

University of São Paulo  
“Luiz de Queiroz” College of Agriculture

Genome-wide association study for reproduction traits in a Brazilian Holstein  
population

**Graciela Vieira Paulin**

Dissertation presented to obtain the degree of Master in  
Science. Area: Animal Science and Pastures

Piracicaba  
2022

Graciela Vieira Paulin  
Zootecnista

Genome-wide association study for reproduction traits in a Brazilian Holstein  
population

Advisor:  
Prof. Dr. **LUIZ LEHMANN COUTUNHO**

Dissertation presented to obtain the degree of Master in  
Science. Area: Animal Science and Pastures

Piracicaba

2022

**Dados Internacionais de Catalogação na Publicação  
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP**

Paulin, Graciela Vieira

Genome-wide association study for reproduction traits in a Brazilian Holstein population / Graciela Vieira Paulin. - - Piracicaba, 2022.

31 p.

Dissertação - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Bovinocultura de leite 2. Características reprodutivas 3. Raça holandesa 4. GWAS 5. Interação gênica I. Título

## DEDICATION

To my parents, Maria Isabel e José Carlos, for all their devotion and support.

To my sister, Glaucia, for always incentivizing me to become a better version of myself.

To my godparents, Angela and Joaquim, for always being by my side.

To my boyfriend, Guilherme, for always motivating and giving me strength in every single moment.

## ACKNOWLEDGEMENTS

To everyone that contributed somehow during this journey, especially God, for my life.

To my parents, José Carlos and Maria Isabel; to my sister, Gláucia, for supporting my choices.

To my grandparents, Angela and Joaquim, for being by my side since ever.

To my boyfriend, Guilherme, for being with me in all moments, supporting and always willing to help.

To my advisor, Dr. Luiz Lehmann Coutinho, for the opportunity, patience, and support on my decisions.

To Dr. Victor Pedrosa and Dr. Laiza Lung, for the companionship and availability, to assist in the elaboration of this work.

To CRV for sharing the data required for the development of this project and for believing in this project.

To the ESALQ/USP for the opportunity and structure.

To the employees of ESALQ/USP who gave support in all the procedures for completing this work.

To all the laboratory colleagues who have always supported me, and have always been ready to help me with my project.

## SUMMARY

RESUMO .....	6
ABSTRACT .....	7
1. BACKGROUND .....	9
1.1. Objectives .....	10
References .....	10
2. GENOME-WIDE ASSOCIATION STUDY FOR REPRODUCTIVE TRAITS IN A BRAZILIAN HOLSTEIN POPULATION .....	13
Abstract.....	13
2.1. Introduction .....	13
2.2. Methods .....	14
2.2.1. Phenotypic measurement.....	15
2.2.2. Genotypic data selection.....	15
2.2.3. Statistical analysis .....	15
2.3. Results .....	17
2.4. Discussion.....	21
2.5. Conclusion.....	23
References .....	23
3. CONCLUSIONS .....	31

## RESUMO

### **Estudo de associação genômica de características reprodutivas em uma população de raça holandesa**

O Brasil ocupa uma posição importante dentro da cadeia produtiva do leite no mundo, e a eficiência na produção vem sendo discutida em diversas áreas dentro dessa cadeia, entre elas a eficiência reprodutiva. Estudo de melhoria e para o melhor entendimento do manejo, da alimentação, da sanidade e em termos genéticos tem sido aplicados para compreender como otimizar a produção leiteira e torná-la cada vez mais rentável para o produtor. Características reprodutivas fenotípicas como idade ao primeiro parto, intervalo entre partos e a quantidade de inseminações necessárias para vaca ficar prenha são parâmetros para entender e monitorar a atividade reprodutiva de um rebanho. O ideal é que a vaca seja precoce e que ela tenha o menor intervalo entre partos possível. Já se sabe que características reprodutivas tem baixa herdabilidade, sofrendo grande influência do ambiente. Nessa linha se vê necessário conhecer os genes que estão envolvidos nesse processo para se ter cada vez mais clareza. Com os estudos de associação genômica e os avanços da biologia molecular é possível explicar o fundo genético dessas características e com o uso de marcadores moleculares encontramos regiões responsáveis pelas origens da variação biológica que nos mostram essas características. Dessa forma, o objetivo desse estudo foi identificar regiões genômicas associadas à idade ao primeiro parto (IPP) e ao intervalo entre partos (IP) em uma população de bovinos da raça Holandesa. Para identificar as regiões que possivelmente alteram o fenótipo, o GWAS foi aplicado no banco de dados de dados fenotípicos de (IPP) e intervalo entre partos (IP) de 39.336 filhas e genótipos de 555 touros em 187 rebanhos diferentes. Os dados passaram por um controle de qualidade e ao final restaram 75.925 SNPs. Os softwares da família BLUPF90 foram utilizados para análise estatística e o modelo animal para pedigree conhecido, com as janelas fixas de 10 SNPs e a variância explicada maior que 1,5 %. Os genes encontrados totalizaram 34 genes, sendo eles: 24 genes relacionados à IPP e 10 genes para IP. Os principais genes encontrados foram APP, PRSS16 e BTN1A1. Esses genes são ligados à resposta imune e podem interferir na resposta do animal à intemperança pós-parto e na recuperação do sistema reprodutivo. Os genes CYR1, PRSS16 e WEE2 são conhecidos por estarem relacionados à embriogênese, relacionando-os às características reprodutivas dos animais. Alguns dos genes identificados no presente estudo foram associados as características reprodutivas e respostas imunológicas dos animais, corroborando com a hipótese de uma expressão genética diferente nos animais estudados e indicando a necessidade o desenvolvimento de novos estudos nesta área.

Palavras-chave: Bovinocultura de leite, Características reprodutivas, Raça holandesa, GWAS, Interação gênica



## ABSTRACT

### **Genome-wide association study for reproduction traits in a Brazilian Holstein population**

Brazil is in an important position in the world of the dairy productive chain, and the efficiency of the production has been discussed in different fields in this chain, among those, reproductive efficiency. Studies on improvement for better comprehension of the management, dietary intake, sanity, and genetic terms, have been applied to comprehend ways to optimize dairy production and turn it more and more profitable for the producer. Reproductive phenotypic characteristics such as age at first calving (AFC), calving interval (CI), and quantity of inseminations needed to impregnate the cow are parameters for understanding and monitoring the reproductive activity of a herd. The ideal is that the cow is as precocious and has the smaller calving interval as possible. It is known that reproductive characteristics have low heritability, undergoing great environmental influence. In this line, it is necessary to know the genes that are involved in this process to have more and more clarity. With genomic association studies and advances in molecular biology, it is possible to explain the genetic background of these characteristics and with the use of molecular markers, we find regions responsible for the origins of the biological variation that show us these characteristics. Thus, this study aimed to identify genomic regions associated with age at first calving (AFC), to calving interval (CI), in a population of Holstein cattle. To identify regions that possibly alter the phenotype, GWAS was applied to the database of phenotypic data of AFC, IC of 39,336 daughters, and genotype of 555 bulls, in 187 different herds, contemporary groups were formed and analyzed in GWAS. Data underwent quality control and by the end, there were 75,925 SNPs. BLUPF90 family software was used for statistical analysis and known pedigree animal model, with 10 SNPs fixed windows, and explained variance over 1.5%. A total of 34 genes were identified: 24 genes linked to AFC and 10 genes linked to CI. The main genes were APP, PRSS16 e BTN1A1, once linked to the animal immune response to the postpartum intemperance and to the reproductive system recovery. The genes CYYR1, PRSS16 e WEE2 are known to be related to embryogenesis, linking them to the reproductive traits of the animals. Some of the identified genes, in the present study, were associated with reproductive and immunological characteristics of the animals, corroborating the hypothesis of a different genetic expression among the investigated animals, and indicating the need for the development of new studies in this field.

Keywords: Dairy cattle, Reproductive traits, Holstein breed, GWAS, Gene interaction



## 1. BACKGROUND

Brazil is currently the fourth largest milk producer in the world, with record production in 2020 - 25,53 billion liters inspected, and with the activity spread all over the country, in all Brazilian states (Anuario do Leite, 2021). According to IBGE (2021) (Brazilian Geography and Statistics Institute, Instituto Brasileiro de Geografia e Estatística), the production of milk per cow in 2020 increased, achieving 35,4 billion liters, a 1,5% increase in comparison to the previous year, an increase attributed to the individual gain increase, once the effective of 16.2 million cows milked was 0.8% lower than the previous year.

These numbers show an increase in the amount of produced milk and a reduction in the amount of milked cows, resulting in more efficient and productive animals, partially attributed to the adoption of biotechnologies allied to the genetical enhancement and activity management (Barkema et al., 2015; LeBlanc, 2010; Lucy, 2001; Ramírez-Rivera et al., 2019). Marestone (2013) states that the technologies applied to the agricultural field influence some factors, including reproductive efficiency, nutritional and sanitary factors, and estrus identification. These factors impact directly the productivity of the dairy cattle, besides their contribution to the delay in the return to ovarian activity in postpartum, consequently, with larger service time and intervals between calvings, reduction in lactation period, and lower production of calves per year and across cows lifetime. That said, management and reproductive efficiency are capable of controlling the calvings, which are influenced by many reproductive traits, as such: age at first calving (AFC), calving interval (CI), and numbers of inseminations to pregnancy (Nunes, 2018). Among the main factors responsible for the low efficiency of the Brazilian dairy cattle are those related to the genetic effects, environment and management. To obtain the maximum production of milk, animals might initiate their reproductive life the earliest as possible, have a service of 120 days, and a CI of 80.5 days (Baimishev, 2018).

Age at first calving and calving interval are practical measures that translate the reproductive efficiency of cattle. Animals that calve for the first time at older ages and have larger CI present lower lifetime production, thus, these characteristics have a great economic impact on the farm (Atashi, 2021). Eastham, et al. (2018), states that heifers with lower AFC are also more fertile in later stages, and can have CI shortened, enhancing the importance of reproductive management on the farm. A high level of production, in this sense, is associated with high reproductive efficiency, turning the selection of reproductive characteristics of great importance. In genetic terms, reproductive traits usually present low heritability (Shao et al., 2021), so the adoption of modern technologies of enhancement and better techniques of management for a more effective production is fundamental for milk production increase per cow (Mwanga, 2018).

Recent advances in molecular technologies can help with the comprehension of these characteristics and an alternative is the usage of molecular markers as single nucleotide polymorphisms (SNPs). The technological innovations in this field can be considered an outbreak, once it enable methodologies that allows identifying and genotyping SNP type markers- a molecular marker used in phenotype-genotype association studies (Lung, 2014).

In line with that, and with aims to increase the precision of the estimations of genetic values, molecular information can be added to the evaluation models, which tend to elevate the accuracy of the estimative. Moreover, interest genes can be identified by genome-wide association studies (GWAS), which can be incorporated into genomic value estimation processes (Calderón-Chagoya et al., 2019; Jiang et al., 2019; Tam, 2019).

The genomic best linear unbiased prediction single-step method (GBLUP or ssGBLUP) is considered simple and efficient. This method takes into account phenotypic, genealogical, and genomic data simultaneously. The use of extra information from non-genotyped animals and the independency of pseudo-phenotypes is responsible for

precision gain over other genomic methods. (Legarra et al., 2014). According to Wang et al. (2012), the ssGBLUP increases both the power and the precision of the analysis, not impacting costs, once it takes advantage of other subjects' phenotypes.

So, through the genomic association analysis, among which, the ssGBLUP, various genomic regions have been associated with fertility characteristics of females in dairy herds, which assist the comprehension of the molecular mechanisms linked to the physiological development of bovine reproduction (Berry et al., 2014; Yudin and Voevoda, 2015).

## 1.1. Objectives

To identify genomic regions associated with age at first calving and calving interval in a population of Holstein cattle.

## References

- Atashi, H., Asaadi, A., & Hostens, M. (2021). Association between age at first calving and lactation performance, lactation curve, calving interval, calf birth weight, and dystocia in Holstein dairy cows. *PLOS ONE*, *16*(1), e0244825. <https://doi.org/10.1371/journal.pone.0244825>
- Baimishev, K., Baimishev, M., Grigoryev, V., Kokhanov, A., Uskova, I., & Khakimov, I. (2018). Increase in reproductive ability of high-producing cows, and qualitative parameters of their offspring, under conditions of intensive milk production. *Asian Pacific Journal of Reproduction*, *7*(4), 167. <https://doi.org/10.4103/2305-0500.237054>
- Barkema, H. W., von Keyserlingk, M. A. G., Kastelic, J. P., Lam, T. J. G. M., Luby, C., Roy, J.-P., LeBlanc, S. J., Keefe, G. P., & Kelton, D. F. (2015). Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *Journal of Dairy Science*, *98*(11), 7426–7445. <https://doi.org/10.3168/jds.2015-9377>
- Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal*, *8*, 105–121. <https://doi.org/10.1017/S1751731114000743>
- Calderón-Chagoya, R., Hernandez-Medrano, J. H., Ruiz-López, F. J., Garcia-Ruiz, A., Vega-Murillo, V. E., Montano-Bermudez, M., Arechavaleta-Velasco, M. E., Gonzalez-Padilla, E., Mejia-Melchor, E. I., Saunders, N., Bonilla-Cardenas, J. A., Garnsworthy, P. C., & Román-Ponce, S. I. (2019). Genome-Wide Association Studies for Methane Production in Dairy Cattle. *Genes*, *10*(12), 995. <https://doi.org/10.3390/genes10120995>
- Eastham, N. T., Coates, A., Cripps, P., Richardson, H., Smith, R., & Oikonomou, G. (2018). Associations between age at first calving and subsequent lactation performance in UK Holstein and Holstein-Friesian dairy cows. *PLOS ONE*, *13*(6), e0197764. <https://doi.org/10.1371/journal.pone.0197764>
- EMBRAPA. Anuário LEITE 2021. São Paulo, 2021. Available in: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1132875/anuario-leite-2021-saude-unica-e-total>
- EMBRAPA. Anuário LEITE 2020. São Paulo, 2020. Available in: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1124722/anuario-leite-2020-leite-de-vacas-felizes>

- Instituto Brasileiro de Geografia e Estatística. (2020). Produção da Pecuária Municipal 2020. Available in: [https://biblioteca.ibge.gov.br/visualizacao/periodicos/84/ppm\\_2020\\_v48\\_br\\_informativo.pdf](https://biblioteca.ibge.gov.br/visualizacao/periodicos/84/ppm_2020_v48_br_informativo.pdf)
- Iung, L. H. de S. (2014). *Associação aditiva, dominante e epistática de SNPs à produção e composição do leite em vacas da raça Holandesa* [Universidade de São Paulo]. <https://doi.org/10.11606/D.11.2014.tde-08082014-171555>
- Jiang, J., Ma, L., Prakapenka, D., VanRaden, P. M., Cole, J. B., & Da, Y. (2019). A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle. *Frontiers in Genetics*, 10. <https://doi.org/10.3389/fgene.2019.00412>
- LeBlanc, S. (2010). Assessing the Association of the Level of Milk Production with Reproductive Performance in Dairy Cattle. *Journal of Reproduction and Development*, 56(S), S1–S7. <https://doi.org/10.1262/jrd.1056S01>
- Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, 166, 54–65. <https://doi.org/10.1016/j.livsci.2014.04.029>
- Lucy, M. C. (2001). Reproductive Loss in High-Producing Dairy Cattle: Where Will It End? *Journal of Dairy Science*, 84(6), 1277–1293. [https://doi.org/10.3168/jds.S0022-0302\(01\)70158-0](https://doi.org/10.3168/jds.S0022-0302(01)70158-0)
- Marestone, B. S., Santos, É. R. dos, Serra, F. B. S., Muniz, C. A. D. S. D., Marques, C. P., Alves, K. B., Alves, M. V., & Alves, R. C. M. (2013). Características reprodutivas, de crescimento e idade ao primeiro parto em bovinos da raça Holandesa. *Semina: Ciências Agrárias*, 34(6Supl2), 4105. <https://doi.org/10.5433/1679-0359.2013v34n6Supl2p4105>
- Mwanga, G., Mujibi, F. D. N., Yonah, Z. O., & Chagunda, M. G. G. (2019). Multi-country investigation of factors influencing breeding decisions by smallholder dairy farmers in sub-Saharan Africa. *Tropical Animal Health and Production*, 51(2), 395–409. <https://doi.org/10.1007/s11250-018-1703-7>
- Nunes, M. A. (2018). *Análises estruturais para componentes de características do leite e índices reprodutivos em bubalinos leiteiros*. Universidade Federal da Grande Dourados.
- Ramírez-Rivera, E. J., Rodríguez-Miranda, J., Huerta-Mora, I. R., Cárdenas-Cágal, A., & Juárez-Barrientos, J. M. (2019). Tropical milk production systems and milk quality: a review. *Tropical Animal Health and Production*, 51(6), 1295–1305. <https://doi.org/10.1007/s11250-019-01922-1>
- Shao, B., Sun, H., Ahmad, M. J., Ghanem, N., Abdel-Shafy, H., Du, C., Deng, T., Mansoor, S., Zhou, Y., Yang, Y., Zhang, S., Yang, L., & Hua, G. (2021). Genetic Features of Reproductive Traits in Bovine and Buffalo: Lessons From Bovine to Buffalo. *Frontiers in Genetics*, 12. <https://doi.org/10.3389/fgene.2021.617128>
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, 20(8), 467–484. <https://doi.org/10.1038/s41576-019-0127-1>
- Wang, H., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, 94(2), 73–83. <https://doi.org/10.1017/S0016672312000274>



## 2. GENOME-WIDE ASSOCIATION STUDY FOR REPRODUCTIVE TRAITS IN A BRAZILIAN HOLSTEIN POPULATION

### Abstract

**Background:** Genome association studies have contributed to the knowledge of a variety of interesting traits in animal production. The present study aimed to evaluate aspects of reproductive traits to optimize the efficiency of milk production. For these purposes, we investigated the associations of the genotype of bulls with reproductive phenotypic, data of age at first calving (AFC), and calving interval (CI), all known as characteristics with low heritability. **Methods:** To identify the regions that possibly alter the phenotype, GWAS was applied to the Brazilian database of phenotypic data on age at first calving (AFC) and calving interval (CI) of 39,336 daughters and genotypes of 555 bulls in 187 different herds. Contemporary groups were formed to analyze GWAS. After quality control 75,925 SNPs remained for further analysis. Then, by using a software from the BLUPF90 family, statistical analyses were performed. The analyses considered an animal model for known pedigree, with the window size of 10 SNPs. The genomic windows that explained more than 1.5% were considered. Some quantitative trait locus regions associated to both traits here investigated were found, 24 genes related to AFC and 10 genes for CI. DAVID was used to verify the functional enrichment of the identified genes. **Results:** APP, PRSS16, and BTN1A1, were identified, Genlinked to immune response and may interfere in the animal's response to the postpartum intertemperate and recovery of the reproductive system. CYYR1, PRSS16, and WEE2 genes were identified in the present study, and are known to be related to embryogenesis, linking them to the reproductive characteristics of the animals. **Conclusions:** Some of the genes identified in the present study have been previously associated with reproductive and immunological aspects of animals, thus, supporting the hypothesis of a different genetic expression in the studied animals, and corroborating the development of further studies in this area

Keywords: genomic evaluation; single-step GBLUP; reproductive trait; milk production efficiency.

### 2.1. Introduction

Dairy producers look for high production and quality in milk as they are the characteristics that generate revenue and profitability on the farm. Marestone et al. (2013) state that technologies applied to the agricultural environment influence some factors such as reproductive efficiency, directly impacting the productivity of the dairy herd. Elmetwally (2018) said reproductive problems contribute to the delay in the return to ovarian activity in the puerperium resulting in a longer service time and calving interval (CI), in addition to reduced lactation with lower production of calves per year per cow, for to have an efficient production in dairy farming is necessary to have a healthy reproduction. The age at first calving, the interval between calving, and the number of inseminations until reaching pregnancy are parameters used to monitor the reproductive efficiency of a herd as they influence births and milk production. Proper management and reproductive efficiency allow parturition to be controlled (Nunes, 2018). In summary, the decline in fertility reduces the percentage of cows in their production period, which reduces the herd's milk production (Norman et al., 2009).

The main factors responsible for the low reproductive efficiency of dairy cattle herds are those related to genetic, environmental, and management effects, which are commonly reflected in an advanced age at first calving (AFC) (Cerón-Muñoz et al., 2004). The animals must start their reproductive life as early as possible and have a service period of 120 days and a calving interval of 80.5 months (Baimishev et al., 2018). Age at first calving (AFC), defined as the period when a calf reaches sexual maturity and can reproduce for the first time, is an important factor in the

cost in dairy herds (Atashi et al., 2021), as advanced age at first calving (AFC) and long calving intervals (CI), reduce the lifetime production of animals and, consequently, decrease the profitability of the dairy activity.

Holstein is the most used breed in the production of milk in the world. This breed has significant advantages in milk production and some disadvantages, especially in relation to its fitness traits such as the low heritability of reproductive traits, such as fertility. Producers have used crosses to test heterosis and arrive at animals that are more adapted and suitable for each production system (Brade, 2012). The Holstein breed raised in a tropical environment suffers a higher energy expenditure to maintain homeothermy and this can harm production and reproduction. Knowing the correct handling techniques and the genomic influences on the traits it is possible to optimize the production and reproduction of these animals (Rezende et al., 2015).

Eastham et al. (2018) state that heifers that conceive earlier and give birth at a younger age are also more fertile in later stages and may have a shorter calving interval (CI) length, confirming the importance of reproductive management in the profitability of the farm. However, Yokoo et al. (2012) consider that the reduction of age at first calving (AFC) and the calving interval (CI) can significantly affect herd productivity due to its influence on calf production during the life of the dam and on the female's ability to re-conceive as soon as possible after the first delivery. In tropical regions, Holstein animals suffer from the adversities of the climate, so it is important to carefully monitor the reproductive indices regardless of the production system, as Bignardi et al. (2015) concluded heritability increases as the environment becomes more favorable.

A high level of production is associated with high reproductive efficiency, making trait selection for traits important. In genetic terms, assessments of reproduction traits usually show low heritability (Asadollahi et al., 2020), indicating that the use of modern breeding technologies, and improving management practices to obtain a more effective production are of fundamental importance for increasing milk production per cow (Mwanga et al., 2019). Recent advances in molecular technologies can help understand the behavior of reproductive traits and another methodology that works for the same purpose is the use of molecular markers for single nucleotide polymorphism (SNP).

Technological innovations represent a true revolution and have brought methodologies to identify and genotype SNP markers (Single Nucleotide Polymorphism), a type of molecular marker used in phenotype-genotype association studies (L. H. de S. Iung, 2014), allowing access to innovative information. SNPs can be used in genome-wide association studies (GWAS), and this strategy is well established for identifying genomic regions related to traits of interest in distinct species.

This study aimed to identify genomic regions associated with age at first calving (AFC) and calving interval (CI) in a population of Holstein cattle.

## 2.2. Methods

The study was conducted using a dataset of Brazilian Holstein cows obtained from CRV Brazil and bull's genomic data from Cooperative CRV headquarters in the Netherlands. Reproductive data of 39,336 daughters of 555 bulls from 187 herds were considered for the present study analysis. The inclusion of daughters was conditioned to the sires with at least 5 previous progenies.

In this study, data were obtained from 187 herds, composed by 39336 daughters of 555 bulls, all Brazilian Holstein.



### 2.2.1. Phenotypic measurement

Data for phenotypic measures were obtained from the daughters. There were 43,281 records for age at first calving (AFC) with an average of 26.81 months, and calving interval (CI) from the first to the second calving, with an average of 410 days.

Contemporary groups (CG) were formed by concatenating herd; calving season, defined as dry (from April to September) and (rainy from October to March); and calving year, considering only the first calving of each cow. For AFC, 2,103 CG were formed, containing from five to 219 cows. For CI, 2,137 CG were formed, containing from five to 197 cows.

### 2.2.2. Genotypic data selection

A custom Illumina 60K chip and Illumina BovineSNP50 BeadChip were used to obtain the genotypic data of the 555 bulls (de Roos et al., 2009), which were imputed as described by (Druet et al. (2010)). A total number of 76,438 SNPs and 555 bulls were retained for further analysis after quality control. The quality control was performed according to Iung et al., (2019), by removing SNPs with a minor allele frequency (MAF) lower than 0.02 and monomorphic ( $n = 513$ ). No SNPs were removed due to the call rate  $< 90\%$ . In the end, 555 animals and 75,925 SNP remained for the analysis.

### 2.2.3. Statistical analysis

To select only the biologically acceptable data, outliers were removed resulting in the calving interval between 300 to 600 days only from the first to the second calving with a mean of 409 days and a standard deviation of 69,8 days. For age at the first calving, the acceptable data was from 21 to 39 months with a mean of 26,8 months and a standard deviation of 3,42 months.

The variance components were estimated in the AIREMLF90 software from the BLUPF90 (Miszta et al., 2002) program family. For AFC, the model included the CG as the fixed effect of group and animal, and residual as random effects. For CI, the same model was used with the inclusion of age at measurement as a covariate. Estimations were based on the restricted maximum likelihood (REML) method, using the Expectation-Maximization (EM) algorithm in the first ten iterations and the Average Information (AI) algorithm in the others.

The GWAS analyses were based on the single-step GBLUP methodology (Wang et al., 2012), using programs of the BLUPF90 family (Miszta et al, 2014). Estimates for the variance components and genetic variance were obtained with AIREMLF90 software, using a convergence criterion of 10-12 to determine convergence of the estimates with the idea of constructing a hybrid genomic relationship matrix for both genotyped and non-genotyped animals; Subsequently, PREGSF90 was used, and BLUPF90 adopted to solve the mixed model equations. The “SNP moving average” option from the postGSF90 package (Aguilar et al., 2014) was used to back-solve the genomic estimated breeding values and to obtain SNP effects for each trait separately. The window size of 10 SNPs was defined based on the average linkage disequilibrium decay in Holstein cattle and on the density of the SNP panel used, following similar studies (Atashi et al., 2020; Nayeri et al., 2016; Qanbari et al., 2009).

AFC and CI were analyzed using an animal model, described as:

$$y = Xb + Za + e$$

where  $y$  is the vector of phenotypic observations for genotyped and non-genotyped animals;  $X$  is the incidence matrix linking the phenotypic records to the fixed effects;  $b$  is the vector of fixed effects, which included age at the measurement only for CI and the contemporary group;  $Z$  is the incidence matrix linking the phenotypic records to each animal;  $a$  is the vector of animal additive genetic effects; and  $e$  is the vector of residual effects. The inverse of the  $H$  matrix can be described as (Aguilar et al., 2010):

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A^{-1}_{22} \end{bmatrix}$$

where  $A$  is the pedigree relationship matrix for all animals;  $A_{22}$  is the relationship matrix for the genotyped animals; and  $G$  is the genomic relationship matrix, which was calculated as (VanRaden, 2008):

$$G = WDW' \frac{\mathbf{1}}{q}$$

where  $W$  is a matrix of gene content adjusted for allele frequencies,  $D$  is a weight matrix for SNP (initially  $D = I$ ), and  $q = \sum_{i=1}^M 2p_i(1-p_i)$  is a normalizing factor.

The SNP effects as well as their weights were calculated in two iterations, as suggested by Lourenço et al. (2020). The iterations were performed as proposed by Wang et al. (2012):

$$D(t) = I$$

$$G_{(t)} = \frac{WD_{(t)}W'}{\sum_{i=1}^M 2p_i(1-p_i)}$$

where  $t$  is the iteration number.

The SNP effects ( $\hat{u}$ ) were obtained as:

$$\hat{u} = \lambda DW'G^{-1}\hat{a}g = DW'[WDW']^{-1}\hat{a}g$$

where  $\hat{a}g$  is a vector of the animal effects of genotyped animals, which was represented by a function of the SNP effects ( $\hat{a}g = Wu$ );  $W$  is the matrix containing the genotypes for each locus;  $\hat{u}$  is the vector of SNP effects;  $\lambda$  is the variance ratio, calculated according to VanRaden, (2008)

$D$  is the diagonal matrix of weights of the SNP variances; and  $G$  is the genomic relationship matrix, constructed as described above. The following equation was applied to calculate the weights used for the SNPs:

$$d_{i(t=1)} = \hat{u}_{i(t)}^2 / 2p_i(1-p_i)$$

where  $i = \text{SNP } i$ .

Lastly, the program calculates  $G$  with the new marker weights as:

$$G_{(t+1)} = \frac{WD_{(t+1)} W^2}{\sum_{i=1}^M 2p_i(1-p_i)}$$

The ssGWAS results were presented as a proportion of the total additive genetic variance explained by the genomic windows of 10 adjacent SNPs, as follows:

$$\frac{\text{var}(\text{window}_i)}{\sigma_a^2} \times 100\% = \frac{\sum_{j=1}^{10} \text{var}(\hat{u}_j)}{\sigma_a^2} \times 100\%$$

where windows are the proportion of the total additive genetic variance explained by the genomic window  $i$ ;  $\sigma_a^2$  is the total additive genetic variance for the trait, and  $\hat{u}_j$  is the effect of SNP  $j$  within genomic window  $i$ .

### 2.2.3.1. Gene annotation

After performing ssGWAS, the genomic windows that explained more than 1.5% were considered (Guarini et al., 2019; Marques et al., 2018). Using the Ensemble Genes 69 database incorporated in the Bio Mart tool (Kinsella et al., 2011) it was possible to identify the positional candidate genes located in the selected regions. Subsequently, the DAVID Bioinformatics database (Huang et al., 2007) was used to perform the functional annotations.

## 2.3. Results

In this study, outliers were disregarded and only data from daughters of bulls with at least five progenies were kept, totalizing 39,336 daughters of 555 bulls in 187 Brazilian Holstein herds. The final values are described in Table 1., The AFC range was from 21 months to 39 months, in a pool of 39,156 data. CI ranged from 300 to 600 days, based on 39,336 data obtained.

**Table 1.** Summary statistics of cows used in genetic analyses

Trait	N	Mean	Standard Deviation	Minimum	Maximum
AFC	39,156	26.81	3.42	21.00	39.00
CI	39,336	409.99	69.80	300.00	600.00

Abbreviation(s): age first calving (AFC), calving interval (CI), standard deviation (STD Dev), minimum (MIN) and maximum (MAX) obtained for the studied traits

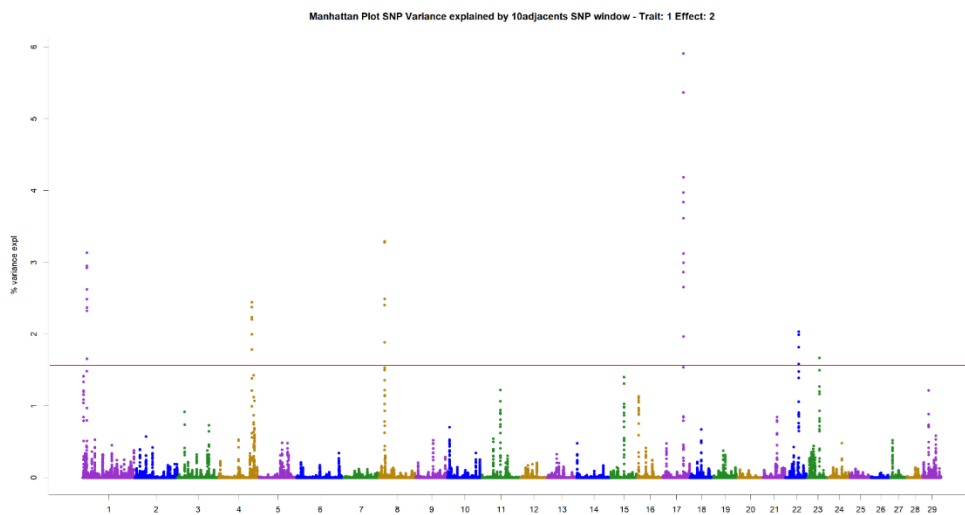
Estimates of heritability and additive, residual, and phenotypic variances of two traits are depicted in Table 2. In this work, we focused on age at first calving and calving interval because as observed in Table 2, heritability estimates were low for CI and moderate for AFC.

**Table 2.** Estimates of heritability and variances of two traits

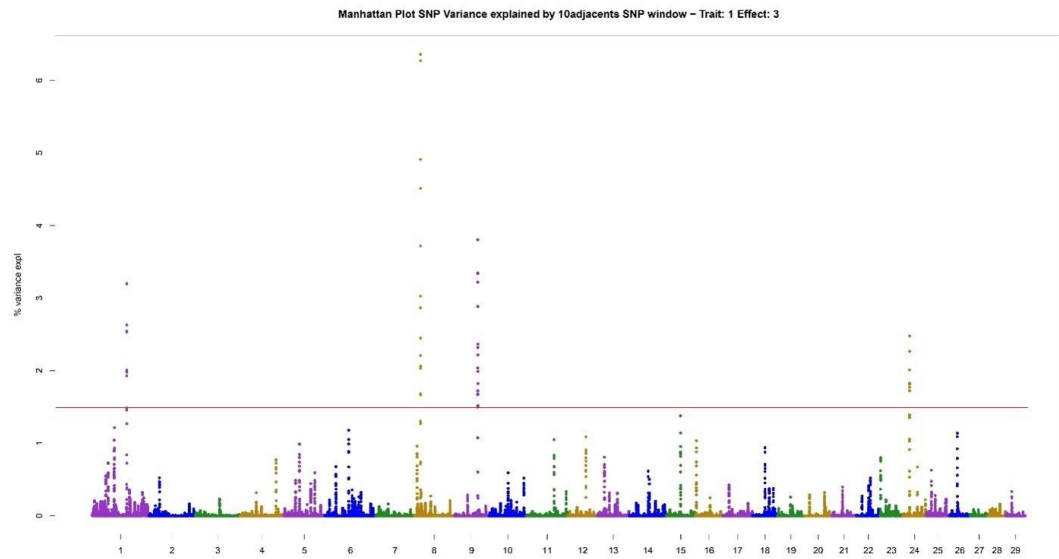
Trait	Variances components			
	Additive (Standard errors)	Residual (Standard errors)	Phenotypic (Standard errors)	Heritability (Standard errors)
AFC	1.01 (0.095)	6.59 (0.090)	7.60 (0.060)	0.13 (0.012)
CI	126.77 (20.282)	2523.30 (25.407)	2650.10 19.637)	0.05 (0.008)

Abbreviations: age first calving (AFC), calving interval (CI)

The GWAS analysis showed the effective chromosomes by the Manhattan plot with the percentages of additive genetic variance explained by windows of 10 adjacent SNPs for AFC and CI, respectively, represented in Figures 1 and 2. The number of relevant SNPs on chromosomes 1, 4, 9, 22, and 23 are presented in Table 3. Only genes inside windows with the highest percentage of additive genetic variance were presented, along with other technical information about the quantitative trait locus (QTL) regions.

**Figure 1.** Manhattan plot of additive genetic variance explained by windows of 10 adjacent SNPs for age first calving (AFC).

**Figure 2.** Manhattan plot of additive genetic variance explained by windows of 10 adjacent SNPs for calving interval (CI).



The quantitative trait locus (QTL) regions associated with the two traits studied had a chromosome in common, these traits were presented in chromosome 1. Moreover, 24 genes were found for AFC and 10 for CI within specific windows, as described in table 3.

**Table 3.** Quantitative trait locus (QTL) regions associated with age first calving (AFC) and calving interval (CI)

Trait	StartSNP	EndSNP	VarExp	Chr	Start position	End position	Gene
AFC	Hapmap44892-BTA-50337	Hapmap25334-BTA-160518	3,13	1	9.975.525	10.345.714	CYYR1, U6, APP
AFC	Hapmap52823-rs29010271	ARS-BFGL-NGS-36185	2,44	4	104.845.288	105.101.872	AGK, DENND11, WEE2, SSBP1, TAS2R3, TAS2R4, PRSS37
AFC	ARS-BFGL-NGS-84349	Hapmap47039-BTA-54432	2,03	22	39.318.252	39.701.144	PTPRG, bta-mir-2369
AFC	BTA-76815	SNP_Z93323_2281	1,67	23	31.269.308	31.703.880	PRSS16, H2AC12, H4C9, H2AC11, ZNF322, OR2M17, OR2M14, ABT1, HMGN4, BTN1A1, H4C8, H2AC10
CI	Hapmap49612-BTA-44210	BTB-01469883	3,20	1	100.474.672	101.006.056	U6
CI	BTA-84290	AAFC03026264_144120	3,80	9	76.290.504	76.998.832	NHSL1, CCDC28A, ECT2L, REPS1, ABRACL, HECA, TXLNB
CI	AAFC03026264_144217	BTA-90741	2,88	9	76.798.928	77.203.568	CITED2, U5, HECA, TXLNB

Abbreviations: VarExp, variance explained by window; Chr, Chromosome; StartSNP, start position; EndSNP, end position.

This list of candidate genes was submitted to DAVID, a tool for functional annotation bioinformatics microarray analysis to perform enrichment analysis to determine biological pathways associated with the genes found for each trait separately (Huang et al., 2009; Sherman et al., 2022) and the results were for AFC with p-value and for CI as suggestive areas in Table 4.

**Table 4.** Results of DAVID to perform enrichment analysis to determine biological pathways associated with the genes found for each trait separately for age first calving (AFC) with p value and for calving interval (CI) as suggestive.

TRAIT	GO	NAME	TERM	p-value
AFC	ENSBTAG00000016880	TAS2R3	detection of chemical stimulus involved in sensory perception of bitter taste	<0.001
AFC	ENSBTAG00000018440	TAS2R4	detection of chemical stimulus involved in sensory perception of bitter taste	<0.001
AFC	ENSBTAG00000054548	H2AC12	chromatin silencing	<0.001
AFC	ENSBTAG00000052895	H2AC10	chromatin silencing	<0.001
CI	ENSBTAG00000001446	ABRACL	ABRA C-terminal like (ABRACL)	-
CI	ENSBTAG00000011224	CITED2	Cbp/p300 interacting transactivator with Glu/Asp rich carboxy-terminal domain 2(CITED2)	-
CI	ENSBTAG00000000245	NHSL1	NHS like 1(NHSL1)	-
CI	ENSBTAG00000012299	REPS1	RALBP1 associated Eps domain containing 1(REPS1)	-
CI	ENSBTAG00000015516	CCDC28A	coiled-coil domain containing 28A(CCDC28A)	-
CI	ENSBTAG00000027312	ECT2L	epithelial cell transforming 2 like (ECT2L)	-
CI	ENSBTAG00000024595	HECA	hdc homolog, cell cycle regulator (HECA)	-
CI	ENSBTAG00000002964	TXLNB	taxilin beta (TXLNB)	-

Abbreviations: Trait, number GO, Name, Term and p-value; age first calving (AFC) and calving interval (CI)

## 2.4. Discussion

We found quantitative trait locus (QTL) regions associated with the two traits studied, for AFC we found 24 genes and 10 for CI within specific windows and they had a chromosome in common, these traits were presented in chromosome 1. Some genes already identified are associated with reproduction characteristics. In some cases, we found support in the literature for a direct link between AFC and CI. However, for some genes the link to the phenotypes was indirect.

The Cysteine and Tyrosine Rich 1 gene (CYR1) presents different functions, as transcriptome analysis impacting the ovoid to tubular transition and tubular to filamentous transition, in differentially expressed genes associated with cellular movement, cell-to-cell signaling, cellular assembly and organization, lipid metabolism, small molecule biochemistry and molecular transport (Ribeiro et al., 2016). The CYR1 was also found in salmon related to its down-regulation activity during embryogenesis (Tackle et al., 2004). And in another study that evaluated the effects of treatment with prostaglandin F2alpha (PGF2A) on the transcriptome responses of the bovine midcycle corpus luteum, luteolysis is essential prior to estrous cycle resumption and is a target for improving fertility (Talbott et al., 2017). In this research, Talbott et al. (2017) showed, that in ruminants CYR1 was related to the regulation of prostaglandin, which directly interferes with reproductive efficiency. That said, it is possible to infer that there is a direct influence on the traits studied in AFC, because only with the correct regulation of sex hormones, such as prostaglandin, do the heifers enter reproductive age.

In Nellore heifers, the APP gene is linked to AFC and early pregnancy, which supports that the regions found in the present study agree with what is found in the literature (Regatieri et al., 2017). Murata and Petersen (2004)

stated that the number of APPs may be associated with early diagnosis, and that is information that allows monitoring the health status of a herd. Furthermore, APPs are called acute phase proteins and when associated with antimicrobial peptides (AMPs) are immunity gauges necessary for the resolution and elimination of inflammation (Brogden et al., 2003; Hancock & Scott, 2000). This is a necessary process for normal endometrial involution, being linked to the process of restoration of a healthy endometrium for the conception of a new pregnancy in cattle. When APP is found in the serum, it is a sign of current or previous uterine inflammation (Chan et al., 2010; Huzzey et al., 2009). In line with this, Chapwanya (2012) concludes that postpartum endometrial inflammation and leukocyte infiltration sustained by increased expression of pro-inflammatory genes, antimicrobial peptides (AMP), and APP, are expected and beneficial responses to the elimination of microorganisms, consequently, restoring ideal conditions for conception.

Reproductive characteristics related to conception are of great concern for producers, because they are extremely important in both herds for meat and for milk production, as they reflect on productive and reproductive efficiency. Productive and reproductive efficiency are the main factors influencing gains or lower financial losses for the producers. These precocious animals will, thus, start generating profit sooner.

Anticipation of AFC and pregnancy interval reduction provide an earlier reduction of the investments needed by the animals, in comparison to those with late AFC and longer pregnancy intervals (Dono et al., 2013).

Yet regarding the APP protein, it has been linked to human diseases. Lim (2016) indicated notable involvement of the APP gene in human muscle degenerative function, but the factors linking these humans' diseases and animal production traits, and if they share the same pathways and sequence variations remain unclear. In bovines, APP was found and recognized as a possible quantitative trait locus (QTL) candidate for further investigation for somatic cell count (SCS) in the dairy cattle population (Kolbehdari et al., 2009).

The PRSS16, serine protease 16, was found linked to the immunological system and consequently with adaptability. Serine protease is responsible for immune cell proliferation and its downregulation in pregnancy maintains the development of an embryo (Beltman et al., 2013). According to Otto (2018), PRSS16 participates in gene interactions consistent with the biology of the immune system, providing adaptability to unpleasant environments. The PRSS16 gene encodes TSSP, a precursor associated with disease susceptibility (Serre et al., 2017), so this gene is of great interest because a disease following calving will delay the subsequent pregnancy of this cow, thus impacting the production.

BTN1A1, a butyrophilin subfamily 1 member A1, is a gene that shows a relationship with some metabolic pathways. In cattle, these metabolic pathways are related not only to milk production but also to the estimated breeding value of milk production traits and reproduction parameters (Rychtářová et al., 2014). Nonetheless, Jonchere (2010) and Smith & Xue (1997) state that BTN1A1 contains immunoglobulin-like (Ig-like) domains involved in cell-cell recognition, cell surface receptors, and immune responses such as the APP gene, and may be involved in the postpartum recovery process. In cows, this gene showed that the heat tolerance coefficient was significant ( $P < 0.05$ ) during the third month of the lactation season (Al-Whaith et al., 2019). Hu et al. (Hu et al., 2009) state that although BTN1A1 function is not completely understood, its expression profile may play a key role in lactation.

The PRSS37, a serine protease 3, has been described in several functions related to reproduction, such as fertility functions, effects in or near genome regions with a signature of selection (Yang, 2015), sperm, including genes related to fertilization (Gòdia et al., 2019), and associated with reproductive processes (Selvaraju et al., 2017).

The gene WEE2 oocyte meiosis inhibiting kinase, in humans, presents a non-hormonal contraception function, preventing fertilization of the oocyte when inhibited (Hanna et al., 2020). In cattle, Jiang et al. (2014) showed that WEE2 does not have a significant role in embryo development. The WEE2 is related to WEE1 and one of its



roles is to inhibit the cell cycle, which is cleaved during apoptosis (Jin & El-Deiry, 2005; Ptak et al., 2011), meaning that they automatically control cell wall deterioration, preventing unwanted cells from multiplying. This is interesting because it involves the control of viable cells to restore the environment when it is needed. In mice, Sang et al (2018) identified that WEE2 deficiency caused fertilization failure and female infertility and that in donkeys, it was involved in the development of the oocyte. Hanna (2010) also indicates that the overexpression of WEE2 can delay oocyte reentry into the meiosis in mice and monkeys.

The NHSL1 (Nance-Horan Syndrome-like 1) found in pathways of fertility results in the bull population (Parker Gaddis et al., 2016) and in female humans, and seemed to be associated with drug resistance in ovarian cancer (Zou et al., 2015). The NHSL1 is linked to cell migration efficiency and cell migration is essential for embryonic development (Law et al., 2021).

In the present study, some of the identified genes were found to be linked to reproduction but not directly to functions that accelerate the process of re-conception or precocity, which is expected due to the traits studied. The APP, PRSS16, and BTN1A1 genes were previously studied and were associated with the immune response, which can be associated with the animal's response to the postpartum intemperance and the rapid recovery of the reproductive system to sustain a new pregnancy. The CYYR1, PRSS16, and WEE2 genes were found in studies that relate to embryogenesis, which we can associate with reproduction and the traits studied, but with care, as more studies are needed to understand the link with the traits of economic interest studied. Most of the genes had already been found but in work with other species, which shows us that they have the potential to be explored further with bovine species.

## **2.5. Conclusion**

We identified some genomic regions that might be associated with the reproductive characteristics of cattle. Moreover, this study showed us that it is possible to increase reproductive efficiency in dairy herds, understanding more and more the importance of the postpartum recovery of a high-production cow. That said, it is important to indicate that genomic studies related to reproduction are on the right path, as some genes had already been found in other works and that the increase in reproductive efficiency is related not only to management but also to the cow's immune system and adequate postpartum recovery.

## **Acknowledgements**

The authors thank Cooperative CRV - Coöperatie Koninklijke in the Netherlands for making the bull data available and CRV Brasil for the daughter data. Also, the authors express their gratitude to the entire team of the Animal Biotechnology Laboratory of ESALQ/USP.

## **Conflict of interest**

The authors declare no conflicts of interest.

## **References**

- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2), 743–752. <https://doi.org/10.3168/jds.2009-2730>
- Aguilar, I., Misztal, I., Tsuruta, S., Legarra, A., & Wang, H. (2014). PREGSF90—POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, 1–3.
- Al-Whaith, H. K., Mahmood, F. A., & Al-Gburi, O. S. (2019). Association between BTN1A1 gene polymorphism and some reproductive efficiency indicator and heat tolerance in Holstein cows. *Plant Archives*, 19(1), 1959–1964.
- Asadollahi, S. S., Asadollahi, S. S., & Hossein-Zadeh, N. G. (2020). Twinning rate is not genetically correlated with production and reproduction traits in Iranian dairy cows. *Animal Production Science*, v. 60(11), 1402-1408–2020 v.60 no.11. <https://doi.org/10.1071/AN19336>
- Atashi, H., Asaadi, A., & Hostens, M. (2021). Association between age at first calving and lactation performance, lactation curve, calving interval, calf birth weight, and dystocia in Holstein dairy cows. *PLOS ONE*, 16(1), e0244825. <https://doi.org/10.1371/journal.pone.0244825>
- Atashi, H., Salavati, M., de Koster, J., Crowe, M., Opsomer, G., & Hostens, M. (2020). A Genome-Wide Association Study for Calving Interval in Holstein Dairy Cows Using Weighted Single-Step Genomic BLUP Approach. *Animals*, 10(3), 500. <https://doi.org/10.3390/ani10030500>
- Baimishev, K., Baimishev, M., Grigoryev, V., Kokhanov, A., Uskova, I., & Khakimov, I. (2018). Increase in reproductive ability of high-producing cows, and qualitative parameters of their offspring, under conditions of intensive milk production. *Asian Pacific Journal of Reproduction*, 7(4), 167. <https://doi.org/10.4103/2305-0500.237054>
- Beltman, M. E., Forde, N., Lonergan, P., & Crowe, M. A. (2013). Altered endometrial immune gene expression in beef heifers with retarded embryos. *Reproduction, Fertility and Development*, 25(6), 966. <https://doi.org/10.1071/RD12232>
- Bignardi, A. B., el Faro, L., Pereira, R. J., Ayres, D. R., Machado, P. F., Albuquerque, L. G. de, & Santana, M. L. (2015). Reaction norm model to describe environmental sensitivity across first lactation in dairy cattle under tropical conditions. *Tropical Animal Health and Production*, 47(7), 1405–1410. <https://doi.org/10.1007/s11250-015-0878-4>
- Brade, W. (2012). Advantages and disadvantages of cross-breeding in dairy cattle, focusing in particular on Holstein cattle. *Berichte Über Landwirtschaft*, 90(3), 380–394.
- Brogden, K. A., Ackermann, M., McCray, P. B., & Tack, B. F. (2003). Antimicrobial peptides in animals and their role in host defences. *International Journal of Antimicrobial Agents*, 22(5), 465–478. [https://doi.org/10.1016/S0924-8579\(03\)00180-8](https://doi.org/10.1016/S0924-8579(03)00180-8)
- Cerón-Muñoz, M. F., Tonhati, H., Costa, C. N., Maldonado-Estrada, J., & Rojas-Sarmiento, D. (2004). Genotype × Environment Interaction for Age at First Calving in Brazilian and Colombian Holsteins. *Journal of Dairy Science*, 87(8), 2455–2458. [https://doi.org/10.3168/jds.S0022-0302\(04\)73369-X](https://doi.org/10.3168/jds.S0022-0302(04)73369-X)
- Chan, J. P.-W., Chang, C.-C., Hsu, W.-L., Liu, W.-B., & Chen, T.-H. (2010). Association of increased serum acute-phase protein concentrations with reproductive performance in dairy cows with postpartum metritis. *Veterinary Clinical Pathology*, 39(1), 72–78. <https://doi.org/10.1111/j.1939-165X.2009.00182.x>

- Chapwanya, A., Meade, K. G., Foley, C., Narciandi, F., Evans, A. C. O., Doherty, M. L., Callanan, J. J., & O'Farrelly, C. (2012). The postpartum endometrial inflammatory response: a normal physiological event with potential implications for bovine fertility. *Reproduction, Fertility and Development*, 24(8), 1028. <https://doi.org/10.1071/RD11153>
- Druet, T., Schrooten, C., & de Roos, A. P. W. (2010). Imputation of genotypes from different single nucleotide polymorphism panels in dairy cattle. *Journal of Dairy Science*, 93(11), 5443–5454. <https://doi.org/10.3168/jds.2010-3255>
- Dono, G., Giraldo, L., & Nazzaro, E. (2013). Contribution of the calving interval to dairy farm profitability: results of a cluster analysis of FADN data for a major milk production area in southern Italy. *Spanish Journal of Agricultural Research*, 11(4), 857. <https://doi.org/10.5424/sjar/2013114-3873>
- Eastham, N. T., Coates, A., Cripps, P., Richardson, H., Smith, R., & Oikonomou, G. (2018). Associations between age at first calving and subsequent lactation performance in UK Holstein and Holstein-Friesian dairy cows. *PLOS ONE*, 13(6), e0197764. <https://doi.org/10.1371/journal.pone.0197764>
- Elmetwally, M. A. (2018). Uterine Involution and Ovarian Activity in Postpartum Holstein Dairy Cows. A Review. *Journal of Veterinary Healthcare*, 1(4), 29–40. <https://doi.org/10.14302/issn.2575-1212.jvhc-18-2447>
- Gòdia, M., Estill, M., Castelló, A., Balasch, S., Rodríguez-Gil, J. E., Krawetz, S. A., Sánchez, A., & Clop, A. (2019). A RNA-Seq Analysis to Describe the Boar Sperm Transcriptome and Its Seasonal Changes. *Frontiers in Genetics*, 10. <https://doi.org/10.3389/fgene.2019.00299>
- Guarini, A. R., Lourenco, D. A. L., Brito, L. F., Sargolzaei, M., Baes, C. F., Miglior, F., Misztal, I., & Schenkel, F. S. (2019). Genetics and genomics of reproductive disorders in Canadian Holstein cattle. *Journal of Dairy Science*, 102(2), 1341–1353. <https://doi.org/10.3168/jds.2018-15038>
- Hancock, R. E. W., & Scott, M. G. (2000). The role of antimicrobial peptides in animal defenses. *Proceedings of the National Academy of Sciences*, 97(16), 8856–8861. <https://doi.org/10.1073/pnas.97.16.8856>
- Hanna, C. B., Mudaliar, D., John, K., Allen, C. L., Sun, L., Hawkinson, J. E., Schönbrunn, E., Georg, G. I., & Jensen, J. T. (2020). Development of WEE2 kinase inhibitors as novel non-hormonal female contraceptives that target meiosis†. *Biology of Reproduction*, 103(2), 368–377. <https://doi.org/10.1093/biolre/ioaa097>
- Hanna, C. B., Yao, S., Patta, M. C., Jensen, J. T., & Wu, X. (2010). WEE2 Is an Oocyte-Specific Meiosis Inhibitor in Rhesus Macaque Monkeys1. *Biology of Reproduction*, 82(6), 1190–1197. <https://doi.org/10.1095/biolreprod.109.081984>
- Hu, H., Wang, J., Bu, D., Wei, H., Zhou, L., Li, F., & Looor, J. J. (2009). In Vitro Culture and Characterization of a Mammary Epithelial Cell Line from Chinese Holstein Dairy Cow. *PLoS ONE*, 4(11), e7636. <https://doi.org/10.1371/journal.pone.0007636>
- Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., & Lempicki, R. A. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research*, 35(suppl\_2), W169–W175. <https://doi.org/10.1093/nar/gkm415>
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57. <https://doi.org/10.1038/nprot.2008.211>
- Huzzey, J. M., Duffield, T. F., LeBlanc, S. J., Veira, D. M., Weary, D. M., & von Keyserlingk, M. A. G. (2009). Short communication: Haptoglobin as an early indicator of metritis. *Journal of Dairy Science*, 92(2), 621–625. <https://doi.org/10.3168/jds.2008-1526>

- Iung, L. H. de S. (2014). Associação aditiva, dominante e epistática de SNPs à produção e composição do leite em vacas da raça Holandesa [Universidade de São Paulo]. <https://doi.org/10.11606/D.11.2014.tde-08082014-171555>
- Iung, L. H. S., Petrini, J., Ramírez-Díaz, J., Salvian, M., Rovadoscki, G. A., Pilonetto, F., Dauria, B. D., Machado, P. F., Coutinho, L. L., Wiggans, G. R., & Mourão, G. B. (2019). Genome-wide association study for milk production traits in a Brazilian Holstein population. *Journal of Dairy Science*, 102(6), 5305–5314. <https://doi.org/10.3168/jds.2018-14811>
- Jiang, Z., Sun, J., Dong, H., Luo, O., Zheng, X., Oberfell, C., Tang, Y., Bi, J., O'Neill, R., Ruan, Y., Chen, J., & Tian, X. (2014). Transcriptional profiles of bovine in vivo pre-implantation development. *BMC Genomics*, 15(1), 756. <https://doi.org/10.1186/1471-2164-15-756>
- Jin, Z., & El-Deiry, W. S. (2005). Overview of cell death signaling pathways. *Cancer Biology & Therapy*, 4(2), 147–171. <https://doi.org/10.4161/cbt.4.2.1508>
- Jonchère, V., Réhault-Godbert, S., Hennequet-Antier, C., Cabau, C., Sibut, V., Cogburn, L. A., Nys, Y., & Gautron, J. (2010). Gene expression profiling to identify eggshell proteins involved in physical defense of the chicken egg. *BMC Genomics*, 11(1), 57. <https://doi.org/10.1186/1471-2164-11-57>
- Kinsella, R. J., Kahari, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., & Flicek, P. (2011). Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database*, 2011(0), bar030–bar030. <https://doi.org/10.1093/database/bar030>
- Kolbehdari, D., Wang, Z., Grant, J. R., Murdoch, B., Prasad, A., Xiu, Z., Marques, E., Stothard, P., & Moore, S. S. (2009). A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. *Journal of Animal Breeding and Genetics*, 126(3), 216–227. <https://doi.org/10.1111/j.1439-0388.2008.00793.x>
- Law, A.-L., Jalal, S., Pallett, T., Mosis, F., Guni, A., Brayford, S., Yolland, L., Marcotti, S., Levitt, J. A., Poland, S. P., Rowe-Sampson, M., Jandke, A., Köchl, R., Pula, G., Ameer-Beg, S. M., Stramer, B. M., & Krause, M. (2021). Nance-Horan Syndrome-like 1 protein negatively regulates Scar/WAVE-Arp2/3 activity and inhibits lamellipodia stability and cell migration. *Nature Communications*, 12(1), 5687. <https://doi.org/10.1038/s41467-021-25916-6>
- Lim, D., Strucken, E. M., Choi, B. H., Chai, H. H., Cho, Y. M., Jang, G. W., Kim, T.-H., Gondro, C., & Lee, S. H. (2016). Genomic Footprints in Selected and Unselected Beef Cattle Breeds in Korea. *PLOS ONE*, 11(3), e0151324. <https://doi.org/10.1371/journal.pone.0151324>
- Lourenço, O., Bhuyan, M., Lenzi, C. H., Dutra, M., Gonzalez-Boquera, C., Centelles, M., & Viñas, X. (2020). GW170817 constraints analyzed with Gogny forces and momentum-dependent interactions. *Physics Letters B*, 803, 135306. <https://doi.org/10.1016/j.physletb.2020.135306>
- Marestone, B. S., Santos, É. R. dos, Serra, F. B. S., Muniz, C. A. D. S. D., Marques, C. P., Alves, K. B., Alves, M. V., & Alves, R. C. M. (2013). Características reprodutivas, de crescimento e idade ao primeiro parto em bovinos da raça Holandesa. *Semina: Ciências Agrárias*, 34(6Supl2), 4105. <https://doi.org/10.5433/1679-0359.2013v34n6Supl2p4105>
- Marques, D. B. D., Bastiaansen, J. W. M., Broekhuijse, M. L. W. J., Lopes, M. S., Knol, E. F., Harlizius, B., Guimarães, S. E. F., Silva, F. F., & Lopes, P. S. (2018). Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. *Genetics Selection Evolution*, 50(1), 40. <https://doi.org/10.1186/s12711-018-0412-z>
- Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). Blupf90 and Related Programs (Bgf90). 7th World Congress on Genetics Applied to Livestock Production., 2001–2002.

- Misztal, Ignacy et al. (2014) Manual for BLUPF90 family of programs. Athens: University of Georgia, 1999.
- Murata, H., Shimada, N., & Yoshioka, M. (2004). Current research on acute phase proteins in veterinary diagnosis: an overview. *The Veterinary Journal*, 168(1), 28–40. [https://doi.org/10.1016/S1090-0233\(03\)00119-9](https://doi.org/10.1016/S1090-0233(03)00119-9)
- Mwanga, G., Mujibi, F. D. N., Yonah, Z. O., & Chagunda, M. G. G. (2019). Multi-country investigation of factors influencing breeding decisions by smallholder dairy farmers in sub-Saharan Africa. *Tropical Animal Health and Production*, 51(2), 395–409. <https://doi.org/10.1007/s11250-018-1703-7>
- Nayeri, S., Sargolzaei, M., Abo-Ismael, M. K., May, N., Miller, S. P., Schenkel, F., Moore, S. S., & Stothard, P. (2016). Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genetics*, 17(1), 75. <https://doi.org/10.1186/s12863-016-0386-1>
- Norman, H. D., Wright, J. R., Hubbard, S. M., Miller, R. H., & Hutchison, J. L. (2009). Reproductive status of Holstein and Jersey cows in the United States. *Journal of Dairy Science*, 92(7), 3517–3528. <https://doi.org/10.3168/jds.2008-1768>
- Nunes, M. A. (2018). Análises estruturais para componentes de características do leite e índices reprodutivos em bubalinos leiteiros. Universidade Federal da Grande Dourados.
- Otto, P. I., Guimarães, S. E. F., Verardo, L. L., Azevedo, A. L. S., Vandenplas, J., Soares, A. C. C., Sevillano, C. A., Veroneze, R., de Fatima A. Pires, M., de Freitas, C., Prata, M. C. A., Furlong, J., Verneque, R. S., Martins, M. F., Panetto, J. C. C., Carvalho, W. A., Gobo, D. O. R., da Silva, M. V. G. B., & Machado, M. A. (2018). Genome-wide association studies for tick resistance in *Bos taurus* × *Bos indicus* crossbred cattle: A deeper look into this intricate mechanism. *Journal of Dairy Science*, 101(12), 11020–11032. <https://doi.org/10.3168/jds.2017-14223>
- Parker Gaddis, K. L., Null, D. J., & Cole, J. B. (2016). Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *Journal of Dairy Science*, 99(8), 6420–6435. <https://doi.org/10.3168/jds.2015-10444>
- Ptak, A., Wróbel, A., & Gregoraszczyk, E. L. (2011). Effect of bisphenol-A on the expression of selected genes involved in cell cycle and apoptosis in the OVCAR-3 cell line. *Toxicology Letters*, 202(1), 30–35. <https://doi.org/10.1016/j.toxlet.2011.01.015>
- Qanbari, S., Pimentel, E. C. G., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A. R., & Simianer, H. (2009). The pattern of linkage disequilibrium in German Holstein cattle. *Animal Genetics*. <https://doi.org/10.1111/j.1365-2052.2009.02011.x>
- Regatieri, I. C., Boligon, A. A., Costa, R. B., de Souza, F. R. P., Baldi, F., Takada, L., Venturini, G. C., de Camargo, G. M. F., Fernandes, G. A., Tonhati, H., de Oliveira, H. N., & de Albuquerque, L. G. (2017). Association between single nucleotide polymorphisms and sexual precocity in Nellore heifers. *Animal Reproduction Science*, 177, 88–96. <https://doi.org/10.1016/j.anireprosci.2016.12.009>
- Rezende, S. R., Munhoz, S. K., Nascimento, M. R. B. de M., & Guimarães, J. L. N. (2015). Características de termorregulação em vacas leiteiras em ambiente tropical: revisão. *Veterinária Notícias*, 21(1), 18–29.
- Ribeiro, E. S., Greco, L. F., Bisinotto, R. S., Lima, F. S., Thatcher, W. W., & Santos, J. E. (2016). Biology of Preimplantation Conceptus at the Onset of Elongation in Dairy Cows<sup>1</sup>. *Biology of Reproduction*, 94(4). <https://doi.org/10.1095/biolreprod.115.134908>
- Rychtářová, J., Sztankóová, Z., Kyselová, J., Zink, V., Štípková, M., Vacek, M., & Štolc, L. (2014). Effect of DGAT1, BTN1A1, OLR1, and STAT1 genes on milk production and reproduction traits in the Czech Fleckvieh breed. *Czech Journal of Animal Science*, 59(No. 2), 45–53. <https://doi.org/10.17221/7228-CJAS>

- Sang, Q., Li, B., Kuang, Y., Wang, X., Zhang, Z., Chen, B., Wu, L., Lyu, Q., Fu, Y., Yan, Z., Mao, X., Xu, Y., Mu, J., Li, Q., Jin, L., He, L., & Wang, L. (2018). Homozygous Mutations in WEE2 Cause Fertilization Failure and Female Infertility. *The American Journal of Human Genetics*, 102(4), 649–657. <https://doi.org/10.1016/j.ajhg.2018.02.015>
- Selvaraju, S., Parthipan, S., Somashekar, L., Kolte, A. P., Krishnan Binsila, B., Arangasamy, A., & Ravindra, J. P. (2017). Occurrence and functional significance of the transcriptome in bovine (*Bos taurus*) spermatozoa. *Scientific Reports*, 7(1), 42392. <https://doi.org/10.1038/srep42392>
- Serre, L., Girard, M., Ramadan, A., Menut, P., Rouquié, N., Lucca, L. E., Mahiddine, K., Leobon, B., Mars, L. T., & Guerder, S. (2017). Thymic-Specific Serine Protease Limits Central Tolerance and Exacerbates Experimental Autoimmune Encephalomyelitis. *The Journal of Immunology*, 199(11), 3748–3756. <https://doi.org/10.4049/jimmunol.1700667>
- Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, 50(W1), W216–W221. <https://doi.org/10.1093/nar/gkac194>
- Smith, D. K., & Xue, H. (1997). Sequence profiles of immunoglobulin and immunoglobulin-like domains. *Journal of Molecular Biology*, 274(4), 530–545. <https://doi.org/10.1006/jmbi.1997.1432>
- Takle, H., Baeverfjord, G., & Andersen, Ø. (2004). Identification of Stress Sensitive Genes in Hyperthermic Atlantic Salmon (*Salmo salar*) Embryos by RAP-PCR. *Fish Physiology and Biochemistry*, 30(3–4), 275–281. <https://doi.org/10.1007/s10695-005-0291-4>
- Talbott, H., Hou, X., Qiu, F., Zhang, P., Guda, C., Yu, F., Cushman, R. A., Wood, J. R., Wang, C., Cupp, A. S., & Davis, J. S. (2017). Early transcriptome responses of the bovine midcycle corpus luteum to prostaglandin F2 $\alpha$  includes cytokine signaling. *Molecular and Cellular Endocrinology*, 452, 93–109. <https://doi.org/10.1016/j.mce.2017.05.018>
- VanRaden, P. M. (2008a). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- VanRaden, P. M. (2008b). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- WANG, H., MISZTAL, I., AGUILAR, I., LEGARRA, A., & MUIR, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, 94(2), 73–83. <https://doi.org/10.1017/S0016672312000274>
- Yang, J. (2015). Haplotype-Based Selection Signature Analysis Using University Of Minnesota And Us Contemporary Holstein Cattle. University of Minnesota.

- Yokoo, M. J., Magnabosco, C. U., Rosa, G. J. M., Lôbo, R. B., & Albuquerque, L. G. (2012). Características reprodutivas e suas associações com outras características de importância econômica na raça Nelore. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 64(1), 91–100. <https://doi.org/10.1590/S0102-09352012000100014>
- Zou, J., Yin, F., Wang, Q., Zhang, W., & Li, L. (2015). Analysis of microarray-identified genes and microRNAs associated with drug resistance in ovarian cancer. *International Journal of Clinical and Experimental Pathology*, 8(6), 6847–6858.





### **3. CONCLUSIONS**

Some genes were identified among the investigated animals regarding immune and reproductive traits – linked to postpartum intemperance and immune system recovery. Also, genes associated with embryogenesis were identified in this population. The findings of the present study shed light on the possibility of selecting animals based on their genetic information, so those not attending the criteria can be dismissed in order to improve dairy production. Nevertheless, further studies are needed to better enhance the knowledge in this field and to enable the practical application of it all.