scientific reports

Check for updates

OPEN

Integrated TWAS, GWAS, and RNAseq results identify candidate genes associated with reproductive traits in cows

Sevda Hosseinzadeh[1✉], Seyed Abbas Rafat[1✉] & Lingzhao Fang[2]

Low fertility in cows leads to early removal from herds. Since reproductive traits are complex and have low heritability, genetic analysis can aid in improving reproduction. This study identified key genes linked to fertility by conducting genome- and transcriptome-wide association studies, RNA-seq analysis, meta-analysis, weighted gene co-expression network analysis, and functional enrichment analysis. Through these methods, we identified candidate genes related to Cow conception rate (CCR), Daughter pregnancy rate (DPR), Heifer conception rate (HCR), and overall fertility traits, helping to improve genetic selection for reproductive success in cows. The identified genes include RPL12, FKBP1B, FZD10, COX10, COX7A2, GAA, ETFBKMT, ACSM5, NUDT9, TIGAR, PAIP2, and PSMB5. Notably, GAA, ETFBKMT, COX10, and COX7A2 are involved in the "generation of precursor metabolites and energy" process. COX10, GAA, ETFBKMT, ACSM5, NUDT9, and TIGAR exhibit significant impacts on CCR, DPR, and HCR. COX7A2, PAIP2, and PSMB5 have been identified as hub genes related to fertility traits. RPL12 plays a role in protein synthesis, essential for gametogenesis and embryo development, while FKBP1B regulates calcium signaling, particularly in oocyte aging and fertility decline, and FZD10 is crucial in Wnt signaling. The identified genes serve as markers for genomic selection aimed at enhancing reproductive traits in cow.

Keywords Cow, Reproduction, RNA-seq, TWAS, WGCNA

Low reproductive capacity is the primary cause of cows being culled early. Lately, the negative relationship between cow production and reproductive qualities has increased, likely due to the pleiotropic effects of genes[1]. This has resulted in reduced cow fertility as a result of strong selective pressure on production traits[2]. Furthermore, the process of reproduction as a complex trait is regulated by numerous genetic loci and influenced by environmental elements, and demonstrates low heritability[3]. Therefore, understanding the genetic architecture of cows with high reproductive potential is beneficial for animal breeding.

Genome-wide association studies (GWAS) identified thousands of associated loci with complex traits[4]. In addition, it is noteworthy that the majority of GWAS loci are non-coding, they may thus indirectly affect complex traits through regulating gene expression[5] and GWAS also has limitation in prioritizing causal genes. GWAS results identify causal genes mainly based on the fine-mapped distance of associated variants in linkage disequilibrium (LD) with the causal marker[6]. So identification of new loci by studies based on physical distance are rare[7]. On the other hand, gene expression serves as a molecular phenotype and enhances the functional interpretation of GWAS discoveries[8].

In addition, RNA sequencing data can serve as a molecular phenotype for identifying expression quantitative trait loci (eQTL) and differentially expressed genes (DEGs). eQTLs are genomic regions where the expression levels are linked to regulatory genotype variants in various tissues[9]. It is worth noting that to obtain gene counts, a multi-step analysis including quality control, trimming, mapping, and assembling of RNA-seq data is necessary. Each step offers a range of options, leading to a complex process that poses a challenge when interpreting the results[10]. Nevertheless, in a meta-analysis strategy, increasing sample size also enhances the statistical power and leads to improved results[11]. Additionally, weighted gene co-expression network analysis (WGCNA) emerges as a suitable technique to identify significant co-expressed genes, their nodes, and hub genes[12]. Moreover, Transcriptome-wide association studies (TWAS) that combine eQTL investigations with GWAS play a role in

[1]Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran. [2]Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus 8000, Denmark. ✉email: sevdahoseinzade@gmail.com; rafata@tabrizu.ac.ir

understanding the genetic impact on transcriptome associations with complex traits[13]. This approach helps in characterizing new genomic regions and determining key genes linked to complex traits in various tissues[14,15].

Therefore, exploring and integrating different genetic data can be helpful for the identification of various effective factors for controlling complex traits. In this study, we aim to interpret the biological mechanisms underlying reproductive traits such as cow conception rate (CCR), daughter pregnancy rate (DPR), heifer conception rate (HCR), and fertility by integrating results from TWAS and GWAS summary analyses, RNA-seq data analysis, meta-analysis, the WGCNA approach, and gene ontology. Through this integration, we seek to identify potential candidate genes and explore their molecular, cellular, and biological functions related to reproductive traits in cattle.

Materials and methods
Data collection
We utilized TWAS summary results from 11,642 Holstein cattle (available at https://cgtex.roslin.ed.ac.uk/downloads/) and GWAS summary statistics from 27,241 Holstein cattle (accessible at https://figshare.com/s/ea726fa95a5bac158ac1) to analyze three traits: Cow Conception Rate (CCR), Daughter Pregnancy Rate (DPR), and Heifer Conception Rate (HCR). HCR is defined as the percentage of inseminated heifers that become pregnant at each service, while CCR represents the percentage of inseminated cows that conceive per service. CCR is typically measured 30 days after artificial insemination or natural mating through ultrasonographic pregnancy diagnosis. DPR reflects the percentage of nonpregnant cows that become pregnant during each 21-day estrous cycle. The primary distinction between cows and heifers lies in their age, with cows generally ranging from 2 to 10 years and heifers being younger, typically under 2 years.The TWAS summary results encompass 23 tissues, including adipose, blood, embryo, hypothalamus, ileum, intramuscular tissue, jejunum, leukocytes, liver, lung, lymph nodes, macrophages, mammary gland, milk cells, monocytes, muscle, ovary, oviduct, pituitary gland, rumen, salivary gland, skin fibroblasts, and uterus. Additionally, gene expression counts were obtained from 40 cattle —20 with high fertility and 20 with low fertility— using samples from endometrial and ovarian tissues, based on mapping to the ARS-UCD1.2 cattle reference genome. This dataset includes 27,607 transcripts and is available at https://cgtex.roslin.ed.ac.uk/downloads/. The details of the pipeline used can be found at https://cgtex.roslin.ed.ac.uk/. For more information, please refer to Table 1.

Analysis method
We utilized TWAS, GWAS, RNA-seq, meta-analysis, WGCNA, GO, and KEGG pathway analysis to investigate reproduction traits. GWAS identifies associations between genetic variants (SNPs) and traits, but often lacks insight into the biological mechanisms, especially when SNPs are in non-coding regions[18]. TWAS bridges this gap by integrating gene expression data with GWAS, connecting genetic variants to gene expression and identifying potential causative genes[19].

Gene expression levels reflect the behavior of genes associated with complex traits[20,21]. To enhance statistical power, meta-analysis aggregates data from multiple RNA-seq studies, resulting in more robust and generalizable findings[22]. WGCNA identifies co-expressed gene modules and correlates them with traits, aiding in the discovery of gene networks[23]. To further elucidate these findings, GO analysis classifies genes into biological processes, molecular functions, and cellular components[24], while KEGG pathway analysis maps genes to established biochemical pathways[25], providing insights into their roles in complex traits such as reproduction.

Tissues correlation based on TWAS result
Adipose, blood, embryo, hypothalamus, ileum, intramuscular, jejunum, leukocyte, liver, lung, lymph node, macrophage, mammary, milk cell, monocyte, muscle, ovary, oviduct, pituitary, rumen, salivary gland, skin fibroblast, and uterus were the 23 tissues whose correlation coefficients were computed to examine the relationship between them. To calculate the correlation coefficient, we used the available Z-score in the TWAS summary statistical results pertaining to CCR, DPR, and HCR. In order to minimize the noise effect in the correlation calculation, we kept genes that have z-score in more than 10 tissues. Also we kept tissues that have z-score in more than 70% of genes. The metan package[26] within the R (v.4.0.5) was utilized to obtain the correlation coefficient. Also candidate genes based on TWAS, and GWAS methods (p-value < 0.01) were introduced related to CCR, DPR, and HCR.

Gene counts analysis of individual datasets and Meta-analysis
The gene counts of 40 cows belonging to endometrial and ovarian tissue were used as input data for the Deaeq2 package[27] in R (v.4.0.5) to identify DEGs between high and low fertility cows for each datasets. The pipeline used by the FarmGTEX consortium to calculate gene counts is available at https://cgtex.roslin.ed.ac.uk/. We used the metaRNASeq package version 1.0.2 [28] in R (v.4.0.5) to find the meta-genes for each tissue. Raw p-values from the datasets for the ovary and endometrium were gathered into a new file and used individually as input data for the metaRNASeq package. We utilized the "invnorm" function to combine p-values, number of repeat equal 6, and 5 for endometrium and 4, and 5 for ovary tissue and Benjamini–Hochberg method used for p-value correction. Finally, genes that showed co-directional expression in all datasets, with p-value ≤ 0.05 at least in one dataset, and meta-analysis p-value ≤ 0.05 were considered as meta-genes for each tissue.

Weighted gene co-expression network analysis
The Deseq2 package's variance stabilizing transformation method was used to normalize the gene counts. The normalized endometrial and ovarian tissue-specific gene counts were used separately as input data for weighted gene co-expression network analysis (WGCNA) (version 1.71) package[29] in R (v.4.0.5) for detection of co-

Group	Trait	Tissue	Run accession	Breed	Age (week)	sequencing platform	Reference
		Dataset1					
Case	Low fertility	Endometrium	SRS951413	Holstein	Unknown	ILLUMINA (Illumina HiSeq 2000)	-
		Endometrium	SRS951423				
		Endometrium	SRS951426				
		Endometrium	SRS951427				
		Endometrium	SRS951431				
		Endometrium	SRS951436				
Control	High fertility	Endometrium	SRS951411				
		Endometrium	SRS951417				
		Endometrium	SRS951420				
		Endometrium	SRS951425				
		Endometrium	SRS951432				
		Endometrium	SRS951437				
		Dataset2					
Case	Low fertility	Endometrium	SRS2749180	Angus	Unknown	ILLUMINA (NextSeq 500)	[16]
		Endometrium	SRS2749186				
		Endometrium	SRS2749193				
		Endometrium	SRS2749197				
		Endometrium	SRS2749207				
Control	High fertility	Endometrium	SRS2749173				
		Endometrium	SRS2749176				
		Endometrium	SRS2749185				
		Endometrium	SRS2749195				
		Endometrium	SRS2749214				
		Dataset3					
Case	Low fertility	Ovary	SRS1115943	Holstein	2.58 years	ILLUMINA (Illumina HiSeq 2000)	[17]
		Ovary	SRS1115944				
		Ovary	SRS1115945				
		Ovary	SRS1115949				
Control	High fertility	Ovary	SRS1115950				
		Ovary	SRS1115951				
		Ovary	SRS1115953				
		Ovary	SRS1115954				
		Dataset4					
Case	Low fertility	Ovary	SRS951450	Holstein	Unknown	ILLUMINA (Illumina HiSeq 2000)	-
		Ovary	SRS951455				
		Ovary	SRS951464				
		Ovary	SRS951465				
		Ovary	SRS951466				
Control	High fertility	Ovary	SRS951439				
		Ovary	SRS951453				
		Ovary	SRS951456				
		Ovary	SRS951462				
		Ovary	SRS951463				

Table 1. Accession numbers and meta-data of gene counts for the analyses.

expressed networks and hub-genes. In order to minimize the impact of noise on the correlation computation, at first step we filtered normalized gene counts and retained genes with gene counts above 10 in more than 90% of cows. In second step, we employed the "removeBatchEffect" function from the limma package (version 3.50) to eliminate batch variations amongst datasets. Then, we investigated missing value and outlier sample using the "goodSamplesGenes" and "hclust" options, respectively. Furthermore, based on the scale-free topology index ($R2$) and mean connectivity for the ovarian and endometrial samples, power equals 10 and 7 were estimated as appropriate amounts, respectively. Thereafter, the "pickSoftThreshold" option was used to calculate the adjacency matrix. An adjacency matrix can be computed based on the correlation coefficient between each pair of genes, allowing genes with only direct connectivity to form networks. Consequently, we computed the Topological Overlap Matrix (TOM) and dissimilarity matrix (1-TOM) in order to add indirect connectivity genes to the weighted network. We used the dynamic tree cut (DTC) algorithm to cluster the genes using the dissimilarity

3

matrix. Ultimately, gene co-expression modules were created using a hybrid approach, a cut height of 0.975, a minimum module size of 30 genes, and a DeepSplit of 2. Following the acquisition of the gene co-expression modules, we used the "moduleEigengenes" function to determine the eigengenes for each module. The outputs obtained up to this point could eventually yield the following outputs: (a) identification of the modules that exhibit a significant correlation with the fertility trait in cow, (b) relationships among the modules, and (c) hub genes in each module. It is noteworthy that each module's hub genes were identified using two different statistical techniques: "chooseOneHubInEachModule", and "chooseTopHubInEachModule" functions.

Functional enrichment analysis of potential candidate genes

The STRING database (https://string-db.org/) was used for the purpose of creating the PPI network and identifying the hub-genes by meta genes. Following this, we obtained common significant genes between the TWAS results of the HCR, the DPR, and the CCR, as well as hub genes, modules and meta-genes of the ovary and endometrium tissues via (https://bioinformatics.psb.ugent.be/webtools/Venn/). Furthermore, we employed the ClueGO plugin within Cytoscape software (version 3.7.2) to conduct enrichment analyses based on the KEGG pathway and Gene Ontology (GO) to interpret the meta genes and common significant genes between the TWAS results, hub genes, modules, and meta-genes individually. Only pathways with p-values below 0.05, following Bonferroni step-down correction, were considered significantly enriched.

Results
Correlations amongst tissues based on TWAS summary results

An analysis of tissue correlations for each trait (CCR, DPR, and HCR) revealed no negative correlation based on the Z score of TWAS results. The TWAS summary for the CCR trait indicated a highly significant positive correlation involving the uterus and 14 tissues. Moreover, the DPR results from TWAS demonstrated significant correlations between the liver and lung, which were associated with 12 and 11 tissues, respectively. Additionally, the TWAS findings for HCR identified significant positive correlations between muscle and 14 tissues. A detailed overview of these results is presented in Fig. 1.

TWAS and GWAS results to identify candidate genes

Based on TWAS and GWAS results, candidate genes for each trait (CCR, DPR, and HCR) were introduced separately. The potential genes for each trait in different tissues, as determined by TWAS, can be found in Supplementary File 1. Additionally, certain candidate genes based on TWAS summary results are shared across multiple tissues with a p-value less than 0.01, and these are summarized in Fig. 2 for each trait to facilitate understanding. Supplementary File 2 contains candidate genes identified through GWAS results. Furthermore, the analysis of both TWAS and GWAS data revealed common potential genes related to each trait, such as AFF1 and ZNF613 for HCR, UMPS for DPR, and UMPS and ZNF613 for CCR trait.

Identification of meta-genes for each tissues

Four datasets were examined separately utilizing DESeq2. Genes demonstrating co-directional expression patterns across all datasets, with a p-value of ≤0.05 in at least one dataset and a meta-analysis p-value of ≤0.05, were identified as meta-genes for each tissue individually. A list of meta-genes related to ovarian and endometrial tissues can be found in **Supplementary File 3**. The expression levels of 90 and 168 meta-genes associated with endometrial and ovarian tissues are represented in a volcano plot across four datasets. This visualization highlights the expression patterns of these meta-genes in the context of the individual dataset analyses, emphasizing the significance of conducting a meta-analysis with multiple comparable datasets instead of merely presenting all identified DEGs from each dataset. Notably, most meta-genes did not show differential expression in any of the individual analyses, as depicted in Fig. 3 of the volcano plot. This finding underscores the critical role of meta-analysis in detecting differential transcripts that exhibit subtle variations.
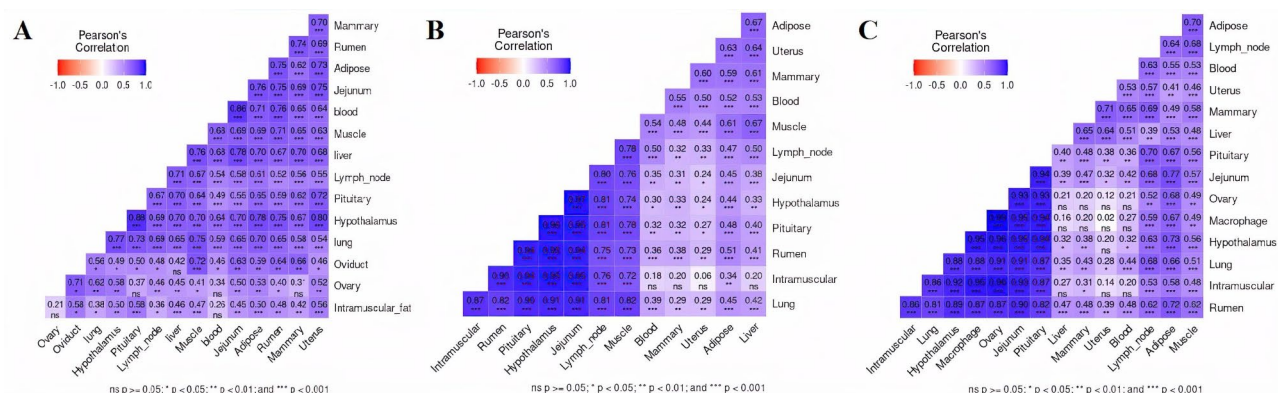

Fig. 1. The correlation plot derived from the three traits' TWAS summary results between 23 tissues. The traits associated with cow conception rate, daughter pregnancy rate, and heifer conception rate are represented by Plots A, B, and C, respectively.
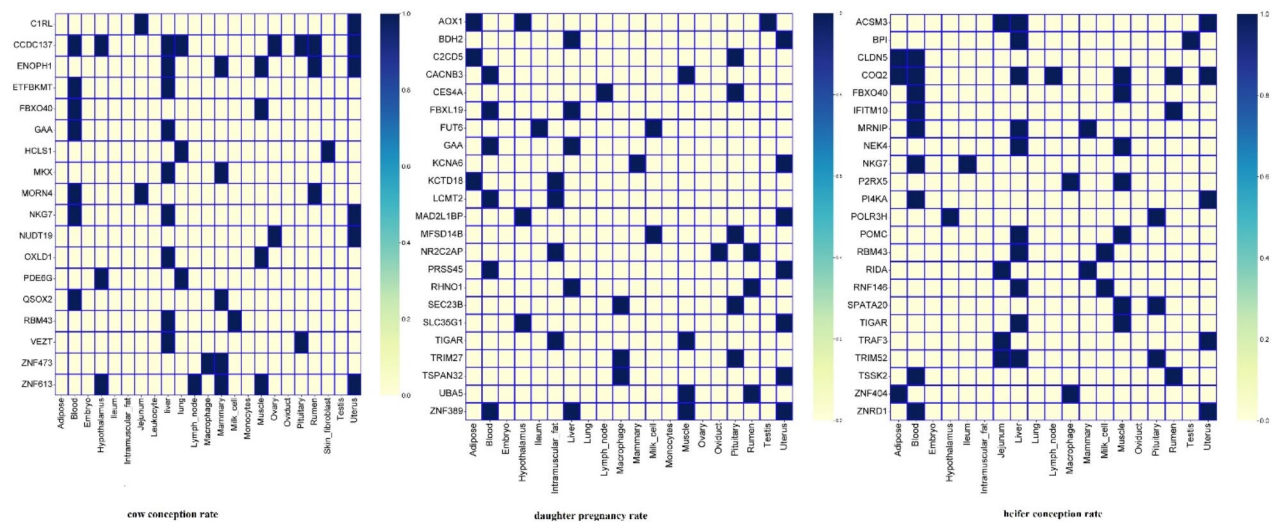
Fig. 2. Discovery of shared genes across tissues for cow conception rate, daughter pregnancy rate, and heifer conception rate using TWAS summary results with a p-value below 0.01.
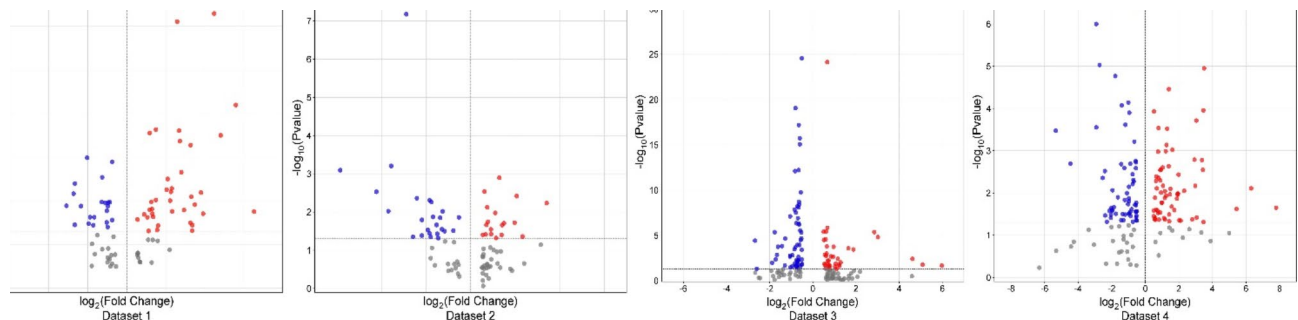

Fig. 3. The meta-gene expressions of the four datasets are represented in a volcano plot. The datasets are arranged from left to right as follows: dataset1, dataset2, dataset3, and dataset4.

Weighted gene co-expression network analysis and identification of the hub-genes

WGCNA was run independently using the expression values of 2,132 and 2,177 genes from the endometrium and ovary tissues, respectively (**Supplementary File 5**). After examining outlier samples and excessively missing values in the first step, a soft-threshold power of 8 was found for the ovary and 10 for the endometrium. (**Supplementary Fig. 1**). Our approach for building the modules was the dynamic tree cut algorithm. For the ovarian tissue, 14 modules were identified, with sizes varying from 52 to 418 genes, alongside 92 genes that were classified as unassigned due to their inability to fit into any of the established modules (Fig. 4-A). Figure 4–B illustrate the hierarchical clustering of genes based on the topological overlap matrix (TOM). Among the 14 identified modules, seven modules were particularly significant: blue ($r = -0.74$, p-value = 4e-04), turquoise ($r = -0.83$, p-value = 2e-05), yellow ($r = -0.78$, p-value = 1e-04), black ($r = -0.74$, p-value = 5e-04), green ($r = -0.82$, p-value = 4e-05), pink ($r = -0.62$, p-value = 0.006), and red ($r = -0.66$, p-value = 0.003). The gene names corresponding to these seven significant modules can be found in **Supplementary File 6-A**. Additionally, Fig. 4-C presents the p-values and pearson correlation coefficients of the identified modules in relation to the fertility trait. In input genes for endometrium tissue, 12 modules with sizes ranging from 33 to 216 genes were identified, in addition to 689 genes that were deemed unassigned since they could not be placed in any of the modules (Fig. 4-D). Figure 4–E depict the hierarchical clustering of genes using the topological overlap matrix (TOM). Two out of the 12 modules including pink ($r = 0.46$, p-value = 0.03), purple ($r = 0.46$, p-value = 0.03) modules were found to be especially significant. The gene names for each of the two major modules are listed in **Supplementary File 6-B**. The identified modules' p-values and Pearson correlation coefficient are presented in Fig. 4-F.

Subsequently, hub genes for each module were identified utilizing the "chooseTopHubInEachModule" and "chooseOneHubInEachModule" methods, as presented in Table 2.
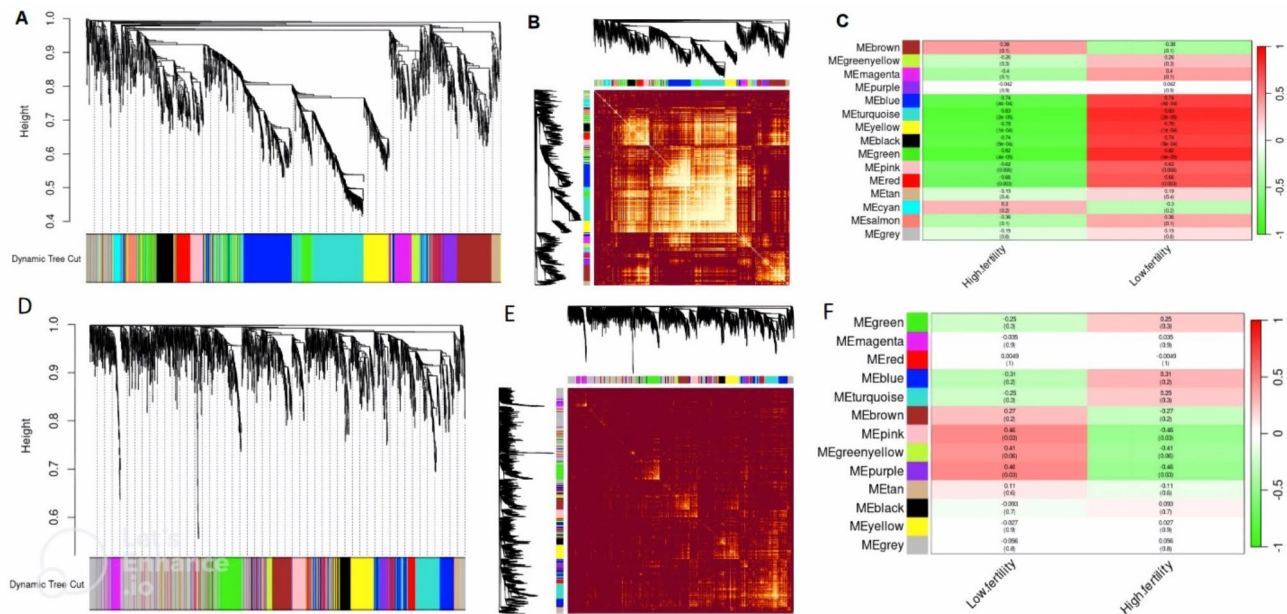
5

Fig. 4. Gene co-expression modules, which were clustered based on a read counts of the endometrial and ovarian samples. In (**A**) and (**D**), the y-axis represents the co-expression distance, the x-axis corresponds to genes, and colors indicate modules, with the gray module representing genes not assigned to any module. (**B**) and (**E**) illustrates a heatmap plot of topological overlap in the gene network. (**C**) and (**F**), the relationship between modules and treatment groups (high fertility and low fertility groups) was examined by calculation of correlating coefficient between traits and module eigengenes. Rows represent module eigengenes, columns represent treatment groups, and each cell contains the correlation value and significance level (p-value) of the relationship.

Ovary			Endometrium		
Significant modules	One hub gene (Gene symbol)	Top hub gene (Gene symbol)	Significant modules	One hub gene (Gene symbol)	Top hub gene (Gene symbol)
blue	ENSBTAG00000012726 (PSMB5)	ENSBTAG00000005096 (COX7A2)	Purple	ENSBTAG00000043564 (ATP8)	ENSBTAG00000043564 (ATP8)
turquoise	ENSBTAG00000018320 (RPLP)	ENSBTAG00000018320 (RPLP)	pink	ENSBTAG00000009150 (SPON1)	ENSBTAG00000009150 (SPON1)
yellow	ENSBTAG00000009761 (ACTR2)	ENSBTAG00000002275 (PTP4A1)			
Black	ENSBTAG00000045582 (ARPC5)	ENSBTAG00000045582 (ARPC5)			
green	ENSBTAG00000021455 (CFL1)	ENSBTAG00000019237 (TXNDC5)			
pink	ENSBTAG00000014553 (ATP6V0D1)	ENSBTAG00000014553 (ATP6V0D1)			
red	ENSBTAG00000008224 (PAIP2)	ENSBTAG00000008224 (PAIP2)			

Table 2. Top five hub genes, top hub, and one hub genes in significant modules for potential fertility in cows.

Functional analysis of the potential candidate genes

Following the identification of 91 and 169 meta-genes for endometrium and ovary tissues, respectively, STRING was used to build their PPI network, as shown in Fig. 5. There were many edges between the nodes in the network that was created, which suggested that the meta-genes were significantly related. Furthermore, the meta-genes were subjected to GO and KEGG pathway analyses using the ClueGO plugin of Cytoscape (Fig. 6). As a result, we were able to identify 2, 8, 9, and 2 significant terms for molecular function (MF), cellular components (CC), biological processes (BP), and KEGG pathways, respectively. Among the most significantly enriched BP terms were "humoral immune response," "protein activation cascade," "antibacterial humoral response," "negative regulation of vasculature development," "heart process," "regulation of blood circulation," "heart contraction," cardiac cell development ," and "cardiac muscle cell development ". In addition, Significant CC terms were "sarcoplasm," "sarcoplasmic reticulum," "contractile fiber part," "contractile fiber," "sarcomere," "myofibril," "I band," and "Z disc". Moreover, "sodium ion transmembrane transporter activity," and "integrin binding" were
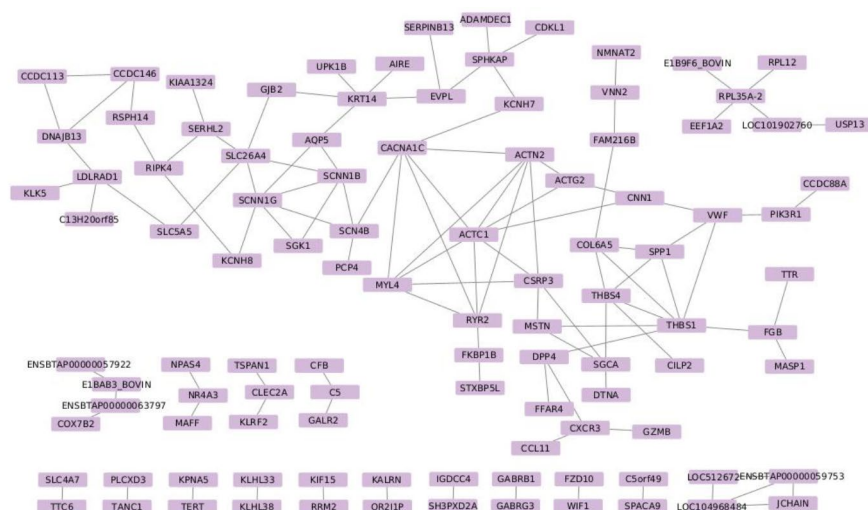
Fig. 5. Examination of ovarian and endometrial meta-genes through protein-protein interaction (PPI) network analysis. Following a meta-analysis of RNA-seq expression data across four datasets, we discovered meta-genes associated with potential fertility traits in cow.
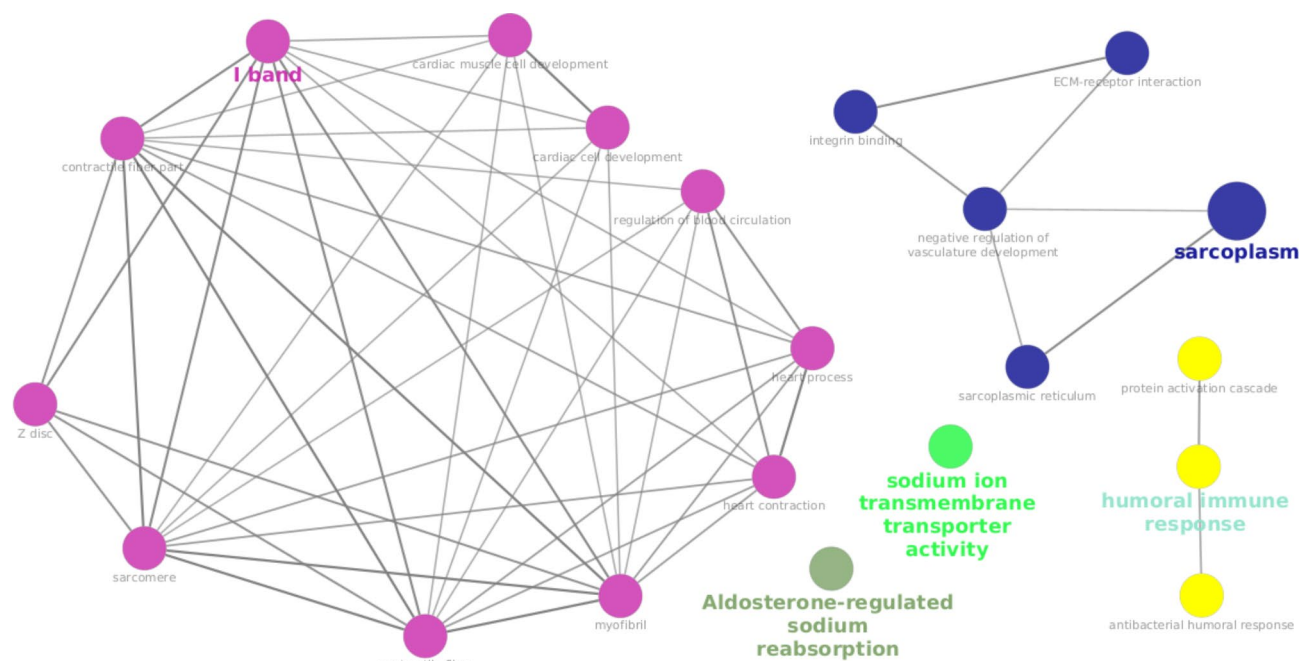

Fig. 6. The significant ontology terms were identified using 258 meta-genes, with a Bonferroni adjusted p-value threshold of less than 0.05. These meta-genes were derived from the results of a meta-analysis conducted on ovarian and endometrial samples.

detected as significant terms for MF. Also, significant KEGG terms were "endoplasmic reticulum chaperone complex" and "iron ion binding" (**Supplementary File 4).**

Ultimately, we obtained shared significant genes between the TWAS results of the HCR, the DPR, and the CCR, as well as hub genes, modules and meta-genes of the ovary and endometrium tissues via (https://bioinform atics.psb.ugent.be/webtools/Venn/). Therefore, common genes between at least three of mentioned results were used for exploration of functional analysis by the ClueGO plugin of Cytoscape software (**Supplementary File 7**).

To identify candidate genes associated with reproductive traits, we employed a comprehensive approach, integrating multiple methodologies. A Venn diagram was generated to visually represent the overlap of significant genes identified through various analyses, including: TWAS, GWAS, RNA-seq, meta-analysis, WGCNA analysis result, and candidate genes involved in significant terms identified by GO and KEGG pathway analysis. By
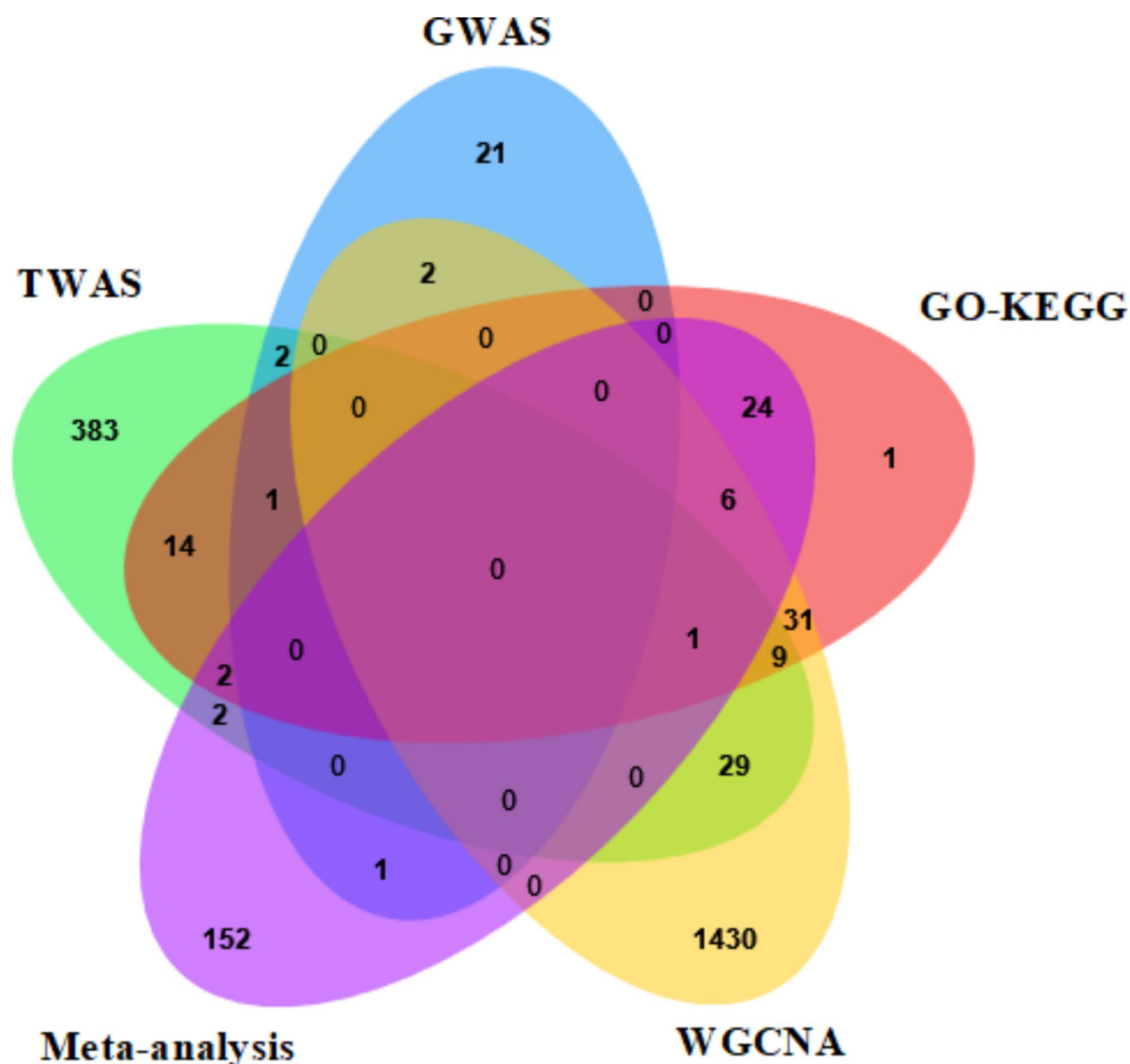
Fig. 7. Venn diagram illustrating the overlap of significant genes identified across multiple analyses, including TWAS, GWAS, and RNA-seq, meta-analysis, and WGCNA results.

combining these approaches, we could capture a more comprehensive set of candidate genes involved in reproductive traits. The Venn diagram highlights the intersections between the significant genes found across these different analyses. This overlap increases confidence in their role in cow fertility.This diagram is presented as Fig. 7, and further supporting details are provided in Supplementary File 9. This combined approach enhances the reliability of our candidate gene identification, offering a robust foundation for future genomic selection aimed at improving reproductive traits in cows.

Discussion
Genetic testing can be most beneficial for fertility traits because they are the most complex, difficult to measure, and low-heritability traits[30]. Also, fertility traits are affected by genetic factors[31]. Therefore, artificial selection and combined with artificial insemination leads to significant improvement in dairy cows fertility[32]. Identifying key genes and variants related to fertility is essential for selective breeding. To achieve this, we analyzed TWAS and GWAS summary results, performed a meta-analysis of RNA sequencing data, WGCNA analysis, and conducted GO and KEGG pathway enrichment analyses using data from FarmGTEX. Our aim was to identify potential candidate genes linked to reproduction traits.

RPL12, UMPS, FKBP1B, FZD10, PIK3R1, THBS1, VWF, TNN, MASP1, SGK1, DSTN, GPI, TGM2, NCOR1, AOX1, ITPR1, SMOC2, RPL11, and APOA1 genes were identified as significant by at least three of the following methods: TWAS, meta-analysis, GO, and WGCNA.

Ribosomal proteins like RPL12 are crucial for protein synthesis and ribosome biogenesis, vital processes in oocyte maturation and early embryonic development. Efficient protein synthesis supports rapid cell division during these stages, influencing folliculogenesis and oocyte quality[33]. Ribosome biogenesis involves the assembly of ribosomal proteins and rRNA, regulated by factors like mTOR signaling and MYC. Disruptions in ribosomal components, such as RPL12 and RPL11, can impair ribosome formation, affecting protein synthesis and cellular development. These proteins are essential for regulating protein production during meiosis, directly impacting oocyte quality and early embryo development[34]. UMPS (Uridine Monophosphate Synthetase) is a key enzyme in pyrimidine biosynthesis, converting orotate to UMP, which is essential for DNA and RNA synthesis. This process is critical for cell division, particularly in oocyte maturation and early embryonic development. Deficiencies in UMPS can cause orotic aciduria, leading to developmental issues and impaired cell function due to disrupted nucleotide production, which in turn affects oocyte quality and embryo viability[35]. FKBP1B plays a crucial role in calcium signaling, which is essential for oocyte activation during fertilization and embryo development. Proper regulation of calcium oscillations influences the transition from the germinal vesicle (GV) stage to the mature oocyte stage, primarily mediated by inositol 1,4,5-trisphosphate receptors (IP3Rs) in the endoplasmic reticulum. Disruptions in these calcium signals can lead to fertilization issues and impair early embryo development[36]. FZD10 participates in Wnt signaling, which is crucial for follicle development and oocyte maturation[37]. PIK3R1 is involved in the PI3K-Akt pathway, impacting ovarian follicular survival and oocyte quality through growth factor signaling[38]. THBS1 and VWF are associated with angiogenesis and vascular health, which are essential for ovarian blood supply and follicle growth[39]. SGK1 regulates ion balance and cell survival, affecting ovarian cells' stress responses[40], while TGM2 is involved in apoptosis and cellular differentiation, both critical for embryo development. NCOR1 modulates gene expression via nuclear hormone receptors, influencing hormonal control over ovarian cycles[41]. ITPR1 is key for calcium signaling, which is vital for oocyte maturation and early embryo development[42]. Other genes, like APOA1, involved in lipid metabolism, affect hormone synthesis and follicle health[43]. Collectively, these genes regulate key pathways related to cell survival, growth, angiogenesis, and signaling that are fundamental to ovarian function, oocyte competence, and successful embryo development.

The significant genes associated with CCR, DPR, and HCR traits based on TWAS are as follows: NLRP12 and COX10 for adipose tissue; FBXO40, ENSBTAG00000037537, and ENSBTAG00000004608 for blood; ZNF613 and LGI4 for the hypothalamus; ENSBTAG00000021433 for the ileum; BPHL for intramuscular fat; ENSBTAG00000054322, GAA, LYSMD3, ETFBKMT, ACSM5, ENSBTAG00000050562, ANKS3, MKX, and NUDT9 for the liver; ENSBTAG00000050865 for the lung; ARRDC3 for lymph nodes; ZNF404 and EMC9 for macrophages; ENSBTAG00000047761 for mammary tissue; ENSBTAG00000050562 for milk cells; and ELOVL6, TIGAR, PARP14, ARHGAP44, DKKL1, and ADAMTS8 for muscle. Additionally, EVA1B was identified for the pituitary gland, while WAC and ENSBTAG00000050562 were significant for the uterus.

It is noteworthy that the COX10 gene is significantly involved in "oxidative phosphorylation," "generation of precursor metabolites and energy," "electron transfer activity," "heme-copper terminal oxidase activity," "oxidoreductase activity, acting on a heme group of donors," "electron transport chain," "hydrogen transport," "oxidoreductase activity, acting on a heme group of donors with oxygen as an acceptor," "cytochrome c oxidase activity," "proton transport," "hydrogen ion transmembrane transporter activity," and "hydrogen ion transmembrane transport." Additionally, previous analysis based on GWAS indicated that the COX10 gene had highly significant effects on DPR, CCR, and HCR traits[44]. Furthermore, it plays a role in the catabolism of carbohydrates and lipids, as well as in the electron transfer to oxygen from reduced cytochrome c[45]. COX10 is vital for the correct assembly of cytochrome c oxidase (complex IV) in the mitochondrial respiratory chain, which is key to ATP production[46]. Mitochondrial function is crucial for follicular development, oocyte maturation, and embryo growth[47]. Disruption of COX10 can impair energy production in oocytes, potentially reducing their quality and fertilization potential. GAA plays a significant role in the "generation of precursor metabolites and energy." The transcription levels of GAA can be altered in sperm under the influence of DNA methylation in humans[48]. The GAA gene encodes an enzyme crucial for breaking down glycogen into glucose within lysosomes, specifically known as acid α-glucosidase (GAA). Glycogen catabolism, including in uterine epithelial cells, is essential for maintaining energy balance[49]. Efficient energy production in ovarian cells supports hormone secretion, follicular development, and oocyte maturation. Disruptions in glycogen metabolism could reduce the energy available to oocytes, compromising their quality and maturation. Furthermore, adequate energy metabolism is vital for the early stages of embryo development and successful implantation[50]. The ETFBKMT and ACSM5 genes have been identified as significant genes related to CCR, DPR, and HCR traits based on TWAS results in the liver. ETFBKMT is essential for biological processes related to the "generation of precursor metabolites and energy." Previous research has shown that during the early luteal phase in heifers, proteins associated with carbohydrate metabolism were observed at elevated levels[51]. This increase is attributed to changes in energy utilization and protein synthesis that occur during embryonic development. Additionally, the bovine embryo depends on the uterine environment for vital nutrients, including growth factors, amino acids, and ions, until it reaches the implantation stage[52]. ETFBKMT is involved in mitochondrial electron transport by methylating the electron transfer flavoprotein[53]. This activity impacts mitochondrial function and energy metabolism. Mitochondria are essential for ATP production, which supports oocyte maturation and early embryonic development. Altered mitochondrial function due to ETFBKMT could impair oocyte quality by reducing energy production and contributing to oxidative stress. The quality of mitochondria in oocytes is also linked to successful fertilization and embryo development[54]. ACSM5 is involved in fatty acid metabolism and the conjugation of carboxylic acids[55] as well as in fat deposition[56], fat synthesis and metabolism[57]. Fatty acid

metabolism is important for energy production during oocyte maturation and early embryo development. Dysfunction of ACSM5 could impair energy metabolism in oocytes, leading to suboptimal oocyte quality and early embryonic development. Ledgard et al. discovered that the endometrium demonstrates heightened metabolic activity during pregnancy[58]. NUDT9 and TIGAR are involved in the biological processes associated with "nucleoside monophosphate metabolic process," "purine nucleoside monophosphate metabolic process," "ribonucleoside monophosphate metabolic process," and "purine ribonucleoside monophosphate metabolic process." Therefore, NUDT9 and TIGAR play roles in metabolism[59], and bovine congestive heart failure[60], respectively NUDT9 is involved in the hydrolysis of ADP-ribose, a molecule important for cellular signaling and energy balance. Disruption of NUDT9 could lead to abnormal signaling and energy imbalances in oocytes, negatively affecting oocyte quality and embryo viability. TIGAR regulates glycolysis and protects cells from oxidative stress by modulating glucose metabolism. Energy metabolism and oxidative stress regulation are crucial for oocyte health and embryo development. TIGAR could protect oocytes from oxidative damage, promoting higher oocyte quality and successful embryo development[61]. The ARHGAP44 gene is significantly downregulated in both individual datasets of ovary and in the meta-analysis results concerning low fertility in cattle. It is involved in the biological process related to the "regulation of actin filament-based processes." ARHGAP44 has been introduced as a candidate gene associated with First Calving[62], CCR, DPR, and HCR traits[63]. ARHGAP44, a member of the Rho GTPase-activating proteins, has been studied for its role in oocyte maturation, ovulation, and early embryo development. Oocyte maturation and embryo development are influenced by genetic and molecular factors, including interactions between granulosa cells and oocytes, as well as the cellular signaling pathways that regulate folliculogenesis and oogenesis. Studies have highlighted the role of genetic factors, including ARHGAP44, in follicle development, oocyte quality, and early embryogenesis. Disruptions in these pathways can lead to defects in reproductive processes, potentially causing infertility or early embryonic arrest.[64] DKKL1 was found to be upregulated in low fertility cattle according to both individual and meta-analysis results involving ovarian tissue. This gene is also associated with weak spermatogenesis[65] and infertility[66]. In contrast, LGI4 was significantly downregulated in the ovary and endometrium, as observed in one of the individual datasets and the meta-analysis results related to low fertility in cattle. It has been shown to play a role in axon segregation and myelin formation in cattle[67,68]. ZNF613 and ZNF404 were significantly downregulated in both the individual dataset and the meta-analysis of ovarian tissue from low fertility cattle compared to high fertility cattle. ZNF613 was identified as a significant gene based on TWAS and GWAS summary results. Additionally, based on QTL analysis, ZNF613 has been recognized as an important gene related to calving difficulty[69], , lengthening of gestation[70], body depth, conception rate, sire calving ease[71–73], fertility[74], and longevity[75]. ZNF613 is thought to play a significant role in gene regulation, specifically impacting processes such as cell growth, differentiation, and apoptosis. These mechanisms are essential for ovarian follicle development, oocyte maturation, and embryo growth. Dysregulation of ZNF613 could potentially impair folliculogenesis or reduce oocyte quality, ultimately influencing fertility[74]. The FBXO40 gene is downregulated in the ovary and endometrium of low fertility cattle. FBXO40 is a candidate gene related to meat production[76], and body weight[77]. NLRP12 gene has a significant role in "negative regulation of interleukin-1 secretion" term of BP. The NLRP12 gene plays a significant role in the biological process of "negative regulation of interleukin-1 secretion." It is crucial in regulating inflammation and immune responses, primarily through its involvement in inflammasome activity. NLRP12 is known to activate caspase-1, leading to the cleavage of pro-inflammatory cytokines like IL-1β. This activity is essential for modulating immune responses, which are critical during reproductive processes such as ovulation, corpus luteum formation, and embryo implantation[78]. Research indicates that NLRP12's regulation of inflammasome activity may influence local inflammatory responses in ovarian and uterine tissues, impacting folliculogenesis and early embryo development. Its activation has also been associated with neutrophil recruitment and other immune responses during viral infections, further supporting its role in managing inflammation across various biological contexts, including reproduction[78].

Based on differential expression and WGCNA analysis results, we identified additional candidate genes, including RPL12, UMPS, FKBP1B, FZD10, PIK3R1, THBS1, VWF, TNN, MASP1, SGK1, DSTN, GPI, TGM2, NCOR1, AOX1, ITPR1, SMOC2, RPL11, APOA1, PSMB5, ACTR2, CFL1, ATP6VOD1, PAIP2, COX7A2, PTP4A1, ARPC5, and TXNDC5. All of these genes were identified as significant based on TWAS, GWAS, individual datasets, WGCNA, or meta-analysis results.

We employed multiple bioinformatics methods to identify these genes, and the results appear to be accurate. However, we strongly recommend conducting further detailed laboratory experiments to clarify the effects of these introduced genes on reproductive traits. Additionally, it would be beneficial to explore the regulatory effects of the identified key genes.

Conclusion

In the present study, we introduced candidate genes potentially related to CCR, DPR, and HCR traits based on GWAS, TWAS, RNA-seq data analysis, meta-analysis, WGCNA, GO, and KEGG pathway analysis for each tissue. The study also explored the correlation between different tissues using findings from TWAS, revealing that the uterus, liver, lung, and muscle show connections with 14, 12, 11, and 14 other tissues, respectively.

We discovered that the RPL12, UMPS, FKBP1B, FZD10, PIK3R1, THBS1, VWF, TNN, MASP1, SGK1, DSTN, GPI, TGM2, NCOR1, AOX1, ITPR1, SMOC2, RPL11, APOA1, COX10, GAA, ETFBKMT, ACSM5, NUDT9, TIGAR, COX7A2, PAIP2, and PSMB5 genes are involved in certain metabolic pathways related to carbohydrates, lipids, and proteins. Irregular metabolism can lead to ovarian dysfunction, resulting in decreased or absent ovulation, which is necessary for fertilization and marks the initial stage of pregnancy. Additionally, the mentioned genes are involved in the immune system, oxidative phosphorylation, and maintaining a functional electron transport chain.

Data availability

References

1. Albarrán-Portillo, B. & Pollott, G. E. The relationship between fertility and lactation characteristics in Holstein cows on United Kingdom commercial dairy farms. *J. Dairy. Sci.* **96**, 635–646 (2013).
2. de Fonseca, P. A. Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS One.* **13**, e0205295 (2018).
3. Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J. & Hansen, P. J. Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. *Anim. Genet.* **47**, 288–297 (2016).
4. Tian, D. et al. GWAS Atlas: a curated resource of genome-wide variant-trait associations in plants and animals. *Nucleic Acids Res.* **48**, D927–D932 (2020).
5. Barbeira, A. N. et al. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol.* **22**, 1–24 (2021).
6. Benner, C. et al. Efficient variable selection using summary data from genome-wide association studies. *Bioinf. 32.* **FINEMAP**, 1493–1501 (2016).
7. Li, M. X., Gui, H. S., Kwan, J. S. H. & Sham, P. C. GATES: a rapid and powerful gene-based association test using extended Simes procedure. *Am. J. Hum. Genet.* **88**, 283–293 (2011).
8. Guo, H. et al. Integration of disease association and eQTL data using a bayesian colocalisation approach highlights six candidate causal genes in immune-mediated diseases. *Hum. Mol. Genet.* **24**, 3305–3313 (2015).
9. Zhang, J. & Zhao, H. eQTL studies: from bulk tissues to single cells. *J. Genet. Genomics* (2023).
10. Chen, J. W., Shrestha, L., Green, G. & Leier, A. Marquez-Lago, T. T. The hitchhikers' guide to RNA sequencing and functional analysis. *Brief. Bioinform.* **24**, bbac529 (2023).
11. de Souza Fonseca, P. A., Suárez-Vega, A. & Cánovas, A. Unrevealing functional candidate genes for bovine fertility through RNA sequencing meta-analysis and regulatory elements networks of co-expressed genes and lncRNAs. *Funct. Integr. Genomics.* **22**, 1361–1376 (2022).
12. Li, W. et al. Integrating sperm cell transcriptome and seminal plasma metabolome to analyze the molecular regulatory mechanism of sperm motility in Holstein stud bulls. *J. Anim. Sci.* **101**, skad214 (2023).
13. Mancuso, N. et al. Integrating Gene expression with Summary Association Statistics To Identify Genes Associated with 30 complex traits. *Am. J. Hum. Genet.* **100**, 473–487 (2017).
14. Wainberg, M. et al. Opportunities and challenges for transcriptome-wide association studies. *Nat. Genet.* **51**, 592–599 (2019).
15. Hosseinzadeh, S., Rafat, S. A., Javanmard, A. & Fang, L. Identification of candidate genes associated with milk production and mastitis based on transcriptome-wide association study. *Anim. Genet.* (2024).
16. Moraes, J. et al. (ed N., G.) Uterine influences on conceptus development in fertility-classified animals. *Proc. Natl. Acad. Sci.* **115** E1749–E1758 (2018).
17. Moore, S. G. et al. Differentially expressed genes in endometrium and corpus luteum of Holstein cows selected for high and low fertility are enriched for sequence variants associated with fertility. *Biol. Reprod.* **94**, 11–19 (2016).
18. Hapala, P. et al. Mapping the electrostatic force field of single molecules from high-resolution scanning probe images. *Nat. Commun.* **7**, 11560 (2016).
19. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
20. Hosseinzadeh, S. & Hasanpur, K. Gene expression networks and functionally enriched pathways involved in the response of domestic chicken to acute heat stress. *Front. Genet.* **14**, 699 (2023).
21. Hosseinzadeh, S. & Hasanpur, K. Whole genome discovery of regulatory genes responsible for the response of chicken to heat stress. *Sci. Rep.* **14**, 6544 (2024).
22. da Silva, M. I. & Ott, T. Effects of conceptus proteins on endometrium and blood leukocytes of dairy cattle using transcriptome and meta-analysis. *bioRxiv* (2024).
23. Guo, H., Xue, R. & Wan, G. Identification of biomarkers associated with ferroptosis in diabetic retinopathy based on WGCNA and machine learning. *Front. Genet.* **15**, 1376771 (2024).
24. Sheet, S., Jang, S. S., Kim, J. H., Park, W. & Kim, D. A transcriptomic analysis of skeletal muscle tissues reveals promising candidate genes and pathways accountable for different daily weight gain in Hanwoo cattle. *Sci. Rep.* **14**, 315 (2024).
25. Garg, P., Vanamamalai, V. K. & Sharma, S. In-silico analysis of cattle blood transcriptome to identify lncRNAs and their role during bovine tuberculosis. *Sci. Rep.* **14**, 16537 (2024).
26. Olivoto, T. & Lúcio, A. D. Metan: an R package for multi-environment trial analysis. *Methods Ecol. Evol.* **11**, 783–789 (2020).
27. Love, M. I., Huber, W. & Anders, S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
28. Marot, G. & Rau, A. metaRNASeq (2013).
29. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
30. Olasege, B. S. et al. Genetic parameter estimates for male and female fertility traits using genomic data to improve fertility in Australian beef cattle. *Anim. Prod. Sci.* (2021).
31. Meier, S. et al. Heifers with positive genetic merit for fertility traits reach puberty earlier and have a greater pregnancy rate than heifers with negative genetic merit for fertility traits. *J. Dairy. Sci.* **104**, 3707–3721 (2021).
32. Seo, D. et al. Long-term artificial selection of Hanwoo (Korean) cattle left genetic signatures for the breeding traits and has altered the genomic structure. *Sci. Rep.* **12**, 6438 (2022).
33. Ni, C. & Buszczak, M. Ribosome biogenesis and function in development and disease. *Development* **150**, dev201187 (2023).
34. Jansova, D., Tetkova, A., Koncicka, M., Kubelka, M. & Susor, A. Localization of RNA and translation in the mammalian oocyte and embryo. *PLoS One.* **13**, e0192544 (2018).
35. Fumagalli, M., Lecca, D., Abbracchio, M. P. & Ceruti, S. Pathophysiological role of purines and pyrimidines in neurodevelopment: Unveiling new pharmacological approaches to congenital brain diseases. *Front. Pharmacol.* **8**, 941 (2017).
36. Sun, B. & Yeh, J. Calcium oscillatory patterns and oocyte activation during fertilization: A possible mechanism for total fertilization failure (TFF) in human in vitro fertilization? *Reprod. Sci.* **28**, 639–648 (2021).
37. Gifford, J. A. H. The role of WNT signaling in adult ovarian folliculogenesis. *Reproduction* **150**, R137–R148 (2015).
38. Li, T. et al. Role of the PI3K-Akt signaling pathway in the pathogenesis of polycystic ovary syndrome. *Reprod. Sci.* **24**, 646–655 (2017).
39. Dompe, C. et al. Human granulosa cells—stemness properties, molecular cross-talk and follicular angiogenesis. *Cells* **10**, 1396 (2021).

40. Lou, Y., Hu, M., Mao, L., Zheng, Y. & Jin, F. Involvement of serum glucocorticoid-regulated kinase 1 in reproductive success. *FASEB J.* **31**, 447–456 (2017).

41. Marie, C. et al. Dysfunction of Human Estrogen Signaling as a Novel Molecular signature of polycystic ovary syndrome. *Int. J. Mol. Sci.* **24**, 16689 (2023).

42. Vicente, J. S. et al. Oocyte quality and in vivo embryo survival after ovarian stimulation in nulliparous and multiparous rabbit does. *Theriogenology* **189**, 53–58 (2022).

43. Liu, K. et al. Proteomics profiling reveals lipid metabolism abnormalities during oogenesis in unexplained recurrent pregnancy loss. *Front. Immunol.* **15**, 1397633 (2024).

44. Liang, Z. et al. A million-cow genome-wide association study of three fertility traits in US Holstein cows. *Int. J. Mol. Sci.* **24**, 10496 (2023).

45. Diaz, F. Cytochrome c oxidase deficiency: patients and animal models. *Biochim. Biophys. Acta (BBA)-Molecular Basis Dis.* **1802**, 100–110 (2010).

46. Mansilla, N., Racca, S., Gras, D. E., Gonzalez, D. H. & Welchen, E. The complexity of mitochondrial complex IV: An update of cytochrome c oxidase biogenesis in plants. *Int. J. Mol. Sci.* **19**, 662 (2018).

47. Yildirim, R. M. & Seli, E. The role of mitochondrial dynamics in oocyte and early embryo development. in *Semin. Cell Dev. Biol.* **159** 52–61 (Elsevier, 2024).

48. Laqqan, M. M. & Yassin, M. M. Influence of tobacco cigarette heavy smoking on DNA methylation patterns and transcription levels of MAPK8IP3, GAA, ANXA2, PRRC2A, and PDE11A genes in human spermatozoa. *Middle East. Fertil. Soc. J.* **26**, 1–16 (2021).

49. Berg, M. D. & Dean, M. The glycogenolytic enzyme acid α-glucosidase is expressed in the bovine uterine endometrium. *Reprod. Domest. Anim.* **59**, e14643 (2024).

50. Sousa, L. C. O. et al. Dietary guanidinoacetic acid as arginine spare molecule for beef cows at late gestation: effects on cow's performance and metabolism, and offspring growth and development. *Anim. Feed Sci. Technol.* **315**, 116047 (2024).

51. Faulkner, S., Elia, G., O'Boyle, P., Dunn, M. & Morris, D. Composition of the bovine uterine proteome is associated with stage of cycle and concentration of systemic progesterone. *Proteomics* **13**, 3333–3353 (2013).

52. Morris, D. & Diskin, M. Effect of progesterone on embryo survival. *Animal* **2**, 1112–1119 (2008).

53. Monné, M., Marobbio, C. M. T., Agrimi, G., Palmieri, L. & Palmieri, F. Mitochondrial transport and metabolism of the major methyl donor and versatile cofactor S-adenosylmethionine, and related diseases: a review. *IUBMB Life.* **74**, 573–591 (2022).

54. Tilly, J. L. & Sinclair, D. A. Germline energetics, aging, and female infertility. *Cell. Metab.* **17**, 838–850 (2013).

55. Criado-Mesas, L. et al. Identification of eQTLs associated with lipid metabolism in Longissimus dorsi muscle of pigs with different genetic backgrounds. *Sci. Rep.* **10**, 9845 (2020).

56. Gu, L. et al. Integrated analysis of lncRNA and gene expression in longissimus dorsi muscle at two developmental stages of Hainan black goats. *PLoS One.* **17**, e0276004 (2022).

57. Du, J., Wang, X., Nie, Q., Yang, J. & Yao, X. Computational study of the binding mechanism of medium chain acyl-CoA synthetase with substrate in Methanosarcina acetivorans. *J. Biotechnol.* **259**, 160–167 (2017).

58. Ledgard, A. M., Lee, R. & Peterson, A. J. Bovine endometrial legumain and TIMP-2 regulation in response to presence of a conceptus. *Mol. Reprod. Dev. Inc Gamete Res.* **76**, 65–74 (2009).

59. Iung, L. H., de Mulder, S., Neves, H. A., Carvalheiro, R. & H. H. de R. & Genomic regions underlying uniformity of yearling weight in Nellore cattle evaluated under different response variables. *BMC Genom.* **19**, 1–13 (2018).

60. Heaton, M. P. et al. Association of ARRDC3 and NFIA variants with bovine congestive heart failure in feedlot cattle. *F1000Research* 11, 385 (2022).

61. Tang, J., Chen, L., Qin, Z. & Sheng, R. Structure, regulation, and biological functions of TIGAR and its role in diseases. *Acta Pharmacol. Sin.* **42**, 1547–1555 (2021).

62. Prakapenka, D., Liang, Z. & Da, Y. Genome-wide association study of age at first calving in US Holstein cows. *Int. J. Mol. Sci.* **24**, 7109 (2023).

63. Luan, M. et al. Genome-wide association study for loin muscle area of commercial crossbred pigs. *Anim. Biosci.* **36**, 861 (2023).

64. Saadeldin, I. M., Tukur, H. A., Aljumaah, R. S. & Sindi, R. A. Rocking the boat: the decisive roles of rho kinases during oocyte, blastocyst, and stem cell development. *Front. cell. Dev. Biol.* **8**, 616762 (2021).

65. Wang, Y. et al. Whole-genome identification and characterization of the DKK Gene Family and its transcription profiles: an analysis of the Chinese Soft-Shell Turtle (Pelodiscus sinensis). *Animals* 14, 931 (2024).

66. Yan, Q. X. et al. Expression of DKKL1 in spermatozoa of men with asthenospermia. *Nan Fang Yi Ke da xue xue bao = J. South. Med. Univ.* **38**, 324–328 (2018).

67. Xue, S. et al. Loss-of-function mutations in LGI4, a secreted ligand involved in Schwann cell myelination, are responsible for arthrogryposis multiplex congenita. *Am. J. Hum. Genet.* **100**, 659–665 (2017).

68. Previtali, S. C. Peripheral nerve development and the pathogenesis of peripheral neuropathy: The sorting point. *Neurotherapeutics* **18**, 2156–2168 (2021).

69. Purfield, D. C., Evans, R. D. & Berry, D. P. Breed-and trait-specific associations define the genetic architecture of calving performance traits in cattle. *J. Anim. Sci.* **98**, skaa151 (2020).

70. Purfield, D. C., Evans, R. D., Carthy, T. R. & Berry, D. P. Genomic regions associated with gestation length detected using whole-genome sequence data differ between dairy and beef cattle. *Front. Genet.* **10**, 1068 (2019).

71. Purfield, D. C., Bradley, D. G., Kearney, J. F. & Berry, D. P. Genome-wide association study for calving traits in holstein–friesian dairy cattle. *Animal* **8**, 224–235 (2014).

72. Mao, X. et al. Fine mapping of a calving QTL on Bos taurus autosome 18 in Holstein cattle. *J. Anim. Breed. Genet.* **133**, 207–218 (2016).

73. Fang, L. et al. Genetic and epigenetic architecture of paternal origin contribute to gestation length in cattle. *Commun. Biol.* **2**, 100 (2019).

74. Ooi, E., Xiang, R., Chamberlain, A. J. & Goddard, M. E. Archetypal clustering reveals physiological mechanisms linking milk yield and fertility in dairy cattle. *J. Dairy. Sci.* (2024).

75. Zhang, Q., Guldbrandtsen, B., Thomasen, J. R., Lund, M. S. & Sahana, G. Genome-wide association study for longevity with whole-genome sequencing in 3 cattle breeds. *J. Dairy. Sci.* **99**, 7289–7298 (2016).

76. Gao, F. et al. Molecular breeding of farm animals through gene editing. *Natl. Sci. Open.* **2**, 20220066 (2023).

77. Shi, J., Luo, L., Eash, J., Ibebunjo, C. & Glass, D. J. The SCF-Fbxo40 complex induces IRS1 ubiquitination in skeletal muscle, limiting IGF1 signaling. *Dev. Cell.* **21**, 835–847 (2011).

78. Wei, B., Billman, Z. P., Nozaki, K., Goodridge, H. S. & Miao, E. A. NLRP3, NLRP6, and NLRP12 are inflammasomes with distinct expression patterns. *Front. Immunol.* **15**, 1418290 (2024).

Acknowledgements

Author contributions

S.H. Main researcher, data analyzer, writer of manuscript. S. A. R. Main researcher, idea, revising the paper, L.F. idea, access guide to data.

Funding

Declarations

Ethics approval and consent to participate

All animal-related experimental protocols conducted in this research study received authorization from the Animal Ethics Committee at the Department of Animal Science, University of Tabriz, Iran. Our adherence to ARRIVE guidelines has been ensured during the submission process.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-82448-x.

Correspondence and requests for materials should be addressed to S.H. or S.A.R.

Reprints and permissions information is available at www.nature.com/reprints.

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