



Article

Weighted Gene Correlation Network Meta-Analysis Reveals Functional Candidate Genes Associated with High- and Sub-Fertile Reproductive Performance in Beef Cattle

Pablo A. S. Fonseca ^{*}, Aroa Suárez-Vega and Angela Cánovas ^{*}

Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, ON N1G 2W1, Canada; asuarezv@uoguelph.ca

* Correspondence: pfonseca@uoguelph.ca (P.A.S.F.); acanovas@uoguelph.ca (A.C.);
Tel.: +1-519-824-4120 (ext. 56295) (A.C.)

Received: 22 April 2020; Accepted: 6 May 2020; Published: 12 May 2020



Abstract: Improved reproductive efficiency could lead to economic benefits for the beef industry, once the intensive selection pressure has led to a decreased fertility. However, several factors limit our understanding of fertility traits, including genetic differences between populations and statistical limitations. In the present study, the RNA-sequencing data from uterine samples of high-fertile (HF) and sub-fertile (SF) animals was integrated using co-expression network meta-analysis, weighted gene correlation network analysis, identification of upstream regulators, variant calling, and network topology approaches. Using this pipeline, top hub-genes harboring fixed variants (HF × SF) were identified in differentially co-expressed gene modules (DcoExp). The functional prioritization analysis identified the genes with highest potential to be key-regulators of the DcoExp modules between HF and SF animals. Consequently, 32 functional candidate genes (10 upstream regulators and 22 top hub-genes of DcoExp modules) were identified. These genes were associated with the regulation of relevant biological processes for fertility, such as embryonic development, germ cell proliferation, and ovarian hormone regulation. Additionally, 100 candidate variants (single nucleotide polymorphisms (SNPs) and insertions and deletions (INDELs)) were identified within those genes. In the long-term, the results obtained here may help to reduce the frequency of subfertility in beef herds, reducing the associated economic losses caused by this condition.

Keywords: meta-analysis; RNA-sequencing; gene network; functional candidate genes; systems biology; subfertility; beef cattle

1. Introduction

Intensive selection pressure has led to a decrease in fertility efficiency in both beef and dairy cattle populations [1,2]. Genetic mechanisms such as pleiotropy, genetic hitchhiking, and epistasis can be the cause of the genetic correlations, and consequent undesirable effects, observed between production and fertility traits [3–7]. Poor fertility and reproductive inefficiency are among the main causes of the negative impact on the profitability of both beef and dairy herds [8,9]. Embryo mortality is the major factor affecting fertility and production costs, with the majority of pregnancy losses occurring in the first month of pregnancy [10,11]. Regarding the comparison between infertility and subfertility, sub-fertile animals are more pervasive in the herds because of the fact that true infertility has a frequency of up to 5% in the herds [12–14]. Additionally, sub-fertile animals can generate progeny, consequently maintaining the causal alleles for this phenotype in the population [15]. On the other hand, the causal alleles for the infertile phenotype tend to reduce its allelic frequency naturally across time owing to the

absence of progeny carrying the alleles. Therefore, sub-fertile animals have a higher probability to have a higher cost in the livestock industry.

Reproductive traits are considered complex phenotypes as they present a high heterogeneity, high environmental impact, and do not follow a Mendelian inheritance pattern [16]. In this sense, investigating the genes involved in complex phenotypes is not a trivial task. Especially when using high-throughput genetic tools, which usually demand a high number of samples and a high accuracy of the phenotype evaluated to obtain a consistent and significant result. RNA-sequencing technology has allowed the identification of several candidate genes and genetic variants associated with fertility traits in cattle in the last decade [17–24]. However, the majority of these studies are focused on the conventional gene by gene differential expression analysis. Other than to provide significant results to understand the genetic basis of complex traits, this approach may result in an underrepresentation of the genetic interactions between genes. The use of co-expression gene networks accounts for the expression profile across multiple samples, leading to the identification of regulatory and functional mechanisms in common [25]. The guilt-by-association (GBA) principle is one of the major measurements to evaluate the quality of co-expression networks, where genes with similar functional activities tend to have a similar expression profile, and consequently, a higher co-expression [26]. Meta-analysis approaches tend to enhance the performance of co-expression networks when the GBA principle is evaluated [27–31]. The application of network meta-analysis using high-throughput expression data is relatively new and can help to improve the detection of differentially expressed genes (DEGs) and to reduce the impact of differences between studies that can be hard to remove, such as bias during the library preparation step, which will implicate in spurious differences between groups [27,32]. This is reinforced by the stronger correlation observed between true log (fold-change) values and the values obtained in the network meta-analyses when these values are compared with those obtained from meta-analyses performed just merging datasets [27]. Consequently, the integration of both approaches (co-expression gene networks and network meta-analysis) can be a good alternative to increase the potential to identify functional candidate genes regulating a complex trait.

In the present work, we performed the integration of network meta-analysis and weighted gene correlation network analysis (WGCNA) approaches in order to scrutinize the co-expression and the genetic basis of high-fertile and sub-fertile phenotypes in beef cattle. Additionally, potential functional candidate variants, fixed in one of the phenotypic groups, were prospected using the RNA-sequencing data.

2. Materials and Methods

2.1. Ethics Approval and Consent to Participate

The current study integrates data from previously published studies. The respective information about the ethics approval can be found in Moraes et al., (2018) [18] and Geary et al. (2016) [20].

2.2. Data Collection

The RNA-sequencing (RNA-seq) data from the endometrium tissue of high- and sub-fertile beef cows (HF and SF, respectively) were retrieved from National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) public database from two previously published studies: GSE81449 and GSE107891 [18,20]. In these studies, the differentially expressed genes between HF and SF beef cows were identified. A total of 20 animals (10 animals per group; HF ($n = 10$) and SF ($n = 10$) from both studies) were analyzed. Briefly, the fertility status of those animals was based on the pregnancy outcome ratio after up to four rounds of successive high-quality embryo transfer protocol of estrus synchronization (PG-6d-CIDR and GnRH), where heifers that did not exhibit standing estrus received GnRH injection on day 0. As described by Moraes et al., (2018) and Geary et al., (2016) [18,20], the pregnancy outcome was detected by ultrasound and those animals with a pregnancy success ratio equal to 100% or 25%–33% were classified as HF and SF, respectively. Additional details about the

breed composition, synchronization protocol, flushes, biopsies, RNA extraction, and sequencing can be found in the original manuscripts [18,20].

2.3. RNA-Sequencing Data Alignment and Variant Calling

The CLC Genomics Workbench 11.0 (CLC bio, Cambridge, MA, US) was used to perform quality control (QC), read alignment, transcript quantification, and variant calling [33–35]. In QC, the PHRED score distribution, GC content, nucleotide contribution, and duplication levels were evaluated as described by Cánovas et al., (2014) [36]. Sequencing reads were aligned against the bovine reference genome ARS-UCD1.2 [37] using the “Map reads to reference” algorithm with the following criteria: match score = 1; mismatch cost = 2, length fraction = 0.5, and similarity fraction = 0.8. Subsequently, we quantified transcript expression (total counts) and only those genes with a fragments per kilobase of exon model per million reads mapped (FPKM) > 0.2 in both conditions (HF and SF) were maintained for the next analyses [38,39]. The variant calling was performed using the fixed ploidy variant detect algorithm (diploidy genome) on CLC Genome Workbench. A required variant probability >90%, a minimum coverage of 10, and a minimum count of 2 were set for the variant detection [34]. The base quality filter was performed using a neighborhood radius = 5, minimum central quality = 20, and minimum neighborhood quality = 15 [24]. Genetic variants (single nucleotide polymorphism, SNP; and insertion and deletion, INDEL) fixed in one of the groups were selected as potential functional variants for further analyses.

2.4. Identification of Genes with Expression Determined by the Study and Outliers

After filtering those genes with an FPKM > 0.2 in both conditions, the raw counts were used to perform a log-likelihood ratio test (LRT) in the DESeq2 package in R [40] in order to estimate the impact of different studies over the gene expression. Those genes with a differential expression significantly affected (adjusted false discovery rate (FDR) 5%, p -value < 0.05) by the different studies (GSE81449 and GSE107891), not by the conditions (HF and SF), were excluded from the analysis. Additionally, the counts for the maintained genes were used to perform a clustering analysis in order to identify potential outliers among the samples. The Manhattan distance among the animals was calculated and used in a multidimensional scaling analysis in order to cluster the animals using the first two principal components. These analyses were performed using the function `cmdscale` in R. The possible outliers were removed from the next steps. In this study, the outliers were classified as those animals that did not cluster following the condition (HF and SF).

2.5. Meta-Analysis of Differentially Expressed Genes

The DEGs were identified using the DESeq2 package in R, where a negative binomial generalized linear model was used using as a fixed effect the condition (HF and SF) of each animal [40]. Initially, this analysis was conducted for each study individually. The threshold to define a gene as DE in each dataset was maintained as described previously by Geary et al., (2016) [20] (GSE81449; adjusted p -value FDR 5% < 0.1 and $|\log(\text{fold-change (FC)})| > 2$) and Moraes et al., (2018) [18] (GSE107891; adjusted p -value FDR 5% < 0.05 and $|\log(\text{FC})| > 2$). Subsequently, the `netmeta` package in R [41] was used to perform a network meta-analysis and calculated the combined test-statistics (p -values and $\log_2(\text{FC})$) for each gene expressed in both datasets. The DEGs in the network meta-analysis approach were identified using a threshold composed by adjusted p -value < 0.1 and $|\log_2(\text{FC})| > 2$, which is the combination of the less stringent threshold from both studies.

2.6. Weighted Correlation Network Analysis

Once the comparability between the datasets was confirmed, the R package WGCNA (Weighted Correlation Network Analysis) [42] was used to identify the differentially co-expressed modules (DcoEx) of genes for HF and SF groups of animals. Briefly, after QC, a soft-thresholding power was chosen based on a criterion of approximate scale-free topology. The first soft-thresholding power to

reach a scale-free topology model fit ≥ 0.8 was selected for each group. Subsequently, the co-expression similarity matrix is raised to this soft-thresholding (8 for HF and 8 for SF) power in order to obtain the adjacency matrix. Consequently, this last matrix was used to calculate the topological overlap measure (TOM). The adjacency matrix was calculated using a signed hybrid network. The module detection was performed using the `blockwiseModules` function using the dynamic tree cut algorithm [43] with a minimum number of genes per module equal to 30 and the maximum size of blocks equal to 9000 (this number was selected to fit all the genes in a single module, which might increase the module detection sensibility). After the module detection by WGCNA package, the R package `km2gcn` [44] was used to reallocate the genes within modules using a k-means clustering approach. Finally, the final modules detected for each group were compared using the following methodology:

- For each sample s in (HF and SF);
- For each module $m(s)$ in s ;
- Apply a Fisher's exact test under the null hypothesis that there is no significant overlapping of $m(\text{HF})$ in SF and $m(\text{SF})$ in HF after a Bonferroni multiple test correction.

At the end of this step, the modules of genes in HF without overlapping in SF, and vice-versa, here called DcoExp modules, were selected. The interconnectivity among the genes within each DcoExp module was plotted using the `igraph` package [45]. In addition, the hub score of the genes within each module was estimated using the principal eigenvector of $A^t(A)$, where A is the adjacency matrix of the module. From these results, the top 10 genes explaining the majority of the module topology were identified.

2.7. Functional Analysis and Annotation of Candidate Genes

The DcoExp modules with top-hub genes harboring variants fixed in one of the conditions (HF and SF) were selected for the functional analysis, here called candidate DcoExp modules. The functional analysis was conducted using the "Core Analysis" function implemented in the ingenuity pathways analysis (IPA—Ingenuity System Inc, Redwood City, CA, USA). Genes without an associated gene symbol or without gene annotation were subjected to an annotation by homology. The BioMart application [46] was used to retrieve the respective associated human gene symbol for those genes. Only the non-annotated genes, with a percentage of identity with the human homolog higher than 80%, were annotated by this approach. The enriched (p -value < 0.05) canonical pathways and diseases and functions for each candidate DcoExp modules were annotated. Additionally, the significant upstream regulators (FDR < 0.05 multiple testing correction) for each candidate DcoExp modules were identified. The functional candidate genes were selected among those genes within candidate DcoExp modules or in the significant upstream regulators that harbor fixed variants and are among those genes in which the expression profile was not determined by the study.

Additionally, a "guilt by association"-based prioritization approach was performed using the GUILDify and ToppGene applications on the functional candidate genes [47,48]. Initially, GUILDify was used to retrieve a "trained-list" of candidate genes associated with pre-selected phenotypes. After this step, the "trained list" obtained using GUILDify and the list of functional candidate genes are used in ToppGene. Briefly, GUILDify uses BIANA knowledge base to identify genes associated with selected phenotypes. In the present study, the phenotypes used on GUILDify were as follows: "fertility", "fertilization", "decidualization", "implantation", "preimplantation", "endometrium", and "embryonic development". BIANA creates a species-specific (human, in this study) interaction network for each gene identified by GUILDify. Subsequently, a prioritization algorithm based on network topology is used to rank the genes. Genes with a GUILD score higher than 0.44 (mean + 3 \times standard deviation of GUILD score) were used to create the "trained" gene list. This list was used on ToppGene and a functional annotation-based prioritization was performed using a fuzzy-based multivariate approach. ToppGene uses the functional information shared among the "trained" and the functional candidate gene lists from several sources including the gene ontology (GO) terms for the three main categories of

molecular function (MF), biological process (BP), and cellular component (CC); human and mouse phenotypes; metabolic pathways; Pubmed publications; co-expression pattern; and diseases. For each functional annotation for each functional candidate gene, a p -value was calculated using a random sampling of 5000 genes from the whole genome. In the next step, using a statistical meta-analysis, the p -values were combined into a final p -value. Subsequently, the genes with a significant p -value, after an FDR of 5% multiple corrections, were selected as prioritized genes. The complete description of the meta-analysis to calculate the final p -value is available in Chen et al., (2009) [48]. Briefly, for each gene G , a p -value is computed by random sampling genes across the whole genome (5000 genes in this study) and a similarity score between G and the *functional annotation* of the trained list using different statistical approaches for categorical (e.g., GO and pathways terms) and numeric terms (expression profile in different tissues). While a fuzzy approach is applied for the categorical terms, a Pearson correlation between the expression vectors of the candidate gene and the genes in the trained list is computed. Finally, a p -value for each *functional annotation* of G is computed using a derivation from the annotation of the genes randomly sampled across the genome using the following formula:

$$p(S_i) = \frac{\binom{\text{Number of random sampled genes with similarity score higher than } G}{}}{\binom{\text{Number of genes in the random sampling processes with functional annotation}}{}} \quad (1)$$

All the p -values obtained for each G are combined using a Fisher's inverse chi-square method, where the p -values are assumed to come from an independent test:

$$-2 \sum_{i=1}^n \log p_i \rightarrow X^2(2n) \quad (2)$$

This metaanalysis approach followed by multiple testing correction substantially reduces the number of false positive functional candidate genes.

3. Results

Figure 1 summarizes the pipeline applied in the present study to identify the functional candidate genes, and candidate genetic variants harboring those genes, within the differentially co-expressed gene modules between HF and SF animals.

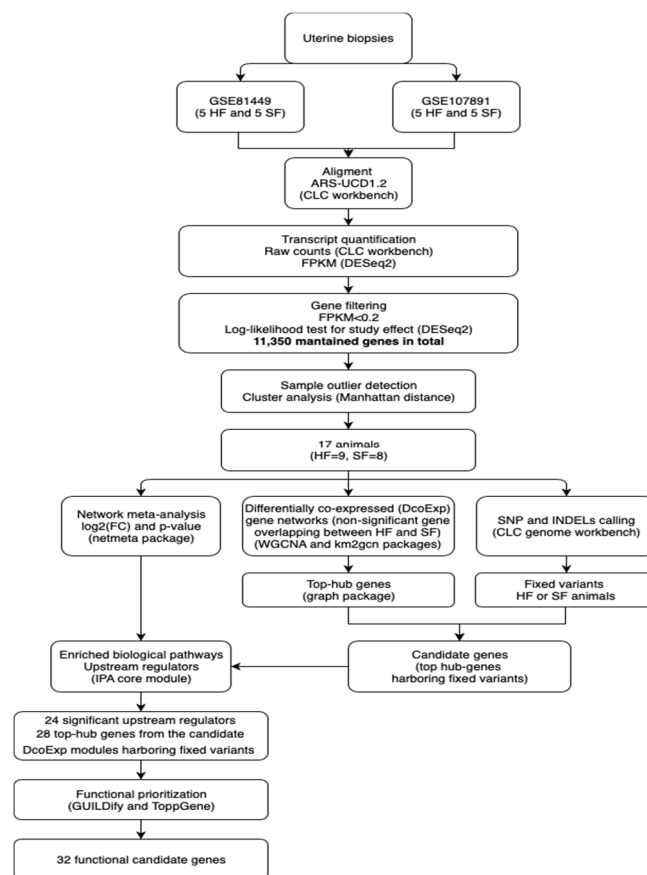


Figure 1. Pipeline for identification of functional candidate genes regulating differentially co-expressed gene networks between high-fertile and sub-fertile beef cows. SF, sub-fertile; HF, high-fertile; INDEL, insertion and deletion; SNP, single nucleotide polymorphism; FPKM, fragments per kilobase of exon model per million reads mapped; WGCNA, weighted gene correlation network analysis; IPA, ingenuity pathways analysis; FC, fold-change.

3.1. RNA-Sequencing and Variant Calling Statistics

The number and percentage of uniquely mapped reads for each sample are shown in Table S1. A total of 13,812 genes were expressed in both studies (GSE81449 and GSE107891) with an FPKM > 0.2. These genes were used for the DEG meta-analysis in order to estimate the combined test-statistics (p -value and $\log_2(\text{FC})$) for each group (HF and SF). Additionally, independently of the DEG meta-analysis, the LRT analysis indicated that 2462 genes had the expression determined by the study, not by the condition. Consequently, 11,350 genes were used to further investigate possible outliers among the samples and initially used in the WGCNA analysis.

Regarding the variant calling analysis, the genomic coordinates, type (SNP or INDEL), and functional impact for the variants fixed in one of the conditions, obtained using Ensembl Variant Effect Predictor (VeP), are shown in Table S2. A total of 2254 variants were identified as uniquely fixed in the HF group (2113 SNPs and 141 INDELS), while 3117 variants were uniquely fixed in the SF group (3034 SNPs and 83 INDELS). The percentage of each functional class for the fixed variants in each group is shown in Figure S1.

3.2. Outlier Detection and Differential Expression Analysis between High-Fertile and Sub-Fertile Animals

The outlier detection analysis resulted in the exclusion of one SF animal in the GSE81449 dataset (SRR3505358) and one animal from each condition (HF: SRX3461001 and SF: SRX3461010) in the GSE107891 dataset (Figure S2). The final sample size used in the analyses was 17 animals (HF = 9

and SF = 8). The overlapping among the DEG genes was calculated for the DEG identified in the meta-analysis performed in this study, including the DEG identified using the alignment with CLC Bio Genomics against the new bovine reference genome ARS-UCD1.2 and the DEG identified originally in the previously published studies [18,20]. A very low overlap was observed across the differential expression analyses between both studies. The network meta-analysis resulted in 14 DEGs (Figure 2a, adjusted p -value < 0.1 and $|\log_2(\text{FC})| > 2$). Additionally, despite the larger number of DEGs, the results obtained with CLC Bio Genomics and DESeq2 using as a reference genome for the mapping step the new bovine assembly ARS-UCD1.2 showed a small overlapping with the original result (maximum overlapping of 11 genes) (Figure 2b). The majority of DEGs in the network meta-analysis (using the non-adjusted p -value) are shared with the results obtained using the GSE107891 dataset when the alignment against the ARS-UCD1.2 reference genome was performed using CLC Bio Genomics. The p -values and $\log_2(\text{FC})$ for all the genes evaluated in the network meta-analysis, as well as the DEGs previously identified in GSE107891 and GSE81449, are presented in Table S3.

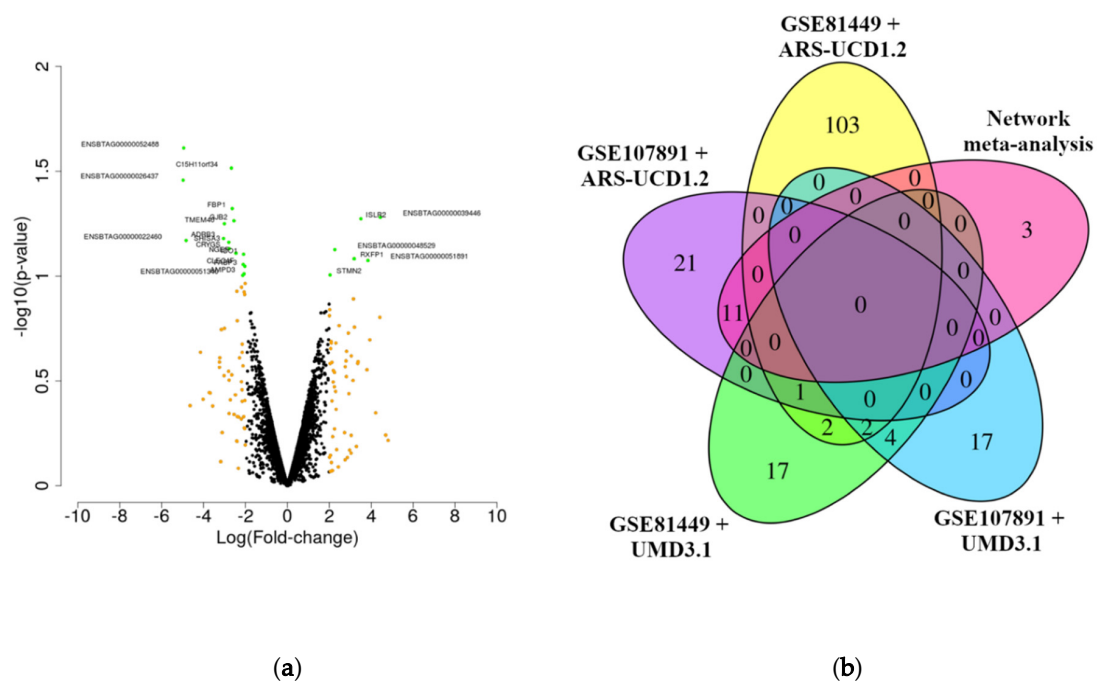


Figure 2. Volcano plot for the differentially expressed genes identified in the network meta-analysis (non-adjusted p -value < 0.1 and $|\log(\text{fold-change})| > 2$) (a) and Venn diagram comparing the results obtained in the different datasets of differentially expressed genes identified using different bovine reference genomes assemblies (b). In (a), the green dots represent differentially expressed genes identified using non-adjusted p -value < 0.1 and $|\log(\text{fold-change})| > 2$. The yellow dots represent differentially expressed genes identified using a $|\log(\text{fold-change})| > 2$. In (b), the differentially expressed genes identified in each dataset are represented in red (network meta-analysis), yellow (GSE107891 data set using ARS-UCD1.2 bovine reference genome), purple (GSE81449 data set using ARS-UCD1.2 bovine reference genome), dark green (GSE107891 data set using UMD 3.1 bovine reference genome; Moraes et al. (2018)), and blue (GSE81449 data set using UMD 3.1 bovine reference genome; Geary et al. (2016)).

3.3. Identification of Candidate Differentially Co-Expressed Gene Modules for High-Fertile and Sub-Fertile Animals

The correlation between the HF and SF gene ranked expression values (FPKM) obtained from DESeq2 package and ranked connectivity, estimated through WGCNA package, was evaluated to estimate the network conservation in the combined dataset composed of the samples from GSE107891 and GSE81449. A correlation of 0.99 ($p < 1 \times 10^{-200}$) was obtained for the ranked expression values and 0.4 (p -value $< 1.7 \times 10^{-191}$) for the ranked connectivity between the HF and SF samples. These results indicate that both datasets are suitable to be analyzed together for the identification of co-expressed gene networks owing to the strong correlation observed. Thirty-two and 34 co-expressed gene modules were identified for HF and SF, respectively. The identification of DcoEx modules (non-overlapping threshold = p -value $< 4.59 \times 10^{-5}$) resulted in 44 modules, 22 for each condition (Table S4). From them, we performed an additional filtering to select the modules with at least 1 of the 10 top-hub genes harboring fixed functional candidate variants (FCVs) identified in the variant calling process (HF = 157 FCV and SF = 214 FCV). The resulting modules, called candidate DcoExp modules, were ten and two in the HF and SF datasets, respectively (Table 1). The list of enriched biological processes, diseases, and functions, as well as upstream regulators obtained for each candidate DcoExp module using IPA core module, can be found in Table S5. It is important to highlight that, for both functional prioritization and canonical metabolic pathway enrichment analysis, we used human, mice, and rat annotations owing to the database availability and the more complete annotation status for these organisms. Consequently, it would be possible to observe small differences in the functions performed by some orthologous gene in different species. Overall, owing to the close evolutionary relationship between these species and cattle, a high level of similarities of functions performed by the orthologous genes in those species is expected.

Table 1. Differentially co-expressed gene modules (DcoExp), and their respective top 10 hub-genes between high-fertile (HF) and sub-fertile (SF) animals.

Module	Number of Genes	Top Hub-Genes
Cyan HF	204	ANKRD65, TIMM17A, ENSBTAG00000046047 , ENSBTAG000000051586, RRP1, PWP2, PSMB5, LTBP2, C11H2orf81, ENSBTAG00000050675
Darkgreen HF	147	SREBF2, DACT2 , SDHA, PPFIA4, ENSBTAG00000047824, ENSBTAG00000052047, ELF3, ENSBTAG00000051421, NPTN, DHRS4
Grey60 HF	177	IARS2, ENSBTAG00000053801, ENSBTAG00000033740 , ENSBTAG00000048975, WRB, ENSBTAG00000052845, FAM214A, EIF2AK3, MPV17, MAPKAP1
Lightgreen HF	136	CEP104, PKP1, PPP1R12B , ENSBTAG00000051541, ENSBTAG00000049133, ARF6, NUMB, SLC25A15 , EEF1AKMT1 , PARP4
Purple HF	227	TIRAP, PYCR2, FMO2 , MIIP, ENSBTAG00000049485, ENSBTAG00000054279, ENSBTAG00000052750, EAF1, SDR39U1, TINF2
Red HF	237	STRADA, ENSBTAG00000054600, MDM4 , MARK1, KLHL20, CACYBP, ABL2, RAB1F , ENSBTAG00000051120, TMEM50B
Saddlebrown HF	114	NME7, CCDC181 , ENSBTAG00000054228, TCTEX1D2, IL20RB, ITGB2, ENSBTAG00000023186, F2RL2 , OIP5, DUT
Tan HF	172	ZMAT2, CPOX, IQCG , WDR53, ENSBTAG00000042475, HACL1 , INTS14, ENSBTAG00000043377, REC8, RPS27L
Turquoise HF	544	SLC45A3, IKBKE, PIGR , PRRX1, TNFRSF1B, SPSB1, CAMTA1, NOL9, TNFRSF25, ARHGEF16
Green HF	240	CTCF, MIA3 , PSEN2, CASZ1, SDF4, COLGALT2 , ENSBTAG00000054874, ENSBTAG00000051836, ENSBTAG00000051084, GOLGB1
Lightgreen SF	244	SOX13, TMEM81, COQ8A, ZBTB48, VWA1, EFHD2, PWP2 , PLEKHO2, NRDE2 , MALL
Paleturquoise SF	130	MTHFR, ENSBTAG00000031572 , JDP2, CCDC142, CD8A, DNAJC27, ENSBTAG00000053045, ENSBTAG00000048432, CABLES2, NECAB3

Note: In bold, those genes harboring functional candidate variants are highlighted.

3.4. Functional Candidate Genes

The functional candidate genes, top-hub genes from the candidate DcoExp modules, and upstream regulators of the DcoExp modules identified with the core analysis from IPA software, harboring FCV, are shown in Tables 2 and 3. This candidate gene list is composed of a total of 52 genes: 24 top-hub genes from the candidate DcoExp modules harboring FCV, and 28 upstream regulators harboring FCV. Additionally, functional candidate genes were using a “guilt by association”-based prioritization analysis (ToppGene software) using a trained dataset of genes related to fertility obtained from GUILDify. The adjusted (FDR) overall p -value of the significantly prioritized candidate genes from the functional prioritization analysis is shown in Table 2. It is important to highlight that the top hub-gene from Cyan module (ENSBTAG00000046047) with a fixed candidate variant was not analyzed using ToppGene owing to the lack of gene symbol annotation (even after the annotation by homology process). The functional prioritization resulted in 32 significantly prioritized functional candidate genes. The enriched terms for the traied dataset and the complete prioritization result for the candidate genes are provided in Table S6. The FCV identified within the genomic coordinates of the prioritized genes and the corresponding functional consequence is shown in Table S7. The relationship between the functional candidate genes and the candidate DcoExp modules is shown in Figure 3. The turquoise HF module showed the highest number of related prioritized functional candidate genes (14), while the pale turquoise SF module showed the smallest number of related genes (1). The percentage of each functional consequence and the number of fixed variants per each prioritized functional candidate gene is shown in Figure 4. In total, 100 FCV were identified in the 32 significantly prioritized functional candidate genes.

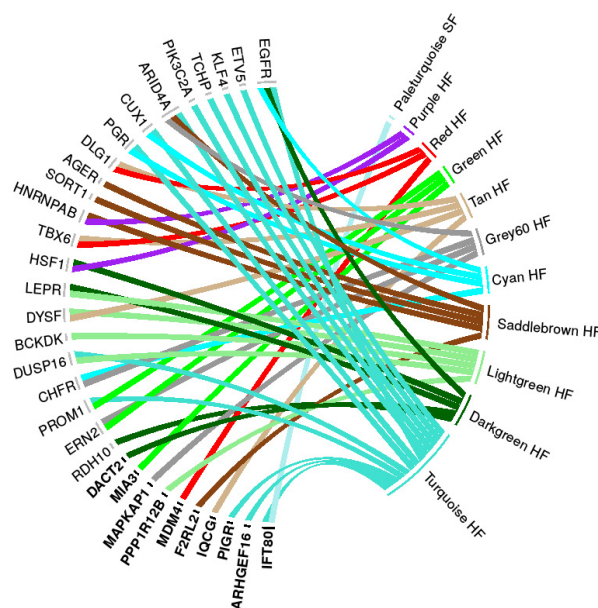


Figure 3. Relationship between the prioritized functional candidate genes (left-hand side) and the candidate differentially co-expressed modules (DcoExp, right-hand side). The genes in bold are the top hub-genes from the DcoExp modules, while the other genes are the significant upstream regulators genes. On the right-hand side, the HF modules are those identified in the high-fertile animals and the SF modules are those identified in the sub-fertile animals.

Table 3. Prioritization result for the significant upstream regulators, harboring functional candidate variants, from differentially co-expressed modules between high-fertile (HF) and sub-fertile (SF) animals.

Gene ID	Mapped Candidate Modules	Regulated Module	Target Genes	Adjusted <i>p</i> -Value (FDR 5%) Upstream Regulation	Overall Adjusted <i>p</i> -Value (FDR 5%) for Prioritization
EGFR	Darkgreen HF	Turquoise HF	BIRC5, CCNA2, CXCL5, E2F1, EXOSC5, FKBP11, FDX3, GFAP, HMGB3, HNRNP1, IGBP1, ITGA6, MYBL2, PDK1, PROM1, PSEN1, RANBP1, SEMA7A, SKP2, TUBA4A, TUBB4A, VEGFA	0.003	0.002 *
EGFR	Darkgreen HF	Cyan HF	CCT5, EIF5A, EPS15, GADD45A, NUTF2, ODC1, PPIA, PSMB5, STAT3, TPST1	0.003	0.002 *
ETV5	-	Turquoise HF	AQP5, CHSY1, KRT19, KRT7, MYB, RAB27A, TJP3, VEGFA	0.016	0.009 *
KLF4	-	Turquoise HF	CCND2, CRABP2, DUSP5, E2F1, HES1, KRT14, KRT19, KRT7, MSX2, PAX2, PROM1, VEGFA, WNT5A	0.032	0.006 *
TCHP	-	Turquoise HF	VEGFA	0.039	0.028 *
COX7A2	-	Turquoise HF	STAR	0.039	0.147
PIK3C2A	-	Turquoise HF	VEGFA	0.039	0.009 *
ARID4A	-	Turquoise HF	E2F1, FOXP3	0.039	0.046 *
ARID4A	-	Saddlebrown HF	HOXB6	0.039	0.046 *
ARID4A	-	Grey60 HF	HOXB3, HOXB5	0.039	0.046 *
CUX1	Cyan HF	Turquoise HF	CCNA2, LTF, RAB36, WNT5A	0.047	0.006 *
PGR	Turquoise HF	Turquoise HF	AK3, HES1, HPGD, ITGA6, MSX2, NPC1, PDGFA, PGR, PPM1H, PRRX1, TAT	0.047	0.006 *
PGR	Turquoise HF	Cyan HF	LIG1, MAP2K3, STAT3, TSC22D3, UCK2, URB2	0.047	0.006 *
IPO9	-	Turquoise HF	PTK2B	0.047	0.09
DLG1	Red HF	Tan HF	KCNJ2	0.045	0.006 *
AGER	-	Saddlebrown HF	CCL4, TJP1	0.046	0.006 *
SORT1	-	Saddlebrown HF	UBE2I	0.047	0.009 *
HNRNPAB	Purple HF	Saddlebrown HF	TJP1	0.047	0.031 *
TBX6	Tan HF	Red HF	HES7	0.049	0.024 *
HSF1	-	Purple HF	CCT4, FKBP4, HSF2, HSP90AA1, HSPA8, HSPH1, KNTC1, RELA, SPHK2, STIP1	0.003	0.009 *
HSF1	-	Darkgreen HF	CSRP2, EFEMP1, INHBB, RPL22	0.003	0.009 *
NUB1	-	Red HF	NEDD8	0.046	0.088
DPH5	-	Red HF	NFKBIA, RELA	0.046	0.077
LEPR	Darkgreen HF	Lightgreen HF	ANGPTL4, CDK2, MMP7, PLP1, SOCS2	0.044	0.006 *
DYSF	Tan HF	Lightgreen HF	CD48, DNABJ1, FCGR2B	0.044	0.012 *
API5	Purple HF	Lightgreen HF	CDK2	0.044	0.077
BCKDK	-	Lightgreen HF	PLP1	0.047	0.041 *
DUSP16	Turquoise HF	Lightgreen HF	VCAM1	0.047	0.014 *
CHFR	Cyan HF	Grey60 HF	PLK1	0.046	0.032 *
PROM1	Turquoise HF	Green HF	DSG2	0.044	0.006 *
ERN2	Grey60 HF	Green HF	XBP1	0.046	0.031 *
UTP3	-	Darkgreen HF	IGLL1/IGLL5	0.043	0.112
RDH10	-	Darkgreen HF	RDH5	0.045	0.032 *

* Significant prioritization at a significance level of 0.05 after false discovery rate correction.

4. Discussion

4.1. Network Meta-Analysis for Identification of Differentially Expressed Genes between High-Fertile and Sub-Fertile Animals

The integration of multiple datasets and genetic information from different levels has been shown to be a powerful strategy for the identification of candidate genes in livestock [3,49–51]. The network meta-analysis performed in the present study reinforced the association between the expression profile of some genes with the high- or sub-fertile condition. The 14 DEGs identified by the network meta-analysis (adjusted p -value < 0.1 and $|\log(\text{fold-change})| > 2$) were not identified as DE in the original results of both previous studies (GSE107891 and GSE81449) described by [18,20]. This result can be explained by the differences in the assemblies, alignment, and quantification algorithms, as well as the detection power improvement observed in the meta-analysis performed in the present study. However, 11 genes were shared with the GSE107891 dataset using the new bovine reference genome ARS-UCD1.2 by CLC BIO genomics. In general, a small overlap was observed in all of the comparison scenarios (Figure 2b). The Simpson's paradox is a common phenomenon observed in biological analysis that can help to address these differences across the analyses. Briefly, the Simpson's paradox occurs when results from combined datasets contradict those from the individual analysis. The impact of Simpson's paradox was already discussed in several fields, such as gene expression network analysis [52–54]. Despite biological bases for the Simpson's paradox still being poorly understood, there are some points that must be highlighted in the present study. First, despite both studies analyzing the transcriptome of the same tissue (endometrium), GSE81449 analyzed a dataset obtained from endometrial biopsy from day 14 post-estrus, while GSE107891 analyzed a dataset from day 17 post-estrus. This difference, together with the population effect, the new mapping strategy (different software and bovine reference genome), and the software used for DE analysis, can affect the transcriptional profile of each sample, consequently resulting in a strong study-dependent effect (as observed in the LRT analysis). Additionally, as shown in Table S3, few genes were identified as DE between HF and SF animals in the original results of both studies, with the majority of these DEGs composed by non-annotated genes (LOCs) or genes with poorly understood biological function. Here, a larger number of DEGs were obtained in the individual datasets. However, the network meta-analysis resulted in a similar number of DEGs, but used a non-adjusted p -value threshold. These results reinforce the difficulty in validating and identifying functional candidate genes using the traditional gene by gene differential expression analysis when complex traits are analyzed without the comparison of extreme groups, such as HF and SF cows. For example, Moraes et al. (2018) [18] identified a significantly larger number of DEGs when comparing high-fertile versus infertile animals. Consequently, new strategies must be applied to better understand the genetic differences between HF and SF animals.

4.2. Differentially Co-Expressed Modules and the Identification of Functional Candidate Genes

The combination of co-expression gene networks, the identification of top hub-genes, the identification of fixed genetic variants in HF or SF group of cows, and the functional analysis using the p -values and $\log_2(\text{FC})$ obtained in the network meta-analysis were used to prospect for functional candidate genes regulating the differences between fertile groups. The integration of different sources of biological information is a powerful tool for the identification of functional candidate genes, which has already resulted in interesting results in livestock species [3,49–51,55,56]. Here, in order to avoid a massive discussion about all the results obtained, only the main achievements will be addressed. Thirty-two prioritized functional candidate genes (22 upstream regulators and 10 top hub-genes), related to 11 candidate DcoExp modules, were identified.

The upstream regulators genes of candidate DcoExp modules, harboring functional candidate variants, were follows: *DLG1* (Tan HF), *AGER* (Saddlebrown HF), *SORT1* (Saddlebrown HF), *HNRNPAB* (Saddlebrown HF), *TBX6* (Red HF), *HSF1* (Purple HF and Darkgreen HF), *LEPR* (Lightgreen HF),

DYSF (Lightgreen HF), *BCKDK* (Lightgreen HF), *CHFR* (Grey60 HF), *ERN2* (Green HF), and *RDH10* (Darkgreen HF). Among these genes, it was possible to identify relevant biological processes associated with fertility, such as oocyte polarization during maturation (*DLG1*), disorders of the müllerian ducts (*TBX6*), regulation of affects gonads or gonadotrophs (*LEPR*), and reproductive success (*HSF1*) [57–62].

In the HF group, the turquoise module showed the largest number of related prioritized functional candidate genes, 14 in total. This module is enriched for relevant biological processes (FDR < 0.05), such as regulation of inositol metabolism (3-phosphoinositide degradation, superpathway of inositol phosphate compounds, D-myo-inositol (1,4,5,6)-tetrakisphosphate biosynthesis, and so on) and estrogen-mediated S-phase entry. The inositol metabolism and the cell cycle mediated by estrogen activity are very important signaling pathways associated with the cellular proliferation in the uterus [63]. Genes located within these modules such as *DUSP16*, *DUSP2*, and *PIK3AP1* are directly associated with the regulation of phosphatidylinositol activity [64,65].

The top hub-genes from turquoise HF module, harboring fixed variants in HF or SF animals, and prioritized in the functional analysis, were *ARHGEF16*, *IFT80*, and *PIGR*. Rho Guanine nucleotide exchange factor 16 (*ARHGEF16*) codifies an ELMO1 interacting protein responsible to promote the clearance of apoptotic cells in a RhoG-dependent and Dock1-independent manner [66]. The relationship between *ARHGEF16* and fertility has yet to be described in the literature, such as by Elliott et al. (2010) and Gong et al. (2018). Interestingly, ELMO1 knockout mice presented multinucleated giant cells, uncleared apoptotic germ cells, and decreased sperm output in Sertoli cells owing to the phagocytic deficiency [67,68]. However, there is no link between Elmo1 and female fertility currently described. The intraflagellar transport 80 (*IFT80*) codifies a intraflagellar transport protein responsible for regulating the Jeune asphyxiating thoracic dystrophy, osteoblast, and chondrocyte differentiation [69–71]. Despite the absence of a direct link between *IFT80* and fertility, the regulation of osteoblast and chondrocyte is essential for embryo survival [72,73]. Interestingly, *VEGFA*, another gene located within the turquoise HF module, is crucial for chondrocyte survival and bone development in the embryonic stage [74]. The polymeric immunoglobulin receptor (*PIGR*) encodes a poly-g receptor in epithelial cells responsible for controlling the transcytosis process that can be regulated by steroids, such as estrogen [75,76]. In uterine epithelial cells, *PIGR* is responsible for transporting polymeric IgA. In rats, the transcriptional levels of *PIGR* are higher in the estrous when compared with proestrus or diestrus [77]. Variations in the immunoglobulin diversity and quantity in the uterus are observed during the ovulatory process, indicating a key regulatory role of ovarian hormones. Consequently, this suggests there is an impact on fertility [78,79]. Additionally, proteomic analysis of fertile and sub-fertile hens suggested that the levels of *PIGR* decreased 24 h after insemination in the uterine fluid, with the main location in the uterovaginal sperm storage tubules (SST), suggesting a response caused by the sperm arrival. These results suggest that *PIGR* responds to sperm arrival in both scenarios in sub-fertile hens, which could be a result of the higher transport activity of IgA and secretory complex to the lumen of SST [80]. It is important to highlight that *PIGR* was the second candidate gene with the highest number of fixed functional candidate variants. All the variants were identified as fixed in HF animals. Five missense variants were identified as fixed in the *PIGR* gene, where four were previously described (rs41790811, rs41790822, rs41790826, and rs41580873) and one is a new variant (c.627A > C or p.Ile162Arg). Two of these variants (rs41790822 and rs41790826) have a predicted deleterious effect based on the Sorting Intolerant from Tolerant (SIFT) score (0.04 and 0.05, respectively). The identification of missense variants with predicted deleterious effect in the HF animals might corroborate the hypothesis that higher activity levels of *PIGR* are associated with the sub-fertile phenotype in hens, as proposed by Riou et al. (2019) [80].

The significant upstream regulators genes identified by the IPA core analysis associated with the turquoise HF module and harboring fixed functional candidate variants were *PGR*, *EGFR*, *PIK3C2A*, *CUX1*, *TCHP*, *ETV5*, *KLF4*, and *ARID4A*. The progesterone receptor (*PGR*) is crucial for the initiation of pregnancy and subsequent preservation of uterus health [81]. Consequently, *PGR* is an interesting marker for uterine receptivity during implantation [82]. In this study, three fixed SNPs in HF animals

were identified in the downstream region of *PGR*. Two of those SNPs were already described (rs208289597 and rs208479533) and one is new (g.- 1458T > C). It is important to highlight that the *PGR* was identified as an upstream regulator of the turquoise module, however, it is also one of the genes that composes this module. Additionally, *PGR* was also identified as an upstream regulator of the cyan HF module. Epidermal growth factor receptor (*EGFR*) is a member of the epidermal growth factor family that plays crucial roles in the regulation of female fertility [83]. Among its functions, *EGFR* regulates puberty, oocyte maturation, uterine development, embryo implantation, and placental overgrowth [83]. Phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 α (*PIK3C2A*) is a member of the PI3-kinases acting in cell proliferation, oncogenic transformation, cell survival, cell migration, and intracellular protein trafficking [84]. The phosphatidylinositol 3 kinase (PI3K) pathway plays a crucial role in the control of mammalian oocyte growth and early follicular development [85]. Cut like homeobox 1 (*CUX1*) codifies a DNA binding protein related to the control of cell cycle progression, cell motility, and invasion [86]. Recently, SNPs on *CUX1* were associated with cow and heifer conception rate in Holstein cattle [87]. Trichoplein keratin filament binding (*TCHP*) is associated with cytoskeleton remodeling neurofilaments and vulvar sarcoma according to MalaCards database (ID: VLV038). However, there is no direct evidence linking this gene with fertility status. ETS variant 5 (*ETV5*) is a transcription factor that plays crucial roles in male fertility, acting in the spermatogonial stem cell self-renewal and maintenance of spermatogonial stem cell niche [88–90]. Female knockout mice for *ETV5* are infertile owing to a decreased ovulation and no interest in mating [91]. Kruppel like factor 4 (*KLF4*) is a transcription factor required for normal development of the barrier function of skin and with the ground state of pluripotent stem cells [92]. Regarding female fertility, *KLF4* mediates the anti-proliferative effects of progesterone during the G0/G1 arrest in endometrial epithelial cells [93]. AT-rich interaction domain 4A (*ARID4A*) is a nuclear binding protein that acts as a transcriptional coactivator of androgen receptor and retinoblastoma-binding protein during the regulation of Sertoli cell function, consequently playing a crucial role in male fertility [94]. However, disregarding the impact of target therapies for ovarian and endometrial cancer, there is no direct link between *ARID4A* and female fertility [95]. Interestingly, dual specificity phosphatase 16 (*DUSP16*) and prominin 1 (*PROM1*) are genes within the turquoise HF modules that act like top hub-genes in other candidate DcoExp modules (lightgreen HF and green HF, respectively). The function of *DUSP16* was described previously, while *PROM1* is a pentaspan transmembrane glycoprotein that was already identified as DE during the window implantation [96]. It is important to highlight that, even without a direct discussion addressed, all these genes harbour fixed variants identified in HF or SF animals.

The other top hub-genes of candidate DcoExp modules in the HF group, harboring fixed functional variants, were as follows: *IQCG* (Tan H), *MAPKAP1* (Grey60 HF), *MDM4* (Red HF), *F2RL2* (Saddlebrown HF), *MIA3* (Green HF), and *PPP1R12B* (Lightgreen HF). IQ motif containing G (*IQCG*) was the gene with the largest number of fixed variants (14 variants, all mapped in the upstream region of the gene and fixed in HF animals). IQ motif containing G is a key regulator of ciliary/flagellar motility that plays a crucial role in the formation of the sperm flagellum and spermiogenesis in mice [97,98]. However, to the best of our knowledge, there is no direct link between female fertility and *IQCG*. The MAPK associated protein 1 (*MAPKAP1*) is a member of FRAP1 complex. The disruption of the FRAP1 complex leads to post implantation lethality caused by the impaired cell proliferation and hypertrophy of embryonic disc and trophoblast [99]. The *MAPKAP1* gene is one of the top hub-genes of the Grey60 HF module, which is enriched for several fertility related biological processes, such as cell cycle regulation, PI3K signaling pathway, and triacylglycerol metabolism. *MDM4* regulator of P53 (*MDM4*) encodes a protein responsible for inhibiting p53 function. Regarding fertility, *MDM4* variants are associated with the susceptibility of ovarian and endometrial cancer [100]. Coagulation factor II thrombin receptor like 2 (*F2RL2*) is a transmembrane G protein-coupled cell surface receptor that is differentially expressed (down-regulated) during pregnancy [101]. Additionally, the Saddlebrown HF module is enriched for fertility-related processes such as androgen signaling, sperm motility, and estrogen-dependent

breast cancer signaling. To the best of our knowledge, MIA SH3 domain ER export factor 3 (*MIA3*) and protein phosphatase 1 regulatory subunit 12B (*PPP1R12B*) do not have a direct link with fertility.

Interestingly, the only candidate DcoExp module identified in the SF animals maintained after the functional analysis was the pale turquoise SF. The top-hub gene identified in this module harboring functional candidate variants was *IFT80*. This gene was also identified as a hub-gene in the turquoise HF module and its association with fertility was mentioned above. The fixed variant identified in *IFT80* is mapped in a splice donor site and it was fixed in the SF animals. The analysis of the enriched canonical pathways, diseases, and functions enriched for this module highlighted a specialization for cell cycle, embryonic development, and cell death and survival. Additionally, an interesting overlap between the canonical pathways enriched in the turquoise HF and the pale turquoise SF is observed regarding the metabolism of inositol, which, as described before, is related to fertility status. The pale turquoise SF was also enriched for ketogenesis and ketolysis (FDR < 0.05), which are relevant processes in animals subjected to a high selective pressure for production traits. Within this module, 3-hydroxybutyrate dehydrogenase 2 (*BDH2*) is the main gene associated with ketone body metabolism [102,103]. Interestingly, a polymorphism in *BDH2* was associated with days open and services per conception in Holstein cows, reinforcing a possible role of this gene with fertility traits in cattle [104].

5. Conclusions

The results obtained here reinforce the increase in detection power for functional candidate genes in the context of data integration. The approach applied in the present study, combining co-expression gene networks, identification of top hub-genes, identification of fixed genetic variants in HF or SF cows, and functional analyses, resulted in the identification of highly relevant functional gene networks and candidate genes associated with fertility status in beef cattle. These results contribute to the better understanding of the genetic architecture of high- and sub-fertility cows. Additionally, candidate functional variants were identified uniquely fixed in one of the fertility groups (HF or SF animals) in the functional prioritized genes. Consequently, these variants could be used in genetic selection programs to validate the contribution of these variants in the fertility status. In the long-term, the results obtained here may help to reduce the frequency of subfertility in beef herds, helping to reduce the economic losses caused by this condition.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/11/5/543/s1>, Figure S1: Percentage of each functional class identified among the fixed variants (SNPs and INDELS) identified in HF (A) or SF (B) animals; Figure S2: Outlier detection before (A) and after (B) log-likelihood ratio test (LRT). The black dots (GSE81449) and squares (GSE107891) represent the high-fertile animals, while the grey dots and squares represent sub-fertile animals; Table S1: Alignment summary statistics; Table S2: Functional impact for the variants fixed in one of the conditions, high- or sub-fertile; Table S3: Differential expression analysis results for all of the samples; Table S4: Differentially co-expressed modules and associated genes; Table S5: Enriched biological processes, diseases, and functions, as well as upstream regulators obtained for each candidate differentially co-expressed modules using IPA core module; Table S6: Functional prioritization analysis results; Table S7: Fixed variants in functional candidate genes.

Author Contributions: P.A.S.F. and A.C. were responsible for the conceptualization, experimental design, RNA-Sequencing analysis, and theoretical discussions. P.A.S.F., A.S.-V., and A.C. participated in the manuscript writing. P.A.S.F. and A.S.-V. were responsible for data curation. P.A.S.F. was responsible for implementing the bioinformatic pipeline, integrating datasets, and performing the functional analysis. A.C. was responsible for funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research (FDE.13.17) was supported by the Sustainable Beef and Forage Science Cluster funded by the Canadian Beef Cattle Check-Off, Beef Cattle Research Council (BCRC), Alberta Beef Producers, Alberta Cattle Feeders' Association, Beef Farmers of Ontario, La Fédération des Producteurs de bovins du Québec, and Agriculture and Agri-Food Canada's Canadian Agricultural Partnership. This study was also supported by NSERC (Natural Sciences and Engineering Research Council); Beef Farmers of Ontario; Ontario Ministry of Agriculture, Food, and Rural Affairs (OMAFRA); Ontario Ministry of Research and Innovation; and Agriculture and Agri-Food Canada (AAFC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Thundathil, J.C.; Dance, A.L.; Kastelic, J.P. Fertility management of bulls to improve beef cattle productivity. *Theriogenology* **2016**, *86*, 397–405. [[CrossRef](#)] [[PubMed](#)]
2. Berry, D.P.; Friggens, N.C.; Lucy, M.; Roche, J.R. Milk Production and Fertility in Cattle. *Annu. Rev. Anim. Biosci.* **2016**, *4*, 269–290. [[CrossRef](#)] [[PubMed](#)]
3. De Souza Fonseca, P.A.; Id-Lahoucine, S.; Reverter, A.; Medrano, J.F.; Fortes, M.S.; Casellas, J.; Miglior, F.; Brito, L.; Carvalho, M.R.S.; Schenkel, F.S.; et al. Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS ONE* **2018**, *13*, e0205295. [[CrossRef](#)]
4. Kadri, N.K.; Sahana, G.; Charlier, C.; Iso-Touru, T.; Guldbrandsen, B.; Karim, L.; Nielsen, U.S.; Panitz, F.; Aamand, G.P.; Schulman, N.; et al. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock. *PLoS Genet.* **2014**, *10*, e1004049. [[CrossRef](#)] [[PubMed](#)]
5. Saatchi, M.; Schnabel, R.D.; Taylor, J.F.; Garrick, D.J. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genom.* **2014**, *15*, 442. [[CrossRef](#)] [[PubMed](#)]
6. Tsuruta, S.; Lourenco, D.A.L.; Misztal, I.; Lawlor, T.J. Genomic analysis of cow mortality and milk production using a threshold-linear model. *J. Dairy Sci.* **2017**, *100*, 7295–7305. [[CrossRef](#)]
7. De Camargo, G.M.F.; Porto-Neto, L.R.; Kelly, M.J.; Bunch, R.J.; McWilliam, S.M.; Tonhati, H.; Lehnert, S.A.; Fortes, M.R.S.; Moore, S.S. Non-synonymous mutations mapped to chromosome X associated with andrological and growth traits in beef cattle. *BMC Genom.* **2015**, *16*, 384. [[CrossRef](#)]
8. Bellows, D.S.; Ott, S.L.; Bellows, R. a Review: Cost of reproductive diseases and conditions in cattle. *Prof. Anim. Sci.* **2002**, *18*, 26–32. [[CrossRef](#)]
9. Shaloo, L.; Cromie, A.; McHugh, N. Effect of fertility on the economics of pasture-based dairy systems. *Animal* **2014**, *8*, 222–231. [[CrossRef](#)]
10. Wiltbank, M.C.; Baez, G.M.; Garcia-Guerra, A.; Toledo, M.Z.; Monteiro, P.L.J.; Melo, L.F.; Ochoa, J.C.; Santos, J.E.P.; Sartori, R. Pivotal periods for pregnancy loss during the first trimester of gestation in lactating dairy cows. *Theriogenology* **2016**, *86*, 239–253. [[CrossRef](#)]
11. Diskin, M.G.; Parr, M.H.; Morris, D.G. Embryo death in cattle: An update. *Reprod. Fertil. Dev.* **2012**, *24*, 244–251. [[CrossRef](#)] [[PubMed](#)]
12. Chebel, R.C.; Santos, J.E.P.; Reynolds, J.P.; Cerri, R.L.A.; Juchem, S.O.; Overton, M. Factors affecting conception rate after artificial insemination and pregnancy loss in lactating dairy cows. *Anim. Reprod. Sci.* **2004**, *84*, 239–255. [[CrossRef](#)] [[PubMed](#)]
13. Bormann, J.M.; Totir, L.R.; Kachman, S.D.; Fernando, R.L.; Wilson, D.E. Pregnancy rate and first-service conception rate in Angus heifers. *J. Anim. Sci.* **2006**, *84*, 2022–2025. [[CrossRef](#)] [[PubMed](#)]
14. Azzam, S.M.; Kinder, J.E.; Nielsen, M.K. Conception rate at first insemination in beef cattle: Effects of season, age and previous reproductive performance. *J. Anim. Sci.* **1989**, *67*, 1405–1410. [[CrossRef](#)] [[PubMed](#)]
15. Id-Lahoucine, S.; Cánovas, A.; Jatón, C.; Miglior, F.; Fonseca, P.A.S.; Sargolzaei, M.; Miller, S.; Schenkel, F.S.; Medrano, J.F.; Casellas, J. Implementation of Bayesian methods to identify SNP and haplotype regions with transmission ratio distortion across the whole genome: TRDscan v.1.0. *J. Dairy Sci.* **2019**, *102*, 3175–3188. [[CrossRef](#)]
16. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)]
17. Butler, S.T.; Pryce, J.E.; Kemper, K.E.; Berry, D.P.; McCabe, M.; Lonergan, P.; Hayes, B.J.; Chamberlain, A.J.; Cormican, P.; 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 Fertility1. *Biol. Reprod.* **2015**, *94*, 1–11.

18. Moraes, J.G.N.; Behura, S.K.; Geary, T.W.; Hansen, P.J.; Neiberghs, H.L.; Spencer, T.E. Uterine influences on conceptus development in fertility-classified animals. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1749–E1758. [[CrossRef](#)]
19. Moran, B.; Butler, S.T.; Moore, S.G.; Machugh, D.E.; Creevey, C.J. Differential gene expression in the endometrium reveals cytoskeletal and immunological genes in lactating dairy cows genetically divergent for fertility traits. *Reprod. Fertil. Dev.* **2017**, *29*, 274–282. [[CrossRef](#)]
20. Geary, T.W.; Burns, G.W.; Moraes, J.G.N.; Moss, J.I.; Denicol, A.C.; Dobbs, K.B.; Ortega, M.S.; Hansen, P.J.; Wehrman, M.E.; Neiberghs, H.; et al. Identification of Beef Heifers with Superior Uterine Capacity for Pregnancy. *Biol. Reprod.* **2016**, *95*, 47. [[CrossRef](#)]
21. Nguyen, L.T.; Reverter, A.; Cánovas, A.; Venus, B.; Islas-Trejo, A.; Porto-Neto, L.R.; Lehnert, S.A.; Medrano, J.F.; Moore, S.S.; Fortes, M.R.S. Global differential gene expression in the pituitary gland and the ovaries of Pre-And postpubertal brahman heifers. *J. Anim. Sci.* **2017**, *95*, 599–615. [[CrossRef](#)] [[PubMed](#)]
22. Fortes, M.R.S.; Nguyen, L.T.; Weller, M.M.D.C.A.; Cánovas, A.; Islas-Trejo, A.; Porto-Neto, L.R.; Reverter, A.; Lehnert, S.A.; Boe-Hansen, G.B.; Thomas, M.G.; et al. Transcriptome analyses identify five transcription factors differentially expressed in the hypothalamus of post- versus prepubertal Brahman heifers. *J. Anim. Sci.* **2016**, *94*, 3693–3702. [[CrossRef](#)] [[PubMed](#)]
23. Nguyen, L.T.; Reverter, A.; Cánovas, A.; Venus, B.; Anderson, S.T.; Islas-Trejo, A.; Dias, M.M.; Crawford, N.F.; Lehnert, S.A.; Medrano, J.F.; et al. STAT6, PBX2, and PBRM1 emerge as predicted regulators of 452 differentially expressed genes associated with puberty in Brahman heifers. *Front. Genet.* **2018**, *9*, 87. [[CrossRef](#)] [[PubMed](#)]
24. Dias, M.M.; Cánovas, A.; Mantilla-Rojas, C.; Riley, D.G.; Luna-Nevarez, P.; Coleman, S.J.; Speidel, S.E.; Enns, R.M.; Islas-Trejo, A.; Medrano, J.F.; et al. SNP detection using RNA-sequences of candidate genes associated with puberty in cattle. *Genet. Mol. Res.* **2017**, *16*, 1–17. [[CrossRef](#)]
25. Eisen, M.B.; Spellman, P.T.; Brown, P.O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14863–14868. [[CrossRef](#)]
26. Oliver, S. Guilt-by-association goes global. *Nature* **2000**, *403*, 601–603. [[CrossRef](#)]
27. Winter, C.; Kosch, R.; Ludlow, M.; Osterhaus, A.D.M.E.; Jung, K. Network meta-analysis correlates with analysis of merged independent transcriptome expression data. *BMC Bioinform.* **2019**, *20*, 144. [[CrossRef](#)]
28. Wren, J.D. A global meta-analysis of microarray expression data to predict unknown gene functions and estimate the literature-data divide. *Bioinformatics* **2009**, *25*, 1694–1701. [[CrossRef](#)]
29. Gillis, J.; Pavlidis, P. The impact of multifunctional genes on guilt “by association” analysis. *PLoS ONE* **2011**, *6*, e17258. [[CrossRef](#)]
30. Blankenburg, H.; Pramstaller, P.P.; Domingues, F.S. A network-based meta-analysis for characterizing the genetic landscape of human aging. *Biogerontology* **2018**, *19*, 81–94. [[CrossRef](#)]
31. Lee, I.; Blom, U.M.; Wang, P.I.; Shim, J.E.; Marcotte, E.M. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res.* **2011**, *21*, 1109–1121. [[CrossRef](#)] [[PubMed](#)]
32. Nygaard, V.; Rødland, E.A.; Hovig, E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics* **2016**, *17*, 29–39. [[CrossRef](#)] [[PubMed](#)]
33. Ayuso, M.; Fernández, A.; Núñez, Y.; Benítez, R.; Isabel, B.; Fernández, A.I.; Rey, A.I.; González-Bulnes, A.; Medrano, J.F.; Cánovas, Á.; et al. Developmental stage, muscle and genetic type modify muscle transcriptome in pigs: Effects on gene expression and regulatory factors involved in growth and metabolism. *PLoS ONE* **2016**, *11*, e0167858. [[CrossRef](#)]
34. Cánovas, A.; Rincon, G.; Islas-Trejo, A.; Wickramasinghe, S.; Medrano, J.F. SNP discovery in the bovine milk transcriptome using RNA-Seq technology. *Mamm. Genome* **2010**, *21*, 592–598. [[CrossRef](#)]
35. Cánovas, A.; Rincón, G.; Islas-Trejo, A.; Jimenez-Flores, R.; Laubscher, A.; Medrano, J.F. RNA sequencing to study gene expression and single nucleotide polymorphism variation associated with citrate content in cow milk. *J. Dairy Sci.* **2013**, *96*, 2637–2648. [[CrossRef](#)] [[PubMed](#)]
36. Cánovas, A.; Rincón, G.; Bevilacqua, C.; Islas-Trejo, A.; Brenaut, P.; Hovey, R.C.; Boutinaud, M.; Morgenthaler, C.; Vanklompberg, M.K.; Martin, P.; et al. Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. *Sci. Rep.* **2014**, *4*, 1–7. [[CrossRef](#)]

37. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, gaaa021. [[CrossRef](#)]
38. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621. [[CrossRef](#)]
39. Marioni, J.C.; Mason, C.E.; Mane, S.M.; Stephens, M.; Gilad, Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **2008**, *18*, 1509–1517. [[CrossRef](#)]
40. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
41. Rucker, G.; Krahn, U.; Jochem, K.; Orestis, E.; Guido, S. Package ‘netmeta’. *Network Meta-Analysis using Frequentist Methods (Version 0.7-0)*. 2015. Available online: <http://kambing.ui.ac.id/cran/web/packages/netmeta/netmeta.pdf> (accessed on 10 May 2020).
42. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
43. Langfelder, P.; Zhang, B.; Horvath, S. Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinformatics* **2008**, *24*, 719–720. [[CrossRef](#)] [[PubMed](#)]
44. Botía, J.A.; Vandrovцова, J.; Forabosco, P.; Guelfi, S.; D’Sa, K.; Hardy, J.; Lewis, C.M.; Ryten, M.; Weale, M.E.; Ramasamy, A.; et al. An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst. Biol.* **2017**, *11*, 47. [[CrossRef](#)] [[PubMed](#)]
45. Csardi, G.; Nepusz, T. The igraph software package for complex network research. *Interf. Complex Syst.* **2006**, *1695*, 1–9.
46. Durinck, S.; Moreau, Y.; Kasprzyk, A.; Davis, S.; De Moor, B.; Brazma, A.; Huber, W. BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **2005**, *21*, 3439–3440. [[CrossRef](#)]
47. Guney, E.; Garcia-Garcia, J.; Oliva, B. GUILDify: A web server for phenotypic characterization of genes through biological data integration and network-based prioritization algorithms. *Bioinformatics* **2014**, *30*, 1789–1790. [[CrossRef](#)]
48. Chen, J.; Bardes, E.E.; Aronow, B.J.; Jegga, A.G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **2009**, *37*, 305–311. [[CrossRef](#)]
49. Fonseca, P.A.d.S.; dos Santos, F.C.; Lam, S.; Suárez-Vega, A.; Miglior, F.; Schenkel, F.S.; Diniz, L.d.A.F.; Id-Lahoucine, S.; Carvalho, M.R.S.; Cánovas, A. Genetic mechanisms underlying spermatic and testicular traits within and among cattle breeds: Systematic review and prioritization of GWAS results. *J. Anim. Sci.* **2018**, *96*, 4978–4999.
50. Cánovas, A.; Reverter, A.; DeAtley, K.L.; Ashley, R.L.; Colgrave, M.L.; Fortes, M.R.S.; Islas-Trejo, A.; Lehnert, S.; Porto-Neto, L.; Rincón, G.; et al. Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. *PLoS ONE* **2014**, *9*, e102551. [[CrossRef](#)]
51. Suárez-Vega, A.; Gutiérrez-Gil, B.; Benavides, J.; Perez, V.; Tosser-Klopp, G.; Klopp, C.; Keennel, S.J.; Arranz, J.J. Combining GWAS and RNA-Seq approaches for detection of the causal mutation for hereditary junctional epidermolysis bullosa in sheep. *PLoS ONE* **2015**, *10*, e0126416. [[CrossRef](#)]
52. Trapnell, C. Defining cell types and states with single-cell genomics. *Genome Res.* **2015**, *25*, 1491–1498. [[CrossRef](#)]
53. Petri, T.; Altmann, S.; Geistlinger, L.; Zimmer, R.; Küffner, R. Addressing false discoveries in network inference. *Bioinformatics* **2015**, *31*, 2836–2843. [[CrossRef](#)]
54. Wu, T.H.; Chu, L.J.; Wang, J.C.; Chen, T.W.; Tien, Y.J.; Lin, W.C.; Ng, W.V. Meta-analytical biomarker search of EST expression data reveals three differentially expressed candidates. *BMC Genom.* **2012**, *13*, 12. [[CrossRef](#)] [[PubMed](#)]
55. Suravajhala, P.; Kogelman, L.J.A.; Kadarmideen, H.N. Multi-omic data integration and analysis using systems genomics approaches: Methods and applications In animal production, health and welfare. *Genet. Sel. Evol.* **2016**, *48*, 1–14. [[CrossRef](#)] [[PubMed](#)]
56. Kadarmideen, H.N. Genomics to systems biology in animal and veterinary sciences: Progress, lessons and opportunities. *Livest. Sci.* **2014**, *166*, 232–248. [[CrossRef](#)]
57. Christians, E.; Davis, A.A.; Thomas, S.D.; Benjamin, I.J. Embryonic development: Maternal effect of hsf1 on reproductive success. *Nature* **2000**, *407*, 693–694. [[CrossRef](#)] [[PubMed](#)]

58. Wang, J.C.; Lv, H.; Wu, K.L.; Zhang, Y.S.; Luo, H.N.; Chen, Z.J. Discs large homologue 1 (Dlg1) coordinates mouse oocyte polarisation during maturation. *Reprod. Fertil. Dev.* **2017**, *29*, 1699–1707. [[CrossRef](#)]
59. Tewes, A.C.; Rall, K.K.; Römer, T.; Hucke, J.; Kapczuk, K.; Brucker, S.; Wieacker, P.; Ledig, S. Variations in RBM8A and TBX6 are associated with disorders of the müllerian ducts. *Fertil. Steril.* **2015**, *103*, 1313–1318. [[CrossRef](#)]
60. White, P.H.; Farkas, D.R.; McFadden, E.E.; Chapman, D.L. Defective somite patterning in mouse embryos with reduced levels of Tbx6. *Development* **2003**, *130*, 1681–1690. [[CrossRef](#)]
61. Rosenbaum, M.; Leibel, R.L. Leptin: A molecule integrating somatic energy stores, energy expenditure and fertility. *Trends Endocrinol. Metab.* **1998**, *9*, 117–124. [[CrossRef](#)]
62. Lu, L.Y.; Yu, X. CHFR is important for the survival of male premeiotic germ cells. *Cell Cycle* **2015**, *14*, 3454–3460. [[CrossRef](#)] [[PubMed](#)]
63. Chen, B.; Pan, H.; Zhu, L.; Deng, Y.; Pollard, J.W. Progesterone inhibits the estrogen-induced phosphoinositide 3-kinase→AKT→GSK-3β→cyclin D1→pRB pathway to block uterine epithelial cell proliferation. *Mol. Endocrinol.* **2005**, *19*, 1978–1990. [[CrossRef](#)] [[PubMed](#)]
64. Patterson, K.I.; Brummer, T.; O'Brien, P.M.; Daly, R.J. Dual-specificity phosphatases: Critical regulators with diverse cellular targets. *Biochem. J.* **2009**, *418*, 475–489. [[CrossRef](#)] [[PubMed](#)]
65. Fruman, D.A.; Chiu, H.; Hopkins, B.D.; Bagrodia, S.; Cantley, L.C.; Abraham, R.T. The PI3K Pathway in Human Disease. *Cell* **2017**, *170*, 605–635. [[CrossRef](#)] [[PubMed](#)]
66. Lee, J.; Park, B.; Kim, G.; Kim, K.; Pak, J.; Kim, K.; Ye, M.B.; Park, S.G.; Park, D. Arhgef16, a novel Elmo1 binding partner, promotes clearance of apoptotic cells via RhoG-dependent Rac1 activation. *Biochim. Biophys. Acta Mol. Cell Res.* **2014**, *1843*, 2438–2447. [[CrossRef](#)]
67. Elliott, M.R.; Zheng, S.; Park, D.; Woodson, R.I.; Reardon, M.A.; Juncadella, I.J.; Kinchen, J.M.; Zhang, J.; Lysiak, J.J.; Ravichandran, K.S. Unexpected requirement for ELMO1 in clearance of apoptotic germ cells in vivo. *Nature* **2010**, *467*, 333–337. [[CrossRef](#)]
68. Gong, P.; Chen, S.; Zhang, L.; Hu, Y.; Gu, A.; Zhang, J.; Wang, Y. RhoG-ELMO1-RAC1 is involved in phagocytosis suppressed by mono-butyl phthalate in TM4 cells. *Environ. Sci. Pollut. Res.* **2018**, *25*, 35440–35450. [[CrossRef](#)]
69. Yang, S.; Wang, C. The intraflagellar transport protein IFT80 is required for cilia formation and osteogenesis. *Bone* **2012**, *51*, 407–417. [[CrossRef](#)]
70. Wang, C.; Yuan, X.; Yang, S. IFT80 is essential for chondrocyte differentiation by regulating Hedgehog and Wnt signaling pathways. *Exp. Cell Res.* **2013**, *319*, 623–632. [[CrossRef](#)]
71. Beales, P.L.; Bland, E.; Tobin, J.L.; Bacchelli, C.; Tuysuz, B.; Hill, J.; Rix, S.; Pearson, C.G.; Kai, M.; Hartley, J.; et al. IFT80, which encodes a conserved intraflagellar transport protein, is mutated in Jeune asphyxiating thoracic dystrophy. *Nat. Genet.* **2007**, *39*, 727–729. [[CrossRef](#)]
72. Gaur, T.; Hussain, S.; Mudhasani, R.; Parulkar, I.; Colby, J.L.; Frederick, D.; Kream, B.E.; van Wijnen, A.J.; Stein, J.L.; Stein, G.S.; et al. Dicer inactivation in osteoprogenitor cells compromises fetal survival and bone formation, while excision in differentiated osteoblasts increases bone mass in the adult mouse. *Dev. Biol.* **2010**, *340*, 10–21. [[CrossRef](#)] [[PubMed](#)]
73. Shapiro, I.M.; Adams, C.S.; Freeman, T.; Srinivas, V. Fate of the hypertrophic chondrocyte: Microenvironmental perspectives on apoptosis and survival in the epiphyseal growth plate. *Birth Defects Res. Part C Embryo Today Rev.* **2005**, *75*, 330–339. [[CrossRef](#)] [[PubMed](#)]
74. Zelzer, E.; Mamluk, R.; Ferrara, N.; Johnson, R.S.; Schipani, E.; Olsen, B.R. VEGFA is necessary for chondrocyte survival during bone development. *Development* **2004**, *131*, 2161–2171. [[CrossRef](#)] [[PubMed](#)]
75. Beagley, K.W.; Gockel, C.M. Regulation of innate and adaptive immunity by the female sex hormones oestradiol and progesterone. *FEMS Immunol. Med. Microbiol.* **2003**, *38*, 13–22. [[CrossRef](#)]
76. Kaetzel, C.S. The polymeric immunoglobulin receptor: Bridging innate and adaptive immune responses at mucosal surfaces. *Immunol. Rev.* **2005**, *206*, 83–99. [[CrossRef](#)]
77. Richardson, J.M.; Kaushic, C.; Wira, C.R. Polymeric Immunoglobulin (Ig) Receptor Production and IgA Transcytosis in Polarized Primary Cultures of Mature Rat Uterine Epithelial Cells. *Biol. Reprod.* **1995**, *53*, 488–498. [[CrossRef](#)]
78. Rachman, F.; Casimiri, V.; Psychoyos, A.; Bernard, O. Immunoglobulins in the mouse uterus during the oestrous cycle. *J. Reprod. Fertil.* **1983**, *69*, 17–21. [[CrossRef](#)]

79. Kutteh, W.H.; Prince, S.J.; Hammond, K.R.; Kutteh, C.C.; Mestecky, J. Variations in immunoglobulins and IgA subclasses of human uterine cervical secretions around the time of ovulation. *Clin. Exp. Immunol.* **1996**, *104*, 538–542. [[CrossRef](#)]
80. Riou, C.; Brionne, A.; Cordeiro, L.; Harichaux, G.; Gargaros, A.; Labas, V.; Gautron, J.; Gérard, N. Proteomic analysis of uterine fluid of fertile and subfertile hens before and after insemination. *Reproduction* **2019**, *158*, 335–356. [[CrossRef](#)]
81. Wetendorf, M.; DeMayo, F.J. The progesterone receptor regulates implantation, decidualization, and glandular development via a complex paracrine signaling network. *Mol. Cell. Endocrinol.* **2012**, *357*, 108–118. [[CrossRef](#)]
82. Lessey, B.A.; Yeh, I.; Castelbaum, A.J.; Fritz, M.A.; Ilesanmi, A.O.; Korzeniowski, P.; Sun, J.; Chwalisz, K. Endometrial progesterone receptors and markers of uterine receptivity in the window of implantation. *Fertil. Steril.* **1996**, *65*, 477–483. [[CrossRef](#)]
83. Hsieh, M.; Lee, D.; Panigone, S.; Horner, K.; Chen, R.; Theologis, A.; Lee, D.C.; Threadgill, D.W.; Conti, M. Luteinizing Hormone-Dependent Activation of the Epidermal Growth Factor Network Is Essential for Ovulation. *Mol. Cell. Biol.* **2007**, *27*, 1914–1924. [[CrossRef](#)] [[PubMed](#)]
84. Martini, M.; De Santis, M.C.; Braccini, L.; Gulluni, F.; Hirsch, E. PI3K/AKT signaling pathway and cancer: An updated review. *Ann. Med.* **2014**, *46*, 372–383. [[CrossRef](#)] [[PubMed](#)]
85. Liu, K.; Rajareddy, S.; Liu, L.; Jagarlamudi, K.; Boman, K.; Selstam, G.; Reddy, P. Control of mammalian oocyte growth and early follicular development by the oocyte PI3 kinase pathway: New roles for an old timer. *Dev. Biol.* **2006**, *299*, 1–11. [[CrossRef](#)]
86. Sansregret, L.; Nepveu, A. The multiple roles of CUX1: Insights from mouse models and cell-based assays. *Gene* **2008**, *412*, 84–94. [[CrossRef](#)]
87. Abdollahi-arpanahi, R.; Carvalho, M.R.; Ribeiro, E.S.; Peñagaricano, F. Association of lipid-related genes implicated in conceptus elongation with female fertility traits in dairy cattle. *J. Dairy Sci.* **2019**, *102*, 10020–10029. [[CrossRef](#)]
88. Chen, C.; Ouyang, W.; Grigura, V.; Zhou, Q.; Carnes, K.; Lim, H.; Zhao, G.Q.; Arber, S.; Kurpios, N.; Murphy, T.L.; et al. ERM is required for transcriptional control of the spermatogonial stem cell niche. *Nature* **2005**, *436*, 1030–1034. [[CrossRef](#)]
89. Oatley, J.M.; Avarbock, M.R.; Brinster, R.L. Glial cell line-derived neurotrophic factor regulation of genes essential for self-renewal of mouse spermatogonial stem cells is dependent on Src family kinase signaling. *J. Biol. Chem.* **2007**, *282*, 25842–25851. [[CrossRef](#)]
90. Alankarage, D.; Lavery, R.; Svingen, T.; Kelly, S.; Ludbrook, L.; Bagheri-Fam, S.; Koopman, P.; Harley, V. SOX9 regulates expression of the male fertility gene Ets variant factor 5 (ETV5) during mammalian sex development. *Int. J. Biochem. Cell Biol.* **2016**, *79*, 41–51. [[CrossRef](#)]
91. Eo, J.; Shin, H.; Kwon, S.; Song, H.; Murphy, K.M.; Lim, J.H. Complex ovarian defects lead to infertility in Etv5^{-/-} female mice. *Mol. Hum. Reprod.* **2011**, *17*, 568–576. [[CrossRef](#)]
92. Guo, G.; Yang, J.; Nichols, J.; Hall, J.S.; Eyres, I.; Mansfield, W.; Smith, A. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **2009**, *136*, 1063–1069. [[CrossRef](#)] [[PubMed](#)]
93. Shimizu, Y.; Takeuchi, T.; Mita, S.; Notsu, T.; Mizuguchi, K.; Kyo, S. Krüppel-like factor 4 mediates anti-proliferative effects of progesterone with G0/G1arrest in human endometrial epithelial cells. *J. Endocrinol. Investig.* **2010**, *33*, 745–750. [[CrossRef](#)] [[PubMed](#)]
94. Wu, R.C.; Jiang, M.; Beaudet, A.L.; Wu, M.Y. ARID4A and ARID4B regulate male fertility, a functional link to the AR and RB pathways. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 4616–4621. [[CrossRef](#)] [[PubMed](#)]
95. Takeda, T.; Banno, K.; Okawa, R.; Yanokura, M.; Iijima, M.; Iriekunitomi, H.; Nakamura, K.; Iida, M.; Adachi, M.; Umene, K.; et al. ARID1A gene mutation in ovarian and endometrial cancers (Review). *Oncol. Rep.* **2016**, *35*, 607–613. [[CrossRef](#)] [[PubMed](#)]
96. Shi, C.; Han, H.J.; Fan, L.J.; Guan, J.; Zheng, X.B.; Chen, X.; Liang, R.; Zhang, X.W.; Sun, K.K.; Cui, Q.H.; et al. Diverse endometrial mRNA signatures during the window of implantation in patients with repeated implantation failure. *Hum. Fertil.* **2018**, *21*, 183–194. [[CrossRef](#)]
97. Harris, T.P.; Schimenti, K.J.; Munroe, R.J.; Schimenti, J.C. IQ motif-containing G (Iqcg) is required for mouse spermiogenesis. *G3 Genes Genomes Genet.* **2014**, *4*, 367–372. [[CrossRef](#)]
98. Li, R.K.; Tan, J.L.; Chen, L.T.; Feng, J.S.; Liang, W.X.; Guo, X.J.; Liu, P.; Chen, Z.; Sha, J.H.; Wang, Y.F.; et al. Iqcg is essential for sperm flagellum formation in mice. *PLoS ONE* **2014**, *9*, e98053. [[CrossRef](#)]

99. Murakami, M.; Ichisaka, T.; Maeda, M.; Oshiro, N.; Hara, K.; Edenhofer, F.; Kiyama, H.; Yonezawa, K.; Yamanaka, S. mTOR Is Essential for Growth and Proliferation in Early Mouse Embryos and Embryonic Stem Cells. *Mol. Cell. Biol.* **2004**, *24*, 6710–6718. [[CrossRef](#)]
100. Gansmo, L.B.; Bjørnslett, M.; Halle, M.K.; Salvesen, H.B.; Dørum, A.; Birkeland, E.; Hveem, K.; Romundstad, P.; Vatten, L.; Lønning, P.E.; et al. The MDM4 SNP34091 (rs4245739) C-allele is associated with increased risk of ovarian—but not endometrial cancer. *Tumor Biol.* **2016**, *24*, 6710–6718. [[CrossRef](#)]
101. Forde, N.; Duffy, G.B.; McGettigan, P.A.; Browne, J.A.; Mehta, J.P.; Kelly, A.K.; Mansouri-Attia, N.; Sandra, O.; Loftus, B.J.; Crowe, M.A.; et al. Evidence for an early endometrial response to pregnancy in cattle: Both dependent upon and independent of interferon tau. *Physiol. Genom.* **2012**, *44*, 799–810. [[CrossRef](#)]
102. Beis, A.; Zammit, V.A.; Newsholme, E.A. Activities of 3-Hydroxybutyrate Dehydrogenase, 3-Oxoacid CoA-Transferase and Acetoacetyl-CoA Thiolase in Relation to Ketone-Body Utilisation in Muscles from Vertebrates and Invertebrates. *Eur. J. Biochem.* **1980**, *104*, 209–215. [[CrossRef](#)] [[PubMed](#)]
103. Koundakjian, P.P.; Snoswell, A.M. Ketone body and fatty acid metabolism in sheep tissues. 3-Hydroxybutyrate dehydrogenase, a cytoplasmic enzyme in sheep liver and kidney. *Biochem. J.* **1970**, *119*, 49–57. [[CrossRef](#)] [[PubMed](#)]
104. Cochran, S.D.; Cole, J.B.; Null, D.J.; Hansen, P.J. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genet.* **2013**, *14*, 49. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).