Ann. Anim. Sci., Vol. 20, No. 3 (2020) 825–851 *DOI: 10.2478/aoas-2020-0031*

sciendo

GENOME-WIDE ASSOCIATION STUDY AND PATHWAY ANALYSIS FOR FEMALE FERTILITY TRAITS IN IRANIAN HOLSTEIN CATTLE

Ali Mohammadi[1], Sadegh Alijani[2♦], Seyed Abbas Rafat[2], Rostam Abdollahi-Arpanahi[3]

[1]Department of Genetics and Animal Breeding, University of Tabriz, Tabriz, Iran
[2]Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
[3]Department of Animal Science, University College of Abureyhan, University of Tehran, Tehran, Iran
♦Corresponding author: Sad-ali@tabrizu.ac.ir

Abstract

Female fertility is an important trait that contributes to cow's profitability and it can be improved by genomic information. The objective of this study was to detect genomic regions and variants affecting fertility traits in Iranian Holstein cattle. A data set comprised of female fertility records and 3,452,730 pedigree information from Iranian Holstein cattle were used to predict the breeding values, which were then employed to estimate the de-regressed proofs (DRP) of genotyped animals. A total of 878 animals with DRP records and 54k SNP markers were utilized in the genome-wide association study (GWAS). The GWAS was performed using a linear regression model with SNP genotype as a linear covariate. The results showed that an SNP on BTA19, ARS-BFGL-NGS-33473, was the most significant SNP associated with days from calving to first service. In total, 69 significant SNPs were located within 27 candidate genes. Novel potential candidate genes include *OSTN*, *DPP6*, *EphA5*, *CADPS2*, *Rfc1*, *ADGRB3*, *Myo3a*, *C10H14orf93*, *KIAA1217*, *RBPJL*, *SLC18A2*, *GARNL3*, *NCALD*, *ASPH*, *ASIC2*, *OR3A1*, *CHRNB4*, *CACNA2D2*, *DLGAP1*, *GRIN2A* and *ME3*. These genes are involved in different pathways relevant to female fertility and other characteristics in mammals. Gene set enrichment analysis showed that thirteen GO terms had significant overrepresentation of genes statistically associated with female fertility traits. The results of network analysis identified *CCNB1* gene as a hub gene in the progesterone-mediated oocyte maturation pathway, significantly associated with age at first calving. The candidate genes identified in this study can be utilized in genomic tests to improve reproductive performance in Holstein cattle.

Key words: cattle, candidate gene, GWAS, gene set enrichment

Reproductive performance is of great importance to profitability in dairy industry. There is an unfavorable genetic correlation between fertility and milk production traits (Kadarmideen et al., 2000; Royal et al., 2002; Walsh et al., 2011). Dairy breeding programs and management strategies focus on milk production which leads to a decrease in the fertility rate (Kulak et al., 1997; Shook, 2006). Undesirable reproductive performance increases calving interval, culling rate and replacement costs

(Ghiasi et al., 2011). Recent studies have shown that the genetic progress for female fertility traits in Iran is still unsatisfying, which might be due to the lack of a national selection breeding objective for these traits (Eghbalsaied, 2011; Ghiasi et al., 2011; Toghiani, 2012; Seyed Sharifi et al., 2017).

Female fertility traits are less responsive to genetic progress through selective breeding due to their low heritability. However, several single nucleotide polymorphism (SNP) panels are available for cattle; therefore, genome-wide association study (GWAS) is the alternative approach to finding genomic regions associated with low heritability traits, such as female fertility traits (Ball, 2005; Goddard and Hayes, 2009). Through identifying the loci associated with trait and understanding its related genes, GWAS could be used to increase conception rate and early pregnancy (Moore et al., 2016; Ortega et al., 2017).

Over the last decade, GWAS has been used as a primary strategy to detect QTL for complex traits (Klein et al., 2005). As a result, many studies have been conducted using GWAS on female fertility traits in cattle (Olsen et al., 2011; Minozzi et al., 2013; Höglund et al., 2015; Nayeri et al., 2016, 2017). Many candidate genes have been identified by genome-wide association studies including *IL6R*, *CACNB2*, *MAN1A2*, *FAM46C*, *ZNF521* and *SLC39A12* for interval between first and last insemination (IFL), *CACNB2*, *CHN2*, *FAM181A* and *TMEM241* for days from calving to first service (DFS), and *NAMPT*, *PIK3CG* and *SLC4A11* for conception rate at first insemination, *ZNF521* for number of inseminations per conception (NS) and *FAM181A* for days open (DO) (Reverter and Fortes, 2013; Nayeri et al., 2016; Frischknecht et al., 2017; Liu et al., 2017; Nayeri et al., 2017). Moreover, several quantitative and QTL regions associated with female fertility traits have been reported in previous studies (Ashwell et al., 2004; Schnabel et al., 2005; Kolbehdari et al., 2009; Pimentel et al., 2011; Nayeri et al., 2017; Cai et al., 2019). Recently, the de-regressed proof (DRP) of female fertility traits has been used as pseudo phenotypes for GWAS on fertility traits of dairy cattle (Liu et al., 2017; Nayeri et al., 2017). Evaluations have moved away from a complicated model to pseudo-observations through the use of a de-regression that can be utilized under a simpler model (Sullivan et al., 2015).

Although, a few GWAS have been conducted on fertility traits in Iranian breeds of sheep including litter size (Abdoli et al., 2018), composite reproductive traits (Abdoli et al., 2019 a) and first lambing age and lambing interval (Abdoli et al., 2019 b). To date, there is no study on GWAS for female fertility traits in Iranian Holstein cattle. Therefore, the aims of this study were: i) to perform GWAS using DRP in order to identify the significant SNPs for female fertility traits in Iranian Holstein cattle, and ii) to conduct a system biology analysis such as gene ontology/pathway and gene network analysis for female fertility traits in Holstein cattle.

Material and methods

Phenotypic data

Pedigree and phenotypic records for insemination and calving dates of Iranian Holstein dairy cattle were obtained from Animal Breeding Center of Iran. Female

fertility traits of heifers and cows (lactation 1 to 3), collected from 1992 to 2018 were used in this study. Summary statistics concerning the pedigree file are presented in Table 1.

Table 1. Brief structure of the pedigree and number of individuals

Information	No. of records
No. of individuals in total	3,452,730
No. of inbred animals in total	2,236,808
No. of sires in total	85,742
No. of dams in total	1,616,057
No. of individuals with progeny	1,701,799
No. of individuals with no progeny	1,750,931
No. of founders	452,122

The following traits were analyzed: age at first service (AFS), age at first calving (AFC), days open (DO), calving interval (CI), gestation length (GL), pregnancy rate (PR), interval between first and last insemination (IFL), days from calving to first service (DFS) and number of services per conception (NS). Due to behavior of these traits, records of AFS and AFC were available only for heifers; GL, IFL and NS traits for both heifers and cows, while all other traits only for cows. Data editing and calculation of the phenotypes were based on Jamrozik et al. (2005) and Ghiasi et al. (2011). The GLM procedure of SAS Version 9.4 (SAS Institute, 2014) was employed to decide which fixed effects should be included in the animal model. The statistical models used for estimation of (Co)variance components and breeding values is shown in Table 2. Female fertility traits (DO, CI, PR, GL, IFL and DFS) measured in three consequent lactations were treated as repeated measurements. It should be noted that we used a classical animal model (Pedigree-BLUP) for estimation of breeding values. (Co)variance components and breeding values were estimated using the Restricted Maximum Likelihood (REML) methodology by BLUPF90 software (Misztal, 2002) based on single-trait analysis.

Table 2. The statistical models used for estimation of (Co)variance components and breeding values for female fertility traits in Iranian Holstein cattle

Trait	Statistical model
1	2
AFS_H	$Y_{tijk} = H_t + YbSb_i + YS_{sj} + a_k + e_{tijk}$
AFC_H	$Y_{tijk} = H_t + YbSb_i + YS_{sj} + a_k + e_{tijk}$
DO[a]	$Y_{tijklmn} = H_t + P_i + Yse_j + MS_k + Age_l + a_m + pe_n + e_{tijklmn}$
GL[a]	$Y_{tijkl} = H_t + P_i + YS_{sj} + a_k + pe_l + e_{tijkl}$
GL_H	$Y_{tijk} = H_t + YbSb_i + YS_{sj} + a_k + e_{tijk}$
PR[a]	$Y_{tijklm} = H_r + P_i + YC_j + MC_k + a_l + pe_m + e_{tijklm}$
IFL_H	$Y_{tijk} = H_t + YbSb_i + MFI_j + a_k + e_{tijk}$
IFL	$Y_{tijklm} = H_t + P_i + YSC_j + Agepc_k + MHI_l + a_m + e_{tijklm}$
IFL[a]	$Y_{tijlk} = H_t + P_i + YSC_j + MS_k + Age_l + a_k + e_{tijk}$

Table 2 – contd.

1	2
CI[a]	$Y_{tijklmn} = H_t + P_i + YSC_j + MS_k + Age_l + a_m + pe_n + e_{tijklmn}$
DFS	$Y_{tijk} = H_t + Agepc_i + YSC_j + a_k + e_{tijk}$
DFS[a]	$Y_{tijlk} = H_t + P_i + YSC_j + Age_l + a_k + e_{tijk}$
NS	$Y_{tijkl} = H_t + YSCp_i + Agepc_j + MFI_k + a_l + e_{tijkl}$
NS_H	$Y_{tijk} = H_t + YbSb_i + MFI_j + a_k + e_{tijk}$

[a]Female fertility traits with "a" suffix include first three lactations together and were treated as repeated measurements in the statistical analysis.

Where: y: the vector of observations; fixed effects include: H_t: herd; P_i: parity; $YbSb_i$: year-season of birth; YSs_j: year-season of insemination; YSC_j: year-season of calving; MS_k: month of insemination; Age_l: age at calving (month); YC_j: year of calving; $YSCp_i$: previous year-season of calving; *Agepc*: previous month of calving; *MFI*: months of first insemination. *a*, *pe* and *e* include additive genetic $a \sim N(0, A\sigma^2_a)$, permanent environment $pe \sim N(0, I\sigma^2_{pe})$, and residual $a \sim N(0, I\sigma^2_e)$ random effects, respectively.

Genotype imputation

Since the animals were genotyped with different SNP panels by different companies (Illumina and GeenSeek Genomic), genotype imputation was conducted for all the genotyped individuals using the FImpute software (Sargolzaei, 2014). According to the manual of FImpute software, the maximum number of chips allowed for imputation is 10 (Sargolzaei, 2014). Therefore, 10 SNP panels with the largest number of genotyped animals were selected and a chip with a density of 54001 was considered as the reference chip. Moreover, a primary quality control was initially performed on the 54001 SNP chip, and SNPs with unspecified physical location or on chromosome X were excluded. As a result, the 54,001 SNP chip decreased to 51,185 markers. Subsequently, the 51,185 SNP chip was considered as the reference SNP chip. Following imputation, all SNP chips reached to density of 51185.

Genotype quality control

Quality control (QC) procedures were applied and SNPs were excluded from the dataset if i) SNP call rate was less than 99%, ii) minor allele frequency (MAF) was less than 1% or iii) deviance from Hardy-Weinberg Equilibrium (HWE) with a P-value was lower than 10^{-6}, and iv) individuals had more than 5% missing genotypes. The number of SNPs and animals remaining for GWAS after QC per each female fertility trait in Holstein cattle are given in Table 3. The sex of genotyped animals was male. There were differences between certain traits regarding the outcome of the QC. Ultimately, the association analysis was implemented for each trait and specific SNPs of each trait on 29 autosomes chromosome in the bovine genome.

Table 3. The number of SNPs and animals remaining for GWAS after quality control (QC)

Trait	After QC	
	Total SNPs	No. of animals
AFS_H	35119	700
AFC_H	31813	874
GL_H	33328	878
IFL_H	34991	700
NS_H	34991	700
IFL_1	34402	551
DFS_1	35967	560
NS_1	36020	560
IFL_2	39405	430
DFS_2	39485	435
NS_2	39245	425
IFL_3	42051	350
DFS_3	42171	355
NS_3	42133	351
DO	36324	555
CI	36342	554
PR	36324	555
GL	36324	555
IFL	43384	270
DFS	43529	270

H: heifer, 1: first lactation, 2: second lactation, 3: third lactation, AFS: age at first service, AFC: age at first calving, GL: gestation length, IFL: interval between first and last insemination, NS: number of services per conception, DFS: days from calving to first service, DO: days open, CI: calving interval, PR: pregnancy rate. Female fertility traits without suffix include first three lactations together and were treated as repeated measurements in the statistical analysis.

Calculating de-regressed proofs

Because an animal estimated breeding value (EBV) includes both pedigree information and daughter phenotypic performance, there is a risk that SNP would be associated on the basis of parent average (PA) rather than own performance. Therefore, de-regressed proofs (DRP) of female fertility traits were used as pseudo phenotypes for association study. The DRP was computed as follows (VanRaden, 2008):

$$DE_{prg} = \frac{Rel_{EBV}}{1 - Rel_{EBV}} - \frac{Rel_{PA}}{1 - Rel_{EBV}}$$

$$Rel_{DD} = \frac{DE_{prg}}{DE_{prg} + 1}$$

$$DRP = PA + \frac{(EBV - PA)}{Rel_{DD}}$$

where: DE_{prg} is the daughter equivalent from progeny information, Rel_{EBV}, Rel_{PA} and Rel_{DD} are the reliabilities of EBV, PA and daughters deviations (DD), respectively.

The statistical model for GWAS analyses was:

$$y_i = \beta g_i + \alpha_i + e_i$$

where: y_i is pseudo phenotype (DRP); β is the linear regression coefficient of the SNP; g_i is the SNP genotypes of the ith bull, α_i is the random additive polygenic effects; and e_i is the residual. Assumptions of the model include $a_i \sim N(0, G\sigma^2_a)$ where G is the genomic relationship matrix and σ^2_a is the polygenic additive genetic variance; and $e_i \sim N(0, R\sigma^2_e)$ where σ^2_e is the residual variance. The GWAS was conducted using a linear regression model with SNP genotype as a linear covariate in the PLINK software (Purcell et al., 2007). The data structure and descriptive statistics are summarized in Table 4.

To account for the multiple comparison problem, Bonferroni correction was applied to control the family-wise error rate (Han et al., 2009). Moreover, principal component analysis (PCA) was carried out for further illustration of genomic composition and differentiation among population structure. The quantile-quantile (QQ) plots were visualized by the distribution of observed versus expected genome-wide P-values, and manhattan plots were constructed using 'qqman' package of R software version 3.5.2 (Turner, 2014).

Table 4. Descriptive statistics of DRP of female fertility traits in Holstein dairy cattle

Trait	No. of records	Mean	Min	Max	SD
AFS_H (d)	700	–12.789	–480.947	239.622	33.469
AFC_H (d)	874	–11.673	–4506.893	958.322	160.078
GL_H (d)	878	–0.246	–11.611	8.0675	1.626
IFL_H (d)	700	8.184	–207.94	5589.803	213.281
NS_H (n)	700	0.0301	–3.163	6.589	0.344
IFL_1 (d)	551	17.516	–1398.885	7332.601	348.746
DFS_1 (d)	560	11.832	–135.029	283.264	40.734
NS_1 (d)	560	0.134	–4.061	2.109	0.302
IFL_2 (d)	430	14.019	–494.429	1393.497	94.431
DFS_2 (d)	435	9.408	–477.030	1080.046	99.473
NS_2 (n)	425	0.109	–9.481	1.882	0.607
IFL_3 (d)	350	3.309	–3009.578	950.547	190.897
DFS_3 (d)	355	2.032	–830.366	771.792	117.795
NS_3 (n)	351	0.119	–2.782	1.165	0.280
DO (d)	555	9.014	–48.903	191.747	20.684
CI (d)	555	9.593	–77.789	169.885	22.929
PR (%)	555	–0.888	–42.232	17.508	6.315
GL (d)	555	–0.511	–8.404	3.951	1.329
IFL (d)	270	2.962	–633.431	708.947	84.689
DFS (d)	270	–2.064	–506.566	64.158	43.428

d: day, n: number.

Gene Ontology, gene network and pathway enrichment analysis

The significant SNPs with P-value less than 0.05 were selected for pathway analysis. The candidate genes marked as significant SNPs or located in 250-Kb flanking regions of SNPs were extracted from Ensemble biomart (https://www.ensembl.org/biomart/) database. The gene set enrichment analysis (GO: Gene Ontology) was conducted using DAVID (https://david.ncifcrf.gov/) database. The biological process of GO category was used for gene set enrichment analysis. In addition, genes network construction was done based on functional interactions among proteins using STRING (https://string-db.org/) database. Constructed network was clustered with the k-means algorithm to define the functional modules (Panahi et al., 2019).

Results

Manhattan plots of all SNPs effects and Q-Q plots of distribution observed against expected P-value concerning female fertility traits in Holstein dairy cattle are presented in Figures 1 a, b, c, d, e and f. Significant SNPs identified for female fertility traits in Iranian Holstein cattle are shown in Table 5. The significant SNP markers were located on 26 different chromosomes (except chromosomes 3, 5 and 27). A total of 268 genes were located within or nearby (250-Kb) 69 significant SNPs in the *Bos taurus* autosomes (BTA). Notably, only one SNP (on BTA22) was associated with both AFS_H and AFC_H traits among the significant SNPs. The highest number of significant SNPs was detected for GL trait (19 SNPs). The most significant SNP located on BTA19 was associated with DFS (ARS-BFGL-NGS-33473). The results of association analysis did not detect any significant SNP for DFS_1, DFS_2, DFS_3, NS_1, NS_3, DO and CI traits. The QQ-plots showed a distribution with $\lambda_{median} = 1$ for AFC_H, DFS_1, IFL_1, IFL_2, DFS_2, NS_2, DFS_3, IFL_3 and IFL the λ_{median} was close to the expected distribution line for IFL_H, GL_H, NS_H, NS_1, NS_3, DFS, DO, PR and CI. However, a strong deviation from one was observed for AFS_H (λ_{median} =1.60) and GL (λ_{median}=1.72). The results of PCA showed that this Holstein cattle population were clustered in different groups (Supplementary file). Using the first two PCA (PCA1 and PCA2) only five groups in this study were clearly distinguishable.

A total of nine SNPs associated with AFS_H were identified. These SNPs were located on BTA1, 4, 6, 13, 21, 22, 23, 29, with the most significant SNP being Hapmap28872-BTA-162555 (on BTA1). Among the significant SNPs for AFS_H, the SNPs on BTA1, 4, 22 and 29 were within the *OSTN*, *SLC37A3*, *CACNA2D2* and ME3 genes, respectively. Moreover, nine SNPs were found to be associated with AFC_H. These SNPs were located on BTA1, 2, 4, 14, 16, 19, 20, 22 and 25. The *CACNA2D2* gene on the BTA22 was detected as significant genes for both AFS_H and AFC_H. In addition, the BTB-01278461 (on BTA4) and BFGL-NGS-117163 (on BTA13) SNPs located within *CADPS2* and *RBPJL* genes, respectively, were associated with IFL_H.

For IFL_1, six significant SNPs were identified, of which two were located on BTA19 and 26 within OR3A1 and SLC18A2 genes, respectively. For traits associated with the second lactation only four significant SNPs were found, among which only BTB-00254334 on BTA6 was associated with IFL_2 which was within EPHA5

gene. Moreover, in the third lactation, two significant SNPs were related to IFL_3 and only ARS-BFGL-NGS-98154 SNP on BTA23 was within the *ZFAND3* gene.

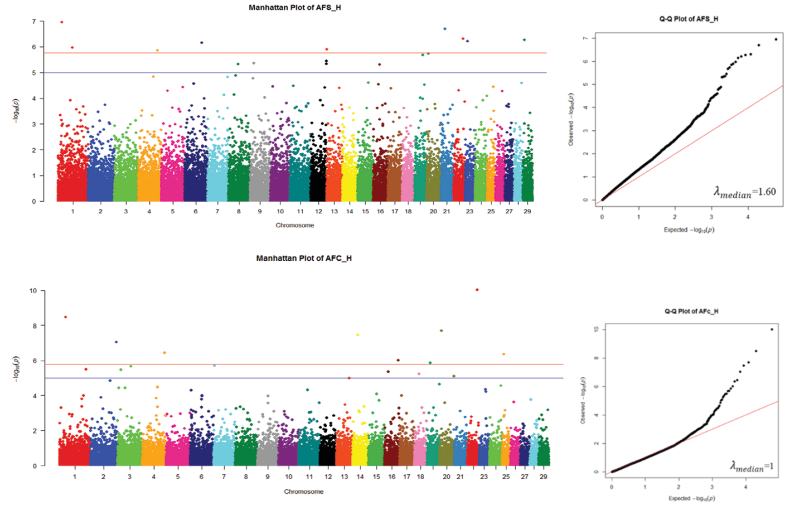

Figure 1 a. Manhattan of GWAS (-log10 P-values) and Q–Q plot for AFS_H, AFC_H traits in Holstein dairy cattle. The dots in manhattan plots represents -log10 (P-value) of the SNPs and blue and red lines denote the threshold for suggestive and significant SNPs, respectively. The red line in Q-Q Plots represents the expected values under the null hypothesis for no association

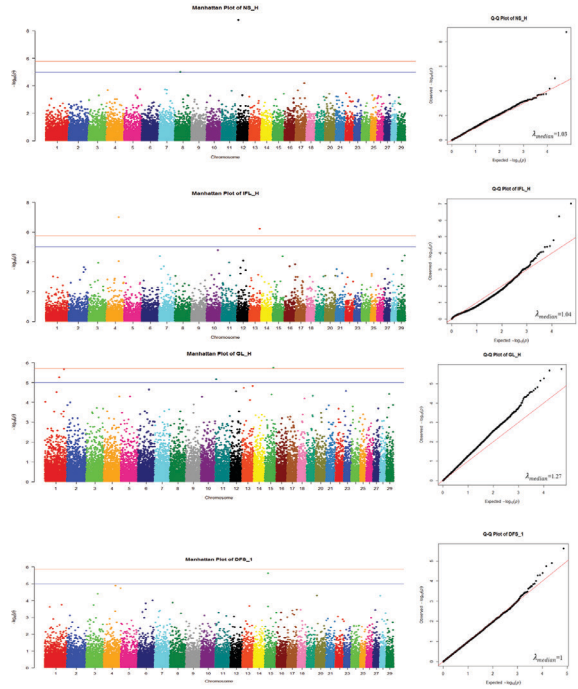

Figure 1 b. Manhattan of GWAS (-log10 P-values) and Q–Q plot for NS_H, IFL_H, GL_H, DFS_1 H traits in Holstein dairy cattle. All other symbols are as in Figure 1a

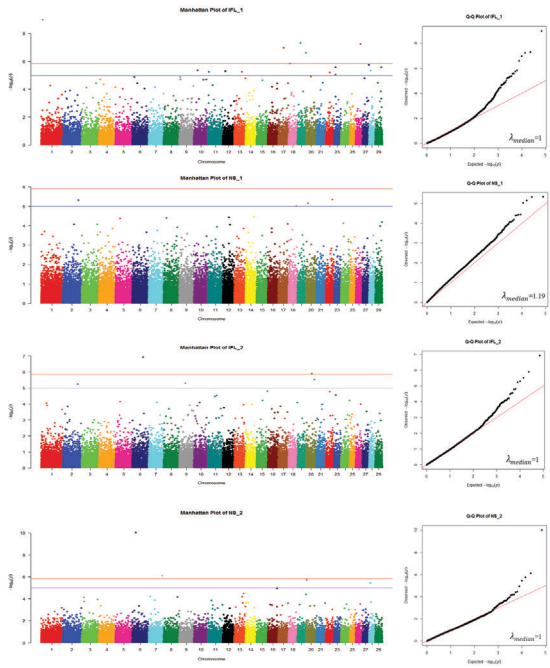

Figure 1 c. Manhattan of GWAS (-log10 P-values) and Q–Q plot for IFL_1, NS_1, IFL_2, NS_2 traits in Holstein dairy cattle. All other symbols are as in Figure 1 a

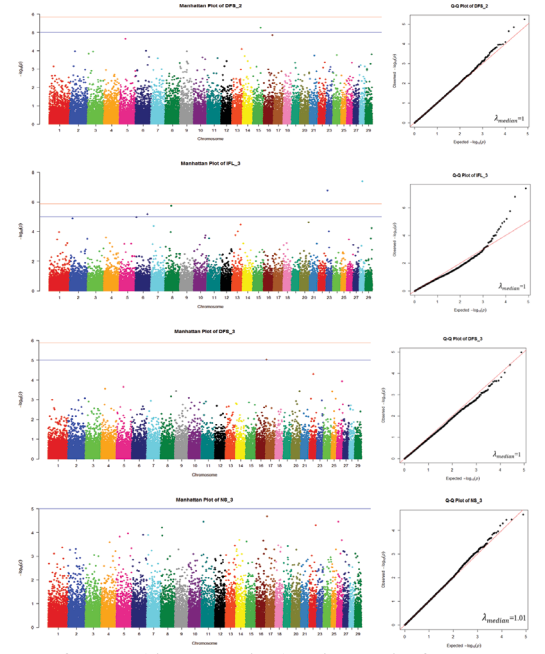

Figure 1 d. Manhattan of GWAS (-log10 P-values) and Q–Q plot for DFS_2, IFL_3, DFS_3, NS_3 traits in Holstein dairy cattle. All other symbols are as in Figure 1 a

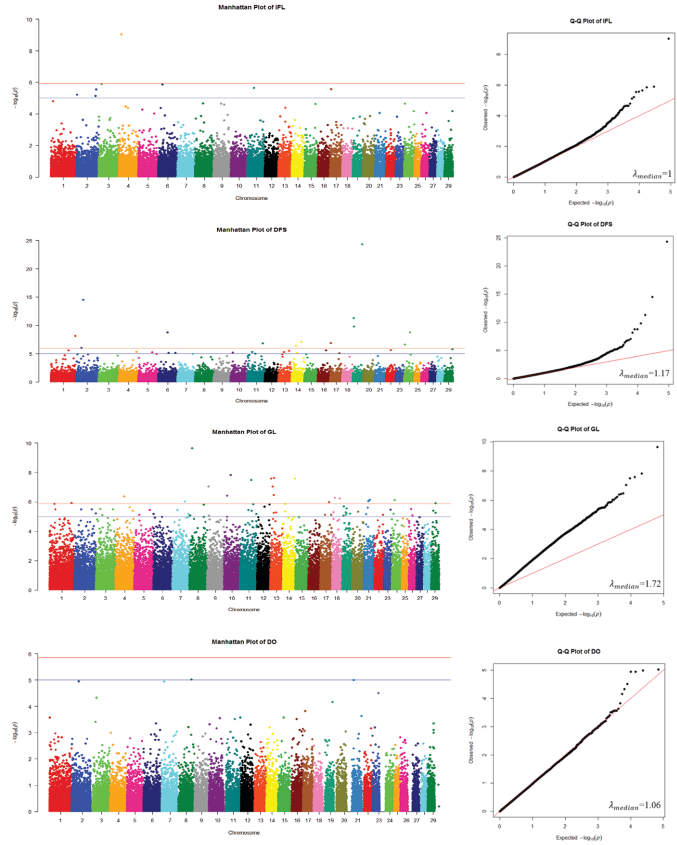

Figure 1 e. Manhattan of GWAS (-log10 P-values) and Q–Q plot for IFL, DFS, GL and DO traits in Holstein dairy cattle. All other symbols are as in Figure 1 a

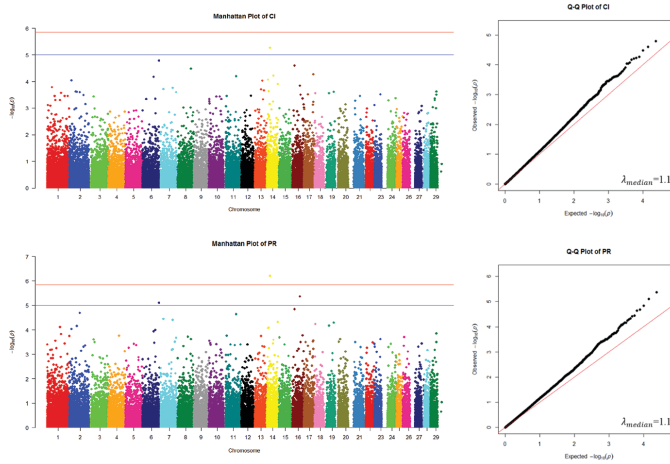

Figure 1 f. Manhattan of GWAS (-log10 P-values) and Q–Q plot for CI and PR traits in Holstein dairy cattle. All other symbols are as in Figure 1 a

Table 5. Significant SNPs identified for female fertility traits by single-marker GWAS method in Iranian Holstein dairy cattle

Nearest Gene name[4]	P-value UNADJ[3]	BONF[2]	Position	Trait[1]	SNP	BTA
1	2	3	4	5	6	7
-	1.093E-07	0.003193	20660542	AFS_H	Hapmap28872-BTA-162555	1
OSTN, UTS2B, CCDC50, ENSBTAG00000050537	1.053E-06	0.03076	76075399	AFS_H	Hapmap26468-BTA-160919	1
-	3.246E-09	9.75e-05	31296369	AFC_H	BTB-01831270	1
-	1.018E-09	3.50e-05	12099726	IFL_1	Hapmap24370-BTA-123717	1
AIRE, ENSBTAG00000014880, ENSBTAG00000014882, ENSBTAG00000007105, GATD3A, PFKL, CFAP410, TRPM2, ENSBTAG00000051579, LRRC3, TSPEAR, TRAPPC10, AGPAT3, ENSBTAG00000051900, ENSBTAG00000037662, ENSBTAG00000052773	1.22E-06	0.03938	144274055	GL	BFGL-NGS-118821	1
TBC1D5	7.23E-09	0.0003147	154495229	DFS	Hapmap48393-BTA-58266	1
PAX7, TAS1R2, KLHDC7A, ALDH4A1, IFFO2, ENSBTAG00000053388	8.9E-08	0.002672	133996973	AFC_H	Hapmap22886-BTA-49773	2
CACNB4, STAM2, ARL5A, NEB	3.25E-15	1.41e-10	44186118	DFS	BFGL-NGS-115659	2
-	9.83E-07	0.04277	32079566	DFS	Hapmap40659-BTA-116165	2
SLC37A3, RAB19, MKRN1, DENND2A, KDM7A, ADCK2, PAR912, NDUFB2	1.378E-06	0.04024	103764826	AFS_H	Hapmap23844-BTA-71970	4
DPP6	3.534E-07	0.01061	116478543	AFC_H	BFGL-NGS-118333	4
CADPS2	9.701E-08	0.00282	87117538	IFL_H	BTB-01278461	4
-	4.181E-07	0.01349	59242828	GL	ARS-BFGL-NGS-99336	4
-	9.267E-10	4.021e-05	17524805	IFL	Hapmap30840-BTA-147842	4
CXCL13, CNOT6L, ENSBTAG000000048934	6.919E-07	0.02021	92325630	AFS_H	Hapmap57216-rs29016177	6
EPHA5	1.166E-07	0.004145	80849148	IFL_2	BTB-00254334	6

Table 5 – contd.

1	2	3	4	5	6	7
ENSBTAG00000015464	9.215E-11	3.261e-06	27833403	NS_2	Hapmap50091-BTA-75608	6
RFC1, TMEM156, KLHL5, WDR19, KLB, RPL9, LIAS, UGDH, SMIM14	1.75E-09	7.63E-05	58458406	DFS	Hapmap22817-BTA-161581	6
-	7.751E-07	0.02743	100877260	NS_2	Hapmap54978-rs29019536	7
ACOT12, ZCCHC9, CKMT2, SSBP2, RASGRF2	9.39E-07	0.03029	81368696	GL	DIAS-163	7
LINGO2, ENSBTAG00000050975	2.254E-10	8.54e-06	16262390	GL	BTA-83017-no-rs	8
ADGRB3, ENSBTAG00000051256	8.98E-08	0.0029	7727243	GL	Hapmap11707-BTA-83854	9
PLEKHO2, PTGDR, PIF1, ENSBTAG00000001423, RBPMS2, ZNF609, OAZ2	1.499E-08	0.000484	45210844	GL	BTB-00420215	10
C10H14orf93, PSMB5, PSMB11, CDH24, AJUBA, HAUS4, PRMT5, C10H14orf119, MMP14, LRP10, SLC7A7, OXA1L, MRPL52, ACIN1, CEBPE, SLC7A8, BCL2L2, PPP1R3E, RBM23, REM2, HOMEZ, RNF212B, ENSBTAG00000052080, ENSBTAG00000052423						
3.818E-07	0.01232	21850545	GL	ARS-BFGL-10 NGS-59079		
DTNB, ENSBTAG00000054828, EFR3B, POMC, DNMT3A, DNAJC27, ADCY3	3.171E-08	0.001023	74106507	GL	Hapmap22700-BTA-126702	11
NOL10, ATP6V1C2, PDIA6, ODC1, HPCAL1	1.419E-06	0.0458	87087584	GL	ARS-BFGL-BAC-7336	11
GARNL3, RALGPS1, SLC2A8, ZNF79, RPL12, LRSAM1, FAM129B, STXBP1, ENSBTAG00000023823	1.57E-07	0.006842	98133723	DFS	BFGL-NGS-118335	11
ENSBTAG00000053445	8.846E-09	0.0002771	9278019	NS_H	ARS-BFGL-NGS-106546	12
PLCB1	1.246E-06	0.03639	1787076	AFS_H	BTB-01795734	13
MYO3A, GAD2, GJD4, FZD8	2.32E-08	0.0008787	26596002	GL	BTA-112222-no-rs	13
FLRT3	2.52E-08	0.000814	8138058	GL	ARS-BFGL-NGS-33115	13

YME1L1, IL2RA, FBH1, ENSBTAG00000007078, ENSBTAG00000053950, ENSBTAG0000048553, ANKRD26, MASTL, ACBD5, ABI1, PDSS1	9.08E-08	0.003439	17739132	GL	ARS-BFGL-NGS-21967	13
KIAA1217	3.448E-07	0.01113	24936426	GL	ARS-BFGL-NGS-66330	13
RBPJL, MATN4, SLPI, WFDC15B, SYS1, SDC4, TP53TG5 ENSBTAG00000014329, ENSBTAG00000046375, ENSBTAG00000037925, ENSBTAG00000051099, PI3, ENSBTAG00000039446, ENSBTAG00000020813, PIGT, SPINT3, DBNDD2, WFDC5, KCNS1, ENSBTAG00000046614, ENSBTAG00000054274, STK4	5.962E-07	0.01733	73655607	IFL_H	BFGL-NGS-117163	13
CLVS1, CHD7	3.443E-08	0.001034	26729709	AFC_H	ARS-BFGL-NGS-107117	14
PPDPFL, SNAI2, EFCAB1, ENSBTAG00000027760	6.252E-07	0.02202	20012067	PR	Hapmap30838-BTA-147387	14
ENSBTAG00000053138, RALYL	2.54E-08	0.0009604	78147166	GL	ARS-BFGL-NGS-28199	14
NCALD, GRHL2	9.09E-08	0.003958	62640592	DFS	Hapmap22954-BTA-35300	14
ASPH, CLVS1						
4.10E-07	0.01783	27059746	DFS	ARS-BFGL-NGS-27940	14	
RCN1, ENSBTAG00000052762, ENSBTAG00000052796, PAX6	1.736E-06	0.04466	62814384	GL_H	BFGL-NGS-118002	15
FLVCR1, VASH2, ANGEL2, SPATA45, TATDN3, NSL1, RPS6KC1, BATF3	9.726E-07	0.0292	70665022	AFC_H	BTB-00661215	16
ENSBTAG00000055004	1.056E-07	0.003633	39291020	IFL_1	BTA-111300-no-rs	17
KSR2, NOS1, FBXO21						
1.06E-06	0.0402	57693459	GL	ARS-BFGL-17		
DCHS2	1.43E-07	0.006243	3504880	DFS	Hapmap48326-BTA-41656	17
FAM92B, KIAA0513, ZDHHC7, CRISPLD2, GSE1	1.469E-06	0.0490	11251876	IFL_1	ARS-BFGL-NGS-22066	18
ABCC11, LONP2, SIAH1, N4BP1	5.26E-07	0.01994	16862635	GL	Hapmap39998-BTA-42612	18

Table 5 – contd.

1	2	3	4	5	6	7
DEDD2, ZNF526, GSK3A, ERF, POU2F2, ZNF547, GRIK5, PRR19, CIC, PAFAH1B3, TMEM145, ATP1A3, RABAC1, CNFN, LIPE, MEGF8, ENSBTAG00000052955, ENSBTAG00000054584, ARHGEF1, CEACAM1, ENSBTAG00000050227, ENSBTAG00000053222						
5.86E-07	0.01892	51081743	GL	Hapmap57492-ss46526490	18	
ASIC2	1.351E-06	0.04056	16879629	AFC_H	ARS-BFGL-NGS-10027	19
-	2.46E-07	0.008464	64622509	IFL_1	ARS-BFGL-NGS-108629	19
-	4.60E-25	2.00e-20	59226339	DFS	ARS-BFGL-NGS-33473	19
ANKFN1	5.16E-12	2.25E-07	6904350	DFS	Hapmap25072-BTA-149354	19
C19H17orf67, DGKE, TRIM25, COIL, SCPEP1, AKAP1, MSI2, ENSBTAG00000052235						
1.56E-10	6.79e-06	7870025	DFS	Hapmap24657-BTA-132711	19	
OR3A1, RAP1GAP2, OR1P1, ENSBTAG00000050539, ENSBTAG00000039633, ENSBTAG00000050394, ENSBTAG00000054909, ENSBTAG00000052339, OR1G1, SPATA22, ENSBTAG00000049401, ASPA, TRPV3, ENSBTAG00000052445, ENSBTAG00000049047,						
ENSBTAG00000050183, ENSBTAG00000049756, ENSBTAG00000050053, ENSBTAG00000054200, ENSBTAG00000054458, ENSBTAG00000037529, ENSBTAG00000049224, ENSBTAG00000054966, ENSBTAG00000049018, ENSBTAG00000054980, ENSBTAG00000047458	4.982E-08	1.71E-03	24025508	IFL_1	UA-IFASA-6091	19

SLC30A5, CCNB1, CENPH, MRPS36, CDK7, CCDC125	2.007E-08	0.000602	10648444	AFC_H	Hapmap54493-rs29018712	20
DROSHA, C20H5orf22, CDH6,PDZD2	1.258E-06	0.04471	42126729	IFL_2	ARS-BFGL-NGS-14869	20
G2E3, SCFD1	7.31E-07	0.02361	41098411	GL	ARS-BFGL-NGS-105488	21
C21H15orf40, RAMAC, ENSBTAG00000050976, HOMER2	1.988E-07	0.005808	23448892	AFS_H	Hapmap27015-BTA-158721	21
CHRNB4, CHRNA3, CHRNA5, PSMA4, HYKK, IREB2, UBE2Q2, ENSBTAG00000048528, FBXO22, NRG4	9.06E-07	0.02924	31005976	GL	ARS-BFGL-NGS-105996	21
CACNA2D2, ZMYND10, RASSF1, NPRL2, IFRD2, CYB561D2, HYAL2, HYAL1, HYAL3, NAA80, LSMEM2, SEMA3B, GNAI2, HEMK1, C22H3orf18, SLC38A3, GNAT1, SEMA3F, CISH, MAP-KAPK3	4.787E-07	0.01398	49958907	AFS_H	Hapmap47041-BTA-54700	22
CACNA2D2, ZMYND10, RASSF1, NPRL2, IFRD2, CYB561D2, HYAL2, HYAL1, HYAL3, NAA80, LSMEM2, SEMA3B, GNAI2, HEMK1, C22H3orf18, SLC38A3, GNAT1, SEMA3F, CISH, MAP-KAPK3	9.493E-11	2.85e-06	49958907	AFC_H	Hapmap47041-BTA-54700	22
KIF6, DAMM2, MOCS1	6.071E-07	0.01774	13656928	AFS_H	Hapmap23651-BTA-161953	23
ZFAND3, MDGA1	1.642E-07	0.006298	11740595	IFL_3	ARS-BFGL-NGS-98154	23
ENSBTAG00000011094						
7.76E-07	0.0294	17674853	GL	ARS-BFGL-24 BAC-2814		
DLGAP1, TGIF1	1.75E-09	7.63E-05	37741062	DFS	Hapmap22828-BTA-90188	24
NETO1	2.43E-07	0.01056	5054635	DFS	Hapmap36637-SCAF-FOLD318693_6319	24
GRIN2A	4.337E-07	0.01302	8631624	AFC_H	BFGL-NGS-112588	25
SLC18A2, PDZD8, KCNK18, VAX1, SHTN1	5.886E-08	0.002025	37571166	IFL_1	BTA-61460-no-rs	26
ENSBTAG00000036111	4.137E-08	0.001587	21786814	IFL_3	Hapmap27165-BTA-139884	28
ME3, ENSBTAG00000046374, CCDC81, HIKESHI, EED	5.23E-07	0.01528	9017512	AFS_H	ARS-BFGL-NGS-31793	29

[1]Female fertility traits without suffix include first three lactations together.
[2]Bonferroni adjusted P-values.
[3]Unadjusted P-value.
[4]Genes in 250-Kb flanking regions of SNP position. Boldface denotes significant SNP was located within this gene.

As shown in Table 5, a total of 35 significant SNPs were associated with the first three lactation traits (GL, DFS, IFL and PR). In total, 19 significant SNPs were observed to be associated with GL, out of which six SNPs were within the *ADGRB3*, *C10H14orf93*, *MYO3A*, *YME1L1*, *KIAA1217* and *CHRNB4* genes. Thirteen significant SNPs were identified to be associated with DFS with the most significant SNP (ARS-BFGL-NGS-33473) located on BTA19. There were no genes in the specified position, while the significant SNPs on DFS were marked by the *TBC1D5*, *CACNB4*, *RFC1*, *GARNL3*, *NCALD*, *ASPH* and *DLGAP1* genes (BTA1, 2, 6, 11, 14, 14 and 24, respectively). Additionally, one significant SNP associated with IFL located on BTA4 and one SNP associated with PR on BTA14 were found.

Gene-Set enrichment and network analysis

The summary of all significant overrepresentation GO terms regarding the studied traits are given in Table 6. Thirteen GO terms (biological process) showed a significant association with female fertility traits (*FDR* < 0.05). Interaction between the identified genes (network analysis) concerning female fertility traits in Holstein cattle is shown in Figures 2 a, b and c. The most significant biological process was mitochondrial electron transport, ubiquinol to cytochrome c (GO: 0006122, *FDR* = 8.26e-10) for DFS. Moreover, the regulation of growth term was significant for both AFS_H and AFC_H traits.

Table 6. Gene ontology (GO) term (biological process) pathways significantly enriched for genes associated with female fertility traits in Holstein dairy cattle

Trait	GO ID	Description	Count in gene set	FDR
AFS_H	GO:0040008	regulation of growth	8 of 147	0.0155
	GO:0034097	response to cytokine	6 of 194	0.0026
	GO:0071345	cellular response to cytokine stimulus	5 of 169	0.0111
AFC_H	GO:0040008	regulation of growth	8 of 147	0.0382
IFL_H	GO:0016255	attachment of GPI anchor to protein	2 of 2	0.0032
GL_H	GO:0034976	response to endoplasmic reticulum stress	4 of 67	3.54e-05
	GO:0006457	protein folding	4 of 91	5.76e-05
DFS	GO:0006122	mitochondrial electron transport, ubiquinol to cytochrome c	6 of 11	8.26e-10
	GO:0046034	ATP metabolic process	6 of 105	1.41e-05
PR	GO:0007155	cell adhesion	5 of 165	4.47e-05
	GO:0045216	cell-cell junction organization	3 of 25	0.00020
GL	GO:0007274	neuromuscular synaptic transmission	3 of 10	0.0093
IFL	GO:0050896	response to stimulus	18 of 1787	0.0041

FDR: False discovery rate.

Through the use of K-means clustering algorithm, the network was divided in three (red, green and blue) clusters. In the first cluster (red) *HYAL2* and *NPRL2*

(AFS_H), *GNAI2* and *GNAT1* (AFC_H), *DLG4* (DFS), *LATS2* (IFL), *ATP1A3* and *TMEM145* (GL), and *CTNNB1* (PR) genes were the hub genes of the network. The hub genes of the second cluster (green) were *CCNB1* and *CDK1* for AFS_H and AFC_H, *TMEM156* for DFS, ENSBTAG00000020813 for IFL, *PSMA4* for GL and *EFCAB1* and *FAM192A* for PR. Moreover, the hub genes of the third cluster (blue) were *GNAI2*, *NPRL2*, *UQCRFS1*, *NOL10* and *GPAA1* genes regarding AFS_H, AFC_H, DFS, GL and IFL, respectively.

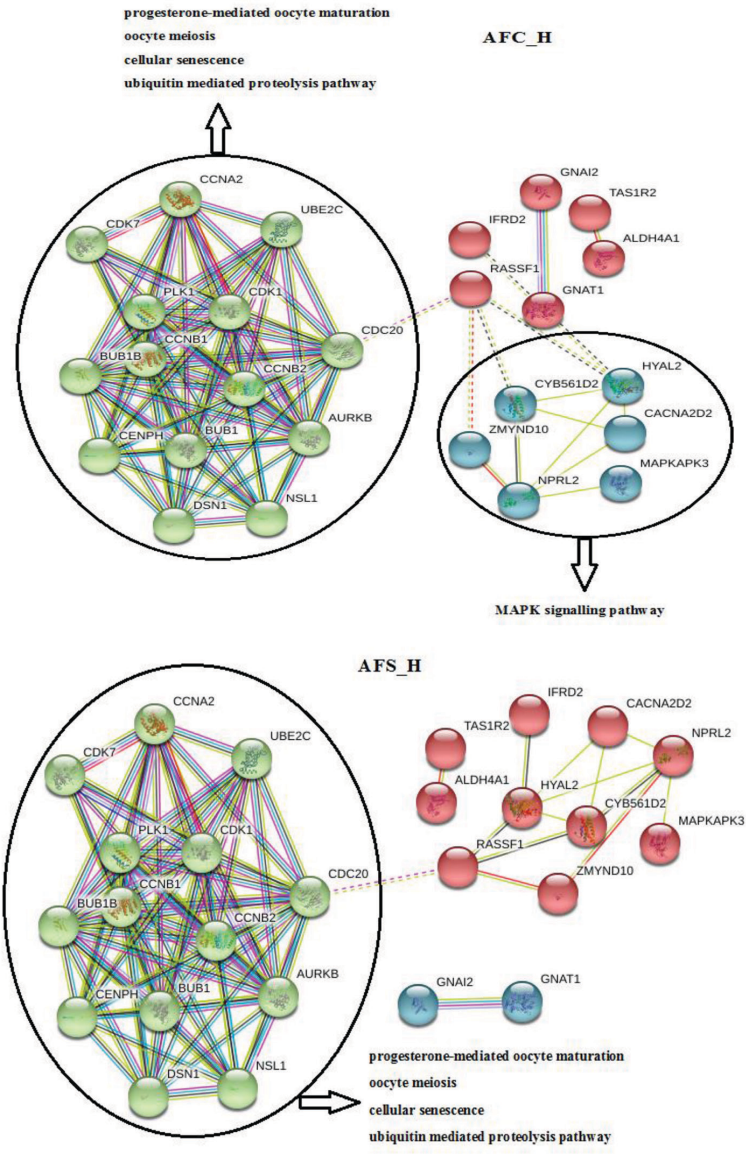

Figure 2 a. Network and interaction between the identified genes for AFS_H and AFC_H traits

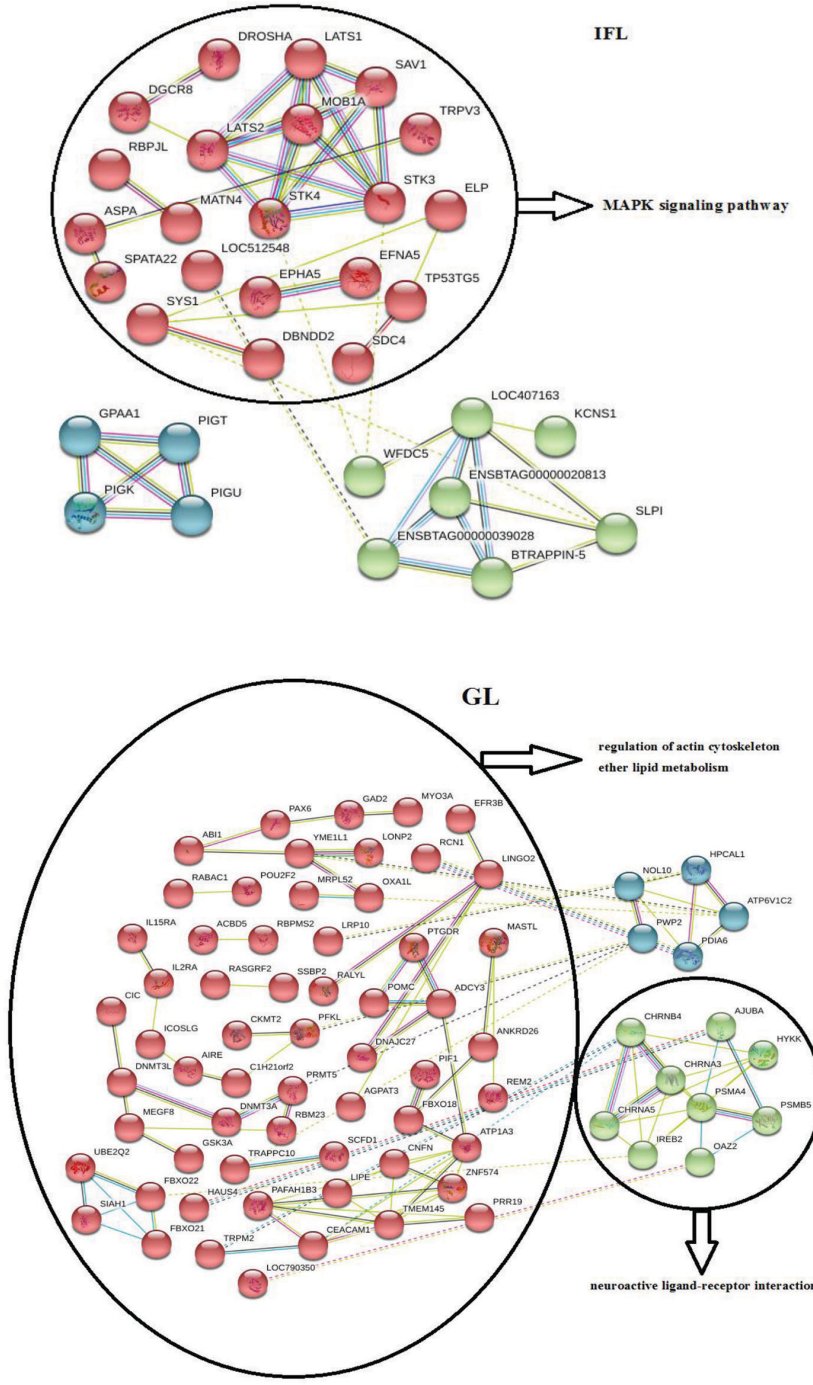

Figure 2 b. Network and interaction between the identified genes for IFL and GL traits

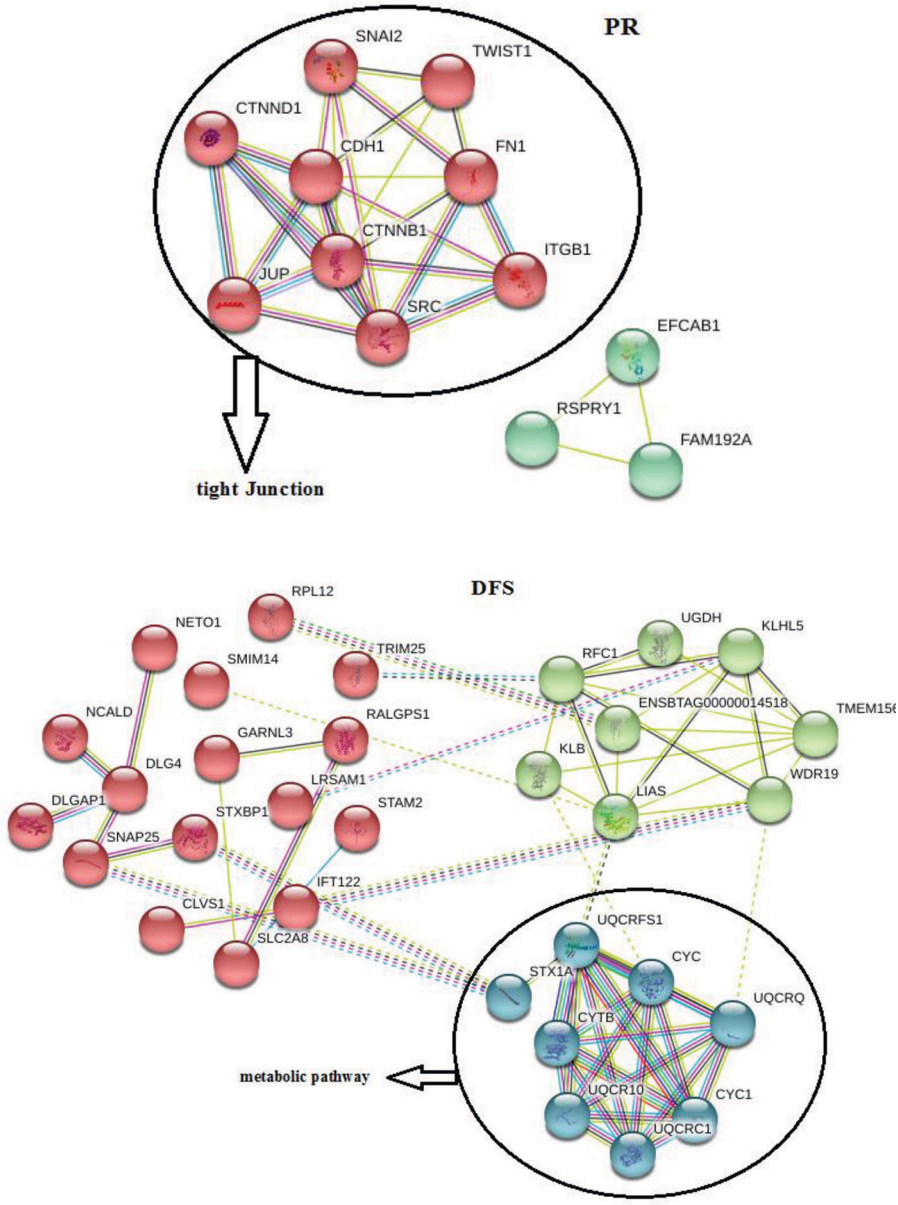

Figure 2 c. Network and interaction between the identified genes for PR and DFS traits

The annotated genes involved in the first cluster were significantly overrepresented in the regulation of actin cytoskeleton and ether lipid metabolism, MAPK signalling pathway and tight junction concerning GL, IFL and PR, respectively. The

genes of the second cluster of the network were involved in the activation of several pathways such as progesterone-mediated oocyte maturation, oocyte meiosis, cellular senescence, ubiquitin mediated proteolysis pathway and neuroactive ligand-receptor interaction. Furthermore, genes involved in the third cluster were enriched in terms of MAPK signalling pathway and metabolic pathway for AFC_H and DFS traits, respectively. The identified genes of metabolic pathway in DFS were *STX1A*, *UQCRFS1*, *CYC*, *CYTB*, *UQCRQ*, *UQCR10*, *CYC1* and *UQCRC1*.

Discussion

Fertility traits have low heritability, and identifying genomic regions affecting these traits can provide novel opportunities for improving fertility in dairy cattle through marker-assisted selection. Accordingly, the GWAS was carried out for nine fertility traits measured on Holstein heifers and cows. Further, gene set enrichment and pathway analysis were done in order to investigate the mechanisms behind the female fertility traits in dairy cattle. The observed genomic inflation factor could be attributed to factors such as sample size, LD structure and a number of causal SNPs affecting the studied traits (Yang et al., 2011; Ghasemi et al., 2019).

To the best of our knowledge, this is the first study to report the association of *OSTN*, *TBC1D5*, *PAX7*, *CACNB4*, *SLC37A3*, *DPP6*, *CADPS2*, *EPHA5*, *RFC1*, *ADGRB3*, *C10H14orf93*, *GARNL3*, *MYO3A*, *YME1L1*, *KIAA1217*, *RBPJL*, *NCALD*, *ASPH*, *ASIC2*, *OR3A1*, *CHRNB4*, *CACNA2D2*, *ZFAND3*, *DLGAP1*, *GRIN2A*, *SLC18A2* and *ME3* genes with female fertility traits in Holstein cattle. Thus, we are able to introduce these genes as candidate genes affecting female fertility traits in Holstein cattle. The significant SNP on BTA2 located within calcium voltage-gated channel auxiliary subunit beta 4 (*CACNB4*) gene is associated with DFS. Moreover, the role of *CACNB4* gene in sire conception rate was reported in US Jersey cattle (Rezende et al., 2018). A SNP at position 154 Kb on BTA1 in *TBC1* Domain Family Member 5 (*TBC1D5*) gene (for DFS) play a role in GTPase activator activity and AP-2 adaptor complex binding (https://www.genecards.org/). It has been further reported that *TBC1D5* gene is significantly associated with pregnancy status (Reverter et al., 2016). *PAX7* gene (for AFC_H) plays a critical role during fetal development (https://www.genecards.org/). The existence of *PAX7* gene, a subset of A_{single} spermatogonia functions, is necessary for maintaining fertility in normal spermatogenesis in healthy mice (Aloisio et al., 2014). The *ZFAND3* gene functions in the development of primary sexual characteristics and establishment of pregnancy (Ma et al., 2016).

Moreover, the *SLC18A2* gene encodes transmembrane protein functions as an ATP-dependent transporter of monoamines (https://www.genecards.org/). Although no functional study has been done on dissecting the role of *SLC18A2* gene in female fertility traits of Holstein cattle, it has been reported that this gene is associated with conception rate, fat yield and net merit (Cochran et al., 2013 a). Also, *SLC18A2* gene was identified as a candidate for the fertilizing ability of sperm and subsequent embryonic development in cattle (Cochran et al., 2013 b).

GO term and network analysis

The genome-wide association analysis was followed by a gene set enrichment analysis in order to identify the potential functional categories and molecular pathways related to female fertility traits. The functional term associated with the regulation of growth (GO:0040008) describes any process that modulates the frequency, rate or extent of cell growth in all or part of an organism (https://www.genecards.org/). Furthermore, GO term contains several subunits, including growth, regulation of developmental growth, developmental growth, cell growth and regulation of cell growth (Mungall et al., 2011). Response to stimulus GO term impacts the state or activity of a cell or an organism as a result of a stimulus (https://www.genecards.org/). No functional study has been done on response to stimulus in female fertility traits. However, the association of response to stimulus with milk, fat and protein yields and mastitis traits in cattle was demonstrated by Fang et al. (2017). Additionally, response to stimulus term was detected for differentially expressed genes in mammary epithelial cells in Kashmiri and Jersey cattle (Bhat et al., 2019). Certain GO terms are closely related to female fertility traits, such as response to cytokine GO term (GO:0034097), also associated with the development of testis and fertility (Loveland et al., 2017).

The *HYAL2* gene (red cluster, AFS_H), located in the proximity of the QTL reported in fertility, played a role in reproduction and production in swine (Rempel et al., 2011). The *CTNNB1* gene in the red cluster (hub gene, PR) was associated with uterine capacity for pregnancy and fertility in beef cattle (Neupane et al., 2017), milk production (Farhadian et al., 2018 a, b) and litter size (Xu et al., 2018) in sheep breeds. Besides, tight junction was the enriched pathway for PR related genes in the red cluster. These results are in accordance with Cochran et al. (2013 a) research on Holstein cattle. The *DLG4* gene was identified as a gene with the highest interaction in the red cluster for DFS trait. In the previous research, *DLG4* gene was the most influential candidate gene affecting body size in sheep and relevant mammalian phenotypes (Kominakis et al., 2017). The genes in the red cluster, concerning GL trait, enriched several pathways such as ether lipid metabolism and regulation of actin cytoskeleton. The ether lipid metabolism pathway was enriched for GL and was involved in cell signalling processes in heifer fertility (Neupane et al., 2017). Ether lipid metabolism genes were involved in chemical reactions with ether lipids (Neupane et al., 2017).

MAPK signalling pathway was activated by red cluster in IFL (for first, second and third lactations) and blue cluster in AFC_H. This pathway is related to complex cellular programs such as cell proliferation and hyperplastic growth (Chang, 2007) and residual feed intake traits (Rolf et al., 2012). Nevertheless, this pathway has not been researched in terms of female fertility traits in Holstein cattle. The proteasome subunit alpha 4 (*PSMA4*) gene is the hub gene of the second cluster (green) for GL, also known as a candidate gene for milk traits or mammary gland in cows (Ibeagha-Awemu et al., 2016).

Among the enriched pathways, progesterone-mediated oocyte maturation pathway, which includes *CCNA2*, *PLK1*, *CCNB1*, *CCNB2* and *BUB1* genes (green cluster), is associated with AFS_H and AFC_H, regulating the uterine receptivity and

maintenance of pregnancy in cattle (Fair and Lonergan, 2012) and oocyte development in sheep breeds (Wang et al., 2015). The effect of progesterone on oocyte quality and embryo development in dairy cattle has further been confirmed (Lonergan, 2011). In addition, Pimentel et al. (2011), reported that CCNB1 (Cyclin B1) gene is significantly associated with heifer and cow 56-day nonreturn rate, IFL, DFS and DO traits. The *CCNB1* and *CDK1* genes (for AFS_H and AFC_H) are hub genes in the green cluster, encoding the proteins of CyclinB1 and *CDK1* respectively. These genes are also involved in cell cycle regulation, centrosome duplication and chromosome segregation (https://www.ncbi.nlm.nih.gov/). Besides, *CDK1* gene is known as a candidate gene for lactation persistency in Canadian Holstein cattle (Do et al., 2017).

Moreover, neuroactive ligand-receptor interaction pathway in GL includes many genes such as follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone receptor, growth hormone releasing hormone receptor, growth hormone and prolactin (Xu et al., 2015). These genes play a crucial role in reproduction, lactation (Doufas and Mastorakos, 2000; Bliss et al., 2010; Breen and Knox, 2012) and mammalian growth. On the other hand, Demiray et al. (2019) reported this pathway and the related genes might play a functional role in pregnancy.

Oocyte meiosis, related to fertility traits, is a pathway in the green cluster for AFS_H and AFC_H (Do et al., 2017). Moreover, oocyte meiosis and progesterone-mediated oocyte maturation participate in reproduction via differentially expressed genes including inositol 1,4,5-trisphosphate receptor type 1, calbindin 2 (*CALM*) and progesterone receptor (Hirose et al., 2013). These pathways play functional roles in pregnancy through regulating hormonal signalling pathways (Demiray et al., 2019). Previous studies have corroborated the pivotal role of progesterone in establishing uterine receptivity and maintenance of pregnancy in cattle (Bazer et al., 2011). The cellular senescence is the other enriched pathway that plays an important role in placental and fetal development during pregnancy (Muñoz-Espín et al., 2013; Storer et al., 2013; Behnia et al., 2015; Velarde and Menon, 2016). Furthermore, Sutovsky (2003) argued that ubiquitin mediated proteolysis pathway has an important role in the quality control of spermatogenesis, fertilization and sperm in mammals. GPAA1 gene has the highest interaction with other genes in the blue cluster (IFL) associated with milk production traits in cattle (Jiang et al., 2014). The metabolic pathway had a significant role in DFS trait in the present analysis. It has been reported that this pathway functions in the reproduction performance in Nellore heifers (Costa et al., 2015).

The present study was designed to identify SNPs associated with female fertility traits in Holstein cattle. The ARS-BFGL-NGS-33473 SNP on BTA19 was detected as the strongest association regarding the studied traits. In total, significant SNPs in the current study were located within 27 genes, 21 of which were novel genes for female fertility traits. *CACNB4*, *TBC1D5*, *PAX7* and *ZFAND3* genes were introduced as candidate genes for female fertility traits in Holstein cattle. Many biological processes and pathways were further identified for female fertility. Interestingly, *CCNB1* gene is significantly associated with female fertility traits and it was detected as a hub gene in the progesterone-mediated oocyte maturation pathway. Overall, the

present GWAS identified candidate genes for fertility traits in Holstein cattle, providing new information regarding the genetic architecture of these traits for enhancing genomic improvement.

Acknowledgments

The authors thank the Animal Breeding Center of Iran for providing genotypes and data. Also, the authors are grateful to Behrouz Mohammad Nazari for his kindly help and support of the present study.

References

Abdoli R., Mirhoseini S.Z., Ghavi Hossein-Zadeh N., Zamani P., Gondro C. (2018). Genome-wide association study to identify genomic regions affecting prolificacy in Lori-Bakhtiari sheep. Anim. Genet., 49: 488–491.

Abdoli R., Mirhoseini S.Z., Ghavi Hossein-Zadeh N., Zamani P., Ferdosi M.H., Gondro C. (2019 a). Genome-wide association study of four composite reproductive traits in Iranian fat-tailed sheep. Reprod. Fert. Develop., 31: 1127–1133.

Abdoli R., Mirhoseini S.Z., Ghavi Hossein-Zadeh N., Zamani P., Moradi M.H., Ferdosi M.H., Gondro C. (2019 b). Genome-wide association study of first lambing age and lambing interval in sheep. Small Rumin. Res., 178: 43–45.

Aloisio G.M., Nakada Y., Saatcioglu H.D., Peña C.G., Baker M.D., Tarnawa E.D., Mukherjee J., Manjunath H., Bugde A., Sengupta A.L. (2014). PAX7 expression defines germline stem cells in the adult testis. J. Clin. Invest., 124: 3929–3944.

Ashwell M., Heyen D., Sonstegard T., Van Tassell C., Da Y., VanRaden P., Ron M., Weller J., Lewin H. (2004). Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. J. Dairy Sci., 87: 468–475.

Bal R.D. (2005). Experimental designs for reliable detection of linkage disequilibrium in unstructured random population association studies. Genetics, 170: 859–873.

Bazer F.W., Spencer T.E., Johnson G.A., Burghardt R.C. (2011). Uterine receptivity to implantation of blastocysts in mammals. Front Biosci. (Schol Ed)., 3: 745–767.

Behnia F., Taylor B.D., Woodson M., Kacerovsky M., Hawkins H., Fortunato S.J., Saade G.R., Menon R. (2015). Chorioamniotic membrane senescence: a signal for parturition? Am. J. Obstet. Gynecol. MFM., 213: 359.

Bhat S.A., Ahmad S.M., Ibeagha-Awemu E.M., Bhat B.A., Dar M.A., Mumtaz P.T., Shah R.A., Ganai N.A. (2019). Comparative transcriptome analysis of mammary epithelial cells at different stages of lactation reveals wide differences in gene expression and pathways regulating milk synthesis between Jersey and Kashmiri cattle. PloS One, 14: 0211773.

Bliss S.P., Navratil A.M., Xie J., Roberson M.S. (2010). GnRH signaling, the gonadotrope and endocrine control of fertility. Front Neuroendocrinol., 31: 322–340.

Breen S.M., Knox R.V. (2012). The impact of dose of FSH (Folltropin) containing LH (Lutropin) on follicular development, estrus and ovulation responses in prepubertal gilts. Anim. Reprod. Sci., 132: 193–200.

Cai Z., Guldbrandtsen B., Lund M.S., Sahana G. (2019). Prioritizing candidate genes for fertility in dairy cows using gene-based analysis, functional annotation and differential gene expression. BMC Genomics, 20: 255.

Chang K. (2007). Key signalling factors and pathways in the molecular determination of skeletal muscle phenotype. Animal, 1: 681–698.

Cochran S.D., Cole J.B., Null D.J., Hansen P.J. (2013 a). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genetics, 14: 49.

Cochran S.D., Cole J.B., Null D.J., Hansen P.J. (2013 b). Single nucleotide polymorphisms in

candidate genes associated with fertilizing ability of sperm and subsequent embryonic development in cattle. Biol. Reprod., 89: 69.
Costa R.B., Camargo G.M., Diaz I.D., Irano N., Dias M.M., Carvalheiro R., Boligon A.A., Baldi F., Oliveira H.N., Tonhati H. (2015). Genome-wide association study of reproductive traits in Nellore heifers using Bayesian inference. Genet. Sel. Evol., 47: 67.
Demiray S.B., Goker E.N.T., Tavmergen E., Yilmaz O., Calimlioglu N., Soykam H.O., Oktem G., Sezerman U. (2019). Differential gene expression analysis of human cumulus cells. Clin. Exp. Reprod. Med., 46: 76.
Do D., Bissonnette N., Lacasse P., Miglior F., Sargolzaei M., Zhao X., Ibeagha-Awemu E. (2017). Genome-wide association analysis and pathways enrichment for lactation persistency in Canadian Holstein cattle. J. Dairy Sci., 100: 1955–1970.
Doufas A.G., Mastorakos G. (2000). The hypothalamic-pituitary-thyroid axis and the female reproductive system. Ann. N. Y. Acad. Sci., 900: 65–76.
Eghbalsaied S. (2011). Estimation of genetic parameters for 13 female fertility indices in Holstein dairy cows. Trop Anim. Health Prod., 43: 811–816.
Fair T., Lonergan P. (2012). The role of progesterone in oocyte acquisition of developmental competence. Reprod. Domest. Anim., 47: 142–147.
Fang L., Sahana G., Su G., Yu Y., Zhang S., Lund M.S., Sørensen P. (2017). Integrating sequence-based GWAS and RNA-Seq provides novel insights into the genetic basis of mastitis and milk production in dairy cattle. Sci. Rep., 7: 455–460.
Farhadian M., Rafat S.A., Hasanpur K., Ebrahimi M., Ebrahimie E. (2018 a). Cross-species meta-analysis of transcriptomic data in combination with supervised machine learning models identifies the common gene signature of lactation process. Front. Genet., 9: 235.
Farhadian M., Rafat S.A., Hasanpur K., Ebrahimie E. (2018 b). Transcriptome signature of the lactation process, identified by meta-analysis of microarray and RNA-Seq data. Biotechnol. Acta, 99: 153–163.
Frischknecht M., Bapst B., Seefried F.R., Signer-Hasler H., Garrick D., Stricker C., Fries R., Russ I., Sölkner J., Bieber A. (2017). Genome-wide association studies of fertility and calving traits in Brown Swiss cattle using imputed whole-genome sequences. BMC Genomics, 18: 910.
Ghasemi M., Zamani P., Vatankhah M., Abdoli R. (2019). Genome-wide association study of birth weight in sheep. Animal, 13: 1797–1803.
Ghiasi H., Pakdel A., Nejati-Javaremi A., Mehrabani-Yeganeh H., Honarvar M., González-Recio O., Carabaño M.J., Alenda R. (2011). Genetic variance components for female fertility in Iranian Holstein cows. Livest Sci., 139: 277–280.
Goddard M.E., Hayes B.J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet., 10: 381.
Han B., Kang H.M., Eskin E. (2009). Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS Genet, 5 (4): e1000456.
Hirose M., Kamoshita M., Fujiwara K., Kato T., Nakamura A., Wojcikiewicz R.J., Parys J.B., Ito J., Kashiwazaki N. (2013). Vitrification procedure decreases inositol 1, 4, 5-trisphophate receptor expression, resulting in low fertility of pig oocytes. J. Anim. Sci., 84: 693–701.
Höglund J.K., Buitenhuis B., Guldbrandtsen B., Lund M.S., Sahana G. (2015). Genome-wide association study for female fertility in Nordic Red cattle. BMC Genetics, 16: 110.
Ibeagha-Awemu E.M., Peters S.O., Akwanji K.A., Imumorin I.G., Zhao X. (2016). High density genome wide genotyping-by-sequencing and association identifies common and low frequency SNPs, and novel candidate genes influencing cow milk traits. Sci. Rep., 6: 31109.
Jamrozik J., Fatehi J., Kistemaker G.J., Schaeffer L.R. (2005). Estimates of genetic parameters for Canadian Holstein female reproduction traits. J. Dairy Sci., 88: 2199–2208.
Jiang L., Liu X., Yang J., Wang H., Jiang J., Liu L., He S., Ding X., Liu J., Zhang Q. (2014). Targeted resequencing of GWAS loci reveals novel genetic variants for milk production traits. BMC Genomics, 15: 1105.
Kadarmideen H., Thompson R., Simm G. (2000). Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. Anim. Sci., 71: 411–419.

Klein R.J., Zeiss C., Chew E.Y., Tsai J.-Y., Sackler R.S., Haynes C., Henning A.K., SanGiovanni J.P., Mane S.M., Mayne S.T. (2005). Complement factor H polymorphism in age-related macular degeneration. Science, 308: 385–389.

Kolbehdari D., Wang Z., Grant J., Murdoch B., Prasad A., Xiu Z., Marques E., Stothard P., Moore S. (2009). A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. J. Anim. Breed. Genet., 126: 216–227.

Kominakis A., Hager-Theodorides A.L., Zoidis E., Saridaki A., Antonakos G., Tsiamis G. (2017). Combined GWAS and 'guilt by association'-based prioritization analysis identifies functional candidate genes for body size in sheep. Genet. Sel. Evol., 49: 41.

Kulak K., Dekkers J., McAllister A., Lee A. (1997). Relationships of early performance traits to lifetime profitability in Holstein cows. Can. J. Anim. Sci., 77: 617–624.

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

Lonergan P. (2011). Influence of progesterone on oocyte quality and embryo development in cows. Theriogenology, 76: 1594–1601.

Loveland K.L., Klein B., Pueschl D., Indumathy S., Bergmann M., Loveland B.E., Hedger M.P., Schuppe H.-C. (2017). Cytokines in male fertility and reproductive pathologies: Immunoregulation and beyond. Front. Endocrinol., 8: 307.

Ma K., Liao M., Liu F., Ye B., Sun F., Yue G.H. (2016). Charactering the ZFAND3 gene mapped in the sex-determining locus in hybrid tilapia (*Oreochromis* spp.). Sci. Rep., 6: 25471.

Minozzi G., Nicolazzi E.L., Stella A., Biffani S., Negrini R., Lazzari B., Ajmone-Marsan P., Williams J.L. (2013). Genome wide analysis of fertility and production traits in Italian Holstein cattle. PLoS One, 8 (11): e80219.

Misztal I., Auvray B., Druet T., Lee D.H. (2002). BLUPF90 and related programs (BGF90). Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpelier, France.

Moore S.G., Pryce J.E., Hayes B.J., Chamberlain A.J., Kemper K.E., Berry D.P., McCabe M., Cormican P., Lonergan P., Fair T. (2016). 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. Rep., 94: 1–11.

Mungall C.J., Bada M., Berardini T.Z., Deegan J., Ireland A., Harris M.A., Hill D.P., Lomax J. (2011). Cross-product extensions of the Gene Ontology. J. Biomed. Inform., 44: 80–86.

Muñoz-Espín D., Cañamero M., Maraver A., Gómez-López G., Contreras J., Murillo-Cuesta S., Rodríguez-Baeza A., Varela-Nieto I., Ruberte J., Collado M. (2013). Programmed cell senescence during mammalian embryonic development. Cell, 155: 1104–1118.

Nayeri S., Sargolzaei M., Abo-Ismail M.K., May N., Miller S.P., Schenkel F., Moore S.S., Stothard P. (2016). Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. BMC Genetics, 17: 75.

Nayeri S., Sargolzaei M., Abo-Ismail M., Miller S., Schenkel F., Moore S., Stothard P. (2017). Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle. J. Dairy Sci., 100: 1246–1258.

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: e0188997.

Olsen H., Hayes B., Kent M., Nome T., Svendsen M., Larsgard A., Lien S. (2011). Genome-wide association mapping in Norwegian Red cattle identifies quantitative trait loci for fertility and milk production on BTA12. Anim. Genet., 42: 466–474.

Ortega M.S., Denicol A.C., Cole J.B., Null D.J., Taylor J.F., Schnabel R.D., Hansen P.J. (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: 3725–3734.

Panahi B., Farhadian M., Dumas J., Hejazi M. (2019). Integration of cross species RNA-seq Meta-analysis and Machine Learning Models identifies the most important salt stress responsive pathways in microalga Dunaliella. Front. Genet., 10: 752.

Pimentel E., Bauersachs S., Tietze M., Simianer H., Tetens J., Thaller G., Reinhardt F., Wolf E., König S. (2011). Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNPs within candidate genes derived by expression profiling. Anim. Genet., 42: 251–262.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., De Bakker P.I., Daly M.J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. AJHG, 81: 559–575.

Rempel L.A., Freking B.A., Miles J.R., Nonneman D.J., Rohrer G.A., Vallet J.L., Schneider J.F. (2011). Association of porcine heparanase and hyaluronidase 1 and 2 with reproductive and production traits in a Landrace–Duroc–Yorkshire population. Front. Genet., 2: 20.

Reverter A., Fortes M. (2013). Building single nucleotide polymorphism-derived gene regulatory networks: towards functional genomewide association studies. J. Anim. Sci., 91: 530–536.

Reverter A., Porto-Neto L., Fortes M., McCulloch R., Lyons R., Moore S., Nicol D., Henshall J., Lehnert S. (2016). Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree. J. Anim. Sci., 94: 4096–4108.

Rezende F., Dietsch G., Peñagaricano F. (2018). Genetic dissection of bull fertility in US Jersey dairy cattle. Anim. Genet., 49: 393–402.

Rolf M., Taylor J., Schnabel R., McKay S., McClure M., Northcutt S., Kerley M., Weaber R. (2012). Genome-wide association analysis for feed efficiency in Angus cattle. Anim. Gen., 43: 367–374.

Royal M., Flint A., Woolliams J. (2002). Genetic and phenotypic relationships among endocrine and traditional fertility traits and production traits in Holstein-Friesian dairy cows. J. Anim. Sci., 85: 958–967.

Sargolzaei M. (2014). SNP1101 User's guide. (Version1.0.0). Semex Alliance, Ontario, Canada.

Schnabel R., Sonstegard T., Taylor J., Ashwell M. (2005). Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families. Anim. Genet., 36: 408–416.

Seyed Sharifi R., Afkan F.N., Eyorigh N.H., Davati J.S. (2017). Estimation of economic value for productive and reproductive traits of Moghan agro-industrial holstein cows by using simulation and bio-economic model (in Persian). IJAS., 9: 15.

Shook G. (2006). Major advances in determining appropriate selection goals. J. Dairy Sci., 89: 1349–1361.

Storer M., Mas A., Robert-Moreno A., Pecoraro M., Ortells M.C., Di Giacomo V., Yosef R., Pilpel N., Krizhanovsky V., Sharpe J. (2013). Senescence is a developmental mechanism that contributes to embryonic growth and patterning. Cell, 155: 1119–1130.

Sullivan P.G., Jamrozik J., Kistemaker G.J. (2015). De-regressing MACE versus domestic EBV for genomics. Interbull. Bulletin, 49.

Sutovsky P. (2003). Ubiquitin-dependent proteolysis in mammalian spermatogenesis, fertilization, and sperm quality control: killing three birds with one stone. Microsc. Res. Tech., 61: 88–102.

Toghiani S. (2012). Genetic relationships between production traits and reproductive performance in Holstein dairy cows. Arch. Anim. Breed., 55: 458–468.

Turner S.D. (2014). qqman: an R package for visualizing GWAS results using QQ and manhattan plots. Biorxiv, 005165.

VanRaden P.M. (2008). Efficient methods to compute genomic predictions. J. Dairy Sci., 91: 4414–4423.

Velarde M.C., Menon R. (2016). Positive and negative effects of cellular senescence during female reproductive aging and pregnancy. J. Endocrinol., 230: 59–76.

Walsh S., Williams E., Evans A. (2011). A review of the causes of poor fertility in high milk producing dairy cows. Anim. Reprod. Sci., 123: 127–138.

Wang H., Zhang L., Cao J., Wu M., Ma X., Liu Z., Liu R., Zhao F., Wei C., Du L. (2015). Genome-wide specific selection in three domestic sheep breeds. PLoS One, 10(6): e0128688.

Xu S., Wang D., Zho D., Lin Y., Che L., Fang Z., Wu D. (2015). Reproductive hormone and transcriptomic responses of pituitary tissue in anestrus gilts induced by nutrient restriction. PloS One, 10 (11): e0143219.

Xu S., Gao L., Xie X., Ren Y., Shen Z., Wang F., Shen M., Eyþórsdóttr E., Hallsson J., Kiseleva T., Juha Kantanen J., Li M. (2018). Genome-wide association analyses highlight the potential for different genetic mechanisms for litter size among sheep breeds. Front. Genet. 9: 118.

Yang J., Weedon M.N., Purcell S., Lettre G., Estrada K., Willer C.J., Smith A.V., Ingelsson E., O'Connell J.R., Mangino M. (2011). Genomic inflation factors under polygenic inheritance. AJHG, 19: 807.

Received: 24 XI 2019
Accepted: 3 III 2020