Identification of genes associated with reproductive function in dairy cattle

M. Sofia Ortega[1]

Division of Animal Sciences, University of Missouri, 65211, Columbia, MO, USA.

Abstract

The use of genomics has improved response to selection for functional traits with low heritability such as fertility traits. Much of the work on fertility traits has been performed through use of genome-wide association studies (GWAS) to identify genetic loci associated with reproductive traits. Under a GWAS approach, the assumption is that the markers on the panel are in linkage disequilibrium with causative mutations. In many cases, identification of the causative mutation is difficult because an associated genetic marker can be in intergenic regions and can be in linkage disequilibrium with variants in several nearby genes. Another approach is to identify candidate genes using knowledge of the biological pathways controlling a trait to search for single nucleotide polymorphism (SNP) in genes in those pathways. This should reveal putative causative markers responsible for genetic variation in biological function, and it is expected that the marker will be more strongly associated with a trait than one in linkage disequilibrium. An example of how a series of candidate gene studies demonstrate that identification of markers in genes involved in reproductive processes can lead to discovery of additional markers associated with genetic variation in reproductive traits is presented. In addition, the inclusion of candidate markers for fertility can improve reliability of genetic estimates for fertility traits, and the repeatability of the effects across a separate population of animals gives confidence that association elucidated by this set of markers is likely to be real. More importantly, the use of candidate genes can provide insights into the biology underpinning genetic variation in fertility, and that this understanding can lead to physiological interventions to improve reproductive function.

Keywords: candidate genes, fertility, genomic selection, reproductive function.

Introduction

Fertility is a complex trait and, it is regulated in part by genetics. In the dairy cow, genetic merit for fertility and production are negatively correlated ranging from 0.35 - 0.60 (Boichard and Manfredi, 1994; VanRaden *et al.*, 2004; Pritchard *et al.*, 2013) and the intense selection for milk production during the last five decades has been one of the causes of a decrease in the genetic merit for fertility in dairy breeds (Butler, 2003). Nevertheless, improvements in reproductive performance of dairy cows has been made during the last decade because of advancements in reproductive management (Royal *et al.*, 2000; Petersson *et al.*, 2008), increased emphasis on genetic selection of reproductive traits (Norman *et al.*, 2014), and incorporation of genomics into genetic selection schemes (García-Ruiz *et al.*, 2016).

Most reproductive traits are controlled by many genes, each of which has a small effect. This is evident from genome wide association studies (GWAS) in which genetic variation in a trait is partitioned into associations with individual single nucleotide polymorphism (SNP). The low heritability characteristic of reproductive traits is indicative that only a small proportion of phenotypic variance is due to additive actions of individual genes and that reliability estimates of breeding values are prone to be low. It does not, however, mean that reproductive traits are not under genetic control, many specific genes have been identified that contain mutations that are associated with reproductive function. Furthermore, clear differences in fertility have been found between genetic lines of animals. In Holstein, for example, cows with higher genetic merit for fertility had fewer services per conception, and shorter intervals from calving to conception compared with cows with low genetic merit for fertility (Cummins *et al.*, 2012a; Ortega *et al.*, 2017a).

Reproductive traits in dairy cattle

In the United States, three main female fertility traits are used in official genetic evaluations of dairy cattle: daughter pregnancy rate (DPR), cow conception rate (CCR) and heifer conception rate (HCR). DPR is defined as the percent of cows eligible for breeding that become pregnant in each 21-day period (i.e., over one estrous cycle). Conceptually, DPR is the product of estrous detection rate (the percent of cows in estrus that are detected in estrus) and pregnancy rate per insemination (the percent of inseminated cows that become pregnant). Practically, DPR is calculated from the term days open, which is the interval from calving to conception. Predicted transmitting ability (PTA) for DPR and days open are nearly linear function of each other. An increase of 1% in PTA for DPR equals a decrease of 4 days in the PTA for days open (VanRaden *et al.*, 2003). Cow conception rate is defined as the percent of lactating cows that become pregnant after each service while HCR is the same variable for heifers (VanRaden *et al.*, 2004). Heritability for these traits in Holsteins range from 0.001 - 0.016 (Pryce *et al.*, 2004; VanRaden *et al.*, 2004; Kuhn *et al.*, 2006). Nonetheless, low heritability has not prevented progress in genetic selection for fertility. During the last 15 years, breeding values for DPR have improved, in part as a result of

including fertility traits into economic indexes such as net merit and by inclusion of genomic information for breeding value calculations.

Types of mutations responsible for genetic variation in reproduction

Genetic variation is the basis of biological diversity in a population. In the bovine genome, the total sequence length is 2,670,123,310 bp (UMD 3.1.1). As of December 2017, there were 102,499,615 SNP and 10,462 other genetic structural variations (>50 base pairs) including deletion/insertions, copy number variant, duplications, inversions, translocations and complex chromosomal arrangements (Aken *et al.*, 2017). Genetic mutations ultimately affect the proteome of the organism either by affecting the structural properties of a protein or by modifying amount of protein in specific tissues. Most of the genetic studies are based in the association of SNP genotypes with a specific phenotype. How the physical location of a mutation relative to the coding and regulatory regions of specific genes can cause variation in phenotype is illustrated in Fig 1.
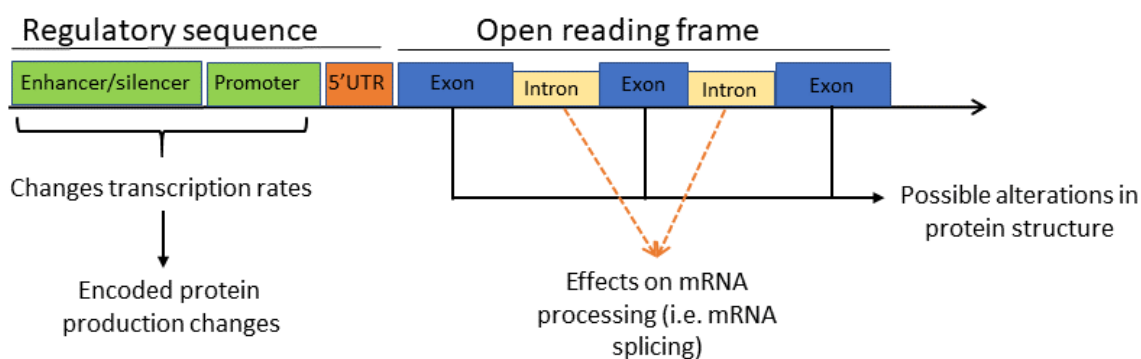

Figure 1. Possible effects of mutations relative to its location within a gene structure.

Approaches for gene discovery and genetic selection for fertility

Genome-wide association studies

GWAS are used to localize genomic regions that contribute to genetic variation of a trait. This approach is based on linkage disequilibrium, which refers to the association of any pair of alleles at different loci. Linkage disequilibrium exists because genes located closely together on a chromosome are more likely to be inherited together; i.e., cross-over events in meiosis are less likely to occur between two loci close together than two loci further apart or on different chromosomes. In a typical GWAS, thousands of SNP are interrogated for association with phenotypic variation in a trait, based on the assumption that SNP studied are in linkage disequilibrium with the causative mutation (Weller and Ron, 2011).

There are advantages and disadvantages of GWAS (Stringer *et al.*, 2011; Riancho, 2012; Frąszczak and Szyda, 2016; Zondervan *et al.*, 2016). The approach is unbiased with respect to previous knowledge of the trait of interest. Moreover, interrogation of a dense number of SNP across the genome can reveal novel markers associated with a trait. On the other hand, the large number of statistical tests performed during the analysis can lead to false positives so that stringent significance thresholds are necessary. One result is that markers with large effects are more likely to be detected and some important markers having a smaller effect may not reach significance. To get around these problems, large sample sizes are required to detect associations, particularly when multiple genes are involved in a trait. A limitation of GWAS is that often markers with significant associations with a trait are located in intergenic regions, and even when they are in linkage with the causative mutation, it is difficult to use this information to understand the basis of the genetic variance of the trait in question (Stringer *et al.*, 2011; Riancho, 2012; Frąszczak and Szyda, 2016; Zondervan *et al.*, 2016). Another limitation of GWAS is poor repeatability. Ioannidis *et al.* (2011) compiled results from a series of GWAS for human disease and found that replication of markers found by GWAS was around 1%. In cattle, the percent of significant SNP found in one population that were repeated in independent populations ranged from 0 (Littlejohn *et al.*, 2012) to 18% (Höglund *et al.*, 2014). Nevertheless, with appropriate sample sizes and statistical testing, GWAS can be very successful at identifying genes and genomic regions associated with specific traits. As an example, Cole *et al.* (2011) used a population of 1654 animals to identify 1586 SNP distributed in 486 genes that were associated with 31 production, reproduction, health and body conformation traits in Holstein cows. As more data becomes available, opportunities to validate these studies across populations become feasible, Liu *et al.* (2017) identified SNP in a Chinese Holstein population and those were later validated in a separate population of Nordic Holsteins. A detailed description of GWAS studies for female fertility can be found in the by Fortes *et al.* (2013), and the meta-assembly by Khatkar *et al.* (2014).

In dairy cattle, VanRaden *et al.* (2008) showed that incorporating data from GWAS into genetic estimates can improve reliability of genetic estimates over those based on parent averages. However, the

amount of improvement depended on the trait, being higher for production traits (increases in reliabilities of 23-43%) than for DPR (17%; Wiggans *et al.*, 2011). The lower increase for DPR probably reflects low heritability and the high degree of polygenicity. The inclusion of genomic information in dairy cattle improvement programs has been of particular importance in the artificial insemination industry, allowing more accurate selection of young bulls and increasing the rate of genetic gain (García-Ruiz *et al.*, 2016). In Holstein cattle, during the last 7 years the inclusion of genomic information derived from GWAS has helped shorten generation interval in sires of bulls from 6.8 to ~2.5 years, boosted genetic gains for milk yield of from 50 to 109 kg per year, and for daughter pregnancy rate from negative values close to 0 to ~0.3 (García-Ruiz *et al.*, 2016).

Candidate gene approach

Another scheme to gene discovery is the use of candidate genes. A candidate gene is any gene thought to contain mutations responsible for a specific phenotype. Identification can be based on several approaches. The first is to search for genes located near genetic markers identified by GWAS. Kirkpatrick and Morris (2015) searched for candidate genes associated with ovulation rate in cattle. A GWAS was followed by Sanger sequencing of the target region in chromosome 10 which included *SMAD3*, *SMAD6* and *IQCH*. A total of 30 SNP in these genes were identified, and a haplotype comprising three SNP (two in *SMAD6* and one in *IQCH*) was associated with increased ovulation rate in daughters of bulls carrying the haplotype. After identification of a deficit of homozygotes for a JH1 haplotype associated with reduced fertility in Jersey using GWAS, sequencing performed in Jersey bulls revealed a nonsense mutation in *CWC15* which is embryonic lethal, as no homozygous individuals are present in the population (Sonstegard *et al.*, 2013). Another approach was presented by Moore *et al.* (2016), where 58 candidate genes for regulation of fertility were identified by searching for genetic variants in differentially expressed genes in the endometrium and corpus luteum of cows with good or poor genetic merit for fertility.

Alternatively, candidate genes can be identified by using existing knowledge of the biological pathways controlling a trait, and search for SNP in genes in those pathways. Work at University of Wisconsin from the Khatib group has focused on using the candidate gene approach to identify genes associated with embryonic development. In one study, SNP were identified in eight genes in the POU1F1 pathway: *POUF1F1*, *GH*, *GHR*, *PRL*, *OPN*, *PRLR*, *STAT5A*, and *UTMP* (Khatib *et al.*, 2009). There were significant associations for a SNP in *OPN* and *STAT5A* with fertilization rate, and for SNP in *GHR*, *STAT5A*, *PRLR* and *UTMP* with development of the embryo to the blastocyst stage. Likewise, Li *et al.* (2012), evaluated 25 genes of the TGFB signaling system. SNP were identified in *IBD3* associated with fertilization rate, and for a SNP in *BMP4* associated with development of the embryo to the blastocyst stage. Khatib *et al.* (2008a), studied the involvement of SNP in *FGF2* on embryonic survival because of the role of the FGF2 in regulation of *IFNT* expression in the trophectoderm (Michael *et al.*, 2006). One SNP in the intron of *FGF2* was identified that was significantly associated with development of the embryo to the blastocyst stage (Khatib *et al.*, 2008a). Likewise, and intronic SNP in *PGR* associated with fertilization rate and embryonic development to the blastocyst stage (Driver *et al.*, 2009).

Tests of association for candidate genes have relatively high statistical power since the number of independent statistical tests is lower than for GWAS (Amos *et al.*, 2011). Unlike GWAS, where genetic markers can change over time or between breeds because of crossover events during meiosis, the allelic association between a functional mutation and a genetically-controlled trait would be stable over time and more likely to extend across breeds. Furthermore, knowledge gained about the role of the gene in control of the trait could lead to improved understanding of the gene's functionality (Zhu and Zhao, 2007; Weller and Ron, 2011). There are limitations to the candidate gene approach. First, it is not easy to determine whether the association of a SNP in a candidate gene is causative or is in linkage disequilibrium with a nearby functional SNP. Increased confidence that a SNP is causative if the same genetic variants have similar effects in an independent population. The best way to verify the functionality of a candidate SNP is often impractical for livestock, namely the use gene editing technology to produce animals with the mutation and evaluate effect on the phenotype of interest. Another problem with the candidate gene approach is that it is most useful for identifying causative mutations in the coding region of genes. However, much genetic variation is located outside the coding region – in the regulatory region of the gene and at distantly located loci involved in epigenetic regulation.

Whole genome sequencing

Whole genome sequencing surveys the entire genetic code of an individual. The advantage of use whole genome sequencing is that it allows identification of complex forms of genetic variation besides SNP, including for example copy number variations. Moreover, by using whole sequencing the reliance on linkage disequilibrium disappears, as the causative mutation is on the generated data (Daetwyler *et al.*, 2014). Haplotypes affecting fertility in dairy breeds previously identified with SNP50 chip (VanRaden *et al.*, 2011), were further studied using whole genome sequence data by Fritz *et al.* (2013); and three novel mutations with damaged protein structure were identified in *GART*, *SHBG* and *SLC37A2* genes. Kadri *et al.* (2014), combining first SNP50 chip genotyping and whole genome sequencing identified a 660-kb deletion in chromosome 12 including four genes which is embryo-lethal in Nordic Red cattle. Using whole genome sequence data on 234 bulls, a mutation in

SMC2 was identified as causative for embryonic loss in cattle (Daetwyler *et al*., 2014). Given the rapidly decreasing cost of sequencing and the increase in number of animals in which whole genome sequences are available, it is likely that whole genome approaches to gene discovery are likely to predominate in the future.

From genotype to function: a fertility story

Identification of genetic variants associated with reproduction can provide clues to understand fertility regulation. Cochran *et al*. (2013a), used a candidate gene approach to identify genes associated with genetic variation in female fertility in Holstein bulls. Genes were identified by searching the literature for two kinds of genes. The first were genes well known to be involved in reproductive processes such as steroidogenesis, follicular development and embryonic development. The second kind, were genes differentially expressed between various physiological conditions in tissues involved in reproductive function. Examples include genes differentially expressed in the endometrium of lactating vs non-lactating cows, and genes differentially expressed between embryos produced *in vitro* compared to embryos produced *in vivo*. In each candidate gene, SNP where identified and only those present in the coding region or regulatory region where selected.

The final list of SNP for analysis included 422 novel candidate SNP (1 SNP per gene) and 12 SNP previously associated with fertility in the literature including *CAST* (Garcia *et al*., 2006), *FGF2* (Khatib *et al*., 2010), *FSHR* (Yang *et al*., 2010), *GHR* (Waters *et al*., 2011), *HSPA1L* (Rosenkrans Jr. *et al*., 2010), *ITGB5* (Feugang *et al*., 2009), *LEP* (Brickell *et al*., 2010), *NLRP9* (Ponsuksili *et al*., 2006), *PAPPA2* (Luna-Nevarez *et al*., 2011), *PGR* (Driver *et al*., 2009), *SERPINA14* (Khatib *et al*., 2007), and *STAT5A* (Khatib *et al*., 2008b). A population of 550 Holstein bulls with divergent genetic merit for DPR, where bulls of low DPR were those with a PTA of -2 or lower, and bulls of high DPR had a PTA of +1.7 or higher was used to test association of SNP with fertility traits (DPR, CCR, and HCR). Significant association were found for 40 SNP with DPR, 22 with HCR, and 33 with CCR. The function of the genes associated with fertility included steroid biosynthesis, genes regulated by estradiol and progesterone and immune function. In a second study, the same SNP were tested in 93 bulls for association with sperm fertilization ability, and subsequent *in vitro* embryonic development (Cochran *et al*., 2013b). There were SNP in 12 genes associated with the percent of cleaved embryos that became blastocysts. From the genes containing SNP associated with percent of cleaved embryos that became blastocyst, *C1QB*, *MON1B*, *PARM1*, *PCCB*, *PMM2*, and *TBC1D24* were associated with DPR, *C1QB* and *PARM1* were associated with HCR, and *C1QB*, *MON1B*, *PARM1*, *PMM2*, *SLC18A2*, *TBC1D24* were associated with CCR.

More recently, SNP with significant associations with fertility found by Cochran *et al*.

(2013a) were tested and validated in a separate population of Holstein cows with divergent genetic merit for fertility, cows were selected to have a high (≥1.5) or low PTA for DPR (≤-1.0). Of 51 genes previously associated with one or more estimates of fertility in bulls, 22 were associated with genotypic estimates of fertility in the cow population (Ortega *et al*., 2016a). In addition, SNP effects were associated with phenotypic measures of fertility, where animals carrying allelic variants associated with higher genetic merit for fertility also exhibited more favorable phenotypic measurements of fertility, having in general higher conception rates, fewer services per conception, and fewer days open (Ortega *et al*., 2017a). Thus, selection for those markers is likely to change actual reproductive performance. The list of SNP found associated with fertility in these studies can be found in Table 1.

There was a modest increase in reliability of genetic estimate for DPR (0.2%) when the SNP were included in the markers currently used for the national genetic evaluation system (Ortega *et al*., 2016a). This increase compares favorably to the 0.5% increase in reliability caused by adding up to 300,000 markers to the 50K bovine SNP chip (VanRaden *et al*., 2013). These findings indicate that the SNP under study here explain genetic variation not fully captured by GWAS and that the SNP are either causative or in higher linkage disequilibrium with the causal mutations than markers distributed across the genome.

The functions that were most represented by those genes containing SNP repeatedly associated with reproductive traits provides an indication of physiological processes important for variation among cows in reproductive function. There were 14 genes containing SNP associated with fertility that were regulated by estradiol and 6 by progesterone (Ortega *et al*., 2017a). Both steroids are essential for reproduction in mammals and there are compelling data indicating the importance of circulating concentrations of steroid hormones for cow fertility. Progesterone concentrations on days 4-7 after AI have been positively associated with pregnancy rate in Holstein heifers (Parr *et al*., 2012), and when follicular development occurs under low progesterone concentrations there is subsequent reduced fertility (Bisinotto *et al*., 2010). Circulating concentrations of steroids may be particularly important in high producing dairy cows, because steroid catabolism is increased and circulating concentrations of estradiol and progesterone are decreased (Wiltbank *et al*., 2006, 2014). It has been shown that cows with high genetic merit for fertility have larger corpora lutea and greater circulating concentrations of progesterone, and improved phenotypic fertility than cows with lower genetic merit for fertility (Cummins *et al*. 2012a, b; Moore *et al*., 2014).

The other function represented by genes with SNP associated with reproduction was immune function. Six genes associated with immune function were associated with genetic and phenotypic measures of fertility (Ortega *et al*., 2017a). Immune function is an important determinant of fertility. Cows that experience

926

Anim. Reprod., v.15, (Suppl.1), p.923-932. 2018

diseases postpartum have reduced reproductive function, are more likely to remain anovular, have decreased pregnancy rates and higher pregnancy losses than healthy cows (Santos *et al.*, 2011, 2016; Ribeiro *et al.*, 2016). In other studies, several of the genes differentially expressed in endometrium, liver, and muscle of Holstein cows with divergent genetic merit for fertility are involved in inflammatory processes (Moran *et al.*, 2015, 2016). There is also evidence that cows can be identified by their immune response (high or low immune responders) and this is associated with the risk of developing diseases including retained placenta and metritis, which directly impact reproductive function (Thompson-Crispi *et al.*, 2012).

Further research on the SNP in *COQ9* provided indirect evidence that function of the protein varied with genotype (Ortega *et al.*, 2017b). *COQ9* was subjected to additional study because the SNP in this gene explained 3% of genetic variation in DPR in Holstein cows (Ortega *et al.*, 2016a). COQ9 is involved in the biosynthesis of COQ10 (Tran and Clarke, 2007; Ben-Meir *et al.*, 2015), which is a critical component of the mitochondrial electron transport system and which is required for mitochondrial ATP synthesis. The missense mutation studied causes a change in the predicted protein structure and was associated with a change in oxidative phosphorylation as reflected in changes in mitochondrial respiratory function. The allele associated with improved fertility was also associated with lower substrate requirements to maintain basal cellular function and reduced proton leaks from the electron transport system. *COQ9* is expressed in reproductive tissues, and these alterations could affect the function of these tissues by improving energy utilization of the cells. Additionally, because of reduced proton leak, the SNP could affect production of reactive oxygen species (Murphy, 2009; Jastroch *et al.*, 2010). Further experimental work in the oocyte revealed that the variant associated with higher fertility was also associated with increased mitochondrial DNA copy number, which is associated with oocyte ATP production, successful oocyte maturation and fertilization (Reynier *et al.*, 2001; May-Panloup *et al.*, 2005; Tsai and St. John, 2016). Therefore, one of the reasons for differences in fertility among *COQ9* genotypes could reside in the allele associated with improved fertility, affects the competence of the oocyte due to higher mitochondrial content.

Another study was performed to understand the possible role of 12 genes containing SNP previously related to embryo competence to become a blastocyst by Cochran *et al.* (2013b). From the 12 genes, only two: *WBP1* and *PARM1* had increased expression at the moment of genome activation. Since the previous associations were based on the paternal SNP genotype, these were the genes most likely to represent actual effects of the SNP on embryonic development. Further

evaluation showed that the SNP in *WBP1* caused changes in predicted protein structure. By reducing transcript abundance of this gene using Gapmer antisense oligonucleotides, it was revealed that *WBP1* plays a critical in trophectoderm formation. WBP1 is a single transmembrane adaptor protein (Pei and Grishin, 2012) that functions to bind a variety of signaling proteins containing the WW1 or WW2 domains. Among these are the proteins KIBRA, SAV1, and YAP involved in the Hippo signaling pathway (Zhao *et al.*, 2010). Hippo signaling has been implicated in differentiation of the blastocyst in the mouse (Nishioka *et al.*, 2009; Lorthongpanich *et al.*, 2013). The transcription factor YAP interacts with TEAD4 to induce transcription of *CDX2* which in turn causes differentiation of the outer cells of the developing blastocyst into trophectoderm (Nishioka *et al.*, 2009). Perhaps the effects of the SNP in *WBP1* modify the interactions of WBP1 with proteins of the hippo signaling pathway.

Evidence was also provided that the SNP in the promoter region of *HSPA1L* improves thermotolerance in the embryo (Ortega *et al.*, 2016b). Previous work has associated this same mutation with increased calf crop in Brahman cattle (Rosenkrans Jr. *et al.*, 2010), and with increased transcription of *HSPA1A/HSPA1L* (primers do not distinguish between the genes) in cells when exposed to high temperatures (Basiricò *et al.*, 2011). Heat stress is known to affect fertility, particularly in dairy cattle, where cows in heat stress conditions show reduced pregnancy rates and pregnancies per AI (Gwazdauskas *et al.*, 1973; Hansen and Aréchiga, 1999; Flamenbaum and Galon, 2010). In this study, expression of *HSPA1A/HSPA1L* was high at the 2-cell stage in the bovine embryo, and when putative zygotes were exposed to heat shock or high oxygen conditions, those embryos inheriting the deletion mutation in *HSPA1L* had greater survival after being exposed to adverse conditions. Perhaps embryonic survival during heat stress could be improved by selecting for thermotolerant genotypes.

Taking all together, this series of studies demonstrated that identification of SNP in genes involved in reproductive processes can lead to discovery of additional markers associated with genetic variation in reproductive traits. Inclusion of these markers in current genomic evaluations also can increase reliability of genetic estimates for fertility. The fact that SNP effects were frequently repeated among two independent populations of animals and that phenotype as well as genotype was affected provides confidence that selection of these markers will improve genetic merit for fertility. As shown for the SNP in *COQ9*, the use of candidate genes can provide insights of the biology underpinning genetic variation in fertility, and that this understanding can lead to physiological interventions to improve reproductive function.

Table 1. SNP associated with fertility traits in more than one candidate gene study[1].

SNP id	Gene	Cow phenotype[2]			Cow genotype[2]			Bull genotype[3]		
		PR	SPC	DO	DPR	HCR	CCR	DPR	HCR	CCR
rs109967779	*ACAT2*				C		C	C		C
rs41766835	*APBB1*				G			G	G	G
rs133700190	*AP3B1*				T	T	T	T	T	T
rs109669573	*BCAS1*			C	C				C	
rs110217852	*BSP3*			A	A		A	A		
rs109332658	*C7H19orf60*				C		C	C		
rs135744058	*CACNA1D*					G		G	G	
rs137601357	*CAST*		T	T	T		T	T		T
rs109621328	*CD14*		C	C				C	C	
rs41711496	*CD40*					G	G	G		
rs133449166	*CSNK1E*				C	C		C	C	C
rs109137982	*FCER1G*	A	A	A	A					
rs43745234	*FSHR*	C							C	
rs41893756	*FUT1*		A	A	A		A	A		A
rs109262355	*FYB*		A	A					A	
rs109830880	*GCNT3*		T			T				
rs109711583	*HSD17B12*				G	G	G	G		
rs110828053	*HSD17B7*		C	C	C	C	C	C	C	C
rs110789098	*IBSP*	T				T	T			
rs111015912	*LDB3*		T					T	T	T
rs41256848	*LHCGR*		G		G					
rs134264563	*OCLN*		G	G	G		G	G		G
rs109813896	*PCCB*		C	C	C		C	C		
rs109629628	*PMM2*	G	G	G	G		G	G		G
rs133729105	*RABEP2*			G				G		G
rs110660625	*TBC1D24*	A	A	A	A			A		A

[1]Shown are genes containing SNP in which a significant association between the SNP and one or more reproductive traits was observed in at least two studies. The letter represents the allele associated with superior reproduction. SNP significant in more than one study but where different alleles were associated with superior reproduction are not included in the table; [2]Based on the population of 2273 Holstein cows. [3]Based on a population of 550 Holstein bulls from Cochran *et al.* (2013a). The table is reproduced from the Journal of Dairy Science (Ortega *et al.*, 2017a).

Concluding remarks

The introduction of genomic selection in dairy cattle has increased rates of genetic gain, particularly for low heritability traits such as fertility. The use of GWAS as a tool for genomic selection has been very successful in improving accuracy of genetic selection in dairy cattle. The pathway to choose for gene discovery will depend on several variants: available information of the phenotype or trait of interest, population size and overall goal of the work. Without previous knowledge of genes involved in the phenotype of interest, GWAS are a powerful tool to identify regions associated with the trait. This also could elucidate candidate genes for further study as GWAS by themselves are not designed primarily to illuminate the underlying biology of the studied phenotype. The use of candidate genes in turn, allow also to improve the SNP panels used for genetic evaluations, by finding markers with stronger associations with the traits of interest that can be included in genomic evaluation schemes. Furthermore, with the identification of candidate genes, functional studies involving gene editing or gene knockout modifications can be developed to understand the tight regulation of reproductive function in cattle. As genotyping cost decrease, more datasets and whole genome sequence data becomes available that can be used to validate markers in different populations; and in the case of sequencing, identifying causal mutations of the phenotypes of interest.

Acknowledgments

References

Aken BL, Achuthan P, Akanni W, Amode MR, Bernsdorff F, Bhai J, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Gil L, Girón CG, Gordon L, Hourlier T, Hunt SE, Janacek SH, Juettemann T, Keenan S, Laird MR, Lavidas I, Maurel T, McLaren W, Moore B, Murphy DN, Nag R, Newman V, Nuhn M, Ong CK, Parker A, Patricio M, Riat HS, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Wilder SP, Zadissa A, Kostadima M, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Cunningham F, Yates A,

Zerbino DR, Flicek P. 2017. Ensembl 2017. *Nucleic Acids Res*, 45:D635-D642.

Amos W, Driscoll E, Hoffman JI. 2011. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proc Biol Sci*, 278:1183-1188.

Basiricò L, Morera P, Primi V, Lacetera N, Nardone A, Bernabucci U. 2011. Cellular thermotolerance is associated with heat shock protein 70.1 genetic polymorphisms in Holstein lactating cows. *Cell Stress Chaperones*, 16:441-448.

Ben-Meir, A, Burstein E, Borrego-Alvarez A, Chong J, Wong E, Yavorska T, Naranian T, Chi M, Wang Y, Bentov Y, Alexis J, Meriano J, Sung H-K, Gasser DL, Moley KH, Hekimi S, Casper RF, Jurisicova A. 2015. Coenzyme Q10 restores oocyte mitochondrial function and fertility during reproductive aging. *Aging Cell*, 14:887-895.

Bisinotto RS, Chebel RC, Santos JEP. 2010. Follicular wave of the ovulatory follicle and not cyclic status influences fertility of dairy cows. *J Dairy Sci*, 93:3578-3587.

Boichard D, Manfredi E. 1994. Genetic analysis of conception rate in French Holstein cattle. *Acta Agric Scand Sect Anim Sci*, 44:138-145.

Brickell JS, Pollott GE, Clempson AM, Otter N, Wathes DC. 2010. Polymorphisms in the bovine leptin gene associated with perinatal mortality in Holstein-Friesian heifers. *J Dairy Sci*, 93:340-347.

Butler WR. 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. *Livest Prod Sci*, 83:211-218.

Cochran SD, Cole JB, Null DJ, Hansen PJ. 2013a. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genet*, 14:49. doi: 10.1186/1471-2156-14-49.

Cochran SD, Cole JB, Null DJ, Hansen PJ. 2013b. Single nucleotide polymorphisms in candidate genes associated with fertilizing ability of sperm and subsequent embryonic development in cattle. *Biol Reprod*, 89:69. doi: 10.1095/biolreprod.113.111260.

Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ Jr, Crooker BA, Van Tassell CP, Yang J, Wang S, Matukumalli LK, Da Y. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. *BMC Genomics*, 12:408. doi.org/10.1186/1471-2164-12-408.

Cummins SB, Lonergan P, Evans ACO, Berry DP, Evans RD, Butler ST. 2012a. Genetic merit for fertility traits in Holstein cows: I. Production characteristics and reproductive efficiency in a pasture-based system. *J Dairy Sci*, 95:1310-1322.

Cummins SB, Lonergan P, Evans ACO, Butler ST. 2012b. Genetic merit for fertility traits in Holstein cows: II. Ovarian follicular and corpus luteum dynamics, reproductive hormones, and estrus behavior. *J Dairy Sci*, 95:3698-3710.

Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, Liao X, Djari A, Rodriguez SC, Grohs C, Esquerré D, Bouchez O,

Rossignol M-N, Klopp C, Rocha D, Fritz S, Eggen A, Bowman PJ, Coote D, Chamberlain AJ, Anderson C, VanTassell CP, Hulsegge I, Goddard ME, Guldbrandtsen B, Lund MS, Veerkamp RF, Boichard DA, Fries R, Hayes BJ. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet*, 46:858-865.

Driver AM, Huang W, Gajic S, Monson RL, Rosa GJM, Khatib H. 2009. Short communication: effects of the progesterone receptor variants on fertility traits in cattle. *J Dairy Sci*, 92:4082-4085.

Feugang JM, Kaya A, Page GP, Chen L, Mehta T, Hirani K, Nazareth L, Topper E, Gibbs R, Memili E. 2009. Two-stage genome-wide association study identifies integrin beta 5 as having potential role in bull fertility. *BMC Genomics*, 10:176. doi: 10.1186/1471-2164-10-176.

Flamenbaum I, Galon N. 2010. Management of heat stress to improve fertility in dairy cows in Israel. *J Reprod Dev*, 56:S37-41.

Fortes MRS, DeAtley KL, Lehnert SA, Burns BM, Reverter A, Hawken RJ, Boe-Hansen G, Moore SS, Thomas MG. 2013. Genomic regions associated with fertility traits in male and female cattle: Advances from microsatellites to high-density chips and beyond. *Anim Reprod Sci*, 141:1-19.

Frąszczak M, Szyda J. 2016. Comparison of significant single nucleotide polymorphisms selections in GWAS for complex traits. *J Appl Genet*, 57:207-213.

Fritz, S Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, Grohs C, Weiss B, Boussaha M, Esquerré D, Klopp C, Rocha D, Boichard D. 2013. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS One*, 8:e65550. doi: 10.1371/journal.pone.0065550.

Garcia MD, Michal JJ, Gaskins CT, Reeves JJ, Ott TL, Liu Y, Jiang Z. 2006. Significant association of the calpastatin gene with fertility and longevity in dairy cattle. *Anim Genet*, 37:304-305.

García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Van Tassell CP. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc Natl Acad Sci USA*, 113:E3995-E4004.

Gwazdauskas FC, Thatcher WW, Wilcox CJ. 1973. Physiological, environmental, and hormonal factors at insemination which may affect conception. *J Dairy Sci*, 56:873-877.

Hansen PJ, Aréchiga CF. 1999. Strategies for managing reproduction in the heat-stressed dairy cow. *J Anim Sci*, 77:36-50.

Höglund JK, Sahana G, Guldbrandtsen B, Lund MS. 2014. Validation of associations for female fertility traits in Nordic Holstein, Nordic Red and Jersey dairy cattle. *BMC Genet*, 15:8. doi: 10.1186/1471-2156-15-8.

Ioannidis JPA Tarone R, McLaughlin JK. 2011. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*, 22:450-456.

Jastroch M, Divakaruni AS, Mookerjee S, Treberg JR, Brand MD. 2010. Mitochondrial proton and

electron leaks. *Essays Biochem*, 47:53-67.

Kadri NK, Sahana G, Charlier C, Iso-Touru T, Guldbrandtsen B, Karim L, Sander Nielsen U, Panitz F, Pedersen Aamand G, Schulman N, Georges M, Vilkki J, Sandø Lund M, Druet T. 2014. 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*, 10:e1004049. doi: 10.1371/journal.pgen.1004049.

Khatib H, Schutzkus V, Chang YM, Rosa GJM. 2007. Pattern of expression of the uterine milk protein gene and its association with productive life in dairy cattle. *J Dairy Sci*, 90:2427-2433.

Khatib H, Maltecca C, Monson RL, Schutzkus V, Wang X, Rutledge JJ. 2008a. The fibroblast growth factor 2 gene is associated with embryonic mortality in cattle. *J Anim Sci*, 86:2063-2067.

Khatib H, Monson RL, Schutzkus V, Kohl DM, Rosa GJM, Rutledge JJ. 2008b. Mutations in the *STAT5A* gene are associated with embryonic survival and milk composition in cattle. *J Dairy Sci*, 91:784-793.

Khatib H, Huang W, Wang X, Tran AH, Bindrim AB, Schutzkus V, Monson RL, Yandell BS. 2009. Single gene and gene interaction effects on fertilization and embryonic survival rates in cattle. *J Dairy Sci*, 92:2238-2247.

Khatib H, Monson RL, Huang W, Khatib R, Schutzkus V, Khateeb H, Parrish JJ. 2010. Short communication: validation of in vitro fertility genes in a Holstein bull population. *J Dairy Sci*, 93:2244-2249.

Khatkar MS, Randhawa IAS, Raadsma HW. 2014. Meta-assembly of genomic regions and variants associated with female reproductive efficiency in cattle. *Livest Sci*, 166:144-157.

Kirkpatrick BW, Morris CA. 2015. A major gene for bovine ovulation rate. *PLoS One*, 10:e0129025. doi: 10.1371/journal.pone.0129025.

Kuhn MT, Hutchison JL, Wiggans GR. 2006. Characterization of Holstein heifer fertility in the United States. *J Dairy Sci*, 89:4907-4920.

Li G, Khateeb K, Schaeffer E, Zhang B, Khatib H. 2012. Genes of the transforming growth factor-beta signalling pathway are associated with pre-implantation embryonic development in cattle. *J Dairy Res*, 79:310-317.

Littlejohn, M, Grala T, Sanders K, Walker C, Waghorn G, Macdonald K, Spelman R, Davis S, Snell R. 2012. Non-replication of genome-wide-based associations of efficient food conversion in dairy cows. *Anim Genet*, 43:781-784.

Liu A, Wang Y, Sahana G, Zhang Q, Liu L, Lund MS, Su G. 2017. Genome-wide association studies for female fertility traits in Chinese and Nordic Holsteins. *Sci Rep*, 7:8487. doi: 10.1038/s41598-017-09170-9.

Lorthongpanich C, Messerschmidt DM, Chan SW, Hong W, Knowles BB, Solter D. 2013. Temporal reduction of LATS kinases in the early preimplantation embryo prevents ICM lineage differentiation. *Genes Dev*, 27:1441-1446.

Luna-Nevarez, P, Rincon G, Medrano JF, Riley DG, Chase CC, Coleman SW, VanLeeuwen DM, DeAtley KL, Islas-Trejo A, Silver GA, Thomas MG. 2011. Single nucleotide polymorphisms in the growth hormone-insulin-like growth factor axis in straightbred and crossbred Angus, Brahman, and Romosinuano heifers: population genetic analyses and association of genotypes with reproductive phenotypes. *J Anim Sci*, 89:926-934.

May-Panloup P, Chrétien MF, Jacques C, Vasseur C, Malthièry Y, Reynier P. 2005. Low oocyte mitochondrial DNA content in ovarian insufficiency. *Hum Reprod*, 20:593-597.

Michael DD, Alvarez IM, Ocón OM, Powell AM, Talbot NC, Johnson SE, Ealy AD. 2006. Fibroblast growth factor-2 is expressed by the bovine uterus and stimulates interferon-tau production in bovine trophectoderm. *Endocrinology*, 147:3571-3579.

Moore SG, Scully S, Browne JA, Fair T, Butler ST. 2014. Genetic merit for fertility traits in Holstein cows: V. Factors affecting circulating progesterone concentrations. *J Dairy Sci*, 97:5543-5557.

Moore SG, Pryce JE, Hayes BJ, Chamberlain AJ, Kemper KE, Berry DP, McCabe M, Cormican P, Lonergan P, Fair T, Butler ST. 2016. Differentially expressed genes in endometrium and corpus luteum of Holstein cows selected for high and low fertility are enriched for sequence variants associated with fertility. *Biol Reprod*, 94:19: 1-11.

Moran B, Butler ST, Moore SG, MacHugh DE, Creevey CJ. 2015. Differential gene expression in the endometrium reveals cytoskeletal and immunological genes in lactating dairy cows genetically divergent for fertility traits. *Reprod Fertil Dev*, 29:274-282.

Moran B, Cummins SB, Creevey CJ, Butler ST. 2016. Transcriptomics of liver and muscle in Holstein cows genetically divergent for fertility highlight differences in nutrient partitioning and inflammation processes. *BMC Genomics*, 17:603. doi:10.1186/s12864-016-2938-1.

Murphy MP. 2009. How mitochondria produce reactive oxygen species. *Biochem J*, 417:1-13.

Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, Yabuta N, Hirahara S, Stephenson R.O, Ogonuki N, Makita R, Kurihara H, Morin-Kensicki E.M, Nojima H, Rossant J, Nakao K, Niwa H, Sasaki H. 2009. The hippo signaling pathway components lats and yap pattern tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell*, 16:398-410.

Norman HD, Wright JR, Hutchison JL, Mattison JM. 2014. Selection changes in the United States due to genomics. *Interbull Bull*, 48:37-39.

Ortega MS, Denicol AC, Cole JB, Null DJ, Hansen PJ. 2016a. Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. *Anim Genet*, 47:288-297.

Ortega MS, Rocha-Frigoni NAS, Mingoti GZ, Roth Z, Hansen PJ. 2016b. Modification of embryonic resistance to heat shock in cattle by melatonin and genetic variation in HSPA1L. *J Dairy Sci*, 99:9152-9164.

Ortega MS, Denicol AC, Cole JB, Null DJ, Taylor

JF, Schnabel RD, Hansen PJ. 2017a. Association of single nucleotide polymorphisms in candidate genes previously related to genetic variation in fertility with phenotypic measurements of reproductive function in Holstein cows. *J Dairy Sci*, 100:3725-3734.

Ortega MS, Wohlgemuth S, Tribulo P, Siqueira LGB, Null DJ, Cole JB, Silva D, Hansen PJ. 2017b. A single nucleotide polymorphism in COQ9 affects mitochondrial and ovarian function and fertility in Holstein cows. *Biol Reprod*, 96:652-663.

Parr MH, Mullen MP, Crowe MA, Roche JF, Lonergan P, Evans ACO, Diskin MG. 2012. Relationship between pregnancy per artificial insemination and early luteal concentrations of progesterone and establishment of repeatability estimates for these traits in Holstein-Friesian heifers. *J Dairy Sci*, 95:2390-2396.

Pei J, Grishin NV. 2012. Unexpected diversity in Shisa-like proteins suggests the importance of their roles as transmembrane adaptors. *Cell Signal*, 24:758-769.

Petersson K-J, Strandberg E, Gustafsson H, Royal MD, Berglund B. 2008. Detection of delayed cyclicity in dairy cows based on progesterone content in monthly milk samples. *Prev Vet Med*, 86:153-163.

Ponsuksili S, Brunner RM, Goldammer T, Kühn C, Walz C, Chomdej S, Tesfaye D, Schellander K, Wimmers K, Schwerin M. 2006. Bovine *NALP5*, *NALP8*, and *NALP9* genes: assignment to a QTL region and the expression in adult tissues, oocytes, and preimplantation embryos. *Biol Reprod*, 74:577-584.

Pritchard T, Coffey M, Mrode R, Wall E. 2013. Genetic parameters for production, health, fertility and longevity traits in dairy cows. *Animal*, 7:34-46.

Pryce JE, Royal MD, Garnsworthy PC, Mao IL. 2004. Fertility in the high-producing dairy cow. *Livest Prod Sci*, 86:125-135.

Reynier P, May-Panloup P, Chrétien M-F, Morgan CJ, Jean M, Savagner F, Barrière P, Malthièry Y. 2001. Mitochondrial DNA content affects the fertilizability of human oocytes. *Mol Hum Reprod*, 7:425-429.

Riancho JA. 2012. Genome-wide association studies (GWAS) in complex diseases: advantages and limitations. *Reumatol Clin*, 8:56-57.

Ribeiro ES, Gomes G, Greco LF, Cerri RLA, Vieira-Neto A, Monteiro Jr PLJ, Lima FS, Bisinotto RS, Thatcher WW, Santos JEP. 2016. Carryover effect of postpartum inflammatory diseases on developmental biology and fertility in lactating dairy cows. *J Dairy Sci*, 99:2201-2220.

Rosenkrans Jr C, Banks A, Reiter S, Looper M. 2010. Calving traits of crossbred Brahman cows are associated with Heat Shock Protein 70 genetic polymorphisms. *Anim, Reprod Sci*, 119:178-182.

Royal M, Mann GE, Flint AP. 2000. Strategies for reversing the trend towards subfertility in dairy cattle. *Vet J*, 160:53-60.

Santos JEP, Bisinotto RS, Ribeiro ES, Lima FS, Greco LF, Staples CR, Thatcher WW. 2011. Applying nutrition and physiology to improve reproduction in dairy cattle. *Reprod Domest Rumin*,

67:387-403.

Santos JEP, Bisinotto RS, Ribeiro ES. 2016. Mechanisms underlying reduced fertility in anovular dairy cows. *Theriogenology*, 86:254-262.

Sonstegard, TS, Cole JB, VanRaden PM, Tassell CPV, Null DJ, Schroeder SG, Bickhart D, McClure MC. 2013. Identification of a nonsense mutation in *CWC15* associated with decreased reproductive efficiency in Jersey cattle. *PLoS One*, 8:e54872. doi: 10.1371/journal.pone.0054872.

Stringer S, Wray NR, Kahn RS, Derks EM. 2011. Underestimated effect sizes in gwas: fundamental limitations of single snp analysis for dichotomous phenotypes. *PLoS One*, 6:e27964. doi: 10.1371/journal.pone.0027964.

Thompson-Crispi KA, Hine B, Quinton M, Miglior F, Mallard BA. 2012. Short communication: association of disease incidence and adaptive immune response in Holstein dairy cows. *J Dairy Sci*, 95:3888-3893.

Tran UC, Clarke CF. 2007. Endogenous synthesis of coenzyme Q in eukaryotes. *Mitochondrion*, 7(suppl):S62-S71.

Tsai T, St John JC. 2016. The role of mitochondrial DNA copy number, variants, and haplotypes in farm animal developmental outcome. *Domest Anim Endocrinol*, 56(suppl):S133-S146.

VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD. 2003. Daughter pregnancy rate evaluation. Available on: http://aipl.arsusda.gov/reference/fertility/DPR_rpt.htm. Accessed on: November 3, 2016.

VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD, Kuhn MT, Wiggans GR. 2004. Development of a national genetic evaluation for cow fertility. *J Dairy Sci*, 87:2285-2292.

VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS. 2008. Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci*, 92:16-24.

VanRaden PM, Olson KM, Null DJ, Hutchison JL. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J Dairy Sci*, 94:6153-6161.

VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, Sonstegard TS, Connor EE, Winters M, van Kaam JBCHM, Valentini A, Van Doormaal BJ, Faust MA, Doak GA. 2013. Genomic imputation and evaluation using high-density Holstein genotypes. *J Dairy Sci*, 96:668-678.

Waters SM, McCabe MS, Howard DJ, Giblin L, Magee DA, MacHugh DE, Berry DP. 2011. Associations between newly discovered polymorphisms in the *Bos Taurus* growth hormone receptor gene and performance traits in Holstein–Friesian dairy cattle. *Anim Genet*, 42:39-49.

Weller JI, Ron M. 2011. Invited review: quantitative trait nucleotide determination in the era of genomic selection. *J Dairy Sci*, 94:1082-1090.

Wiggans GR, VanRaden PM, Cooper TA. 2011. The genomic evaluation system in the United States: past,

present, future. *J Dairy Sci*, 94:3202-3211.

Wiltbank M, Lopez H, Sartori R, Sangsritavong S, Gümen A. 2006. Changes in reproductive physiology of lactating dairy cows due to elevated steroid metabolism. *Theriogenology*, 65:17-29.

Wiltbank MC, Souza AH, Carvalho PD, Cunha AP, Giordano JO, Fricke PM, Baez GM, MG Diskin. 2014. Physiological and practical effects of progesterone on reproduction in dairy cattle. *Animal*, 8:70-81.

Yang W-C, Li S-J, Tang K-Q, Hua G-H, Zhang C-Y, Yu J-N, Han L, Yang L-G. 2010. Polymorphisms in the 5' upstream region of the FSH receptor gene, and their association with superovulation traits in Chinese Holstein cows. *Anim Reprod Sci*, 119:172-177.

Zhao B, Li L, Lei Q, Guan K-L. 2010. The Hippo–YAP pathway in organ size control and tumorigenesis: an updated version. *Genes Dev*, 24:862-874.

Zhu M, Zhao S. 2007. Candidate gene identification approach: progress and challenges. *Int J Biol Sci*, 3:420-427.

Zondervan KT, Rahmioglu N, Morris AP, Nyholt DR, Montgomery GW, Becker CM, Missmer SA. 2016. Beyond endometriosis genome-wide association study: from genomics to phenomics to the patient. *Semin Reprod Med*, 34:242-254.

932

Anim. Reprod., v.15, (Suppl.1), p.923-932. 2018