BTA6, BTA7, BTA15, and BTA22 with IOI; on BTA22 with C-LA, on BTA18 with LA60, and on BTA3 with CLAFS. Some of the BTAs where the QTLs were observed (Table 7) are in line with previous findings. For instance, the same QTL regions found on BTA2 in the present study have been reported in association with conception rate in Holstein heifers (Kiser et al., 2019). A QTL region associated with ovulation rate has been reported on BTA19 (Kirkpatrick et al., 2000). We found a QTL (18 Mb, data not shown) on BTA19 that has embedded *COL1A1* and *NF1* genes associated with IOI in SR cows. Earlier studies have found that the *COL1A1* gene is associated with infertility (Salilew-Wondim et al., 2010) and the *NF1* gene with classical tumor suppression (Walker et al., 2006). In a study on Holstein-Friesian cows, Tenghe et al. (2016) found QTLs on BTA2, BTA3, and BTA17 for LPL; BTA3 for IOI; BTA4 for C-LA; BTA1, BTA2, and BTA23 for CLAFS, and BTA3&25 for LA60. Similarly, Berry et al. (2012) identified two QTL regions in association with C-LA on BTA2 and BTA21 in Holstein-Friesian cows. In the latter study, the SNPs detected on BTA2 and BTA21 explained 0.51% and 0.35% of the genetic variances, respectively in CLA and those SNPs demonstrated 24 and 18 Bayes factor estimates that indicated strong evidence of a QTL, and this report support the QTLs detected in the current study. The authors also indicated that the search for genes within 500 kb from the SNP that was detected in an intergenic region on BTA2 revealed 16 genes (three of them have functions that are related to fertility) whereas only one gene was found for the second SNP that was detected on BTA21.

In the present study on cows of the Holstein and SR breeds, a total of 40 genes associated with endocrine and classical fertility traits were found in the 20 QTL regions identified in the 12 target chromosomes (Table 7). In a study by Nyman et al. (2019) on Holstein-Friesian cows, nine, nine, and five possible candidate genes that control fertility were detected on BTA8, BTA17, and BTA23, respectively. Similarly, Tenghe et al. (2016) detected a total of 83 (61 annotated and 22 uncharacterized) candidate genes by fine-mapping BTA2 and BTA3 in Holstein-Friesian cows, and found that 24 (16 annotated and 8 uncharacterized) of these were associated with C-LA. Variations in terms of sample size, breed, study population, genome coverage, the bovine genome assembly, and the GWAS method employed may explain the differences in results between the present and previous studies (Berry et al., 2012; Tenghe et al., 2016; Nyman et al., 2019).

5. Conclusions

Fertility is a complex trait influenced by many small polygenic effects and by environmental factors. In the present study, we identified a wide range of genomic regions associated with endocrine and classical fertility traits. The LPL, IOI, CI, and CLI traits were found to be influenced by many genes. Moreover, based on the significantly differentiated SNPs, we identified genes with pleiotropic effects that control and regulate two or more endocrine and classical fertility traits. The results obtained suggest that the endocrine fertility traits LPL and IOI may have the potential to improve fertility in dairy cows, particularly SR cows. Potential candidate genes to be further studied in SR and Holstein cows are *EFNA5, GNAQ, NID2, SLC25A17, CREB5, ATF2, CACNAD1, CREB5*, and *KIF3A*. Our findings offer new insights into the genetic basis of endocrine and classical fertility in SR and Holstein dairy breeds. A challenge for future research is to determine why different regions seem relevant for different traits and breeds, and the practical implications of this for genomic selection.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgements

We thank the herd owners in Sweden who contributed data of their herds. This study was performed within the Swedish part of the Nordic research project NorFert (Improving Nordic dairy cow fertility through genetics). The main participant is the Swedish University of Agricultural Sciences (SLU). Industry partners are Växa Sverige (Stockholm, Sweden), where we especially want to acknowledge Jan-Åke Eriksson for valuable contributions, VikingGenetics (Randers, Denmark), Nordic Cattle Genetic Evaluation Ltd. (NAV, Aarhus, Denmark), and Lattec I/S (Hillerod, Denmark). We are grateful for project funding provided by the Swedish Farmers' Foundation for Agricultural Research (Project ID: O-15–20–587).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.livsci.2021.104731.

References

Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., Lawlor, T.J., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci 93, 743–752.

Aguilar, I., Misztal, I., Legarra, A., Tsuruta, S., 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J. Anim. Breed. Genet. 128, 422–428.

Aguilar, I., Misztal, I., Tsuruta, S., Legarra, A., Wang, H., 2014. PREGSF90–POSTGSF90: computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. In: Pages 1-3 in Proc. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, BC, Canada. Champaign, IL. American Society of Animal Science.

Aguilar, I., Legarra, A., Fernando Cardoso, F., Masuda, Y., Lourenco, D., Misztal, I., 2019. Frequentist *p*-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. Genet. Sel. Evol. 51, 28.

Ahmad, N., Townsend, E.C., Dailey, R.A., Inskeep, E.K., 1997. Relationships of Hormonal Patterns and Fertility to Occurrence of Two or Three Waves of Ovarian Follicles, Before and After Breeding, in Beef Cows and Heifers. Anim. Reprod. Sci. 49 (1), 13–28.

Anderson, R.A., Feathergill, K.A., Drisdel, R.C., Rawlins, R.G., Mack, S.R., Zaneveld, L.J. D., 1994. Atrial natriuretic peptide (ANP) as a stimulus of the human acrosome reaction and a component of ovarian follicular fluid: correlation of follicular ANP with in vitro fertilization outcome. J. Androl. 15, 61–70.

Babwah, A.V., 2015. Uterine and placental KISS1 regulate pregnancy: what we know and the challenges that lie ahead. Reproduction 150, 121–128.

Barbero, M.M.D., Santos, D.J.A., Takada, L., de Camargo, G.M.F., Freitas, A.C., Diaz, I.S. D.P., de Souza, F.R.P., Tonhati, H., Albuquerque, L.G., Oliveira, H.N., 2017. Prospecting polymorphisms in the PPP3CA and FABP4 genes and their association with early pregnancy probability in Nellore heifers. Livest Sci 203, 76–81.

Beja-Pereira, A., Luikart, G., England, P.R., Bradley, D.G., Jann, O.C., et al., 2003. Gene-culture coevolution between cattle milk protein genes and human lactase genes. Nat Genet 35, 311–313.

Berry, D.P., Bastiaansen, J.W.M., Veerkamp, R.F., Wijga, S., Wall, E., Berglund, B., Calus, M.P.L., 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.

Bhattacharya, M., Babwah, A.V., 2015. Kisspeptin: beyond the brain. Endocrinology 156, 1218–1227.

Bijl, J., Sauvageau, M., Thompson, A., Sauvageau, G., 2005. High incidence of proviral integrations in the Hoxa locus in a new model of E2a–PBX1-induced B-cell leukemia. Genes and Development; 19, 224–233.

Bruinjé, T.C., Colazo, M.G., Ribeiro, E.S., Gobikrushanth, M., Ambrose, D.J., 2019. Using in-line milk progesterone data to characterize parameters of luteal activity and their association with fertility in Holstein cows. J Dairy Sci 102 (1), 780–798.

Christensen, O.F., Lund, M.S., 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42, 2.

Daetwyler, H., Villanueva, B., Woolliams, J.A., 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3, 1–8.

Dias, M.M., Souza, F.R.P., Takada, L., Feitosa, F.L.B., Costa, R.B., Diaz, I.D.P.S., Cardoso, D.F., Tonussi, R.L., Baldi, F., Albuquerque, L.G., Oliveira, H.N., 2015. Study of lipid metabolism-related genes as candidate genes of sexual precocity in Nellore cattle. Genet. Mol. Res. 14 (1), 234–243.

Dias, M.M., Cánovas, A., Mantilla-Rojas, C., Riley, D.G., Luna-Nevarez, P., Coleman, S.J., Speidel, S.E., Enns, R.M., Islas-Trejo, A., Medrano, J.F., Moore, S.S., Fortes, M.R.S., Nguyen, L.T., Venus, B., Diaz, I.S.D.P., Souza, F.R.P., Fonseca, L.F.S., Baldi, F., Albuquerque, L.G., Thomas, M.G., Oliveira, H.N., 2017. SNP detection using RNA-sequences of candidate genes associated with puberty in cattle. Genetics Mol. Res. 16, 1.

Diskin, M.G., Morris, D.G., 2008. Embryonic and early foetal losses in cattle and other ruminants. Reprod Domest Anim; 43 (2), 260–267.

Dorfman, M.D., Garcia-Rudaz, C., Alderman, Z., Kerr, B., Lomniczi, A., Dissen, G.A., Castellano, J.M., Garcia-Galiano, D., Gaytan, F., Xu, B., Tena-Sempere, M., Ojeda, S. R., 2014. Loss of Ntrk2/Kiss1r signaling in oocytes causes premature ovarian failure. Endocrinology 155 (8), 3098–3111.

Duque-Afonso, J., Feng, J., Scherer, F., Lin, C.H., Wong, S.H.K., Wang, Z., Iwasaki, M., Cleary, M.L., 2015. Comparative genomics reveals multistep pathogenesis of E2A-PBX1 acute lymphoblastic leukemia. JCI 125 (9).

Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conc. Genet. Resour; 4, 359–361.

Edwards, C.J., Ginja, C., Kantanen, J., Peŕez-Pardal, L., Tresset, A., et al., 2011. Dual Origins of Dairy Cattle Farming - Evidence from a Comprehensive Survey of European Y-Chromosomal Variation. PLoS ONE 6 (1), e15922.

Ezzat Ahmed, A., Saito, H., Sawada, T., Yaegashi, T., Yamashita, T., Hirata, T., Sawai, K., Hashizume, T., 2009. Characteristics of the Stimulatory Effect of kisspeptin-10 on the Secretion of Luteinizing Hormone, Follicle-Stimulating Hormone and Growth Hormone in Prepubertal Male and Female Cattle. J Reprod Dev 55 (6), 650–654.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol; 14, 2611–2620.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res; 10, 564–567.

Fernández, J.C., Pérez, J.E., Herrera, N., Martínez, R., Bejarano, D., Rocha, J.F., 2019. Genomic association study for age at first calving and calving interval in Romosinuano and Costeño con Cuernos cattle. Genetic. Mol. Res. 18 (2).

Forde, N., Mehta, J.P., Minten, M., Crowe, M.A., Roche, J.F., Spencer, T.E., Lonergan, P., 2012. Effects of low progesterone on the endometrial transcriptome in cattle. Biol. Reproduct.; 87 (5), 1–11, 124.

Gaytan, F., Garcia-Galiano, D., Dorfman, M.D., Manfredi-Lozano, M., Castellano, J.M., Dissen, G.A., Ojeda, S.R., Tena-Sempere, M., 2014. Kisspeptin receptor haplo-insufficiency causes premature ovarian failure despite preserved gonadotropin secretion. Endocrinology 155 (8), 3088–3097.

Goy, M.F., 1991. cGMP: the wayward child of the cyclic nucleotide family. Trends Neurosci 14, 293–298.

Grealy, M., Glynn, M.A., Sreenan, J.M., 1997. Cyclic AMP and cyclic GMP concentrations in, and efflux from, preimplantation cattle embryos Animal Reproduction. Science 48, 175–185.

Hatzirodos, N., Irving-Rodgers, H.F., Hummitzsch, K., Rodgers, R.J., 2014a. Transcriptome profiling of the theca interna from bovine ovarian follicles during Atresia. PLoS ONE 9 (6), e99706.

Hatzirodos, N., Hummitzsch, K., Irving-Rodgers, H.F., Harland, M., Morris, S.E., Rodgers, R.J., 2014b. Transcriptome profiling of granulosa cells from bovine ovarian follicles during atresia. BMC Genomics (40), 15.

Ji, Y., Han, Z., Shao, L., Zhao, Y., 2016. Ultrasound-targeted microbubble destruction of calcium channel subunit α 1D siRNA inhibits breast cancer via G protein-coupled receptor 30. Oncol. Rep. 36, 1886–1892.

Kim, S.K., Selleri, L., Lee, J.S., Zhang, A.Y., Gu, X., Jacob, Y., Cleary, M.L., 2002. Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus. Nature Genetics; 30, 430–435.

Kirkpatrick, B.W., Byla, B.M., Gregory, K.E., 2000. Mapping quantitative trait loci for bovine ovulation rate. Mammalian Genome 11, 136–139.

Kirkpatrick, B.W., Morris, C.A., 2015. A major gene for bovine ovulation rate. PLOS ONE; 10, e0129025.

Kiser, N.J., Keuter, E.M., Seabury, C.M., Neupane, M., Moraes, J.G.N., Dalton, J., Burns, G.W., Spencer, T.E., Neibergs, H.L., 2019. Validation of 46 loci associated with female fertility traits in cattle. BMC Genomics 20, 576.

Lee, S.J., McPherron, A.C., 2001. Regulation of myostatin activity and muscle growth. PNA Sci 98, 9306–9311.

León, S., Fernadois, D., Sull, A., Sull, J., Calder, M., Hayashi, K., Bhattacharya, M., Power, S., Vilos, G.A., Angelos, G., Vilos, A.G., Tena-Sempere, M., Babwah, A.V., 2016. Beyond the brain - Peripheral kisspeptin signaling is essential for promoting endometrial gland development and function. Nat.Sci. Reports 6, 29073.

Liu, A., Wang, Y., Sahana, G., Zhang, Q., Liu, L., Sandø Lund, M., Su, G., 2017. Genome-wide association studies for female fertility traits in Chinese and Nordic Holsteins. Sci. Rep. 7, 8487.

Lopez-Bergami, P., Lau, E., Ronai, Z., 2010. Emerging roles of ATF2 and the dynamic AP1 network in cancer. Nat. Rev. Cancer 10, 65–76.

Mamo, S., Mehta, J.P., Forde, N., McGettigan, P., Lonergan, P., 2012. Conceptus-endometrium crosstalk during maternal recognition of pregnancy in cattle. Biol. of Reprod 87 (1), 1–9.

Marszalek, J.R., Ruiz-Lozano, P., Roberts, E., Chien, K.R., Goldstein, L.S., 1999. Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. PNASci 96, 5043–5048.

Medeiros de Oliveira Silva, R., Bonvino Stafuzza, N., de Oliveira Fragomeni, B., Miguel Ferreira de Camargo, G., Matos Ceacero, T., Noely dos Santos GoncËalves Cyrillo, J., et al., 2017. Genome-Wide Association Study for Carcass Traits in an Experimental Nelore Cattle Population. PLoSONE 12 (1), e0169860.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., Lee, D.H., 2002. BLUPF90 and related programs (BGF90). In: Proc. 7th World Cong. Gen. Appl. Livest. Prod., 28 Communication, Montpellier, France, pp. 21–22.

Misztal, I., Legarra, A., Aguilar, I., 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J. Dairy Sci. 92, 4648–4655.
Misztal, I., Samuel, E.A., Muir, W.M., 2013. Experiences with a single-step genome evaluation. Poult. Sci. 92, 2530–2534.
Naniwa, Y., Nakatsukasa, K., Setsuda, S., Oishi, S., Fujii, N., Matsuda, F., Uenotama, Y., Tsukamura, H., Maeda, K., Ohkura, S., 2013. Effects of Full-Length Kisspeptin Administration on Follicular Development in Japanese Black Beef Cows. J. Reprod. Dev. 59, 588–594.
NAV (Nordic Cattle Genetic Evaluation), 2021. NAV Routine Genetic Evaluation of Dairy Cattle – Data and Genetic Models. Accessed Aug. 18, 2021. https://www.nordicebv.info/.pdf.
Neupane, M., Geary, T.W., Kiser, J.N., Burns, G.W., Hansen, P.J., Spencer, T.E., Neibergs, H.L., 2017. Loci and pathways associated with uterine capacity for pregnancy and fertility in beef cattle. PLoS ONE 12 (12), e0188997.
Nyman, S., Johansson, K., de Koning, D.J., Berry, D.P., Veerkamp, R.F., Wall, E., Berglund, B., 2014. Genetic analysis of atypical progesterone profiles in Holstein-Friesian cows from experimental research herds. J. Dairy Sci. 97, 7230–7239.
Nyman, S., Gustafsson, H., Berglund, B., 2018. Extent and pattern of pregnancy losses and progesterone levels during gestation in Swedish Red and Swedish Holstein dairy cows. Acta Vet Scand 60, 68.
Nyman, S., Duchemin, S.I., de Koning, D.J., Berglund, B., 2019. Genome-wide association study of normal and atypical progesterone profiles in Holstein-Friesian dairy cows. J. Dairy Sci. 102, 3204–3215.
Oliveira, H.R., Lourenco, D.A.L., Masuda, Y., Misztal, I., Tsuruta, S., Jamrozik, J., Brito, L.F., Silva, F.F., Cant, J.P., Schenkel, F.S., 2019. Single-step genome-wide association for longitudinal traits of Canadian Ayrshire, Holstein, and Jersey dairy cattle. J. Dairy Sci. 102 (11), 9995–10011.
Ortega, M.S., 2018. Identification of genes associated with reproductive function in dairy cattle. In: Proceedings of the 10th International Ruminant Reproduction Symposium (IRRS 2018). Foz do Iguaçu, PR, Brazil. September 16th to 20th, 2018.
Ortega, M.S., Denicol, A.C., Cole, J.B., 2017. 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. 100 (5), 3725–3734.
Parborell, F., Abramovich, D., Tesone, M., 2008. Intrabursal Administration of the Antiangiopoietin 1 Antibody Produces a Delay in Rat Follicular Development Associated with an Increase in Ovarian Apoptosis Mediated by Changes in the Expression of BCL2 Related Genes. Biol. Reprod 78, 506–513.
Parker-Gaddis, K.L., Null, D.J., Cole, J.B., 2016. Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. J. Dairy Sci. 99 (8), 6420–6435.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics; 155, 945–959.
Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., et al., 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 9, 1–9, 202010.1093/gigascience/giaa021.
Roxström, A., Strandberg, E., Berglund, B., Emanuelson, U., Philipsson, J., 2001. Genetic and environmental correlations among female fertility traits and the ability to show oestrus, and milk production. Acta Agric. Scand. Sect. A. Animal Sci. 51, 192–199.
Salilew-Wondim, D., Hölker, M., Rings, F., Ghanem, N., Ulas-Cinar, M., Peippo, J., et al., 2010. Bovine pretransfer endometrium and embryo transcriptome fingerprints as predictors of pregnancy success after embryo transfer. Physiol. Genomics 42, 201–218.
Sangsritavong, S., Combs, D.K., Sartori, R., Armentano, L.E., Wiltbank, M.C., 2002. High feed intake increases liver blood flow and metabolism of progesterone and estradiol-17beta in dairy cattle. J Dairy Sci; 85, 2831–2842.
Sargolzaei, M., Chesnais, J.P., Schenkel, F.S., 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15, 478.
Satheesh, P., Bhattacharya, T.K., Kumar, P., Chatterjee, R.N., Dhara, S.K., Paswan, C., Shukla, R., Dushyanth, K., 2016. Gene expression and silencing of activin receptor type 2A (ACVR2A) in myoblast cells of chicken. Brit. Poultry Sci. 57 (6), 763–770.
Schulz, L.C., Bahr, J.M., 2003. Glucose-6-phosphate isomerase is necessary for embryo implantation in the domestic ferret. PNAS 100 (14), 8561–8566.
Shin, J., Prescott, M., Mair, J., Campbell, R.E., 2014. Roles for Primary Cilia in Gonadotrophin-Releasing Hormone Neurones in the Mouse. J. Neuroendocrinol. 26, 18–25.
Song, B.S., Yoon, S.B., Kim, J.S., Sim, B.W., Kim, Y.H., Cha, J.J., Choi, S.A., Min, H.K., Lee, Y., Huh, J.W., et al., 2012. Induction of autophagy promotes preattachment development of bovine embryos by reducing endoplasmic reticulum stress. Biol. Reprod 87 (8), 1–11.
Song, Z.H., Yu, H.Y., Wang, P., Mao, G.K., Liu, W.X., Li, M.N., Wang, H.N., Shang, Y.L., Liu, C., Xu, Z.L., et al., 2015. Germ cell-specific Atg7 knockout results in primary ovarian insufficiency in female mice. Cell Death. Dis. 6, e1589.
Stranden, I., Garrick, D.J., 2009. Technical note: derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. J. Dairy Science 92, 2971–2975.
Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., von Mering, C., 2019. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47 (D1), D607–D613.
Tahir, M.S., Nguyen, L.T., Schulz, B.L., Boe-Hansen, G.A., Thomas, M.G., Moore, S.S., Lau, L.Y., Fortes, M.R.S., 2019. Proteomics Recapitulates Ovarian Proteins Relevant to Puberty and Fertility in Brahman Heifers (Bos indicus L.). Genes (Basel) 10 (11), 923.
Tarekegn, G.M., Gullstrand, P., Strandberg, E., Båge, R., Rius-Vilarrasa, E., Christensen, J.M., Berglund, B., 2019. Genetic parameters of endocrine fertility traits based on in-line milk progesterone profiles in Swedish Red and Holstein dairy cows. J Dairy Sci 102 (12), 11207–11216.
Tenghe, A.M.M., Bouwman, A.C., Berglund, B., Strandberg, E., Blom, J.Y., Veerkamp, R. F., 2015. Estimating genetic parameters for fertility in dairy cows from in-line milk progesterone profiles. J. Dairy Sci. 98, 5763–5773 https://doi.org/10.3168/jds.2014-8732.
Tenghe, A.M.M., Bouwman, A.C., Berglund, B., Strandberg, E., de Koning, D.J., Veerkamp, R.F., 2016. Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. J Dairy Sci (7), 99.
Teyssèdre, S., Dupuis, M.C., Guérin, G., Schibler, L., Denoix, J.M., Elsen, J.M., Ricard, A., 2012. Genome-wide association studies for osteochondrosis in French Trotter horses. J Anim. Sci. 90, 45–53.
Ticianelli, J.S., Emanuelli, I.P., Satrapa, R.A., Castilho, A.C.S., Loureiro, B., Sudano, M.J., Fontes, P.K., Pinto, R.F.P., Razza, E.M., Surjus, R.S., Sartori, R., Assumpçã, M.E.O.A., Visintin, J.A., Barros, C.M., Paula-Lopes, F.F., 2017. Gene expression profile in heat-shocked Holstein and Nelore oocytes and cumulus cells. Reprod. Fertility Dev. 29, 1787–1802.
Valcarce, D.G., Cartón-García, F., Herráez, M.P., Robles, V., 2013. Effect of cryopreservation on human sperm messenger RNAs crucial for fertilization and early embryo development. Cryobiology 67, 84–90.
VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. J Dairy Sci 91, 4414–4423.
VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F., et al., 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92, 16–24, 16±24.
Vitezica, Z.G., Aguilar, I., Misztal, I., Legarra, A., 2011. Bias in genomic predictions for populations under selection. Genet Res (Camb) 93 (5), 357–366.
Walker, L., Thompson, D., Easton, D., Ponder, B., Ponder, M., Frayling, I., et al., 2006. A prospective study of neurofibromatosis type 1 cancer incidence in the UK. Br. J. Cancer 95 (2), 233–238.
Wang, H., Misztal, I., Aguilar, I., Legarra, A., Muir, W.M., 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res (Camb) 94 (2), 73–83.
Wang, H., Misztal, I., Aguilar, I., Legarra, A., Fernando, R., Vitezica, Z., Okimoto, R., Wing, T., Hawken, R., Muir, W.M., 2014. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front Genet 5 (2).
Wang, L., Shao, Z., Chen, S., Shi, L., Li, L., 2017. A SLC24A2 gene variant uncovered in pancreatic ductal adenocarcinoma by whole exome sequencing. Tohoku J. Exp. Med. 241, 287–295.
Wigglesworth, K., Lee, K.B., Emori, C., Sugiura, K., Eppig, J.J., 2015. Transcriptomic diversification of developing cumulus and mural granulosa cells in mouse ovarian follicles. Biol of Repr 92 (1), 23, 1–14.
Wiltbank, M.C., Souza, A.H., Carvalho, P.D., Bender, R.W., Nascimento, A.B., 2011. Improving fertility to timed artificial insemination by manipulation of circulating progesterone concentrations in lactating dairy cattle. Reprod Fertil Dev; 24, 238–243.
Worku, T., Wang, K., Ayers, D., Wu, D., Rehman, Z.U., Zhou, H., Yang, L., 2018. Regulatory roles of ephrinA5 and its novel signaling pathway in mouse primary granulosa cell apoptosis and proliferation. Cell cycle 17 (7), 892–902.
Yang, S., Wang, S., Luo, A., Ding, T., Lai, Z., Shen, W., Ma, X., Cao, C., Shi, I., Jiang, J., Rong, F., Ma, I., Tian, Y., Du, X., Lu, Y., Li, Y., Wang, S., 2013. Expression patterns and regulatory functions of microRNAs during the initiation of primordial follicle development in the neonatal mouse ovary. Biol. Reprod. 89 (126), 1–11.
Yu, M., Du, G., Xu, Q., Huang, Z., Huang, X., Qin, Y., Han, L., Fan, Y., Zhang, Y., Han, Y., Jiang, Z., Xia, Y., Wang, X., Lu, C., 2018. Integrated analysis of DNA methylome and transcriptome identified CREB5 as a novel risk gene contributing to recurrent pregnancy loss. EBioMedicine 35, 334–344.
Zielak-Steciwko, A.E., Browne, J.A., McGettigan, P.A., Gajewska, M., Dzie¸cioł, M., Szulc, T., Evans, A.C.O., 2014. Expression of microRNAs and their target genes and pathways associated with ovarian follicle development in cattle. Physiol Genomics 46, 735–745.