



Identification of genomic regions associated with total and progressive sperm motility in Italian Holstein bulls

J. Ramirez-Diaz,^{1,2*} S. Cenadelli,³ V. Bornaghi,³ G. Bongioni,³ S. M. Montedoro,³ A. Achilli,⁴ C. Capelli,⁵
J. C. Rincon,⁶ M. Milanese,⁷ M. M. Passamonti,¹ L. Colli,¹ M. Barbato,¹ J. L. Williams,¹ and P. Ajmone Marsan¹

¹Department of Animal Sciences, Food and Nutrition (DIANA), Università Cattolica del Sacro Cuore, Piacenza, Italy 29122

²Institute of Agricultural Biology and Biotechnology (IBBA), Consiglio Nazionale di Ricerca, Milano, Italy

³Institute Lazzaro Spallanzani, Rivolta d'Adda (CR), Cremona, Italy

⁴Department of Biology and Biotechnology, Università degli Studi di Pavia, Pavia, Italy

⁵Department of Chemical, Life and Environmental Sustainability Sciences, Università degli Studi di Parma, Parma, Italy

⁶Department of Animal Science, Universidad Nacional de Colombia, Palmira, Valle del Cauca, Colombia

⁷Department for Innovation in Biological, Agri-food and Forestry Systems (DIBAF), Università degli Studi della Tuscia, Viterbo, Italy

ABSTRACT

Sperm motility is directly related to the ability of sperm to move through the female reproductive tract to reach the ovum. Sperm motility is a complex trait that is influenced by environmental and genetic factors and is associated with male fertility, oocyte penetration rate, and reproductive success of cattle. In this study we carried out a GWAS in Italian Holstein bulls to identify candidate regions and genes associated with variations in progressive and total motility (PM and TM, respectively). After quality control, the final data set consisted of 5,960 records from 949 bulls having semen collected in 10 artificial insemination stations and genotyped at 412,737 SNPs (call rate >95%; minor allele frequency >5%). (Co)variance components were estimated using single trait mixed models, and associations between SNPs and phenotypes were assessed using a genomic BLUP approach. Ten windows that explained the greatest percentage of genetic variance were located on *Bos taurus* autosomes 1, 2, 4, 6, 7, 23, and 26 for TM and *Bos taurus* autosomes 1, 2, 4, 6, 8, 16, 23, and 26 for PM. A total of 150 genes for TM and 72 genes for PM were identified within these genomic regions. Gene Ontology enrichment analyses identified significant Gene Ontology terms involved with energy homeostasis, membrane functions, sperm-egg interactions, protection against oxidative stress, olfactory receptors, and immune system. There was significant enrichment of quantitative trait loci for fertility, calving ease, immune response, feed intake, and carcass weight within the candidate windows. These results contribute

to understanding the architecture of the genetic control of sperm motility and may aid in the development of strategies to identify subfertile bulls and improve reproductive success.

Key words: candidate genes, dairy bulls, genome-wide association, semen quality

INTRODUCTION

In 2020, the world market for AI was valued ~\$3.96 billion with an expected annual growth of ~5.94% by 2028. Among other factors, the quantity, quality, and sanitary status of the semen will determine the reproductive success of a bull. Although fertility of semen is highly influenced by age, season, nutrition, management, and the interval between ejaculations (Fuerst-Waltl et al., 2006; Suchocki and Szyda, 2015), it is also known that genetics and biological processes involved in sperm production and sperm-ovum interactions play a significant role (Druet et al., 2009; Yin et al., 2019b; Sweatt et al., 2020).

Sperm motility (SM) affects the ability of the sperm to reach and penetrate the zona pellucida, fertilize the ovum, and initiate embryogenesis (Alves et al., 2020). Sperm motility is routinely measured by insemination centers using computer-assisted sperm analysis. However, it has been shown that SM is a complex phenotype which is affected by the environmental (Druet et al., 2009) and both additive and nonadditive genetic effects (VanRaden et al., 2011; Peñagaricano et al., 2012). Consequently, a better understanding of the genomic architecture of semen-related complex traits is expected to help improve fertility through the assisted selective breeding of bulls. The heritability estimates for SM vary widely among studies on dairy and beef cattle, from very low values of 0.01–0.13 (Gredler et al., 2007; Corbet et al., 2013; Suchocki and Szyda, 2015) to

Received December 13, 2021.

Accepted August 10, 2022.

*Corresponding author: johanna.ramirezdz@ibba.cnr.it

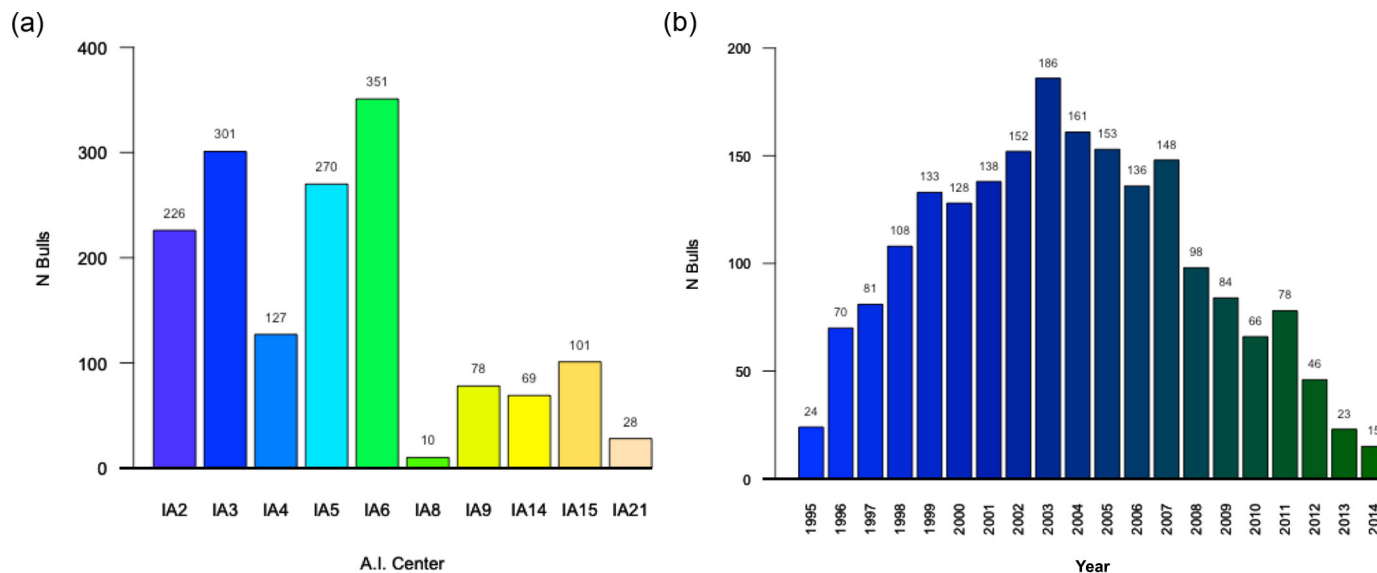


Figure 1. (a) Number of bulls collected at each AI center identified by anonymous codes and (b) number of bulls sampled per year of birth.

moderate and high values of 0.43–0.60 (Hering et al., 2014).

A large number of genes, some having pleiotropic effects on different semen traits and cell types, are involved in the production, maturation, and sperm viability and can affect semen quality and fertility (Ballou et al., 2006; Zhang et al., 2016). Here, we identify novel genomic regions and candidate genes putatively involved in total and progressive motility (**PM**) in Italian Holstein bulls using a GWAS approach.

MATERIALS AND METHODS

Phenotypic and Genotypic Data

We used 5,960 records of total and progressive SM from 949 Italian Holstein bulls with ages between 1 and 11 yr from 10 AI centers (Figure 1a), collected between 1995 and 2014 (Figure 1b). The number of measurements per sire varied between 3 and 78 for both total motility and PM traits. The procedures for semen collection are standardized across AI centers. Frozen-thawed sperm are routinely analyzed in the collecting centers using a CASA Integrated Visual Optical System (IVOS; CASA System, Hamilton Thorne Inc.) and a standardized protocol to assess SM. Total motility (**TM**) is the percentage of sperm moving irrespective of direction, and PM is the percentage of sperm that exhibit linear movement. Sperm phenotypic data were retrieved from the database of the Istituto Lazzari Spallanzani, in charge of official sperm quality controls of national bulls in Italy. The DNA extracted from se-

men was already available at the Biobank of Università Cattolica (see Mancini et al., 2014).

Illumina BovineSNP50 genotypes for 636 animals were imputed to high-density (777,962) with Beagle 4.1 (Browning and Browning, 2016) using a reference set of 1,006 animals, genotyped with the Illumina BovineHD array. Both data sets were previously updated to the ARS-UCD1.2 (Rosen et al., 2020) bovine assembly version using Plink 1.9 (Chang et al., 2015). Quality control excluded indels, SNPs with minor allele frequency lower than 0.05 (Wiggans et al., 2009), and duplicated markers from both data sets. For the BovineSNP50 data set, only SNPs in common with HD data set were used. Finally, the filtered reference data set, comprising 569,630 SNPs, was phased and missing genotypes imputed through Beagle 5.2 (Browning et al., 2021). Low quality imputed variants, with DR2 (dosage R^2 , which is the estimated squared correlation between the estimated allele dose and the true allele dose), lower than 0.3 were excluded (Wu et al., 2019; Chen et al., 2021), giving a final imputed data set of 417,256 SNPs. As 91 animals were genotyped both with 50k and HD panels, these were removed from the imputed data set and their HD data were considered for the downstream analysis. Therefore, the GWAS analysis was performed using 949 animals with phenotype and genotype information (404 animals genotyped with BovineHD and 545 animals imputed from BovineSNP50). A further quality control was carried out to exclude SNPs with minor allele frequency <0.05 . After quality control, 412,737 SNPs and 949 animals were used for further analyses.

Genome-Wide Association Study

The genomic BLUP using individuals with repeated measurements were analyzed with a repeatability single trait animal model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_u\mathbf{u} + \mathbf{W}_{pe}\mathbf{pe} + \mathbf{e},$$

where \mathbf{y} is a vector of observed phenotypes (TM or PM) for all sires, $\boldsymbol{\beta}$ is the vector of fixed effects (AI center, contemporary group: year-season, quadratic, and linear effects of age), \mathbf{u} is the random vector of additive genetic effects, \mathbf{pe} is the random vector of permanent environmental effects, and \mathbf{e} is the random vector of residual effects. The matrices \mathbf{X} , \mathbf{Z} , and \mathbf{W} are the incidence matrices of fixed effects, additive genetic effects, and permanent environmental effects, respectively. Random effects were assumed to follow a multivariate normal distribution, as follows:

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{pe} | \sigma_u^2, \sigma_{pe}^2, \sigma_e^2 \\ \mathbf{e} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \mathbf{G}\sigma_u^2 & 0 & 0 \\ 0 & \mathbf{I}\sigma_{pe}^2 & 0 \\ 0 & 0 & \mathbf{I}\sigma_e^2 \end{pmatrix} \right],$$

where σ_u^2 , σ_{pe}^2 , and σ_e^2 are the additive genetic variance, the permanent environment variance, and the residual variance, respectively; \mathbf{I} is an identity matrix. \mathbf{G} is a genomic relationship matrix derived from marker information (VanRaden, 2008). The heritability was estimated as the ratio of additive genetic variance to the total phenotypic variance ($\sigma_u^2 + \sigma_{pe}^2 + \sigma_e^2$) and the repeatability was calculated as sum of individual variance ($\sigma_u^2 + \sigma_{pe}^2$) divided by phenotypic variance.

The percentage of additive genetic variance explained by the i th SNP window was then calculated as

$$\frac{\text{Var}(u_i)}{\sigma_u^2} \times 100 = \frac{\text{Var}\left(\sum_{j=1}^x \mathbf{Z}_j u_j\right)}{\sigma_u^2} \times 100,$$

where u_i is the genetic value of the i th genomic region under consideration, σ_u^2 is the total additive genetic variance, x is the total number of adjacent SNPs within the 1 Mb region, \mathbf{Z}_j is the vector of gene content of the j th SNP for all windows, and u_j is the effect of the j th SNP within i th window. GWAS results are reported as the percentage of genetic additive variance explained by windows of 1 Mb and plotted using CMplot package in R v.1.3.1073 (R core team 2020). Additionally, we calculated the P -value of individual SNP according to

Aguilar et al. (2019). All procedures were performed using BLUPF90 family programs (Miszta et al., 2014). To control false positives, we applied the false discovery rate (FDR) method for multiple testing according to Benjamini and Hochberg (1995) as follows:

$$FDR = m \times Pmax/n,$$

where m is the number of times to be tested, $Pmax$ is the genome-wide significance level empirical P -value of FDR adjusted, and n is the number of significant SNPs at the assigned FDR level (i.e., 0.05).

Gene Annotation and Functional Enrichment

Our results indicated that single SNPs with the smallest P -values did not explain a large percentage of genetic variance. Therefore, we used the percentage of variance explained by SNPs within consecutive 1 Mb windows to identify genomic regions of interest for further investigation. Genes within the 10 windows explaining the largest amount of variance were annotated using the GALLO v.0.99 R package (Fonseca et al., 2020) and the ENSEMBL annotation of the *Bos taurus* ARS-UCD 1.2 genome assembly (<http://bovinegenome.org>). The genes identified from ENSEMBL were collected to perform a functional annotation by DAVID v6.8 (<https://david.ncifcrf.gov/home.jsp>) and REVIGO (Supek et al., 2011) using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (<https://www.genome.jp/kegg>) databases for *B. taurus*. The annotation and enrichment analyses of QTL in adjacent windows (1 Mb) were assessed for the effects of neighboring SNPs, which can be captured due to linkage disequilibrium. The QTL enrichment was evaluated for all the annotated QTL and the P -values were calculated and corrected for multiple testing, using a value of FDR < 0.05 to identify the enriched QTL. The annotation and enrichment analyses were performed using the GALLO v.0.99. package and the Animal QTLdb for the ARS-UCD 1.2 bovine genome (<http://bovinegenome.org>).

RESULTS

Descriptive Statistics and Genetic Parameters

The bulls had a mean of 55.33% (12.70 SD) for TM and 43.21% (11.04 SD) for PM. The fixed effect of the contemporary group and the quadratic and linear effect of age (covariable) were highly significant in the model for both traits ($P < 0.001$). The effect of the AI center was significant ($P < 0.01$) and the estimates of herita-

Table 1. Additive genetic ($\hat{\sigma}_a^2$), permanent environmental ($\hat{\sigma}_{ep}^2$), and residual ($\hat{\sigma}_e^2$) variance and genetic parameters heritability (\hat{h}^2), genetic correlation (\hat{r}_g^2), and repeatability (\pm SE) for total motility (TM) and progressive motility (PM) in Italian Holsteins bulls¹

Trait	$\hat{\sigma}_a^2$	$\hat{\sigma}_{ep}^2$	$\hat{\sigma}_e^2$	\hat{h}^2	\hat{r}_g^2	r
TM	14.96 \pm 2.02	20.85 \pm 2.27	78.80 \pm 1.13	0.13 \pm 0.0062	0.62 \pm 0.010	0.31 \pm 0.0027
PM	13.05 \pm 2.28	24.36 \pm 2.51	82.43 \pm 1.17	0.10 \pm 0.0045		0.33 \pm 0.0019

¹Repeatability: $r = \sigma_a^2 + \sigma_{ep}^2 / \sigma_a^2 + \sigma_{ep}^2 + \sigma_e^2$.

bility were low for both TM (0.13 \pm 0.0062) and PM (0.10 \pm 0.0045). The genetic correlation between TM and PM was high (0.62 \pm 0.010) and the repeatability was similar for the 2 traits: 0.31 and 0.33 for TM and PM, respectively (Table 1).

QTL Regions and Candidate Genes

As previously noted, our results indicated that the SNPs with the smallest P -values did not explain a large percentage of genetic variance (Figure 2–3), contrary to expectation. P -values are influenced by sample size (Greenland et al., 2016) and, in particular, by allele frequencies (Aguilar et al., 2019).

The 10 windows that explained the greatest percentage of genetic variance were located on BTA 1, 2, 4, 6, 7, 23, and 26 for TM and BTA 1, 2, 4, 6, 8, 16, 23, and 26 for PM (Figure 4). These regions explained 6.91% and 6.07% of the additive genetic variance for TM and PM, respectively (Table 2). A total of 150 genes for TM and 107 genes for PM were annotated by GALLO within these significant windows (Table 2). Among them, 130 genes for TM and 90 genes for PM were found in the *Bos taurus* DAVID database.

The DAVID annotations and their relative P -values were grouped using REVIGO into biological processes (BP), cellular components (CC), and molecular functions (MF), each comprising 13, 23, and 45 genes for TM and 24, 18, and 61 for PM (Table 3). A total of 42 genes are in common between TM and PM.

The BoLA and *JSP.1* genes on BTA 23 were associated with both PM and TM, whereas the OR (olfactory receptor) family genes were only related to PM. The glutathione S-transferase alpha genes, involved with the metabolism of glutathione, which is an antioxidant, were only associated with TM.

Some QTL within the genomic windows explaining the greatest percentage of genetic variance were associated with reproductive traits: these included BTA 1 and 6 for TM, and BTA 1 and 23 for PM. The enrichment analysis of the QTL (FDR-corrected $P < 0.05$) for TM identified 28 QTL related to the ease of calving and 6 QTL associated with the immune response on BTA 6 and BTA 23. In addition, QTL involved with feed intake (7), carcass weight (24), and fat thickness (10) were associated with TM on BTA 6. For PM, 2 QTL related to reproductive efficiency and with milk yield were identified on BTA 23 (Table 4).

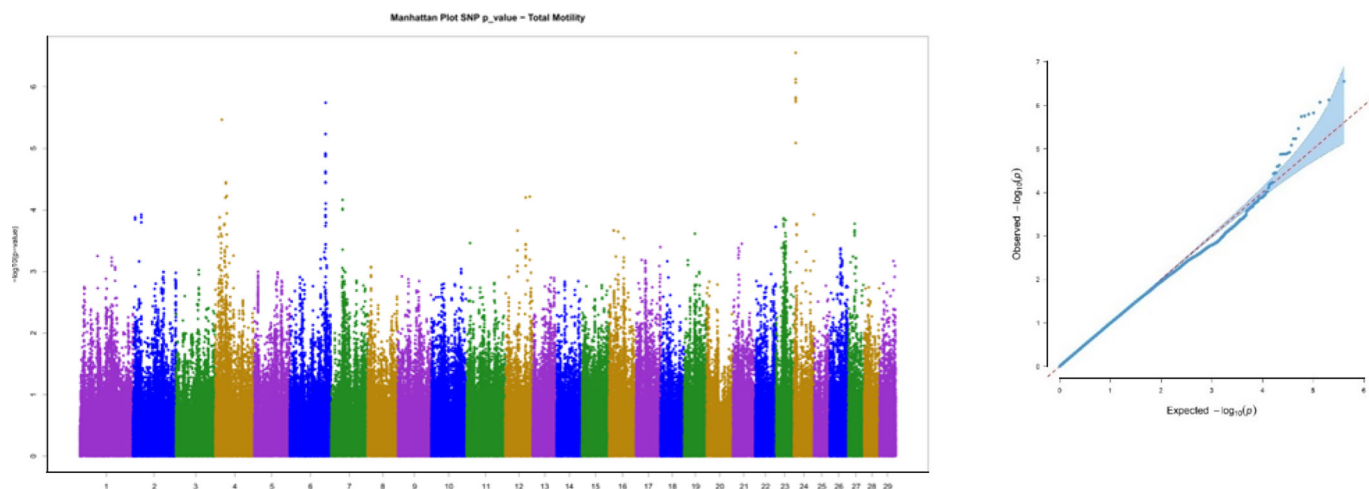


Figure 2. Manhattan plot displaying the association of all SNP for progressive motility (PM; P -values of individual SNP effects) and quantile-quantile plots (right), which show the observed distribution of P -values against the expected P -values under the null hypothesis of no association. The horizontal axis shows $-\log_{10}$ transformed expected P -values and the vertical axis indicates $-\log_{10}$ transformed observed P -values.

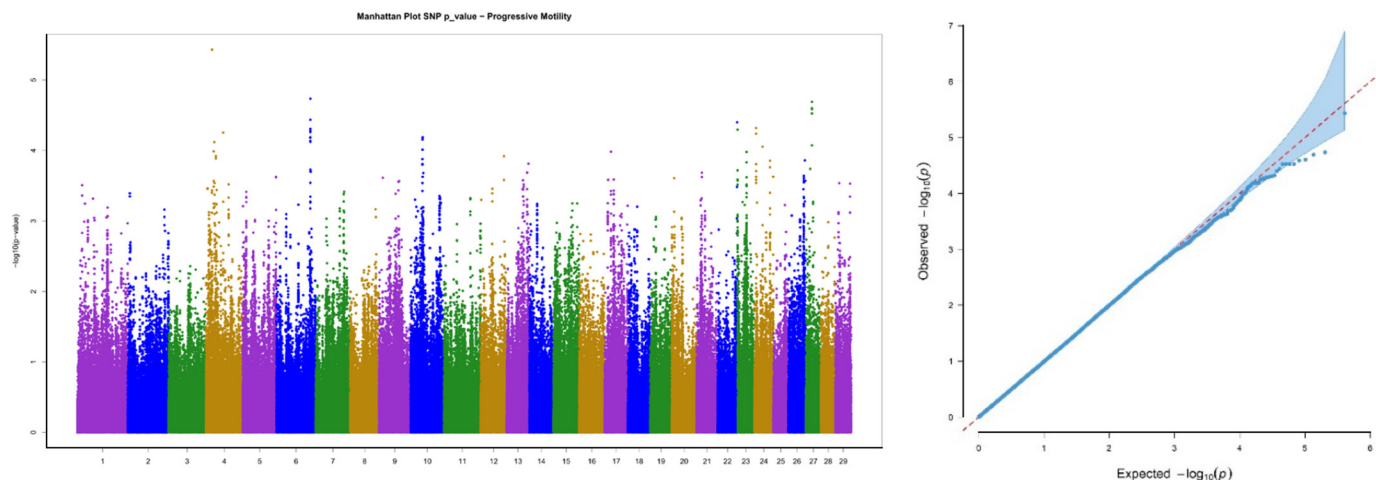


Figure 3. Manhattan plot displaying the association of all SNP for progressive motility (PM; P -values of individual SNP effects) and quantile-quantile plots (right), which show the observed distribution of P -values against the expected P -values under the null hypothesis of no association. The horizontal axis shows $-\log_{10}$ transformed expected P -values and the vertical axis indicates $-\log_{10}$ transformed observed P -values.

DISCUSSION

Genetic Parameters

The heritability for TM and PM estimated in this study was low (TM: 0.14 ± 0.0062 and PM: 0.10 ± 0.0045) but comparable to estimates obtained by others using the pedigree relationship matrix A (0.13, Druet et al., 2009; 0.12, Yin et al., 2019a) for PM. High heritability estimates, of up to 0.43 have been reported for motility in Holstein bulls (Al-Kanaan et al., 2015; Butler et al., 2020); however, estimates for PM in this breed are low (0.13–0.15, Butler et al., 2020). The differences in heritability estimates may be the result of environmental and management differences that we found to have a significant effect on these traits. In our study estimated repeatability was moderate (0.31 and 0.33 for TM and PM, respectively) which suggests an effect of the permanent environment and an even higher effect of the residual variance.

Candidate Genes and QTL Annotation

In the present study, no loci with significant P -values were found using a single SNP analysis. The SNPs with the smallest P -values would be expected to explain a relatively higher proportion of genetic variation. However, this is not always the case, as P -values are affected by sample size, the magnitude of the effect detected (Greenland et al., 2016), and in particular, by allele frequencies (Aguilar et al., 2019). Liu et al. (2021), in a single-step GWAS in swine, found no overlap between the first 20 SNPs with lowest P -values and those explaining the largest variance. These authors attribute

this discrepancy to linkage disequilibrium between SNPs and the causal variants and interaction between SNPs. Therefore, we used the variance explained by sliding window intervals to identify genomic regions involved in sperm motility.

The ten 1-Mb windows with highest additive genetic variance for TM and PM explained, respectively, 6.91% and 6.07% of the additive genetic variance, indicating the highly polygenic nature of these traits. Other studies have also found that several regions across the genome each make small contributions to the total genetic variation for SM. For example, Yin et al. (2019b) found that the most significant regions, located on BTA 1, 3, 4, 5, 11, 17, and 19, explained less than 0.1% of the genetic variance for SM in Holstein bulls. In beef cattle, Sweett et al. (2020) identified 10 windows on BTA 9, 13, 20, and 24 that explained 7.17% of the genetic variance of this trait. There is little consistency in the regions identified across studies, even for signals on the same chromosomes; Rezende et al. (2018) identified genes associated with SM on BTA 5, 11, 22, and 25 in Jersey cattle, with no regions shared with other studies.

The window with highest additive genetic variance we identified was on BTA1 (102,482,169–103,484,977), which explained 1.29% of the genetic variance for TM and 1.62% for PM. This window included the gene *E1BHH0*, which codes for profilin. Profilins are ubiquitous proteins present in mammals that regulate actin polymerization and are essential for cell viability (Witke et al., 2001), fertilization, and normal embryo development in cattle (Rawe et al., 2006). Four profilin isoforms are expressed in mammalian testis, including spermatogenic cells (Behnen et al., 2009). Profilin-1 and profilin-3 play an important role in mouse sperm

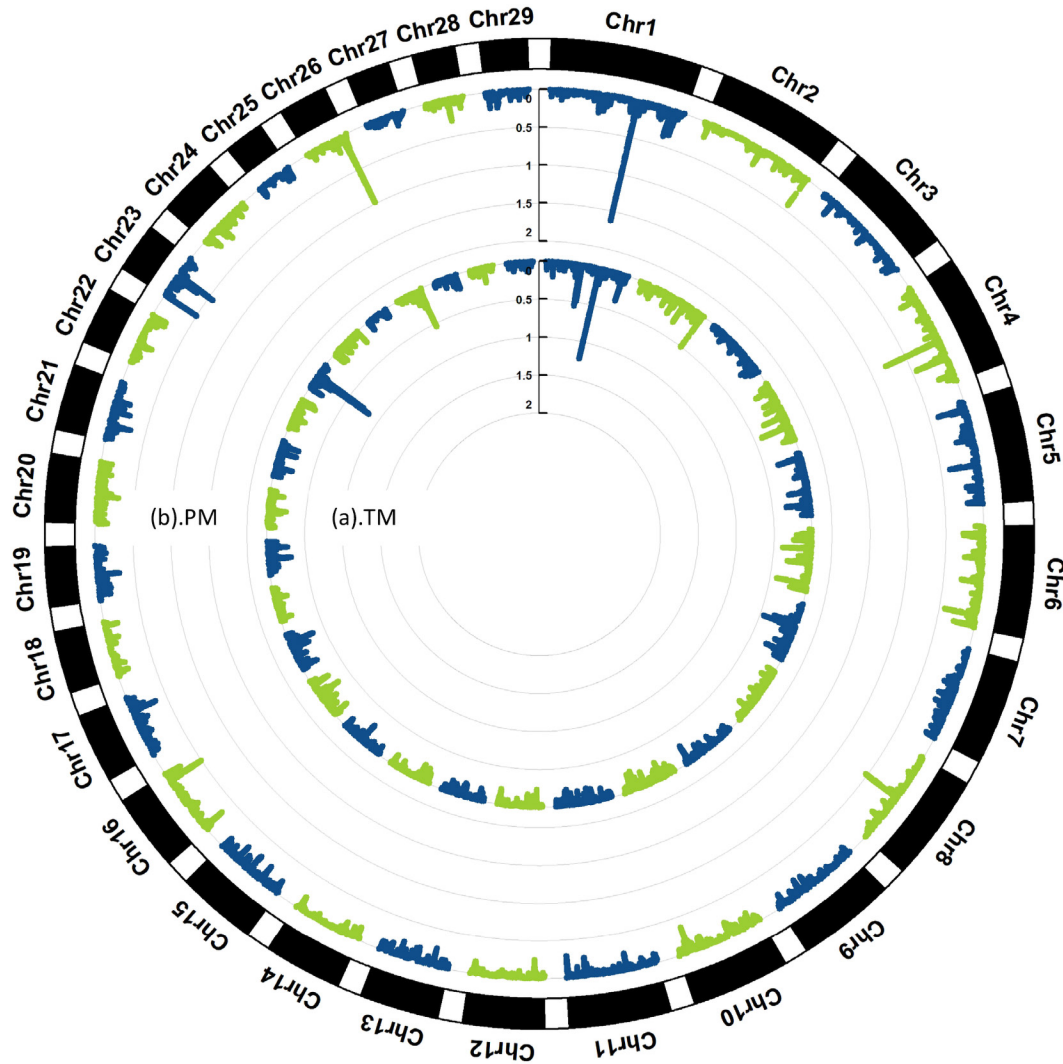


Figure 4. Circular Manhattan plot for the proportion of genetic variance explained by nonoverlapping windows of 1 Mb for (a) total motility (TM) and (b) progressive motility (PM) in Italian Holstein Bulls. The scale bar represents the percentage of variance explained by each window. Chr = chromosome.

metabolism and motility (Vicens et al., 2014) and the loss of profilin-3 has been associated with impairment of spermiogenesis (Umer et al., 2021). The second most important region (0.98% variance) for PM was on BTA26 (46,055,585–47,048,342) and included cytokinesis dedicator 1 (*DOCK1*) which encodes a member of the migration and invasion dedicator of the cytokinesis protein superfamily which is involved in cell motility (Valente et al., 2016). With regard to TM, the second most important region (0.83% variance) was on BTA 23 (24,376,529–25,372,486) and included *IL17A* and *IL17F* which have an immunoregulatory role in the male reproductive tract (Bagri et al., 2022). Many of the genes in the other significant windows that fell into BP, CC, and MF categories (Table 3) are associ-

ated with spermatogenesis and male fertility [e.g., the peptidylarginine deiminase family (**PAD**), on BTA 2]. The PAD are a class of Ca^{+2} dependent enzymes that are widely expressed in mammalian tissues including ovules, ovaries, testes, and in early embryos (Wang and Wang, 2013). The PAD enzymes perform the citrullination of arginine residues, which affects the stability and degradation of proteins, and play a crucial role during cell development, differentiation, and fertility (Fichtner et al., 2020). The enzyme PADI6 is involved in decondensation of sperm chromatin (Wright et al., 2003; Alghamdi et al., 2019). In humans, sperm chromatin organization has been shown to affect sperm function during fertilization and early embryo development (Miller et al., 2010) and has been implicated

Table 2. Genomic regions and genes in the 10 windows explaining the highest percentage of genetic variance for total motility (TM) and progressive motility (PM) in Italian Holstein Bulls

Chr ¹	Window region		TM variance (%)	PM variance (%)	Gene
	Start	End			
1	65,216,449	66,211,180	0.55	*	<i>ENSBTAG00000048371, FBXO40, GOLGB1, GTF2E1, HCLS1, HGD, NDUFB4, POLQ, RABL3, STXBP5L</i>
	68,304,713	69,303,189	0.48	*	<i>CCDC14, ENSBTAG00000053239, ITGB5, KALRN, ROPN1, UMPS</i>
	102,624,551	103,622,454	1.29	1.62	<i>E1BHJ0</i>
	148,523,433	149,497,374	0.36	*	<i>CHAF1B, CLDN14, DOP1B, HLCS, MORC3, PIGP, RIPPLY3, SIM2, TTC3, VPS26C</i>
2	134,981,487	135,980,997	0.50	0.43	<i>ARHGEF10L, ARHGEF19, ATP13A2, CPLANE2, CROCC, EPHA2, FBXO42, MFAP2, NECAP2, PADI1, PADI2, PADI3, PADI4, PADI6, RCC2, SDHB, SPATA21, SZRD1</i>
4	69,538,518	70,533,461		0.78	<i>CBX3, HNRNPA2B1, NFE2L3, SNX10</i>
	90,956,171	91,949,079	0.39	0.44	<i>ARF5, FSCN3, GCC1, GRM8, PAX4, SND1, ZNF800</i>
6	37,145,522	38,144,195	0.38	*	<i>DCAF16, FAM184B, LAP3, LCORL, MED28, NCAPG</i>
	102,406,084	103,403,193	0.41	0.42	<i>CRMP1,^a DMP1,^c DSPP,^c ENSBTAG00000006508^b ENSBTAG00000006509^b ENSBTAG00000017829^b ENSBTAG00000017830^b ENSBTAG00000021730^b ENSBTAG00000054995^c JAKMIP1,^a PPP2R2C,^c SPARCL1,^c WFS1^c</i>
7	44,213,448	45,212,096	0.44	*	<i>AFF4, BTBD2, CSNK1G2, FSTL4, GDF9, HSPA4, LEAP2, SCAMP4, SHROOM1, UQCRQ, ZCCHC10</i>
8	54,914,798	55,908,878	*	0.52	<i>TLE4</i>
16	70,458,024	71,454,118	*	0.49	<i>ANGEL2, ATF3, BATF3, DTL, FLVCR1, NENF, NSL1, PACC1, PPP2R5A, RPS6KC1, SPATA45, TATDN3, VASH2</i>
23	336,748	1,330,072	*	0.49	<i>KHDRBS2</i>
	24,376,529	25,372,486	0.83	*	<i>CILK1, EFHC1, ENSBTAG00000025486 ENSBTAG00000055193, FBXO9, GCM1, GSTA1, GSTA2, GSTA3, GSTA4, GSTA5, IL17A, IL17F, MCM3, PAQR8, PKHD1, TMEM14A, TRAM2</i>
26	27,974,488	28,973,510	0.74	*	<i>ABCF1, ATAT1, BOLA, BOLA-NC1, C23H6orf136, C23H6orf15, CCHCR1, CDSN, DDR1, DHX16 ENSBTAG00000004491 ENSBTAG00000005147 ENSBTAG00000007076 ENSBTAG00000037422 ENSBTAG00000038620 ENSBTAG00000054589 FLOT1, GNL1, GTF2H4, IER3, JSP.1, MDC1, MOG, MRPS18B, NRM, POLR1H, POU5F1, PPP1R10, PPP1R11, PPP1R18, PRR3, PSORS1C2, RNF39, RPP21, SFTA2, TCF19, TRIM10, TRIM15, TRIM26, TRIM31, TRIM39, TRIM40, TUBB, VARS2, ZFP57</i>
	28,014,882	29,014,696	*	0.56	<i>ABCF1, ATAT1, BOLA BOLA-NC1, C23H6orf136, C23H6orf15, CDSN, DDR1, DHX16 ENSBTAG00000005147 ENSBTAG00000007075 ENSBTAG00000017240 ENSBTAG00000037422 ENSBTAG00000038619 ENSBTAG00000054440 ENSBTAG00000054589 FLOT1, GABBR1, GNL1, GTF2H4, IER3, JSP.1, MDC1, MOG, MRPS18B, NRM, POLR1H, PPP1R10, PPP1R11, PPP1R18, PRR3, RNF39, RPP21, SFTA2, TRIM10, TRIM15, TRIM26, TRIM31, TRIM39, TRIM40, TUBB OR5V1, OR2H1D, OR2H1, OR10C1, OR11W11, VARS2, ZFP57</i>
	46,055,585	47,048,342	0.52	0.98	<i>DOCK1, INSN2</i>

¹Chr = chromosome.*Regions not identified for the trait; ^agenes associated only with TM, ^bgenes associated only PM.*Regions not identified for the trait; ^agenes associated only with TM, ^cgenes associated with TM and PM.

Table 3. Gene Ontology terms significantly associated with biological processes (BP), cellular components (CC), and molecular functions (MF) related to total motility (TM) and progressive motility (PM)

Category	Group	GO: Term	Gene name	BTA	TM FDR ¹	PM FDR	
Biological process	Alpha-AA biosynthetic process	GO: 0036414	<i>PADI1, PADI2</i>	2	9.74×10^{-9}	1.5×10^{-6}	
		Histone citrullination	<i>PADI6, PADI4</i>				
		GO: 0018101	<i>PADI3</i>		9.74×10^{-9}	1.5×10^{-7}	
		Protein citrullination					
		GO: 0018195 Peptidyl-arginine modification			2.35×10^{-4}	3.8×10^{-4}	
		GO: 0006749	<i>GSTA3, GSTA4</i>	23	2.23×10^{-3}	*	
		Glutathione metabolic process	<i>GSTA2, GSTA1</i>				
			<i>GSTA5</i>				
		Antigen processing and presentation	GO: 0002428 antigen processing and presentation	<i>JSP.1, BOLA</i>	23	0.0016	0.0046
			<i>BOLA-NC1</i>				
Cellular components	Immune system process	GO: 0002376	<i>EPHA2, PADI2</i>	2	*	0.0192	
		immune system process	<i>SNX10</i>				
			<i>BOLA, BOLA-NC1, FLOT1, JSP.1, MOG</i>	23			
			<i>RNF39, TRIM10</i>				
			<i>TRIM15, TRIM31</i>				
			<i>TRIM40</i>				
			<i>DOCK1</i>	26			
		Sensory perception of chemical stimulus	GO: 0003008 system process	<i>OR12D2, OR10C1, OR2H1,</i>	23	*	4.8×10^{-10}
			GO: 0007606 sensory perception of chemical stimulus	<i>OR5V1, OR12D3</i>			
			GO: 0051606 detection of stimulus	<i>OR12D2, OR12D3</i>	23	*	5.1×10^{-6}
	G protein-coupled receptor signaling pathway	GO: 0007186	<i>GABBR1, OR12D2,</i>	23	*	1.3×10^{-8}	
		G protein-coupled receptor signaling pathway	<i>OR10C1, OR2H1</i>				
		GO: 0015630	<i>OR5V1, OR12D3</i>				
	Microtubule cytoskeleton	GO: 0015630 microtubule cytoskeleton	<i>CCHCR1, CILK1, EFHC1,</i>	23	0.0037	*	
			<i>MCM3, PKHD1</i>				
			<i>CBX3, SNX10</i>	4	*	0.0058	
			<i>EVC</i>	6			
			<i>PPP2R5A</i>	16			
			<i>CROCC</i>	2	0.0037	0.0058	
			<i>CRMP1, JAKMIP1</i>	6			
			<i>ATAT1, FLOT1, TUBB</i>	23			
		GO: 0072686 mitotic spindle	<i>ATAT1, EFHC1</i>	23	0.0002	*	
		GO: 0005856 cytoskeleton	<i>PKHD1</i>				
			<i>CCHCR1</i>	23	0.0320	*	
			<i>EFHC1, MCM3, PKHD1</i>				
			<i>CBX3</i>	4	*	0.0310	
			<i>EVC</i>	6			
			<i>PPP2R5A</i>	16			
			<i>CPLANE2, CROCC</i>	2	0.0320	0.0310	
			<i>PADI6, RCC2</i>				
			<i>FSCN3</i>	4			
			<i>CRMP1, JAKMIP1</i>	6			
			<i>ATAT1, FLOT1</i>	23			
			<i>TUBB</i>				
	Protein serine/threonine phosphatase complex	GO: 1903293 phosphatase complex	<i>PPP1R10, PPP1R11</i>	23	*	0.0045	
		GO: 0008287 protein serine/threonine phosphatase complex	<i>PPP2R2C</i>				
	Focal adhesion	GO: 0005925 focal adhesion	<i>EPHA2, NECAP2</i>	2	0.0713	*	
			<i>ATP13A2</i>				
			<i>MDC1, ATAT1</i>	23			
			<i>FLOT1</i>				

Continued

Table 3 (Continued). Gene Ontology terms significantly associated with biological processes (BP), cellular components (CC), and molecular functions (MF) related to total motility (TM) and progressive motility (PM)

Category	Group	GO: Term	Gene name	BTA	TM FDR ¹	PM FDR			
Molecular Functions	Hydrolase activity acting on carbon-nitrogen (but not peptide) bonds	GO: 0004668 protein-arginine deiminase activity GO: 0016810 hydrolase activity acting on carbon-nitrogen (but not peptide) bonds	<i>PADI2</i>	2	3.6×10^{-10}	2.91×10^{-9}			
			<i>PADI1</i>						
			<i>PADI6</i>						
			<i>PADI4</i>						
			<i>PADI3</i>						
	Binding	GO: 0043167 ion binding	GO: 0005488 binding	<i>FBXO40, GTF2E1, HGD, KALRN, MORC1</i>	1	1.8×10^{-4}	*		
				<i>LAP3</i>	6				
				<i>CSNK1G2, FSTL4, HSPA4</i>	7				
				<i>CILK1, EFHC1</i>	23				
				<i>GCM1, MCM3</i>					
			<i>TCF19</i>						
			<i>CROCC, MFAP2</i>	2	*	0.00098			
			<i>RCC2</i>						
			<i>CBX3, FSCN3</i>	4					
			<i>HNRNPA2B1</i>						
			<i>NFE2L3</i>						
			<i>PAX4, SND1</i>						
			<i>SNX10</i>						
			<i>CRMP1, DSPP</i>	6					
			<i>JAKMIP1, WFS1</i>						
			<i>TLE4</i>	8					
			<i>PPP2R5A,</i>	16					
			<i>RPS6KC1</i>						
			<i>CDSN, FLOT1</i>	23					
			<i>GTF2H4</i>						
			<i>KHDRBS2</i>						
			<i>MDC1, MOG</i>						
			<i>NRM, PPP1R11</i>						
			<i>PPP1R18, RPP21</i>						
			<i>ATP13A2</i>	2	1.8×10^{-4}	0.00098			
			<i>CPLANE2, EPHA2</i>						
			<i>PADI1, PADI2</i>						
			<i>PADI3, PADI4, PADI6,</i>						
			<i>SDHB, SPATA21</i>						
			<i>ARF5</i>	4					
			<i>SPARCL1</i>	6					
			<i>ABCF1, DDR1, DHX16</i>	23					
			<i>GNL1, POLR1H</i>						
			<i>PPP1R10, PRR3, RNF39,</i>						
			<i>TRIM10, TRIM15, TRIM31</i>						
			<i>TRIM40, TUBB, VARS2,</i>						
			<i>ZFP57</i>						
			<i>GSTA5, GSTA3, GSTA4,</i>	23	2.3×10^{-8}	*			
			<i>GSTA2</i>						
			<i>GSTA3</i>						
			Olfactory reception activity	GO: 0004984 olfactory reception activity	GO: 0004984 olfactory reception activity	<i>OR12D12</i>	23	*	9.1×10^{-11}
						<i>OR10C1</i>			
						<i>OR2H1</i>			
						<i>OR5V1</i>			
						<i>OR12D3</i>			
Molecular transducer activity	GO: 0060089 molecular transducer activity	GO: 0060089 molecular transducer activity	<i>EPHA2</i>	2	*	6.8×10^{-7}			
			<i>OR12D12</i>	23					
			<i>OR10C1</i>						
			<i>OR2H1</i>						
			<i>OR5V1</i>						
			<i>OR12D3</i>						

¹FDR = false discovery rate.

*Genome Ontology not identified for the trait.

Table 4. Quantitative trait loci annotated and significant in the enrichment analysis for total (TM) and progressive (PM) motility

Trait	QTL type	QTL class	BTA	Observed number of QTL	Expected number of QTL	P-value	FDR ¹
TM	Production	Average daily feed intake	6	7	9	2.75×10^{-9}	6.48×10^{-8}
	Reproduction	Calving ease (direct)	6	28	37	2.83×10^{-11}	1.33×10^{-9}
	Meat and carcass	Carcass weight	6	24	137	3.01×10^{-9}	3.53×10^{-4}
		Fat thickness at the 12th rib	6	10	21	5.15×10^{-7}	8.07×10^{-8}
	Health	Antibody-mediated immune response	23	4	10	3.76×10^{-4}	3.53×10^{-3}
		Cell-mediated immune response	23	2	3	4.41×10^{-4}	3.45×10^{-3}
PM	Milk	Milk yield, and percentage of protein and fat in milk (EBV)	1	1	7	0.00674	0.00549
	Reproduction	Reproductive efficiency	23	1	7	0.01097	0.00974

¹FDR = false discovery rate.

in subfertility (Souza et al., 2018). The *BoLA* class 1 genes and *JSP.1* are associated with both PM and TM. These genes map on BTA 23 and are part of the major histocompatibility complex involved in antigen presentation (Tijjani et al., 2019). The major histocompatibility complex antigens play a key role in immune response and affect production, diseases, and fertility traits, in particular the sperm-ovum interaction (Ellis and Hammond, 2014). *BoLA-NC1* is essential for the establishment and maintenance of pregnancy and in the maternal immune response to the fetus (Shu et al., 2012).

Five genes in the most significant GO biological processes term (GO:BP), were associated exclusively with TM. Glutathione S-transferase alpha (*GSTA*), which is within one of the regions with the highest effect on BTA 23, belongs to the GST superfamily of multigenic and multifunctional isoenzymes that catalyze glutathione-dependent reactions (GO:BP Glutathione metabolic process and GO:MF Glutathione transferase activity). Expression of the GST has been reported in male reproductive tissues in cattle (Juyena and Stelletta, 2012), pigs (Pérez-Patiño et al., 2019), and humans (Dutta et al., 2021) including sperm, seminal plasma, and in testicular somatic cells, particularly in Leydig and Sertoli cells (Benbrahim-Tallaa et al., 2002). *GSTA* also plays a role in spermatogenesis, sperm capacitation and sperm-oocyte binding, and is involved in the nuclear decondensation of the sperm chromatin (Kumar et al., 2014; Kwon et al., 2014; Llavanera et al., 2020).

Genes on BTA 23 associated with TM included tripartite motifs (*TRIM31* and *TRIM10*) and interleukins (*IL17A*, *IL17F*) that code for proteins involved in immune processes in mammals (GO:BP: antigen processing and presentation and immune system process). *TRIM31* is a member of the E3 ubiquitin ligases protein family that is involved with innate immunity (Ozato et al., 2008; Tijjani et al., 2019) and *IL17A* is a member of the interleukin family that initiate a potent

inflammatory response (Zenobia and Hajishengallis, 2015). In humans the level of IL17 in seminal fluid has been shown to inversely correlate with SM and mortality (Qian, 2012) and to play an important role in successful embryo implantation. In addition, IL-17 regulates the secretion of cytokines (IL-1 β and TNF- α), chemokines, and promotes the recruitment of neutrophils (Song et al., 2016; Wattedegera et al., 2017) which are related to the modulation of the immune response to sperm antigens and possibly the protection of sperm from the response of the female immune system (Vera et al., 2003).

The GO:CC terms contained the α -tubulin (*ATAT1*), tubulin beta class I (*TUBB*), and flotillin 1 (*FLOT1*) genes, which code for proteins that have been proposed as biomarkers of male fertility (Gilbert et al., 2007; Fu et al., 2019; Patil et al., 2020). Alpha-tubulin is a cytoskeletal protein that contributes to microtubule organization, sperm morphology, and flagellar function, and is present in the cytoplasm of Sertoli cells, Leydig cells, spermatogonia, primary spermatocytes, secondary spermatocytes, and spermatids of seminiferous tubules (Kalebic et al., 2013). Tubulin plays a role in microtubule formation in sperm cells of buffaloes (Fu et al., 2019) and is expressed in spermatogonia, Sertoli cells, spermatids, and spermatozoa of pig testes (Luo et al., 2010). In buffalo sperm, α -tubulin has been shown to provide structural support, facilitate cell motility, and to participate in oocyte binding and activation (Patil et al., 2020). *JAKMIP1* and *PKDH1* are related to sperm function and cell polarization (Hatzirodos et al., 2014), embryo development and survival (Gigarel et al., 2008), respectively. *MDC1*, *EPHAH2*, *ATAT1*, and *FLOT1* code for proteins that are associated with focal adhesion, a process involved in fertilization, ovum transport, and sperm-ovum interaction (Talbot et al., 2003). Additionally, *PPP1R10*, *PPP1R11*, *PPP2R2C*, and *FSCN3* genes were identified and code for proteins related to spermatogenesis, SM, and fertility (Han et

al., 2007, 2008; Cheng et al., 2009; Gao et al., 2019), *PPP2R5A* codes for protein phosphatase 2 regulatory subunit B alpha and is correlated with spermatogonial mitosis and it is related to sphingolipid signaling pathway which is associated with SM (Sarakul et al., 2018; Cai et al., 2021). Genes within the region of BTA 23, which showed an association with PM in the current study, belong to the OR family and *KHDRBS2*. The OR are involved in chemotaxis and determine speed, strength, and direction of movement of the spermatozoa toward the ovum (Ali et al., 2021). Of the 20 to 66 testicular OR known in mammals, at least 8 are expressed in sperm cells, testis, and epididymus (Milardi et al., 2018). In humans, the OR influence spermatogenesis, epididymal growth, acrosome reaction, initiate flagella movement and facilitate cell-to-cell communication (Olaniyan et al., 2021). The OR in the sperm cell activate proteins that increase intracellular calcium ions and hence motility (Vanderhaeghen et al., 1997). *KHDRBS2*, has been associated with fertility in San Martinero cattle (De León et al., 2019) and pregnancy status in Brahman beef cattle (Reverter et al., 2016).

The QTL enrichment analysis for TM revealed QTL associated with feed intake, calving ease, carcass weight, and fat thickness on BTA 6, and antibody-mediated immune response and cell-mediated immune response on BTA 23. Previous studies have identified other QTL on BTA 23 related to genes such as *IL4* (interleukin-4), antibody-mediated immune response, and DMI that may be associated with SM (Wang et al. 2012).

CONCLUSIONS

In this study, we identified a large number of genes associated with SM and involved with sperm and sperm-related traits in Italian Holsteins bulls. These findings suggest that many genes controlling SM could have pleiotropic effects and be functionally correlated with other traits related to fertility. Our results show the highly polygenic nature of SM, with each of numerous candidate genes responsible for a small proportion of the genetic variance. SM influences male fertility, a trait that has probably experienced strong selective pressure even before cattle domestication, about 10,000 years ago. As such, beneficial alleles of genes with major effects on this trait are likely to have reached fixation, leaving only variants with small effect in the domesticated populations, which were identified by GWAS as presented here. The characterization of selection signatures across different timescales, as well as the analysis of ancient genomes from domestic and wild animals, may help in identifying variants related to fertility that have experienced selection before and after domestication events.

ACKNOWLEDGMENTS

J.L.W. was supported by the project LEO Livestock Environment Open-data, 16.2 – PSRN 2014-2020 funded through Fondo Europeo Agricolo per lo Sviluppo Rurale (FEASR). A.A. and P.A.M. received support from the Italian Ministry of Education, University and Research (MUR; Rome, Italy) for Progetti PRIN2017 20174BTC4R. The authors have not stated any conflicts of interest.

REFERENCES

- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol.* 51:28. <https://doi.org/10.1186/s12711-019-0469-3>.
- Al-Kanaan, A., S. König, and K. Brügemann. 2015. Effects of heat stress on semen characteristics of Holstein bulls estimated on a continuous phenotypic and genetic scale. *Livest. Sci.* 177:15–24. <https://doi.org/10.1016/j.livsci.2015.04.003>.
- Alghamdi, M., D. Alasmari, A. Assiri, E. Mattar, A. A. Aljaddawi, S. G. Alattas, and E. M. Redwan. 2019. An overview of the intrinsic role of citrullination in autoimmune disorders. *J. Immunol. Res.* 2019:7592851. <https://doi.org/10.1155/2019/7592851>.
- Ali, M. A., Y. Wang, Z. Qin, X. Yuan, Y. Zhang, and C. Zeng. 2021. Odorant and Taste Receptors in sperm chemotaxis and cryopreservation: Roles and implications in sperm capacitation, motility, and fertility. *Genes (Basel)* 12:488. <https://doi.org/10.3390/genes12040488>.
- Alves, M. B. R., E. C. C. Celeghini, and C. Belleannée. 2020. From sperm motility to sperm-borne microRNA signatures: New approaches to predict male fertility potential. *Front. Cell Dev. Biol.* 8:791. <https://doi.org/10.3389/fcell.2020.00791>.
- Bagri, P., V. C. Anipindi, and C. Kaushic. 2022. The role of IL-17 during infections in the female reproductive tract. *Front. Immunol.* 13:861444. <https://doi.org/10.3389/fimmu.2022.861444>.
- Ballow, D., M. L. Meistrich, M. Matzuk, and A. Rajkovic. 2006. *Sohlh1* is essential for spermatogonial differentiation. *Dev. Biol.* 294:161–167. <https://doi.org/10.1016/j.ydbio.2006.02.027>.
- Behnen, M., K. Murk, P. Kursula, H. Cappallo-Obermann, M. Rothkegel, A. L. Kierszenbaum, and C. Kirchhoff. 2009. Testis-expressed profilins 3 and 4 show distinct functional characteristics and localize in the acroplaxome-manchette complex in spermatids. *BMC Cell Biol.* 10:34. <https://doi.org/10.1186/1471-2121-10-34>.
- Benbrahim-Tallaa, L., E. Tabone, G. Tosser-Klopp, F. Hatay, and M. Benahmed. 2002. Glutathione S-Transferase Alpha expressed in porcine sertoli cells is under the control of follicle-stimulating hormone and testosterone. *Biol. Reprod.* 66:1734–1742. <https://doi.org/10.1095/biolreprod66.6.1734>.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Browning, B. L., and S. R. Browning. 2016. Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* 98:116–126. <https://doi.org/10.1016/j.ajhg.2015.11.020>.
- Browning, B. L., X. Tian, Y. Zhou, and S. R. Browning. 2021. Fast two-stage phasing of large-scale sequence data. *Am. J. Hum. Genet.* 108:1880–1890. <https://doi.org/10.1016/j.ajhg.2021.08.005>.
- Butler, M. L., J. M. Bormann, R. L. Weaver, D. M. Grieger, and M. M. Rolf. 2020. Selection for bull fertility: A review. *Transl. Anim. Sci.* 4:423–441. <https://doi.org/10.1093/tas/txz174>.
- Cai, X., S. Wu, T. D. Mipam, H. Luo, C. Yi, C. Xu, W. Zhao, H. Wang, and J. Zhong. 2021. Testis transcriptome profiling identified lncRNAs involved in spermatogenic arrest of cattleyak. *Funct.*

- Integr. Genomics 21:665–678. <https://doi.org/10.1007/s10142-021-00806-8>.
- Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4:7. <https://doi.org/10.1186/s13742-015-0047-8>.
- Chen, L., H. J. Abel, I. Das, D. E. Larson, L. Ganel, K. L. Kanchi, A. A. Regier, E. P. Young, C. J. Kang, A. J. Scott, C. Chiang, X. Wang, S. Lu, R. Christ, S. K. Service, C. W. K. Chiang, A. S. Havulinna, J. Kuusisto, M. Boehnke, M. Laakso, A. Palotie, S. Ripatti, N. B. Freimer, A. E. Locke, N. O. Stitzel, and I. M. Hall. 2021. Association of structural variation with cardiometabolic traits in Finns. *Am. J. Hum. Genet.* 108:583–596. <https://doi.org/10.1016/j.ajhg.2021.03.008>.
- Cheng, L., S. Pilder, A. C. Nairn, S. Ramdas, and S. Vijayaraghavan. 2009. PP1 γ 2 and PPP1R11 are parts of a multimeric complex in developing testicular germ cells in which their steady state levels are reciprocally related. *PLoS One* 4:e4861. <https://doi.org/10.1371/journal.pone.0004861>.
- Corbet, N. J., B. M. Burns, D. J. Johnston, M. L. Wolcott, D. H. Corbet, B. K. Venus, Y. Li, M. R. McGowan, and R. G. Holroyd. 2013. Male traits and herd reproductive capability in tropical beef cattle. 2. Genetic parameters of bull traits. *Anim. Prod. Sci.* 53:101. <https://doi.org/10.1071/AN12163>.
- De León, C., C. Manrique, R. Martínez, and J. F. Rocha. 2019. Research article genomic association study for adaptability traits in four Colombian cattle breeds. *Genet. Mol. Res.* 18. <https://doi.org/10.4238/gmr18373>.
- Druet, T., S. Fritz, E. Sellem, B. Basso, O. Gérard, L. Salas-Cortes, P. Humblot, X. Druart, and A. Eggen. 2009. Estimation of genetic parameters and genome scan for 15 semen characteristics traits of Holstein bulls. *J. Anim. Breed. Genet.* 126:269–277. <https://doi.org/10.1111/j.1439-0388.2008.00788.x>.
- Dutta, S., P. Sengupta, P. Slama, and S. Roychoudhury. 2021. Oxidative stress, testicular inflammatory pathways, and male reproduction. *Int. J. Mol. Sci.* 22:10043. <https://doi.org/10.3390/ijms221810043>.
- Ellis, S. A., and J. A. Hammond. 2014. The functional significance of cattle major histocompatibility complex class I genetic diversity. *Ann. Rev. Anim. Biosci.* 2:285–306. <https://doi.org/10.1146/annurev-animal-022513-114234>.
- Fichtner, T., F. Kotarski, U. Gärtner, I. Conejeros, C. Hermosilla, C. Wrenzycki, and A. Taubert. 2020. Bovine sperm samples induce different NET phenotypes in a NADPH oxidase-, PAD4-, and Ca⁺⁺-dependent process. *Biol. Reprod.* 102:902–914. <https://doi.org/10.1093/biolre/iaaa003>.
- Fonseca, P. A. S., A. Suárez-Vega, G. Marras, and Á. Cánovas. 2020. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *Gigascience* 9:giaa149. <https://doi.org/10.1093/gigascience/giaa149>.
- Fu, Q., L. Pan, D. Huang, Z. Wang, Z. Hou, and M. Zhang. 2019. Proteomic profiles of buffalo spermatozoa and seminal plasma. *Theriogenology* 134:74–82. <https://doi.org/10.1016/j.theriogenology.2019.05.013>.
- Fuerst-Waltl, B., H. Schwarzenbacher, C. Perner, and J. Sölkner. 2006. Effects of age and environmental factors on semen production and semen quality of Austrian Simmental bulls. *Anim. Reprod. Sci.* 95:27–37. <https://doi.org/10.1016/j.anireprosci.2005.09.002>.
- Gao, Y., S. Li, Z. Lai, Z. Zhou, F. Wu, Y. Huang, X. Lan, C. Lei, H. Chen, and R. Dang. 2019. Analysis of long non-coding RNA and mRNA expression profiling in immature and mature bovine (*Bos taurus*). *Front. Genet.* 10:646. <https://doi.org/10.3389/fgene.2019.00646>.
- Gigarel, N., N. Frydman, P. Bulet, V. Kerbrat, G. Tachdjian, R. Fanchin, C. Antignac, R. Frydman, A. Mummich, and J. Steffann. 2008. Preimplantation genetic diagnosis for autosomal recessive polycystic kidney disease. *Reprod. Biomed. Online* 16:152–158. [https://doi.org/10.1016/S1472-6483\(10\)60569-X](https://doi.org/10.1016/S1472-6483(10)60569-X).
- Gilbert, I., N. Bissonnette, G. Boissonneault, M. Vallée, and C. Robert. 2007. A molecular analysis of the population of mRNA in bovine spermatozoa. *Reproduction* 133:1073–1086. <https://doi.org/10.1530/REP-06-0292>.
- Gredler, B., C. Fuerst, B. Fuerst-Waltl, H. Schwarzenbacher, and J. Sölkner. 2007. Genetic parameters for semen production traits in Austrian dual-purpose Simmental bulls. *Reprod. Domest. Anim.* 42:326–328. <https://doi.org/10.1111/j.1439-0531.2006.00778.x>.
- Greenland, S., S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman, and D. G. Altman. 2016. Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations. *Eur. J. Epidemiol.* 31:337–350. <https://doi.org/10.1007/s10654-016-0149-3>.
- Han, Y., X.-X. Song, H.-L. Feng, C.-K. Cheung, P.-M. Lam, C.-C. Wang, and C. J. Haines. 2008. Mutations of t-complex testis expressed gene 5 transcripts in the testis of sterile t-haplotype mutant mouse. *Asian J. Androl.* 10:219–226. <https://doi.org/10.1111/j.1745-7262.2008.00323.x>.
- Han, Y. B., H. L. Feng, C. K. Cheung, P. M. Lam, C. C. Wang, and C. J. Haines. 2007. Expression of a novel T-complex testis expressed 5 (*Tctex5*) in mouse testis, epididymis, and spermatozoa. *Mol. Reprod. Dev.* 74:1132–1140. <https://doi.org/10.1002/mrd.20631>.
- Hatzirodos, N., H. F. Irving-Rodgers, K. Hummitzsch, M. L. Harland, S. E. Morris, and R. J. Rodgers. 2014. Transcriptome profiling of granulosa cells of bovine ovarian follicles during growth from small to large antral sizes. *BMC Genomics* 15:24. <https://doi.org/10.1186/1471-2164-15-24>.
- Hering, D. M., K. Olenski, and S. Kaminski. 2014. Genome-wide association study for poor sperm motility in Holstein-Friesian bulls. *Anim. Reprod. Sci.* 146:89–97. <https://doi.org/10.1016/j.anireprosci.2014.01.012>.
- Juyena, N. S., and C. Stelletta. 2012. Seminal plasma: An essential attribute to spermatozoa. *J. Androl.* 33:536–551. <https://doi.org/10.1016/j.jandrol.110.012583>.
- Kalebic, N., S. Sorrentino, E. Perlas, G. Bolasco, C. Martinez, and P. A. Heppenstall. 2013. α TAT1 is the major α -tubulin acetyltransferase in mice. *Nat. Commun.* 4:1962. <https://doi.org/10.1038/ncomms2962>.
- Kumar, R., V. K. Singh, and S. K. Atreja. 2014. Glutathione-S-transferase: Role in buffalo (*Bubalus bubalis*) sperm capacitation and cryopreservation. *Theriogenology* 81:587–598. <https://doi.org/10.1016/j.theriogenology.2013.11.012>.
- Kwon, W.-S., M. S. Rahman, J.-S. Lee, J. Kim, S.-J. Yoon, Y.-J. Park, Y.-A. You, S. Hwang, and M.-G. Pang. 2014. A comprehensive proteomic approach to identifying capacitation related proteins in boar spermatozoa. *BMC Genomics* 15:897. <https://doi.org/10.1186/1471-2164-15-897>.
- Liu, H., H. Song, Y. Jiang, Y. Jiang, F. Zhang, Y. Liu, Y. Shi, X. Ding, and C. Wang. 2021. A single-step genome wide association study on body size traits using imputation-based whole-genome sequence data in Yorkshire pigs. *Front. Genet.* 12:629049. <https://doi.org/10.3389/fgene.2021.629049>.
- Llavanera, M., Y. Mateo-Otero, S. Bonet, I. Barranco, B. Fernández-Fuertes, and M. Yeste. 2020. The triple role of glutathione S-transferases in mammalian male fertility. *Cell. Mol. Life Sci.* 77:2331–2342. <https://doi.org/10.1007/s00018-019-03405-w>.
- Luo, J., J. R. Rodriguez-Sosa, L. Tang, A. Bondareva, S. Megee, and I. Dobrinski. 2010. Expression pattern of acetylated α -tubulin in porcine spermatogonia. *Mol. Reprod. Dev.* 77:348–352. <https://doi.org/10.1002/mrd.21153>.
- Mancini, G., M. Gargani, G. Chillemi, E. L. Nicolazzi, P. A. Marsan, A. Valentini, and L. Pariset. 2014. Signatures of selection in five Italian cattle breeds detected by a 54K SNP panel. *Mol. Biol. Rep.* 41:957–965. <https://doi.org/10.1007/s11033-013-2940-5>.
- Milardi, D., C. Colussi, G. Grande, F. Vincenzoni, F. Pierconti, F. Mancini, S. Baroni, M. Castagnola, R. Marana, and A. Pontecorvi. 2018. Olfactory receptors in semen and in the male tract: From proteome to proteins. *Front. Endocrinol. (Lausanne)* 8:379. <https://doi.org/10.3389/fendo.2017.00379>.
- Miller, D., M. Brinkworth, and D. Iles. 2010. Paternal DNA packaging in spermatozoa: More than the sum of its parts? DNA, histones,

- protamines and epigenetics. *Reproduction* 139:287–301. <https://doi.org/10.1530/REP-09-0281>.
- Misztal, I., S. Tsuruta, D. A. L. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs. University of Georgia. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf.
- Olaniyan, O. T., A. Dare, G. E. Okotie, C. O. Adetunji, B. O. Ibitoye, O. Eweoya, J. B. Dare, and B. J. Okoli. 2021. Ovarian odorant-like biomolecules in promoting chemotaxis behavior of spermatozoa olfactory receptors during migration, maturation, and fertilization. *Middle East Fertil. Soc. J.* 26:3. <https://doi.org/10.1186/s43043-020-00049-w>.
- Ozato, K., D.-M. Shin, T.-H. Chang, and H. C. Morse III. 2008. TRIM family proteins and their emerging roles in innate immunity. *Nat. Rev. Immunol.* 8:849–860. <https://doi.org/10.1038/nri2413>.
- Patil, S. K., L. Somashekar, S. Selvaraju, K. V. Jamuna, S. Parthipan, B. K. Binsila, R. V. Prasad, and J. P. Ravindra. 2020. Immunohistological mapping and functional association of seminal proteins in testis and excurrent ducts with sperm function in buffalo. *Reprod. Domest. Anim.* 55:998–1010. <https://doi.org/10.1111/rda.13742>.
- Peñagaricano, F., K. A. Weigel, and H. Khatib. 2012. Genome-wide association study identifies candidate markers for bull fertility in Holstein dairy cattle. *Anim. Genet.* 43:65–71. <https://doi.org/10.1111/j.1365-2052.2012.02350.x>.
- Pérez-Patiño, C., I. Parrilla, J. Li, I. Barranco, E. A. Martínez, H. Rodríguez-Martínez, and J. Roca. 2019. The proteome of pig spermatozoa is remodeled during ejaculation. *Mol. Cell. Proteomics* 18:41–50. <https://doi.org/10.1074/mcp.RA118.000840>.
- Qian, L. 2012. The relationship between IL-17 and male infertility: Semen analysis. *Afr. J. Microbiol. Res.* 6:5672–5677.
- Rawe, V. Y., C. Payne, and G. Schatten. 2006. Profilin and actin-related proteins regulate microfilament dynamics during early mammalian embryogenesis. *Hum. Reprod.* 21:1143–1153. <https://doi.org/10.1093/humrep/dei480>.
- Reverter, A., L. R. Porto-Neto, M. R. S. Fortes, R. McCulloch, R. E. Lyons, S. Moore, D. Nicol, J. Henshall, and S. A. Lehnert. 2016. Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree. *J. Anim. Sci.* 94:4096–4108. <https://doi.org/10.2527/jas.2016-0675>.
- Rezende, F. M., G. O. Dietsch, and F. Peñagaricano. 2018. Genetic dissection of bull fertility in US Jersey dairy cattle. *Anim. Genet.* 49:393–402. <https://doi.org/10.1111/age.12710>.
- Rosen, B. D., D. M. Bickhart, R. D. Schnabel, S. Koren, C. G. Elsik, E. Tseng, T. N. Rowan, W. Y. Low, A. Zimin, C. Coudrey, R. Hall, W. Li, A. Rhie, J. Ghurye, S. D. McKay, F. Thibaud-Nissen, J. Hoffman, B. M. Murdoch, W. M. Snelling, T. G. McDanel, J. A. Hammond, J. C. Schwartz, W. Nandolo, D. E. Hagen, C. Dreischer, S. J. Schultheiss, S. G. Schroeder, A. M. Phillippy, J. B. Cole, C. P. Van Tassel, G. Liu, T. P. L. Smith, and J. F. Medrano. 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* 9:giaa021. <https://doi.org/10.1093/gigascience/giaa021>.
- Sarakul, M., M. A. Elzo, S. Koonawootrittriron, T. Suwanasopee, D. Jattawa, and T. Laodim. 2018. Characterization of biological pathways associated with semen traits in the Thai multibreed dairy population. *Anim. Reprod. Sci.* 197:324–334. <https://doi.org/10.1016/j.anireprosci.2018.09.002>.
- Shu, L., X. Peng, S. Zhang, G. Deng, Y. Wu, M. He, B. Li, C. Li, and K. Zhang. 2012. Non-classical major histocompatibility complex class makes a crucial contribution to reproduction in the dairy cow. *J. Reprod. Dev.* 58:569–575. <https://doi.org/10.1262/jrd.2011-018>.
- Song, X., X. He, X. Li, and Y. Qian. 2016. The roles and functional mechanisms of interleukin-17 family cytokines in mucosal immunity. *Cell. Mol. Immunol.* 13:418–431. <https://doi.org/10.1038/cmi.2015.105>.
- Souza, E. T., C. V. Silva, B. A. N. Travençolo, B. G. Alves, and M. E. Beletti. 2018. Sperm chromatin alterations in fertile and subfertile bulls. *Reprod. Biol.* 18:177–181. <https://doi.org/10.1016/j.repbio.2018.04.001>.
- Suchocki, T., and J. Szyda. 2015. Genome-wide association study for semen production traits in Holstein-Friesian bulls. *J. Dairy Sci.* 98:5774–5780. <https://doi.org/10.3168/jds.2014-8951>.
- Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. <https://doi.org/10.1371/journal.pone.0021800>.
- Sweett, H., P. A. S. Fonseca, A. Suárez-Vega, A. Livernois, F. Miglior, and A. Cánovas. 2020. Genome-wide association study to identify genomic regions and positional candidate genes associated with male fertility in beef cattle. *Sci. Rep.* 10:20102. <https://doi.org/10.1038/s41598-020-75758-3>.
- Talbot, P., B. D. Shur, and D. G. Myles. 2003. Cell adhesion and fertilization: Steps in oocyte transport, sperm-zona pellucida interactions, and sperm-egg fusion. *Biol. Reprod.* 68:1–9. <https://doi.org/10.1095/biolreprod.102.007856>.
- Tijjani, A., Y. T. Utsunomiya, A. G. Ezekwe, O. Nashiru, and O. Hanotte. 2019. Genome sequence analysis reveals selection signatures in endangered trypanotolerant West African Muturu Cattle. *Front. Genet.* 10:442. <https://doi.org/10.3389/fgene.2019.00442>.
- Umer, N., L. Arevalo, I. Kosterin, S. Phadke, K. Lohanadan, G. Kirfel, D. Sons, W. Witke, and H. Schorle. 2021. Loss of profilin3 impairs spermiogenesis by affecting acrosome biogenesis, autophagy, manchette development, and mitochondrial organization. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3839761>.
- Valente, T. S., F. Baldi, A. C. Sant'Anna, L. G. Albuquerque, and M. J. R. Paranhos da Costa. 2016. Genome-wide association study between single nucleotide polymorphisms and flight speed in Nelore cattle. *PLoS One* 11:e0156956. <https://doi.org/10.1371/journal.pone.0156956>.
- Vanderhaeghen, P., S. Schurmans, G. Vassart, and M. Parmentier. 1997. Specific repertoire of olfactory receptor genes in the male germ cells of several mammalian species. *Genomics* 39:239–246. <https://doi.org/10.1006/geno.1996.4490>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- VanRaden, P. M., K. M. Olson, D. J. Null, and J. L. Hutchison. 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* 94:6153–6161. <https://doi.org/10.3168/jds.2011-4624>.
- Vera, O., L. A. Vásquez, and M. Gladys Muñoz. 2003. Semen quality and presence of cytokines in seminal fluid of bull ejaculates. *Theriogenology* 60:553–558. [https://doi.org/10.1016/S0093-691X\(03\)00031-1](https://doi.org/10.1016/S0093-691X(03)00031-1).
- Vicens, A., L. Lüke, and E. R. S. Roldan. 2014. Proteins involved in motility and sperm-egg interaction evolve more rapidly in mouse spermatozoa. *PLoS One* 9:e91302. <https://doi.org/10.1371/journal.pone.0091302>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* 94:73–83. <https://doi.org/10.1017/S0016672312000274>.
- Wang, S., and Y. Wang. 2013. Peptidylarginine deiminases in citrullination, gene regulation, health, and pathogenesis. *Biochim. Biophys. Acta.* 1829:1126–1135. <https://doi.org/10.1016/j.bbgram.2013.07.003>.
- Wattegedera, S. R., Y. Corripio-Miyar, Y. Pang, D. Frew, T. N. McNeilly, J. Palarea-Albaladejo, C. J. McInnes, J. C. Hope, E. J. Glass, and G. Entrican. 2017. Enhancing the toolbox to study IL-17A in cattle and sheep. *Vet. Res.* 48:20. <https://doi.org/10.1186/s13567-017-0426-5>.
- Wiggans, G. R., T. S. Sonstegard, P. M. VanRaden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, F. S. Schenkel, and C. P. Van Tassel. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92:3431–3436. <https://doi.org/10.3168/jds.2008-1758>.
- Witke, W., J. D. Sutherland, A. Sharpe, M. Arai, and D. J. Kwiatkowski. 2001. Profilin I is essential for cell survival and cell division in early mouse development. *Proc. Natl. Acad. Sci. USA* 98:3832–3836. <https://doi.org/10.1073/pnas.051515498>.

- Wright, P. W., L. C. Bolling, M. E. Calvert, O. F. Sarmiento, E. V. Berkeley, M. C. Shea, Z. Hao, F. C. Jayes, L. A. Bush, J. Shetty, A. N. Shore, P. P. Reddi, K. S. Tung, E. Samy, M. M. Allietta, N. E. Sherman, J. C. Herr, and S. A. Coonrod. 2003. ePAD, an oocyte and early embryo-abundant peptidylarginine deiminase-like protein that localizes to egg cytoplasmic sheets. *Dev. Biol.* 256:74–89. [https://doi.org/10.1016/S0012-1606\(02\)00126-4](https://doi.org/10.1016/S0012-1606(02)00126-4).
- Wu, D., J. Dou, X. Chai, C. Bellis, A. Wilm, C. C. Shih, W. W. J. Soon, N. Bertin, C. B. Lin, C. C. Khor, M. DeGiorgio, S. Cheng, L. Bao, N. Karnani, W. Y. K. Hwang, S. Davila, P. Tan, A. Shabir, A. Moh, E.-K. Tan, J. N. Foo, L. L. Goh, K. P. Leong, R. S. Y. Foo, C. S. P. Lam, A. M. Richards, C.-Y. Cheng, T. Aung, T. Y. Wong, H. H. Ng, J. Liu, C. Wang, M. A. Ackers-Johnson, E. Aliwarga, K. H. K. Ban, D. Bertrand, J. C. Chambers, D. L. H. Chan, C. X. L. Chan, M. L. Chee, M. L. Chee, P. Chen, Y. Chen, E. G. Y. Chew, W. J. Chew, L. H. Y. Chiam, J. P. C. Chong, I. Chua, S. A. Cook, W. Dai, R. Dorajoo, C.-S. Foo, R. S. M. Goh, A. M. Hillmer, I. D. Irwan, F. Jaufeerally, A. Javed, J. Jeyakani, J. T. H. Koh, J. Y. Koh, P. Krishnaswamy, J. L. Kuan, N. Kumari, A. S. Lee, S. E. Lee, S. Lee, Y. L. Lee, S. T. Leong, Z. Li, P. Y. Li, J. X. Liew, O. W. Liew, S. C. Lim, W. K. Lim, C. W. Lim, T. B. Lim, C. K. Lim, S. Y. Loh, A. W. Lok, C. W. L. Chin, S. Majithia, S. Maurer-Stroh, W. Y. Meah, S. Q. Mok, N. Nargarajan, P. Ng, S. B. Ng, Z. Ng, J. Y. X. Ng, E. Ng, S. L. Ng, S. Nusinovic, C. T. Ong, B. Pan, V. Pedergrana, S. Poh, S. Prabhakar, K. M. Prakash, I. Quek, C. Sabanayagam, W. Q. See, Y. Y. Sia, X. Sim, W. C. Sim, J. So, D. K. N. Soon, E. S. Tai, N. Y. Tan, L. C. S. Tan, H. C. Tan, W. L. W. Tan, M. Tandiono, A. Tay, S. Thakur, Y. C. Tham, Z. Tiang, G. L.-X. Toh, P. K. Tsai, L. Veeravalli, C. S. Verma, L. Wang, M. R. Wang, W.-C. Wong, Z. Xie, K. K. Yeo, L. Zhang, W. Zhai, and Y. Zhao. 2019. Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. *Cell* 179:736–749.e15. <https://doi.org/10.1016/j.cell.2019.09.019>.
- Yin, H., L. Fang, C. Qin, and S. Zhang. 2019a. Estimation of the genetic parameters for semen traits in Chinese Holstein bulls. *BMC Genet.* 20:51. <https://doi.org/10.1186/s12863-019-0752-x>.
- Yin, H., C. Zhou, S. Shi, L. Fang, J. Liu, D. Sun, L. Jiang, and S. Zhang. 2019b. Weighted single-step genome-wide association study of semen traits in Holstein bulls of China. *Front. Genet.* 10:1053. <https://doi.org/10.3389/fgene.2019.01053>.
- Zenobia, C., and G. Hajishengallis. 2015. Basic biology and role of interleukin-17 in immunity and inflammation. *Periodontol.* 2000 69:142–159. <https://doi.org/10.1111/prd.12083>.
- Zhang, J., F. Liu, J.C. Reif, and Y. Jiang. 2021. On the use of GBLUP and its extension for GWAS with additive and epistatic effects. *G3 Genes|Genomes|Genetics* 11:7. <https://doi.org/10.1093/g3journal/jkab122>.
- Zhang, T., J. Oatley, V. J. Bardwell, and D. Zarkower. 2016. DMRT1 is required for mouse spermatogonial stem cell maintenance and replenishment. *PLoS Genet.* 12:e1006293. <https://doi.org/10.1371/journal.pgen.1006293>.