# Genome-Wide Association Analyses of Fertility Traits in Beef Heifers

Morgan R. Stegemiller [1], Gordon K. Murdoch [1,2], Troy N. Rowan [3], Kimberly M. Davenport [1], Gabrielle M. Becker [1], John B. Hall [1,4,*] and Brenda M. Murdoch [1,5,*]

1 Department of Animal, Veterinary & Food Sciences, University of Idaho, Moscow, ID 83843, USA; mstegemiller@uidaho.edu (M.R.S.); gordon.murdoch@wsu.edu (G.K.M.); kmdavenport@uidaho.edu (K.M.D.); gbecker@uidaho.edu (G.M.B.)
2 Department of Animal Sciences, Washington State University, Pullman, WA 99164, USA
3 Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA; tnr343@mail.missouri.edu
4 Nancy M. Cummings Research, Education, and Extension Center, University of Idaho, Carmen, ID 83462, USA
5 Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA
* Correspondence: jbhall@uidaho.edu (J.B.H.); bmurdoch@uidaho.edu (B.M.M.); Tel.: +1-208-756-2749 (J.B.H.); +1-208-885-2088 (B.M.M.)

**Abstract:** The ability of livestock to reproduce efficiently is critical to the sustainability of animal agriculture. Antral follicle count (AFC) and reproductive tract scores (RTS) can be used to estimate fertility in beef heifers, but the genetic mechanisms influencing variation in these measures are not well understood. Two genome-wide association studies (GWAS) were conducted to identify the significant loci associated with these traits. In total, 293 crossbred beef heifers were genotyped on the Bovine GGP 50K chip and genotypes were imputed to 836,121 markers. A GWAS was performed with the AFC phenotype for 217 heifers with a multi-locus mixed model, conducted using the year, age at time of sampling and principal component analysis groupings as the covariates. The RTS GWAS was performed with 289 heifers using an additive correlation/trend test comparing prepubertal to pubertal heifers. The loci on chromosomes 2, 3 and 23 were significant in the AFC GWAS and the loci on chromosomes 2, 8, 10 and 11 were significant in the RTS GWAS. The significant region on chromosome 2 was similar between both analyses. These regions contained genes associated with cell proliferation, transcription, apoptosis and development. This study proposes candidate genes for beef cattle fertility, although future research is needed to elucidate the precise mechanisms.

**Keywords:** GWAS; crossbred beef heifers; imputed; antral follicle count; reproductive tract score

## 1. Introduction

Cattle producers benefit from animals that reproduce reliably and efficiently, making fertility a critical trait in the cattle industry. Improving reproductive efficiency can be accomplished by selecting replacement heifers with higher fertility and a longer reproductive life span. There are several measures used to estimate fertility, including Anti-Müllerian hormone concentration, days open and calving performance: this study focuses on the antral follicle count (AFC) and reproductive tract scores (RTS).

Antral follicle count is highly repeatable within individuals although it can vary within a population [1]. The fertility measurement of AFC has been utilized in cattle and requires only a single, non-invasive ultrasound examination [2]. The heritability of AFC in cattle is 0.25 and is positively correlated with other indirect measures of fertility, such as endometrial thickness, super ovulatory response and herd longevity [1–6]. In addition, cattle with a higher AFC tend to have ovaries of significantly greater length, size and weight [7]. Ovaries with a greater number of antral follicles indicate a larger ovarian reserve (OR) [8]. A smaller OR has been associated with reduced ovarian function [9]. Although the proportion of healthy to total number of follicles is the same between animals

with either high or low AFC, the total number of healthy follicles is greater in those with high AFC (>25) [7]. Furthermore, *Bos taurus*, *Bos indicus* and *Bos indicus-taurus* animals that have higher AFC also produce more oocytes during superovulation [2,4,5]. Previous studies in *Bos taurus* crossbred beef heifers and Holstein-Friesian dairy cows have shown that animals with a high AFC are more likely to be pregnant at the end of breeding season [8,10].

In addition to AFC, producers use RTS as a semi-objective measurement of pubertal status and to determine the age of puberty that has a heritability of 0.43 [11]. Briefly, scoring is based on palpation of follicular development, corpus luteum presence and reproductive tract tone. A scale from one to five is used, in which a score of one indicates the animal is immature or in anestrous while five shows the animal is mature and cycling, as described previously [12]. Unsurprisingly, heifers with a higher RTS are more likely to conceive and to conceive earlier in the breeding season than heifers with a lower RTS [13].

Understanding the biological mechanisms contributing to increased AFC and earlier reproductive tract development will allow producers to select for more reproductively efficient animals as replacements in their herd. This study used genome-wide association studies (GWAS) to investigate if AFC and RTS exhibit significant genetic associations with genetic variation and has identified the potential biological pathways involved. Previous GWAS in cattle, such as Neupane et al. and Cole et al., have identified the genes associated with fertility and reproductive traits in cattle; however, neither examined AFC or RTS [14,15]. Determining the mechanisms contributing to follicular and reproductive tract development will provide a basis for future fertility studies.

## 2. Materials and Methods

### 2.1. Animals and Phenotype Collection of Antral Follicle Count and Reproductive Tract Scores

This study examined a total of 293 crossbred heifers over a two-year period, 139 from year one and 154 from year two. The heifers were sired by Angus, Hereford, Simmental, Simmental-Angus (SimAngus) or Shorthorn bulls. These animals were raised at the University of Idaho Nancy M. Cummings Research, Education and Extension Center in Carmen, Idaho. The heifers ranged from 10.5 to 13.5 months of age and had a body condition score between five and seven at the time of evaluating the antral follicles and reproductive tracts. AFC data were collected on heifers ($n$ = 220) by performing ovarian ultrasound imaging with an Ibex EVO portable ultrasound with a 7.5 MHz linear probe [7,10]. The recorded ultrasounds were examined to identify follicles $\geq$3 mm, which were counted for the total AFC [16]. Reproductive tracts were scored in heifers ($n$ = 293) using palpation and confirmed with ultrasound [12,16]. Data from the first year of heifers have been previously published by Reynolds et. al in 2018 [16].

### 2.2. Genotyping

Blood was collected from 293 heifers at NMCREEC and shipped to the University of Idaho where DNA was isolated using the phenol chloroform method as previously described [17]. DNA for each animal was genotyped with the Bovine GGP 50K chip that consisted of 47,843 Single Nucleotide Polymorphism (SNP) markers (Neogen, Lincoln, NE, USA). In total, 293 samples were genotyped; however, four samples were removed as they had a call rate <0.9. Therefore, 289 samples were analyzed; however, only a subset of these samples ($n$ = 217) had AFC data. Further, non-autosomal markers and those with a call rate <0.9 were removed. The remaining 45,436 variants were phased using Eagle (v2.4.1) [18], and then imputed up to 836,121 SNP markers with Minimac3 [19] using the methods and reference panels described in Rowan et al. [20]. Briefly, the imputation reference panel contained 9629 animals genotyped on the Illumina HD array (777K SNPs), 28,183 animals genotyped on the GGP-F250 array (~227K SNPs) and 2718 animals genotyped on both high-density assays. The multi-breed imputation reference contained between 354 (Shorthorn) and 16,703 (Angus) high-density individuals from the component breeds represented in the genotyped dataset. Rowan et al. (2019) observed high individual imputation accuracies ($r^2$ > 0.99) for crossbred animals using this same multi-breed reference panel [20]. After

the imputation, SNPs with a minor allele frequency (MAF) <0.01 were discarded and the remaining 712,666 SNPs were used in the subsequent analyses.

### 2.3. Genotypic Analyses

A principal component analysis (PCA) was performed to examine the genetic relatedness of heifers. A Fisher's Exact Test was run in R version 3.6.2 and used to determine if the PCA groups had a difference in proportion of heifers classified in the high, medium and low AFC groups [21]. An association analysis was performed for 217 heifers with AFC data. A kinship matrix was first created to correct for any population structure that may exist in the sample set. Year of collection, age at time of AFC (to the nearest half month) and PCA groupings were used as covariates. A multi-locus, mixed-model additive association test was performed [22]. A separate association test with 289 heifers and RTS phenotype was also performed. For this GWAS, RTS of 1 and 2 (prepubertal heifers) were classified as cases, and RTS of 3–5 were classified as controls. Subsequent to a genetic relationship matrix correction, an additive correlation/trend association test was performed. The SNP & Variation Suite^{TM} software, version 8.7.2, was used for the PCA and GWAS analyses (Golden Helix, Inc., Bozeman, MT, USA; www.goldenhelix.com). Genome-wide significance threshold of $p < 1.00 \times 10^{-5}$ was set based on previous species-specific research [23–25].

## 3. Results

### 3.1. Principal Component Analysis

In order to examine the genetic relationship of the heifers used in the AFC analysis, we performed a principal component analysis (PCA) and plotted the first two eigenvalues. The PCA plot depicts that the heifers separate into four distinct groups (Figure 1). Group 1 consists of heifers that were sired by a single SimAngus sire and Group 2 are heifers sired by six Hereford bulls. Group 3 consists of heifers sired by a single Angus bull and Group 4, the largest group, consists of heifers out of 18 Angus, SimAngus, Simmental and Shorthorn sires. Groups 1 and 3 are comprised of heifers that belong in the year two cohort, and their sires were not used previously for year 1. Groups 2 and 4 consist of heifers sired by animals that were used for both years. The individual heifers in this plot are colored based on AFC categories of low ($\leq$15), medium (16–24) and high ($\geq$25) [16]. Interestingly, heifers in the high AFC category are not evenly distributed between the PCA groups. Over half (55%) of all high AFC heifers are grouped in PCA Group 2, and although PCA Group 4 contains 50% of the population, only 22% of the Group 4 heifers had a high AFC. Using a Fisher's Exact Test, there was a significant ($p = 0.004$) difference in the proportion of heifers classified as high AFC vs. those classified as medium and low between the PCA groups.

### 3.2. GWAS with Antral Follicle Count

A GWAS was performed to test for genetic associations in 217 heifers with different AFCs. Chromosomes 2, 3 and 23 exhibited loci that are significantly associated with differences in AFC (Table 1). In total there were 14 significant SNPs, with seven on chromosome 2, five on chromosome 3 and two SNPs on chromosome 23. The results from the GWAS are displayed in a Manhattan plot in Figure 2.
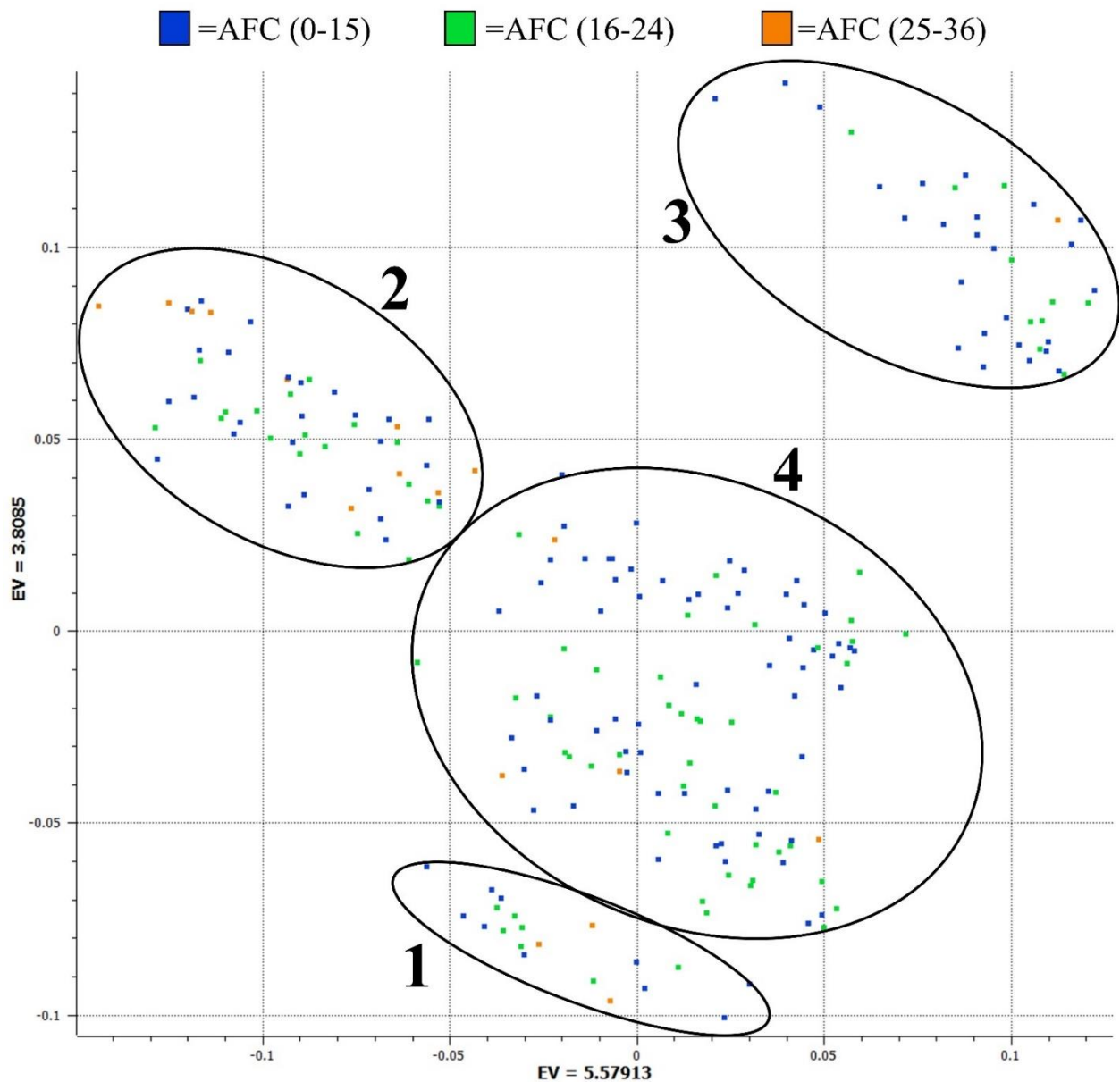
**Figure 1.** Principal component analysis plot to group animals by genotype by plotting the two largest eigenvectors (EV). Group 1 are heifers of a single SimAngus sire; Group 2 are all Hereford-sired heifers; Group 3 are heifers of a single Angus sire; and Group 4 are the heifers of the remining Angus, Simmental, SimAngus and Shorthorn sires. The individual heifers are color coded based on the number of the antral follicle count (AFC) observed, with blue exhibiting lower (0–15), green medium (16–24) and orange the highest number of (25–36) the AFC observed.

**Table 1.** Chromosome, rs number, position, *p*-value, −log(*p*-value), percentage of variance explained, and the candidate genes for significant SNPs associated with the AFC.

| Chromosome | rs Number | Position (bp) | *p*-value | −log(*p*-value) | Percentage of Variance Explained (%) | Candidate Genes |
|---|---|---|---|---|---|---|
| 23 | 42333762 | 48495249 | $7.20 \times 10^{-8}$ | 7.14 | 12.9 | *BMP6, RREB1* |
| | 42333752 | 48497351 | $8.03 \times 10^{-8}$ | 7.10 | 12.8 | |
| 2 | 110145174 | 96774781 | $2.53 \times 10^{-7}$ | 6.60 | 11.9 | *CREB1, FZD5, PLEKHM3, IDH1, PIKFYVE, MAP2* |
| | 109967601 | 96714406 | $3.67 \times 10^{-6}$ | 5.43 | 9.7 | |
| | 109066756 | 96804827 | $3.68 \times 10^{-6}$ | 5.43 | 9.7 | |
| | 109574474 | 96807982 | $3.68 \times 10^{-6}$ | 5.43 | 9.7 | |
| | 110145277 | 96801078 | $4.16 \times 10^{-6}$ | 5.38 | 9.6 | |
| | 109609461 | 96801273 | $4.16 \times 10^{-6}$ | 5.38 | 9.6 | |
| | 135894326 | 96802055 | $4.16 \times 10^{-6}$ | 5.38 | 9.6 | |
| 3 | 133573457 | 104450767 | $7.58 \times 10^{-6}$ | 5.12 | 9.1 | *FOXO6* |
| | 43366810 | 104451813 | $7.58 \times 10^{-6}$ | 5.12 | 9.1 | |
| | 43367756 | 104420695 | $9.79 \times 10^{-6}$ | 5.01 | 8.9 | |
| | 110027403 | 104422767 | $9.79 \times 10^{-6}$ | 5.01 | 8.9 | |
| | 43367746 | 104423898 | $9.79 \times 10^{-6}$ | 5.01 | 8.9 | |
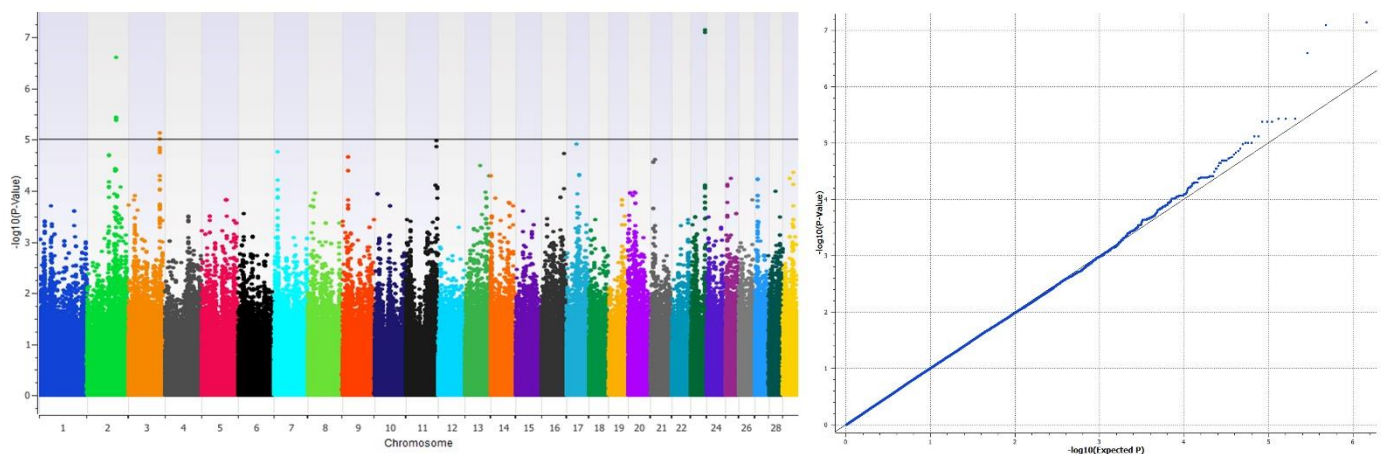


**Figure 2.** Manhattan plot of the GWAS for variation of the antral follicle count (AFC) in heifers in the left panel and the corresponding qq plot in the right panel. The solid line denotes genome-wide significance at −log(*p*-value) = 5.

The percentage of variance explained (PVE) identified in Table 1 was calculated in SVS to show how much variation can be explained by the effects of the marker [22]. The two significant SNPs on chromosome 23 explained most of the variation at 12.9% with the seven SNPs on chromosome 2 explaining 11.9% of the variation. Further, the five SNPs on chromosome 3 explained 9.1%. In relation to the reference genome ARS-UCD 1.2, the homozygous reference markers on chromosomes 2 and 23 are associated with increased AFC while the reference markers on chromosome 3 are associated with decreased AFC [26].

Subsequently, annotated genes within 1 Mb upstream and downstream of significant SNPs were identified. An enlarged view of the genomic regions containing significant SNPs as well as all markers within 1 Mb upstream and downstream are illustrated in Figure 3. Within this region, there are a total of 12 genes located on chromosome 2, 22 genes on chromosomes 3 and 13 genes on chromosome 23 (Supplemental Table S1).
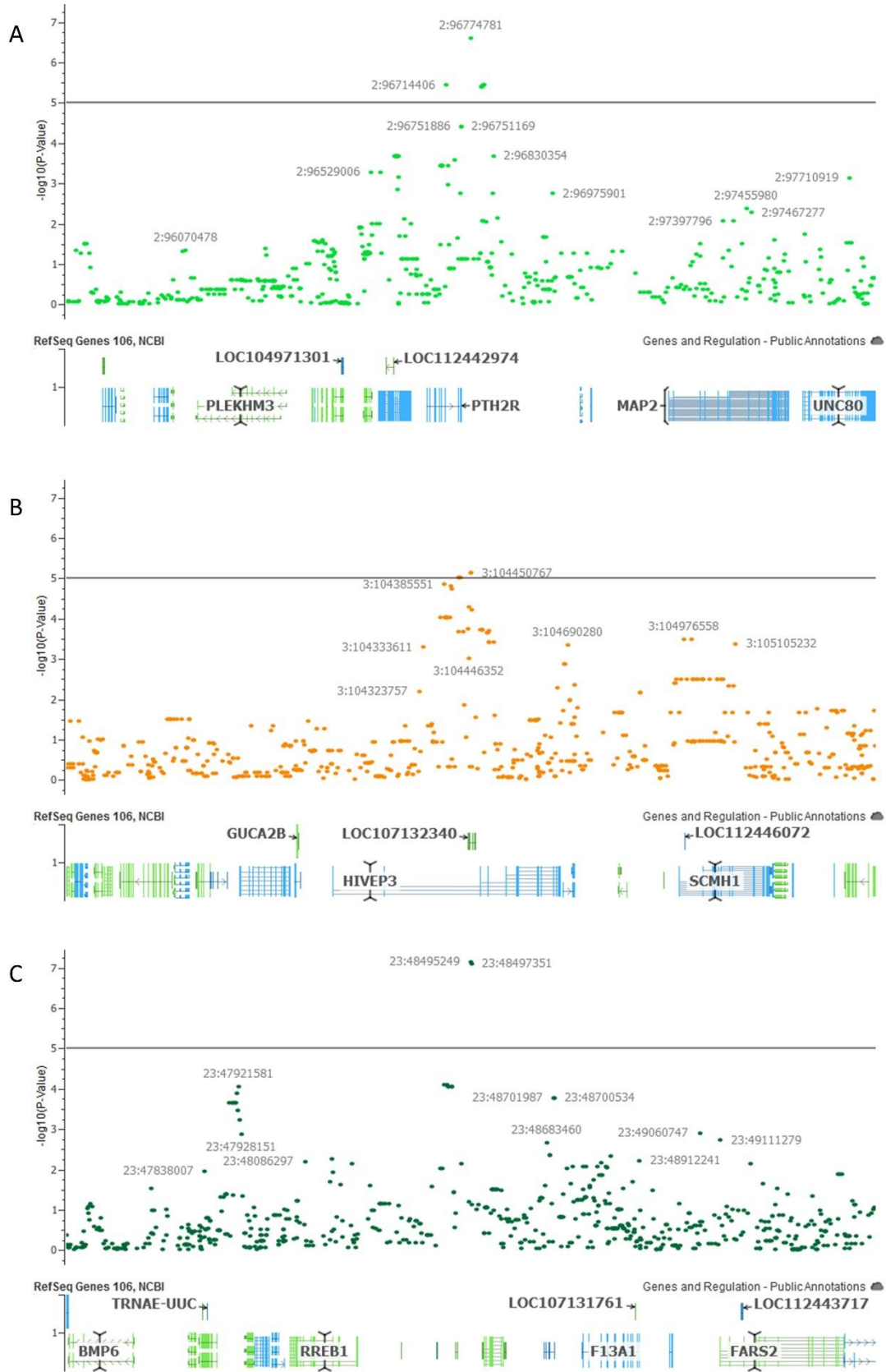
**Figure 3.** Regions within 1 Mbp of the most significant SNP for the antral follicle count GWAS results for (**A**) chromosome 2; (**B**) chromosome 3; (**C**) chromosome 23.

The genes in the regions of interest identified as candidate genes are noted in Table 1. The genes were categorized as candidate genes because they have been identified in pathways that have known effects or the potential to affect follicle development. The most significant region on chromosme 23 had two candidate genes identified. On chromosome 2, a total of six candidate genes were identified while on chromosome 3 one gene was identified. Of these nine identified genes, six are currently known to contribute to or affect the process of three well-known biological pathways linked to reproduction. These pathways are the PI3K/AKT, WNT signaling and MAP kinase (MAPK) pathways, all of which have been identified as affecting folliculogenesis and demonstrates prior knowledge of candidate gene pathways.

### 3.3. GWAS with Reproductive Tract Scores

A separate GWAS was performed using 289 heifers to investigate genetic associations with the RTS phenotypes (Figure 4). The focus of this analysis was to examine the pre-pubertal animals in comparison to the pubertal heifers. There are four significant loci above the genome-wide significance threshold of $p < 1.00 \times 10^{-5}$. Chromosomes 2 and 8 had two significant SNPs each and chromosomes 10 and 11 had one significant SNP each (Table 2). There was the same area of significance between the two fertility traits on chromosome 2 at 96.8 Mb.
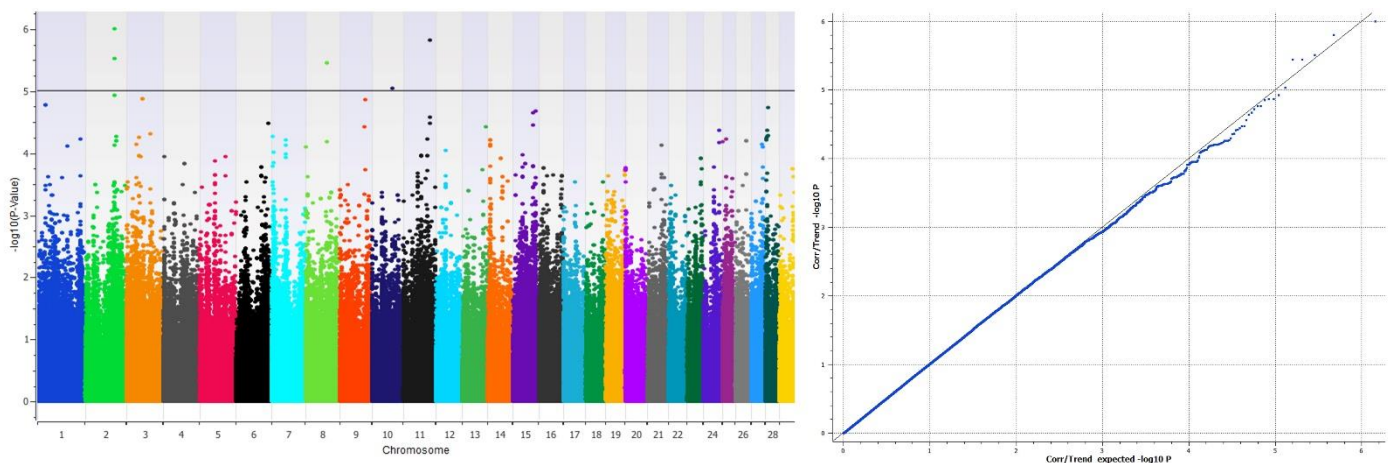


**Figure 4.** Manhattan plot of the GWAS for the variation in reproductive tract scores (RTS) in heifers in the left panel and the corresponding qq plots in the panel on the right. The solid line is at significance level –log(*p*-value) = 5.

**Table 2.** Chromosome, rs number, position, *p*-value, –log(*p*-value) and candidate genes for significant SNPs associated with the RTS.

| Chromosome | rs Number | Position (bp) | *p*-value | $-\log(p\text{-value})$ | Candidate Genes |
|:---:|:---:|:---:|:---:|:---:|:---:|
| 2 | 110658876 | 96879125 | $9.99 \times 10^{-7}$ | 6.00 | *CREB1, FZD5, PLEKHM3, IDH1, PIKFYVE, MAP2* |
| | 136255286 | 96870517 | $3.05 \times 10^{-6}$ | 5.52 | |
| 8 | 110172413 | 71343554 | $3.56 \times 10^{-6}$ | 5.45 | *LOXL2, STC1* |
| | 137380598 | 71359794 | $3.56 \times 10^{-6}$ | 5.45 | |
| 10 | 134739799 | 68971489 | $9.11 \times 10^{-6}$ | 5.04 | *PELI2, TMEM260, OTX2* |
| 11 | 111004666 | 90023995 | $1.55 \times 10^{-6}$ | 5.81 | *RNF144A, CMPK2, SOX11* |

From these four locations, regions of significance were defined as 1 Mb upstream and downstream of each significant SNP. These regions were investigated using the ARS-UCD 1.2 reference genome (Figure 5) [26]. The significant SNPs identified were in gene-rich regions. In the significant regions there are a total of 14 genes on chromosome 2, 11 genes

on chromosome 8, nine genes on chromosome 10 and four genes on chromosome 11 (Supplemental Table S2).
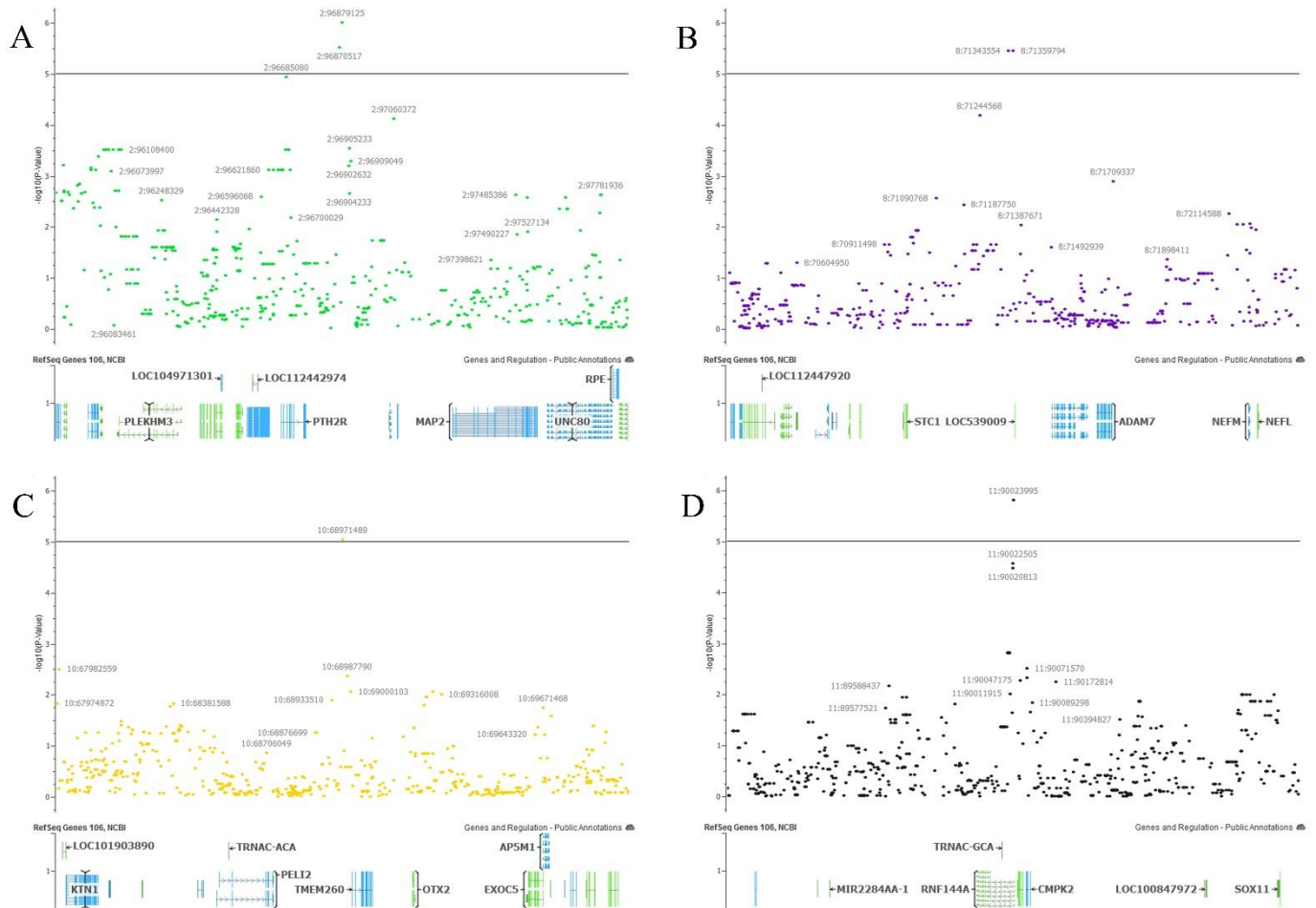


**Figure 5.** Regions within 1 Mbp of the most significant SNP for the reproductive tract scores (RTS) GWAS results for (**A**) chromosome 2; (**B**) chromosome 8; (**C**) chromosome 10; (**D**) chromosome 11.

Of the genes listed in the regions of interest, 15 genes were identified as candidate genes (Table 2). The genes were identified as candidate genes because they have been identified in pathways that have known effects or the potential to affect the development of the reproductive tract. From the 15 candidate genes, six are located in proximity with the most significant loci on chromosome 2. These six genes on chromosome 2 are located in the proximity of significant SNPs identified in both the AFC and RTS GWASs. The other significant regions on chromosome 8, 10 and 11 have three candidate genes each for a total of nine candidate genes. Two of these candidate genes are involved with pathways previously mentioned, the WNT signaling and the MAPK pathway.

## 4. Discussion

Fertility is a highly variable trait but is of major importance in the cattle industry. Cattle fertility can be measured by AFC and RTS, although little is understood of the genetic variation associated with these measurements. In order to examine the genetic associations with these traits we conducted two GWAS using crossbred heifers.

### 4.1. GWAS Related to Antral Follicle Count

The genes located within the significant regions of the AFC GWAS that have previously been described in important biological pathways are reported in Figure 6. One notable pathway is the PI3K/AKT pathway, as it is known to be involved in the activation of AKT in granulosa cells in rats and we can predict a synonymous role in the granulosa cells of cattle [27]. Two candidate genes identified in this study, pleckstrin homology domain containing M3 (*PLEKHM3*) and forkhead box O6 (*FOXO6*), are closely related to the PI3K/AKT pathway. *PLEKHM3* codes for a scaffold protein for AKT that localizes AKT to the plasma membrane, an essential step in AKT activation, whereas FOXO6 is a transcription factor that is regulated by AKT [28,29]. Little is known about FOXO6 in tissues other than the brain; however, other genes in this family are involved in cell metabolism and death of oocytes [30,31].
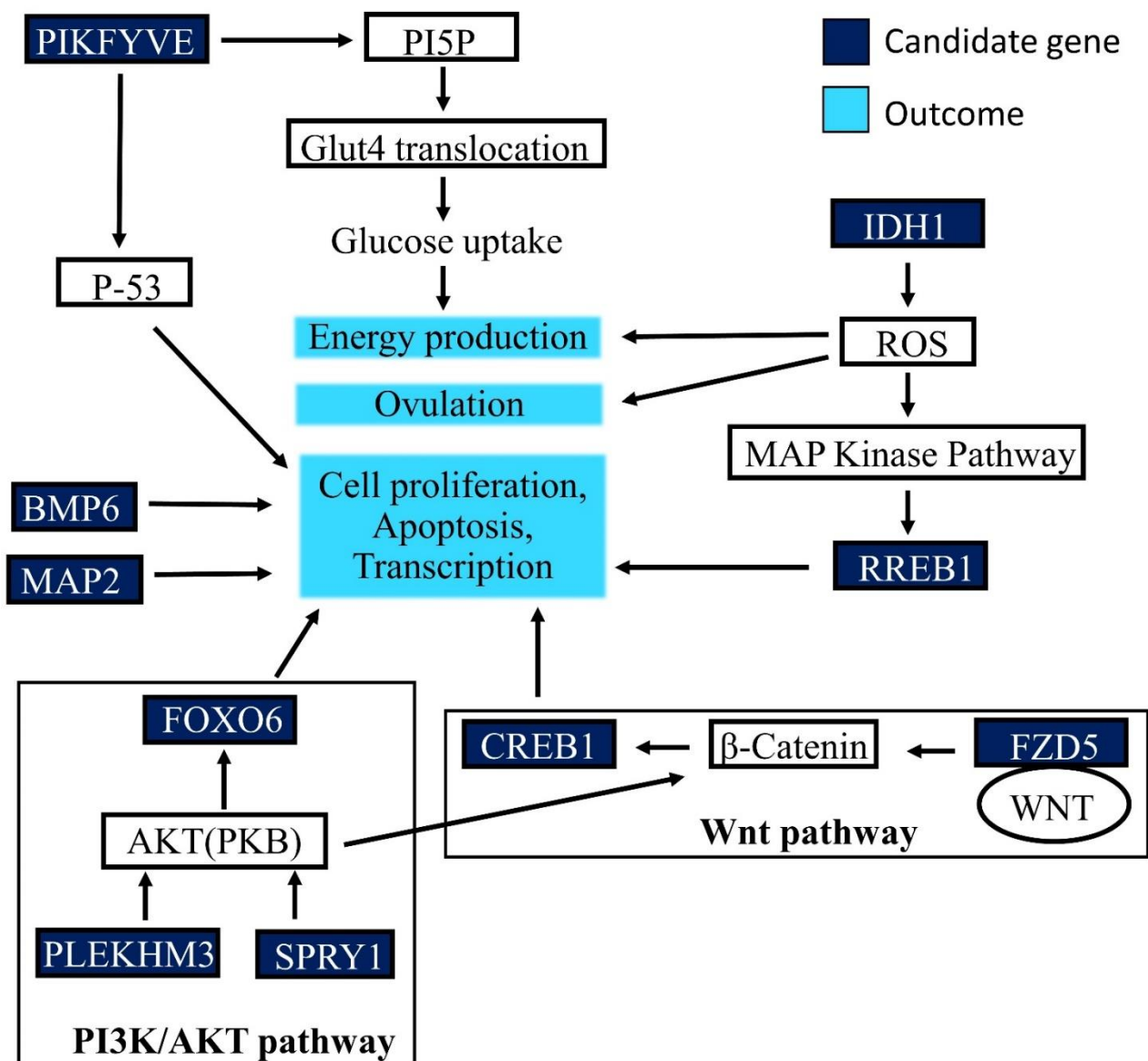


**Figure 6.** Proposed pathway(s) for the candidate genes in regions of the significant SNPs associated with antral follicle count (AFC). Arrows indicate an effect not increasing or decreasing.

The candidate gene phosphoinositide kinase, FYVE-type zinc finger containing protein (*PIKFYVE*) directly affects the level of the cellular phosphoinositides, PI(3,5)P2 and

PI5P [32]. When the PI5P levels increase, this triggers the translocation of glucose transporters to the cell membrane, which leads to an increase in intracellular glucose (energy) levels [33]. PI5P further regulates the p-53 dependent apoptotic pathway, and it is plausible that preventing apoptosis supports folliculogenesis and maturation [33]. *PIKFYVE* has been identified as a putative target of microRNA that is differentially expressed in dominant theca cells [34]. It is conceivable that the association of this pathway with reproductive traits relates to the influence on cellular energy levels as well as regulation of apoptosis.

The WNT signaling pathway contributes to oocyte development [35]. The canonical WNT pathway uses β-catenin to affect gene transcription [30]. Previous studies show that WNT proteins are required for ovarian follicle development via regulating FSH and LH signaling [36]. WNT2 induces an accumulation of β-catenin that through the canonical pathway causes an increase in estradiol production [35]. Alternatively, for normal female fertility, WNT5a suppresses the canonical pathway [36]. The critical role of Wnt5a in folliculogenesis is supported by the report that Wnt5a-null mice experience an increased rate of follicular atresia [36]. Two candidate genes are involved in this pathway, frizzled receptor 5 (*FZD5*) and cAMP responsive element binding protein (*CREB1*). FZD5 is a receptor for several WNT ligands, including WNT5a and WNT2. β-catenin is also affected by the PI3K/AKT pathway mentioned previously. β-catenin can be phosphorylated by AKT, which increases its transcriptional activity [37]. CREB1 is associated with the canonical pathway and is shown to have decreased levels in this process when the canonical pathway is suppressed [36]. *CREB1* is also a putative target of microRNA that is differentially expressed in the dominant granulosa cells [34].

The candidate gene isocitrate dehydrogenase 1 (*IDH1*) codes for a protein involved in the TCA cycle and regulates oxidoreductase activity as well as metabolism. A study in rats engineered to mimic PCOS saw a reduction in *IDH1* levels, indicating that *IDH1* is involved in follicular growth and ovulation [38]. Due to its involvement in the TCA cycle, *IDH1* may affect the energy and reactive oxygen species (ROS) levels [39]. ROS are necessary to induce ovulation [39]. In addition, ROS are mediators of the MAPK pathway, which has been shown to affect follicle maturation [39,40]. Another candidate gene ras responsive element binding protein (*RREB1*) is a downstream effector of the MAPK [41]. Overall, changes in ROS levels can affect the outcomes of the MAPK pathway, thereby influencing cell proliferation and the likelihood of ovulation.

The last two candidate genes are bone morphogenic protein-6 (*BMP6*) and microtubule-associated protein 2 (*MAP2*). Several bone morphological proteins and their receptors are involved of cell proliferation, apoptosis and cell differentiation [42]. In the ovary, *BMP6* works to regulate FSH through decreasing the cAMP levels [43]. The last gene, *MAP2*, codes for a protein that helps to stabilize microtubules during growth, a step which is important for proper separation during mitosis [44]. This protein is also shown to be upregulated as follicles mature normally from primordial to secondary follicles in mice [45].

A study by Fortes et. al examined genes associated with puberty in tropical breeds of cattle [46]. The authors of this study identified two genes, HIVEP zinc finger 3 (*HIVEP3*) and *RREB1*, which are also found in our regions of significance for our AFC GWAS on chromosome 3 and chromosome 23, respectively. The previous association of these regions with onset of puberty supports our findings with fertility traits in beef cattle.

Identifying the underlying genetic variation that contributes to the biological mechanisms of AFC can help improve knowledge of the aspects of germ cell development that are currently not understood. Previous studies have enumerated the benefits of animals that present a high AFC. However, more recent studies in various dairy and *Bos indicus* animals have shown contrasting results. In a study with Holsteins, heifers in the high AFC group had more days open and less percent pregnant at the end of lactation than either the low or medium groups [47]. Similarly, a study in Nelore cattle reports the highest conception rates in animals with the lowest AFC, and the lowest conception rates in animals with the highest AFC [48].

One possible explanation is that some animals with high AFC have polycystic ovary syndrome (PCOS), which negatively impacts fertility and skews the relationship between follicle number and fertility [47]. Females with PCOS may have a higher AFC, but chronic anovulation results in subfertility [49,50]. In addition, the question remains as to whether increased AFC indicates a longer reproductive lifespan or alternatively that the OR is depleted more quickly [51]. This study has identified nine candidate genes that are associated with differences in AFC. These genes also warrant further investigation with respect to their influence and associations with PCOS.

*4.2. GWAS Related to Reproductive Tract Scores (RTS)*

This study identifies regions on four chromosomes that are significantly associated with the RTS. The region on chromosome 2 is the same region also associated with AFC. A component of RTS determination is ovarian structure and either the presence or absence of ovulating follicles. Thus, it is reasonable to have the same region be significant as both GWAS include follicle development. Two candidate genes identified in the RTS GWAS influence the pathways previously discussed the context of AFC. These include Pellino2 (*PELI2*), which is associated with the MAPK pathway, and sex-determining region Y-box transcription factor 11 (*SOX11*), which is associated with the WNT signaling pathway [52]. PELI2 can activate the MAPK pathway, which, as previously discussed, can affect the maturation of ovarian follicles [40,53]. *SOX11* is expressed in ovarian cells and can dampen the WNT signaling pathway [52,54]. As mentioned previously, WNT proteins are crucial to regulate FSH and LH signaling for follicle development [36].

Multiple candidate genes also affect follicle maturation in pathways not previously discussed. Although first identified in maintenance of mineral levels, Stanniocalcin (*STC1*) has been shown to be highly detectable in the ovaries of mature mice [55]. It has been reported that overexpression of *STC1* in mice is significantly associated with a reduction in litter size [56]. Although this mechanism remains unclear, it is evident that *STC1* affects reproduction in mammals. RNF144A is a transmembrane ligase that interacts with epidermal growth factor receptor (EGFR) and thereby affects EGFR-stimulated cell proliferation [57]. Previous studies in mice demonstrate that EGFR is required for cumulous-oocyte complex maturation [58]. Through this interaction, changes in RNF144A can affect the maturation and development of follicles.

Several of these candidate genes affect mammalian reproductive ability. Lysyl oxidase-like protein-2 (*LOXL2*) functions in extracellular matrix production by collagen IV assembly [59]. *LOXL2* has been identified in human female reproductive tracts and in cells with advanced aging so its presence in mammalian reproductive tracts is important in our study when comparing heifers that mature less quickly than their cohorts [60,61]. Orthodenticle homeobox 2 (*OTX2*) affects the GnRH, LHB and FSHB levels in pituitary tissues of mice. In addition, mice without Otx2 had delayed vaginal opening and fewer litters than mice with a functioning Otx2 [62]. It is known that mice without Otx2 have smaller litters, abnormal estrous cycles and lack of corpus lutea [62]. Any change in OTX2 could affect not only follicle development and ovulation but also reproductive physiology.

Two final genes of note in our significant region: transmembrane protein 260 (*TMEM260*) and cytodine/uridine monophosphate kinase 2 (*CMPK2*). *TMEM260* is a putative transmembrane protein with an unknown function. According to NCBI, the expression of *TMEM260* is high in ovaries, implying a function in mammal reproduction [63]. *CMPK2* is a mitochondrial protein that phosphorylates dCMP and dUMP during mtDNA synthesis; however, it has not been observed in all tissues [64,65]. *CMPK2*, as it functions in the mitochondria, may be involved in the regulation of essential nucleotide during mitochondrial biogenesis and influence energy production and growth [64].

## 5. Conclusions

The GWAS identified three regions significantly associated with the AFC and four associated with the RTS in crossbred heifers. One region on chromosome 2 is significantly

associated with both fertility phenotypes and contains six candidate genes. In addition to these six genes, three candidate genes relating to AFC are associated with pathways that contribute to follicle growth through energy production, cell proliferation, transcription or apoptosis. In addition to the six candidate genes on chromosome 2, nine more candidate genes are in proximity to the regions associated with RTS and could affect the reproductive development of heifers. This study reports an association between these regions and AFC or RTS; however, additional work with larger sample sizes is needed to establish causal variants and the mechanisms of the effects they have on follicle development and fertility in cattle.

# References

1. Jimenez-Krassel, F.; Folger, J.K.; Ireland, J.L.H.; Smith, G.W.; Hou, X.; Davis, J.S.; Lonergan, P.; Evans, A.C.O.; Ireland, J.J. Evidence That High Variation in Ovarian Reserves of Healthy Young Adults Has a Negative Impact on the Corpus Luteum and Endometrium During Estrous Cycles in Cattle. *Biol. Reprod.* **2009**, *80*, 1272–1281. [CrossRef]
2. Ireland, J.J.; Ward, F.; Jimenez-Krassel, F.; Ireland, J.L.H.; Smith, G.W.; Lonergan, P.; Evans, A.C.O. Follicle numbers are highly repeatable within individual animals but are inversely correlated with FSH concentrations and the proportion of good-quality embryos after ovarian stimulation in cattle. *Hum. Reprod.* **2007**, *22*, 1687–1695. [CrossRef]
3. Walsh, S.W.; Mossa, F.; Butler, S.T.; Berry, D.P.; Scheetz, D.; Jimenez-Krassel, F.; Ireland, J.J. Heritability and impact of environmental effects during pregnancy on antral follicle count in cattle. *J. Dairy Sci.* **2014**, *97*, 4503–4511. [CrossRef]
4. Silva-Santos, K.C.; Siloto, L.S.; Santos, G.M.G.; Morotti, F.; Marcantonio, T.N.; Seneda, M.M. Comparison of Antral and Preantral Ovarian Follicle Populations Between Bos indicus and Bos indicus-taurus Cows with High or Low Antral Follicles Counts. *Reprod. Domest. Anim.* **2014**, *49*, 48–51. [CrossRef] [PubMed]
5. Dos Santos, G.M.G.; Silva-Santos, K.C.; Barreiros, T.R.R.; Morotti, F.; Sanches, B.V.; de Moraes, F.L.Z.; Blaschi, W.; Seneda, M.M. High numbers of antral follicles are positively associated with in vitro embryo production but not the conception rate for FTAI in Nelore cattle. *Anim. Reprod. Sci.* **2016**, *165*, 17–21. [CrossRef] [PubMed]
6. Jimenez-Krassel, F.; Scheetz, D.M.; Neuder, L.M.; Ireland, J.L.H.; Pursley, J.R.; Smith, G.W.; Tempelman, R.J.; Ferris, T.; Roudebush, W.E.; Mossa, F.; et al. Concentration of anti-Müllerian hormone in dairy heifers is positively associated with productive herd life. *J. Dairy Sci.* **2015**, *98*, 3036–3045. [CrossRef]
7. Ireland, J.L.H.; Scheetz, D.; Jimenez-Krassel, F.; Themmen, A.P.N.; Ward, F.; Lonergan, P.; Smith, G.W.; Perez, G.I.; Evans, A.C.O.; Ireland, J.J. Antral Follicle Count Reliably Predicts Number of Morphologically Healthy Oocytes and Follicles in Ovaries of Young Adult Cattle. *Biol. Reprod.* **2008**, *79*, 1219–1225. [CrossRef]

8. Mossa, F.; Walsh, S.W.; Butler, S.T.; Berry, D.P.; Carter, F.; Lonergan, P.; Smith, G.W.; Ireland, J.J.; Evans, A.C.O. Low numbers of ovarian follicles ≥3 mm in diameter are associated with low fertility in dairy cows. *J. Dairy Sci.* **2012**, *95*, 2355–2361. [CrossRef]

9. Mossa, F.; Ireland, J.J. Physiology and Endocrinology Symposium: Anti-Müllerian hormone: A biomarker for the ovarian reserve, ovarian function, and fertility in dairy cows. *J. Anim. Sci.* **2019**, *97*, 1446–1455. [CrossRef]

10. Cushman, R.A.; Allan, M.F.; Kuehn, L.A.; Snelling, W.M.; Cupp, A.S.; Freetly, H.C. Evaluation of antral follicle count and ovarian morphology in crossbred beef cows: Investigation of influence of stage of the estrous cycle, age, and birth weight. *J. Anim. Sci.* **2009**, *87*, 1971–1980. [CrossRef]

11. Brinks, J.S. Genetic influences on reproductive performance of two-year-old beef females. In *Factors Affecting Calf Crop*; Fields, M.J., Sand, R.J., Eds.; CRC Press: Boca Raton, FL, USA, 1994; pp. 45–53.

12. Martin, L.C.; Brinks, J.S.; Bourdon, R.M.; Cundiff, L.V. Genetic effects on beef heifer puberty and subsequent reproduction. *J. Anim. Sci.* **1992**, *70*, 4006–4017. [CrossRef] [PubMed]

13. Gutierrez, K.; Kasimanickam, R.; Tibary, A.; Gay, J.M.; Kastelic, J.P.; Hall, J.B.; Whittier, W.D. Effect of reproductive tract scoring on reproductive efficiency in beef heifers bred by timed insemination and natural service versus only natural service. *Theriogenology* **2014**, *81*, 918–924. [CrossRef] [PubMed]

14. Neupane, M.; Geary, T.W.; Kiser, J.N.; Burns, G.W.; Hansen, P.J.; Spencer, T.E.; Neibergs, H.L. Loci and pathways associated with uterine capacity for pregnancy and fertility in beef cattle. *PLoS ONE* **2017**, *12*, e0188997. [CrossRef]

15. Cole, J.B.; Wiggans, G.R.; Ma, L.; Sonstegard, T.S.; Lawlor, T.J.; Crooker, B.A.; Van Tassell, C.P.; Yang, J.; Wang, S.; Matukumalli, L.K.; et al. Genome-Wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genom.* **2011**, *12*, 408. [CrossRef]

16. Reynolds, M.; Chibisa, G.; Ahmadzadeh, A.; Hall, J. Reproductive development and fertility traits among heifers in different residual feed intake groups1. *Transl. Anim. Sci.* **2018**, *2*, S175–S179. [CrossRef] [PubMed]

17. Sambrook, J.; Fritsch, E.F.; Maniatis, T. *Molecular Cloning: A laboratory Manual*; Cold Spring Harbor Laboratory Press: New York, NY, USA, 1989.

18. Loh, P.-R.; Danecek, P.; Palamara, P.F.; Fuchsberger, C.A.; Reshef, Y.K.; Finucane, H.; Schoenherr, S.; Forer, L.; McCarthy, S.; Abecasis, G.R.; et al. Reference-Based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **2016**, *48*, 1443–1448. [CrossRef]

19. Das, S.; Forer, L.; Schönherr, S.; Sidore, C.; Locke, A.E.; Kwong, A.; Vrieze, S.I.; Chew, E.Y.; Levy, S.; McGue, M.; et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **2016**, *48*, 1284–1287. [CrossRef]

20. Rowan, T.N.; Hoff, J.L.; Crum, T.E.; Taylor, J.F.; Schnabel, R.D.; Decker, J.E. A multi-breed reference panel and additional rare variants maximize imputation accuracy in cattle. *Genet. Sel. Evol.* **2019**, *51*, 77. [CrossRef]

21. Fisher, R.A. Statistical Methods for Research Workers. In *Breakthroughs in Statistics: Methodology and Distribution*; Kotz, S., Johnson, N.L., Eds.; Springer: Berlin/Heidelberg, Germany, 1992; pp. 66–70. [CrossRef]

22. Segura, V.; Vilhjálmsson, B.J.; Platt, A.; Korte, A.; Seren, Ü.; Long, Q.; Nordborg, M. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet.* **2012**, *44*, 825–830. [CrossRef]

23. Seabury, C.M.; Oldeschulte, D.L.; Saatchi, M.; Beever, J.E.; Decker, J.E.; Halley, Y.A.; Bhattarai, E.K.; Molaei, M.; Freetly, H.C.; Hansen, S.L.; et al. Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genom.* **2017**, *18*, 386. [CrossRef]

24. Higgins, M.G.; Fitzsimons, C.; McClure, M.C.; McKenna, C.; Conroy, S.; Kenny, D.A.; McGee, M.; Waters, S.M.; Morris, D.W. GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle. *Sci. Rep.* **2018**, *8*, 14301. [CrossRef] [PubMed]

25. Zhang, F.; Wang, Y.; Mukiibi, R.; Chen, L.; Vinsky, M.; Plastow, G.; Basarab, J.; Stothard, P.; Li, C. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: Feed efficiency and component traits. *BMC Genom.* **2020**, *21*, 36. [CrossRef] [PubMed]

26. 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.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience* **2020**, *9*. [CrossRef] [PubMed]

27. Gonzalez-Robayna, I.J.; Falender, A.E.; Ochsner, S.; Firestone, G.L.; Richards, J.S. Follicle-Stimulating Hormone (FSH) Stimulates Phosphorylation and Activation of Protein Kinase B (PKB/Akt) and Serum and Glucocorticoid-Induced Kinase (Sgk): Evidence for A Kinase-Independent Signaling by FSH in Granulosa Cells. *Mol. Endocrinol.* **2000**, *14*, 1283–1300. [CrossRef] [PubMed]

28. Virtanen, C.; Paris, J.; Takahashi, M. Identification and Characterization of a Novel Gene, dapr, Involved in Skeletal Muscle Differentiation and Protein Kinase B Signaling. *J. Biol. Chem.* **2009**, *284*, 1636–1643. [CrossRef] [PubMed]

29. Tzivion, G.; Dobson, M.; Ramakrishnan, G. FoxO transcription factors; Regulation by AKT and 14-3-3 proteins. *Biochim. Biophys. Acta Mol. Cell Res.* **2011**, *1813*, 1938–1945. [CrossRef] [PubMed]

30. Maiese, K.; Chong, Z.Z.; Shang, Y.C.; Hou, J. Rogue proliferation versus restorative protection: Where do we draw the line for Wnt and Forkhead signaling? *Expert Opin. Ther. Targets* **2008**, *12*, 905–916. [CrossRef] [PubMed]

31. Jacobs, F.M.J.; van der Heide, L.P.; Wijchers, P.J.E.C.; Burbach, J.P.H.; Hoekman, M.F.M.; Smidt, M.P. FoxO6, a Novel Member of the FoxO Class of Transcription Factors with Distinct Shuttling Dynamics. *J. Biol. Chem.* **2003**, *278*, 35959–35967. [CrossRef]

32. Zolov, S.N.; Bridges, D.; Zhang, Y.; Lee, W.-W.; Riehle, E.; Verma, R.; Lenk, G.M.; Converso-Baran, K.; Weide, T.; Albin, R.L.; et al. In vivo, Pikfyve generates PI(3,5)P2, which serves as both a signaling lipid and the major precursor for PI5P. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17472–17477. [CrossRef]

33. Hasegawa, J.; Strunk, B.S.; Weisman, L.S. PI5P and PI(3,5)P2: Minor, but Essential Phosphoinositides. *Cell Struct. Funct.* **2017**, *42*, 49–60. [CrossRef]

34. Zielak-Steciwko, A.E.; Browne, J.A.; McGettigan, P.A.; Gajewska, M.; Dzięcioł, M.; Szulc, T.; Evans, A.C.O. Expression of microRNAs and their target genes and pathways associated with ovarian follicle development in cattle. *Physiol. Genom.* **2014**, *46*, 735–745. [CrossRef]

35. Castañon, B.I.; Stapp, A.D.; Gifford, C.A.; Spicer, L.J.; Hallford, D.M.; Gifford, J.A.H. Follicle-Stimulating hormone regulation of estradiol production: Possible involvement of WNT2 and β-catenin in bovine granulosa cells. *J. Anim. Sci.* **2012**, *90*, 3789–3797. [CrossRef]

36. Abedini, A.; Zamberlam, G.; Lapointe, E.; Tourigny, C.; Boyer, A.; Paquet, M.; Hayashi, K.; Honda, H.; Kikuchi, A.; Price, C.; et al. WNT5a is required for normal ovarian follicle development and antagonizes gonadotropin responsiveness in granulosa cells by suppressing canonical WNT signaling. *FASEB J.* **2016**, *30*, 1534–1547. [CrossRef] [PubMed]

37. Fang, D.; Hawke, D.; Zheng, Y.; Xia, Y.; Meisenhelder, J.; Nika, H.; Mills, G.B.; Kobayashi, R.; Hunter, T.; Lu, Z. Phosphorylation of β-Catenin by AKT Promotes β-Catenin Transcriptional Activity. *J. Biol. Chem.* **2007**, *282*, 11221–11229. [CrossRef]

38. Salilew-Wondim, D.; Wang, Q.; Tesfaye, D.; Schellander, K.; Hoelker, M.; Hossain, M.M.; Tsang, B.K. Polycystic ovarian syndrome is accompanied by repression of gene signatures associated with biosynthesis and metabolism of steroids, cholesterol and lipids. *J. Ovarian Res.* **2015**, *8*, 24. [CrossRef] [PubMed]

39. Shkolnik, K.; Tadmor, A.; Ben-Dor, S.; Nevo, N.; Galiani, D.; Dekel, N. Reactive oxygen species are indispensable in ovulation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1462–1467. [CrossRef]

40. Fan, H.-Y.; Liu, Z.; Shimada, M.; Sterneck, E.; Johnson, P.F.; Hedrick, S.M.; Richards, J.S. MAPK3/1 (ERK1/2) in Ovarian Granulosa Cells Are Essential for Female Fertility. *Science* **2009**, *324*, 938–941. [CrossRef]

41. Deng, Y.-N.; Xia, Z.; Zhang, P.; Ejaz, S.; Liang, S. Transcription Factor RREB1: From Target Genes towards Biological Functions. *Int. J. Biol. Sci.* **2020**, *16*, 1463–1473. [CrossRef]

42. Wagner, D.O.; Sieber, C.; Bhushan, R.; Börgermann, J.H.; Graf, D.; Knaus, P. BMPs: From Bone to Body Morphogenetic Proteins. *Sci. Signal.* **2010**, *3*. [CrossRef]

43. Otsuka, F.; Moore, R.K.; Shimasaki, S. Biological Function and Cellular Mechanism of Bone Morphogenetic Protein-6 in the Ovary. *J. Biol. Chem.* **2001**, *276*, 32889–32895. [CrossRef]

44. Kalcheva, N.; Rockwood, J.M.; Kress, Y.; Steiner, A.; Shafit-Zagardo, B. Molecular and functional characteristics of MAP-2a: Ability of MAP-2a versus MAP-2b to induce stable microtubules in COS cells. *Cell Motil. Cytoskelet.* **1998**, *40*, 272–285. [CrossRef]

45. Yoon, S.-J.; Kim, K.-H.; Chung, H.-M.; Choi, D.-H.; Lee, W.-S.; Cha, K.-Y.; Lee, K.-A. Gene expression profiling of early follicular development in primordial, primary, and secondary follicles. *Fertil. Steril.* **2006**, *85*, 193–203. [CrossRef] [PubMed]

46. Fortes, M.R.S.; Reverter, A.; Nagaraj, S.H.; Zhang, Y.; Jonsson, N.N.; Barris, W.; Lehnert, S.; Boe-Hansen, G.B.; Hawken, R.J. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J. Anim. Sci.* **2011**, *89*, 1669–1683. [CrossRef] [PubMed]

47. Jimenez-Krassel, F.; Scheetz, D.M.; Neuder, L.M.; Pursley, J.R.; Ireland, J.J. A single ultrasound determination of ≥25 follicles ≥3 mm in diameter in dairy heifers is predictive of a reduced productive herd life. *J. Dairy Sci.* **2017**, *100*, 5019–5027. [CrossRef]

48. Morotti, F.; Zangirolamo, A.F.; Silva, N.C.; Silva, C.B.; Rosa, C.O.; Seneda, M.M. Antral follicle count in cattle: Advantages, challenges, and controversy. *Anim. Reprod.* **2017**, *14*, 514–520. [CrossRef]

49. Zawadzki, J.K.; Dunaif, A. Diagnostic Criteria for Polycystic Ovary Syndrome: Towards a Rationale Approach. In *Polycystic Ovary Syndrome*; Dunaif, A., Givens, J.R., Haseltine, F.P., Merriam, G.R., Eds.; Blackwell Scientific Pubications: Boston, MA, USA, 1992; pp. 377–384.

50. Wiser, A.; Shalom-Paz, E.; Hyman, J.H.; Sokal-Arnon, T.; Bantan, N.; Holzer, H.; Tulandi, T. Age-Related normogram for antral follicle count in women with polycystic ovary syndrome. *Reprod. BioMed. Online* **2013**, *27*, 414–418. [CrossRef]

51. Cushman, R.A.; Perry, G.A.; Britt, J.H. 39 Current understanding of factors influencing antral follicle count and applications to reproductive management in cattle. *J. Anim. Sci.* **2018**, *96*, 21–22. [CrossRef]

52. Liu, Z.; Zhong, Y.; Chen, Y.J.; Chen, H. SOX11 regulates apoptosis and cell cycle in hepatocellular carcinoma via Wnt/β-catenin signaling pathway. *Biotechnol. Appl. Biochem.* **2019**, *66*, 240–246. [CrossRef]

53. Jensen, L.E.; Whitehead, A.S. Pellino2 activates the mitogen activated protein kinase pathway. *FEBS Lett.* **2003**, *545*, 199–202. [CrossRef]

54. Fang, G.; Liu, J.; Wang, Q.; Huang, X.; Yang, R.; Pang, Y.; Yang, M. MicroRNA-223-3p Regulates Ovarian Cancer Cell Proliferation and Invasion by Targeting SOX11 Expression. *Int. J. Mol. Sci.* **2017**, *18*, 1208. [CrossRef]

55. Deol, H.K.; Varghese, R.; Wagner, G.F.; DiMattia, G.E. Dynamic Regulation of Mouse Ovarian Stanniocalcin Expression during Gestation and Lactation. *Endocrinology* **2000**, *141*, 3412–3421. [CrossRef]

56. Varghese, R.; Gagliardi, A.D.; Bialek, P.E.; Yee, S.-P.; Wagner, G.F.; Dimattia, G.E. Overexpression of Human Stanniocalcin Affects Growth and Reproduction in Transgenic Mice. *Endocrinology* **2002**, *143*, 868–876. [CrossRef]

57. Ho, S.-R.; Lin, W.-C. RNF144A sustains EGFR signaling to promote EGF-dependent cell proliferation. *J. Biol. Chem.* **2018**, *293*, 16307–16323. [CrossRef]

58. Su, Y.-Q.; Sugiura, K.; Li, Q.; Wigglesworth, K.; Matzuk, M.M.; Eppig, J.J. Mouse Oocytes Enable LH-Induced Maturation of the Cumulus-Oocyte Complex via Promoting EGF Receptor-Dependent Signaling. *Mol. Endocrinol.* **2010**, *24*, 1230–1239. [CrossRef]

59. Bignon, M.; Pichol-Thievend, C.; Hardouin, J.; Malbouyres, M.; Bréchot, N.; Nasciutti, L.; Barret, A.; Teillon, J.; Guillon, E.; Etienne, E.; et al. Lysyl oxidase-like protein-2 regulates sprouting angiogenesis and type IV collagen assembly in the endothelial basement membrane. *Blood* **2011**, *118*, 3979–3989. [CrossRef] [PubMed]

60. Shynlova, O.; Bortolini, M.A.T.; Alarab, M.; Shynlova, O.; Bortolini, M.A.T.; Alarab, M. Genes responsible for vaginal extracellular matrix metabolism are modulated by women's reproductive cycle and menopause. *Int. Braz. J. Urol.* **2013**, *39*, 257–267. [CrossRef] [PubMed]

61. Murano, S.; Thweatt, R.; Shmookler Reis, R.J.; Jones, R.A.; Moerman, E.J.; Goldstein, S. Diverse gene sequences are overexpressed in werner syndrome fibroblasts undergoing premature replicative senescence. *Mol. Cell. Biol.* **1991**, *11*, 3905–3914. [CrossRef] [PubMed]

62. Diaczok, D.; DiVall, S.; Matsuo, I.; Wondisford, F.E.; Wolfe, A.M.; Radovick, S. Deletion of Otx2 in GnRH Neurons Results in a Mouse Model of Hypogonadotropic Hypogonadism. *Mol. Endocrinol.* **2011**, *25*, 833–846. [CrossRef] [PubMed]

63. National Center for Biotechnology Information. Gene. Available online: https://www.ncbi.nlm.nih.gov/gene/ (accessed on 14 December 2020).

64. Xu, Y.; Johansson, M.; Karlsson, A. Human UMP-CMP Kinase 2, a Novel Nucleoside Monophosphate Kinase Localized in Mitochondria. *J. Biol. Chem.* **2008**, *283*, 1563–1571. [CrossRef]

65. Gandhi, V.V.; Samuels, D.C. Enzyme Kinetics of the Mitochondrial Deoxyribonucleoside Salvage Pathway Are Not Sufficient to Support Rapid mtDNA Replication. *PLoS Comput. Biol.* **2011**, *7*, e1002078. [CrossRef]