GENOMICS OF HEIFER PREGNANCY, DAYS OPEN, AND DAYS TO CONCEPTION IN

RED ANGUS HEIFERS

A Thesis

Presented to

The Faculty of the Graduate School

At the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for Degree

Master of Science

By

Esdras Tuyishimire

Dr. Jared Decker, Thesis Supervisor

May 2022

The undersigned, appointed by the dean of the Graduate School, have examined the **thesis**

entitled

GENOMICS OF HEIFER PREGNANCY, DAYS OPEN, AND DAYS TO CONCEPTION

IN RED ANGUS HEIFERS

presented by **Esdras Tuyishimire**,

a candidate for the degree of **Master of Science**,

and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. Jared Decker

_____

Dr. Robert Schnabel

_____

Dr. Elizabeth King

_____

Dr. Ruthie Angelovici

ACKNOWLEDGMENTS

Table of Contents

List of Tables

List of Figures

0.   Abstract

Although it is challenging to collect information on reproductive traits, fertility traits are believed to play a significant role in the growth of the beef industry and bring revenue to the producers. It has been challenging to make genetic improvement of fertility traits because of scarce records and their low heritability. The joint analyses of economic and indicator traits provided much information for the analysis of other production traits, resulting in favorable improvement. The same method can be applied to reproductive traits to increase prediction accuracy. Furthermore, the ubiquity of high-throughput genotypes and the development of advanced computational methods give hope that there is a possibility of determining the genetic influence on fertility traits. The advancement in phenotypic collection technology eases the keeping of records on both economic and indicator fertility traits. Also, it allows the development of new traits to be analyzed while reducing the cost of data availability.

Heifer Pregnancy (HP) - the ability for a heifer to conceive by the end of the breeding season; Days Open (DO) - days of breeding season a heifer remained open; and Days to Conception (DC) – the number of days it took for a heifer to get pregnant, are easy to measure fertility traits with enough information to make beef reproductive genetic improvement in the industry. The objective of the current study is to identify DNA markers tagging genes influencing HP, DC, and DC among 18,039 Red Angus heifers, find the genetic relationship among those traits, and creating the genomic predictions of days to conception and days open.

The results show a small heritability of 0.119 to 0.131, 0.10 to 0.102, and 0.0749 to 0.112 for HP, DO, and DC, respectively. The DC model resulted in higher accuracy than other models. There was a high correlation between HP and DO (r = - 0.61 and 0.85 from the linear model and the

liability scale model, respectively). The de-regressed estimated breeding value (EBV) genome-wide association (GWAS) yielded 58 and 2 significant SNPs at suggestive significant level (p-value < 1.0e-05) for HP and DC, respectively. This study found GRID2 and ZMIZ1 genes to be associated with heifer pregnancy, and it has been speculated that the central nervous system related genes ontology and the hormones it controls might suggest the physiology behind some of the reproductive differences.

The present study confirms the low heritability status of fertility traits and shows the possibilities of genetic enhancement based on the obtained accuracies. The identified genes and gene terms will serve as starting points for future studies that might focus on different phenotypes while reducing the cost of phenotyping and improving the accuracy of fertility traits genomic predictions.

Chapter 1 Literature Review

1.1. The contrast between production and fertility traits

Since the introduction of *Bos taurus* in the US around the 16th century [1, 2], livestock producers were already applying artificial selection focusing on visual features. The involvement of academic researchers and government agencies in the 20th century revolutionized the industry as they put much effort into increasing productivity [3–5]. The early work focused on the growth traits like weaning weight and yearling weight, then on carcass traits like carcass marbling score [6–9]. The records on growth traits are abundant because they are quite easy to measure, and their quantitative nature facilitates the analysis, hence enabling genetic progress. The more the beef industry grew, the researchers were able to juxtapose the indicator traits in association with some economic traits. It has been possible to improve some of the hard-to-measure traits because additional information would come from indicator traits [10]. Examples of indicator traits include birth weight, which is associated with weaning weight, yearling weight, and calving ease; ultrasound measures give the status for carcass marbling score.

The application of statistical analysis and genetics allowed to estimate some of the parameters that breeders based on for selection decisions of production traits; as a result, there was great progress in beef growth and carcass traits (Figure 1a, b). When a high-density SNP array was introduced [11], the industry changed forever as some traits like weaning weight, yearling weight, carcass weight, and back fat could be studied at the genomic level [12, 13], hence the increased accuracy of prediction [14]. It did not take very long to realize that the reproductive traits were left behind while analyzing production and carcass characteristics. Primarily in the dairy industry, they identified that increased production impaired fertility [15, 16]. Wiltbank, 1970 called for research in beef cattle reproduction and identified the loss of calf crop production and the increased calving

season as the primary concerns; thus, he posed some questions that the scientific community would answer to solve the problem [17]. The Wiltbank and other scientists' call allowed the transition to add the fertility traits among the economic traits that should be studied.

While many factors were identified to affect fertility traits, people in different domains tried to have input. They developed some techniques, including artificial insemination, embryo transfer, improved management practices, and selection on fertility; consequently, the currently observed improvement (Figure 1c), yet the problem remains not fully solved. There is still a need for further studies to understand the genes governing reproduction in beef heifers.
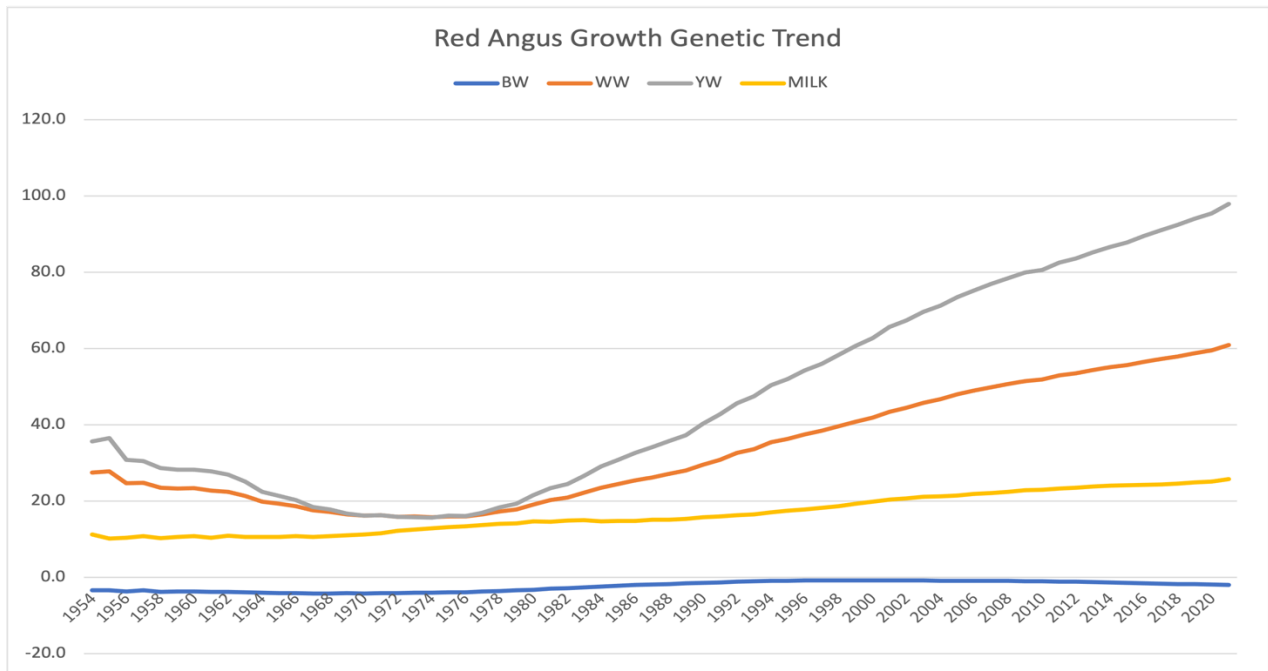
(a)
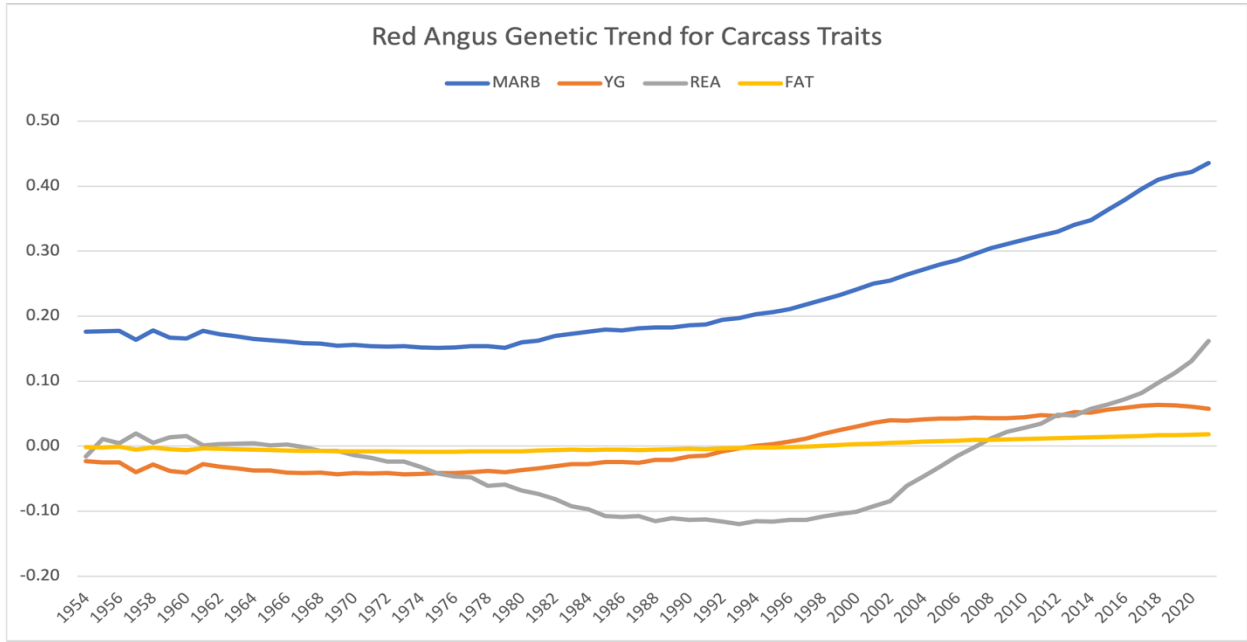
(b)
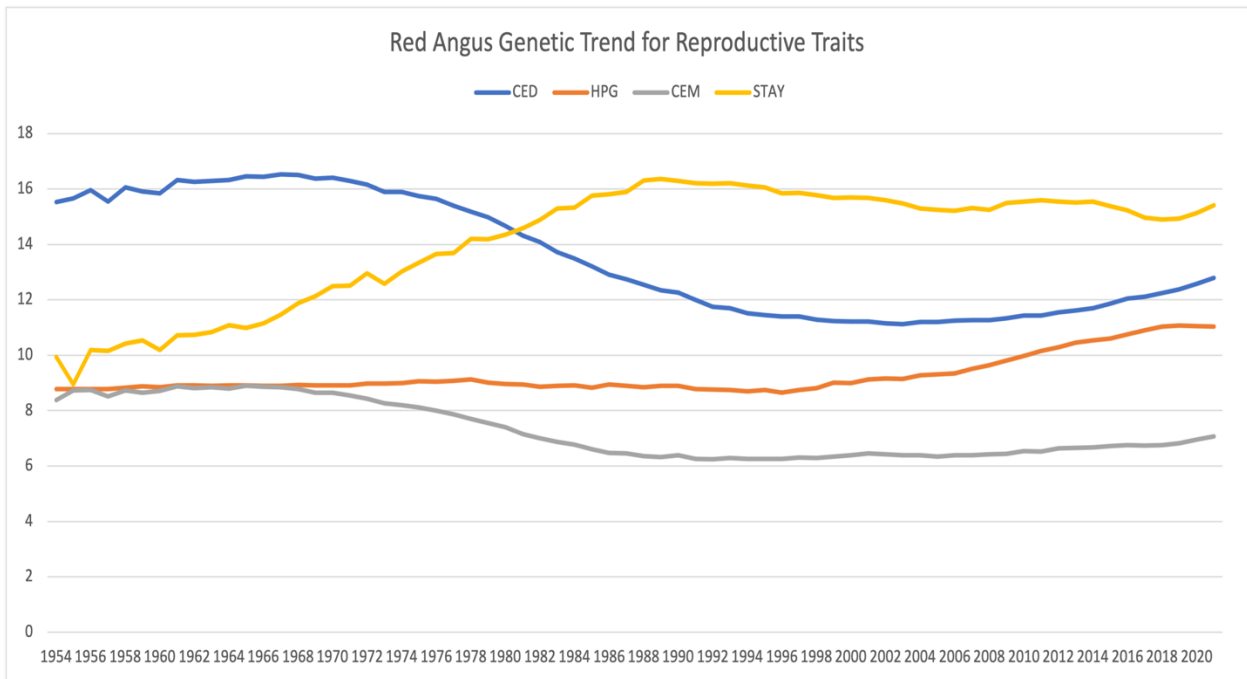
(c)

Figure 1 Genetic trends in growth (a), carcass (b), and reproduction (c) traits in Red Angus cattle [18].

From the 1980s to 2004, the growth and carcass traits increased while the fertility traits except stayability decreased. The improvement in fertility traits after 2004 might indicate researchers' effort after they started to include those traits in selection decisions, although that selection is in Red Angus and a very few breeds, as it has been indicated by Rowan et al. 2021 [19].

1.2. Importance of fertility traits

Although reproductive traits are notorious for being lowlily heritable [17, 20], which restricts quick genetic progress, they are pivotal in sustaining the animal industry. It might be very hard to ignore the importance of fertility in the beef industry, especially for those in a cow-calf operation. In fact, the whole industry is based on calf production; therefore, it is crucial to ensure that a heifer is and remains pregnant until it calves a healthy calf. Some of the fertility traits that have been identified to be essential to the industry are heifer pregnancy, stayability, and calving ease [21– 23]. On the other side, to have successful fertility, puberty onset should happen; thus, puberty-related traits are considered fertility indicators. Not only that a heifer that reaches puberty early in the breeding season weans a heavier calf, but in their study, Roberts et al. found that there was an 11 to 13 % less calving rate in the first 21 days among the heifers that failed to reach puberty during the breeding season than their contemporaries [24]. Therefore, interest is made for every pound increased on a weaned calf whose dam reached puberty early in the breeding season. In addition, while there is a cost related to raising and maintaining a heifer, culling comes with a loss to the producers. Fertility problems are among the top reasons for culling in the livestock industry [25]. In the dairy industry, the increased insemination rate, linked to long days open and pregnancy rate, reduced the profit by $205 per year per cow in 2004 [26]. Also, in 2008, it was shown that in Florida, $75 was lost per failed pregnancy during the breeding season [27], and that loss might

even be higher if adjusted for inflation. Despite the increase in technology, reproductive failure and loss of embryo issues are still around, and an estimated loss of $2.8 billion in the US beef industry per year might happen if nothing is done [28]. Given the identified failures and what could be gained if a heifer gets pregnant early in the breeding season, it is essential to study the fertility traits further to increase the number of calves born alive and reduce the cost related to reproductive failure.

1.3. Previously studied reproductive traits at the genetic level.

Burns et al. reviewed factors that affect reproductive performance in beef cattle, and they classified them into four main categories: management, environmental, physiological, and genetics [29]. Nowadays, technology and scientific research growth helped improve management, environmental, and physiological conditions. At the same time, genetic improvement remains a challenge, and little is known about the genetics underlying reproductive traits. Most fertility traits are hard to measure, and due to their binary nature, it is also challenging to make accurate predictions because of their low heritability.

After realizing a decrease in fertility, scientists put effort into improving these traits; hence, there are some noticeable improvements. Some of the studied traits in the beef industry include heifer pregnancy, stability, calving ease, age at puberty, first pregnancy rate, and to less extent, days open [20, 21, 23]. As suggested, the economic traits might be analyzed jointly with indicator traits (Table 1) to provide enough information, but a very few studies performed those joint analyses.

Table 1 Economic fertility traits and their respective indicator traits

Economic Traits	Indicator Trait
Calving ease	Birth weight
	Gestation length
Stayability	Days to calving
	Days open
	Day to conception
	Calving records
Heifer pregnancy	Age at puberty
	Follicular diameter
	Antral follicle count
	Postpartum interval
	Scrotal circumference
	Reproductive tract score

     Due to the complexity of fertility traits, it is challenging to perform across breed predictions [5, 14]; consequently, because of its various breeds, the beef industry runs behind dairy. On the other hand, there are some accomplished within breed research on HP, stayability, calving ease, days open, and scrotal circumference [30–37]. In those studies, the reported heritability was 0.085 to 0.27; 0.059 to 0.57; 0.19; 0.091 to 0.192; and 0.41 to 0.71 for HP, stayability, calving ease, days open, and scrotal circumference, respectively. Speidel et al. found some quantitative traits loci

(QTL) that are presumed to be associated with heifer pregnancy and stayability [38]. Some additional fertility traits that have been studied include antral follicular count (AFC) and reproductive tract score (RTS), with a sample size of 293 heifers [39]. Even though all those studies were able to find some loci and candidate genes associated with reproduction, the researchers concluded that validation studies were needed as they were not sure about the results. In addition to uncertain results, those studies used a small sample size which might not be reliable when studying complex heritable traits, and the genomic information was from low density markers, while high densities with much genomic information are currently available.

Even though their higher heritability can explain the success made in production and carcass than some reproductive traits, the quantitative nature, and the joint analyses included indicator traits added advantages to that success. For instance, the birth weight would be regarded as an indicator trait for weaning and yearling weight, and ultrasound measures for intramuscular fat, marbling score, and ribeye area; thus, they were usually analyzed together, which provides enough information for statistical analysis, hence improved accuracy [10, 13, 40, 41].

Despite the small improvement in HP and Stayability in Red Angus since 2012, there is still room for progress given the advancement in statistical and genomic tools that increase prediction accuracy [42]. The ubiquity of sequencing technology coupled with the development of advanced computational tools in the 21st century gives hope that there is a possibility of determining the genetic influence of fertility traits. The present studies should perform analyses that blend both pedigree and genomic information to increase fertility traits' accuracy while scrutinizing economic and indicator traits to maximize available information. Cushman et al. provide a list of some fertility traits and their corresponding indicators as some of them are presented in table 1 [43], but most of those traits are expensive to collect and require much work.

While the overall goal in the beef industry is to maximize profit while reducing production costs, the increase in phenotyping technology would help select the most relevant and introduce new traits for the analysis, which does not cost much and with less intensive work.

Chapter 2 Genomics of Heifer pregnancy, days open, and days to conception in Red Angus heifers

2.1. Introduction

Despite the previously reported data scarcity and low heritability status of reproductive traits that hinder the genetic progress in beef cattle [44–46], these traits are believed to have a pivotal role in the development of the beef industry. Mercadante et al. predicted a loss of $21 billion of revenue due to the failure to produce a live calf at weaning and the cost of breeding [28]. Morrey and Biased predicted that the beef industry would lose over $210 million due to late breeding heifers [47] as it is phenotypically related to cows continued reproductive success (i.e., stayability). Given that loss, much effort is required to improve reproductive traits. Previous studies have found reproductive problems associated with management, environment, physiology, and genetics [20, 29, 48, 49]. Some of those have been addressed; consequently, there has been improvement in reproductive traits among beef cattle. Yet, the genetic part of the reproductive traits is not fully understood. In the past, researchers made conclusions about fertility traits on heritability estimates only without any consideration of genetic effect and response to selection which led them to conclude that genetic improvement was almost impossible for reproductive traits [17, 50].

With the advancement in computational and genomic technology, many studies were performed to predict some genetic parameters but very few incorporated genomic data with a small sample size. Cammack et al. reviewed most of the studies performed using only pedigree information [20], and most of those studies resulted in low heritability as expected for fertility traits [21, 23, 31, 37]. The reduced cost of genome sequencing and SNP array development [11] helped scientists collect genotypic information and analyze some of the reproductive traits at the genomic level. A couple of studies have used genomic information to estimate *Bos taurus* beef

reproductive traits heritability and find the associated single nucleotide polymorphism (SNP) and candidate genes [38, 51–54]. Nevertheless, the sample size and the number of SNPs in those studies were small. Reproductive traits are polygenic and require a significant sample size and high-density markers [52] to find the true SNPs in association and the genes that have an impact on the traits. Evans et al. 1999 found that the accuracy of heifer pregnancy would increase when analyzed with scrotal circumference [21], and Toghiani found that multi-trait genomic evaluation would elaborate on the accuracy of selection of fertility traits when both quantitative and discrete variables are analyzed jointly [52].

Not only a large dataset is needed to make accurate predictions on reproductive traits, but also high-quality phenotypes that are aligned with genotype are required [48]. Imputation allows inferring from low to high marker density with a large reference population of genotyped animals, which increases the ability to detect variants associated with the phenotype [55, 56]. This study aims to use high-density markers to find genes influencing fertility traits. Heifer Pregnancy (HP) – the ability for a heifer to conceive by the end of the breeding season and remain pregnant, Days Open (DO) – days of a breeding season a heifer remained open; and Days to Conception (DC) – numbers of days it took for a heifer to get pregnant, are promising, easy to measure fertility traits. We hypothesize these traits have sufficient information to improve the genetic prediction of reproduction in beef cattle, especially when analyzed jointly.

Eler et al. mentioned that it is irrelevant to use estimates from indicator traits while direct reproductive measures are available [57]. But in some circumstances, estimates from fertility traits are not available because they demand intensive labor with a high cost of data collection; as a result, the indicator traits might be used for indirect prediction of fertility traits. The first objective of the current study is to estimate the variance components of HP, DO, and DC using three software

suites by fitting a linear and threshold model for a binary trait (HP). The second objective is to estimate the accuracy of genomic prediction for each trait and identify DNA markers associated with the observed phenotypes and the genetic relationship among the traits. It has been presumed that genes in common influence both heifer pregnancy and days intervals(DO and DC).

2.2. Materials and Methods

2.2.1. Phenotypes and data quality control

The Red Angus Association of America provided phenotypes, genotypes, and pedigree records. Information was provided for 4,097 genotyped heifers with heifer pregnancy records, as well as 13,942 that were in the same contemporary groups as the genotyped heifers. The full phenotype dataset had 18,485 individuals born between 1988 to 2018. There were records on heifer pregnancy, start and end dates of the breeding season, contemporary group, calf birthdate, and sex. Animals in a contemporary group share information on the farm, calving season, and birth year. Heifer Pregnancy (**HP**) was recorded as 1 for pregnant and 0 for non-pregnant. In contrast, Days Open (**DO**) and Days to Conception (**DC**) were calculated as continuous variables. Both DO and DC were calculated based on the mean gestation length of 283 days, as has reported in [58]. Thus, DC was expressed as the difference in calf birth date, conception length, and the start of the breeding season; $DC = Calf\ birth\ date - conception\ length - start\ of\ the\ breeding\ season$. The same applies for DO, but heifers who did not become pregnant by the end of the breeding season were assigned the season length as their Days Open value. The season length was calculated as the difference between the end and the start of the breeding season. Appendix 2.6.1. shows distribution of quantitative variables before quality control.

While it is possible to have some noise during data collection, it was necessary to perform quality control on the phenotype data. The phenotype quality control was done by removing all individuals who were 300 days old or younger at the start of the breeding season. Heifers whose DO records were more than 365 or less than -30 days were removed from the dataset. Again, heifers with more than 250 or less than -21 days on DC and DO were considered missing (only DO and DC) based on the data distribution (Figure 2). Heifers are typically bred at 15 months of age[59–62] ; thus, 300 days of age was set as cutoff (Figure 3 (c) in Appendix 2.6.2.). While some of the tools used during this study consider -9 as missing, DC and DO with -9 days were assigned -9.1 days to avoid that problem.
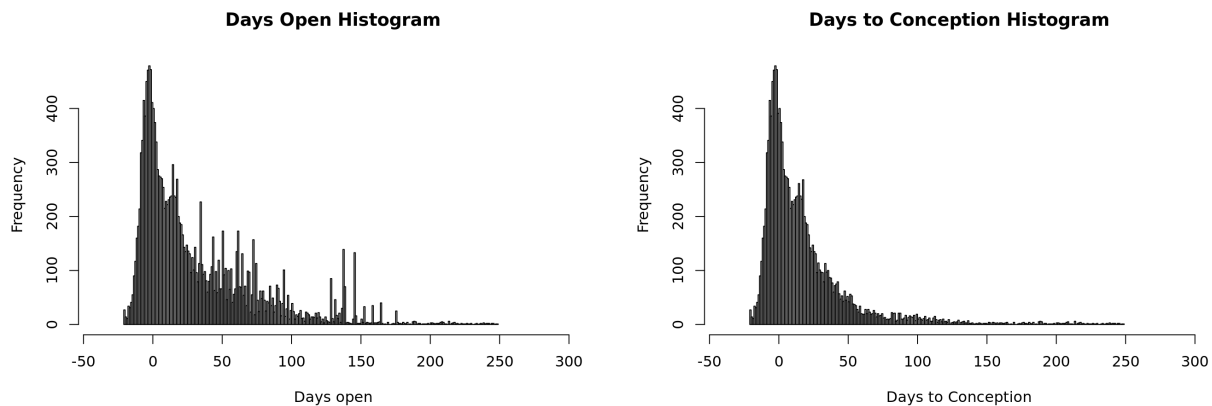
Figure 2 Distribution of days open and days to conception after data filtering. Most heifers were pregnant by 100 days, while very few heifers went beyond that (DC histogram).

Days open include heifers who reached the end of the breeding season without becoming pregnant, and most of those heifers are between 40 days and 150 days (DO histogram), indicating that most of the herds followed 90 days breeding season length.

After data quality control, 18,039 individuals remained for the analysis. Below is a summary table for heifers with variable's number of records, minimum, maximum, mean, and standard deviation (Table 2).

Table 2 Summary statistics for Heifer Pregnancy, Days to Conception, Days Open, Heifer's Age (in days), and Season Length (in days). n stands for sample size, and SD is a standard deviation.

Trait	n	minimum	maximum	mean	SD
Heifer Pregnancy	18,039	0	1	0.8	0.39
Days to Conception	14,231	-21	249	18	33
Days Open	17,938	-21	249	31	42
Age	17,949	301	546	422	30
Season Length	17,763	0	244	86	37

2.2.2. Genotypes

The genotype information on the individuals used in this study came from the Red Angus Association of America. The ARS-UCD1.2 bovine reference genome was used to identify the SNP positions [63]. Heifers were genotyped using a variety of arrays used for genomic-enhanced EPDs. The genotypes passed through the University of Missouri multibreed imputation pipeline described by Rowan et al. [56]. Rowan et al. performed quality control on autosomal variants only using PLINK software [64] and removed individuals and variants with less than 90% call rate [56]. Among the Red Angus, 811,679 SNPs were included in the imputed data. Apart from the initial quality control done during imputation, SNPs with MAF < 0.01 were removed from this study.

2.2.3. Variance components estimation, and genome-wide association with Genome-wide Complex Trait Analysis (GCTA)

The GCTA software (v1.93.3beta2) was used to estimate the variance components of HP (binary), DO, and DC. Both univariate and bivariate linear mixed model analyses were performed for variance estimation. Only genotyped heifers (n = 4,097) were analyzed with GCTA [65]. Among those heifers, the number of analyzed individuals varied from trait to trait as the missing phenotypes were filtered out. Each model fitted the contemporary group and Genomic Relationship Matrix (GRM) to have a random effect. Contemporary group was fit as random due to the small number of animals within many contemporary groups [66]. In these models, heifer's age and breeding season length were fit as fixed effects (Equation.1). The genomic relationship matrix (GRM) was calculated using autosomal SNPs, with the method of Yang et al [65]. The variance components were calculated by genome-based restricted maximum likelihood using the average information [65]. The mathematical representation of the univariate model to estimate variance and heritability is in equation.1. In addition, the estimated breeding values were calculated using the same equation. The heifer pregnancy heritability was adjusted on the liability scale considering the prevalence of 81.3% (The value obtained as the pregnancy rate among 18,039 heifers). Equation 2 represents the bivariate model for genetic correlation between HP and DO.

$$Y = Hh + Ss + Cp + Zu + e \quad \text{(Equation.1)}$$

Where Y is a vector of the phenotype of interest (can take the values of HP, DO, or DC)

h is the vector of heifer's age fixed effect

s is the vector of season length fixed effect

p is the random effect of the contemporary group with $p \sim N(0, I\sigma_p^2)$

u is the random additive genetic effect with $u \sim N(0,\ A\sigma_u^2)$

e is the random residual effect with $e \sim N(0,\ I\sigma_e^2)$

H, S, C, and Z are the incidence matrices relating traits observation in Y to the vectors h, s, p, and u, respectively.

A and I being genetic relationship and identity matrices, respectively, among the individuals.

Only HP and DO were used for the bivariate model since there was no record of DC for non-pregnant heifers, leading to convergence problems. Also, DO and DC are highly correlated; thus, there was no need to fit them together. Equation.2 shows the mathematical representation of the bivariate model.

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

(Equation.2)

All letters are as described in the first equation. In contrast, numbers 1 and 2 in the second equation represent heifer pregnancy and days open, respectively.

The random genetic effects $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N(0, A \otimes V)$ in which $V = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

V being the genetic variance-covariance matrix, σ_1^2 is the genetic variance for HP, σ_2^2 is the genetic variance for DO, and $\sigma_1\sigma_2$ is the genetic covariance of HP and DO.

Also, the random contemporary group effect $\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \sim N(0, I \otimes P)$ in which $P = \begin{bmatrix} p_1^2 & p_1p_2 \\ p_1p_2 & p_2^2 \end{bmatrix}$

P is the contemporary group variance-covariance matrix; p_1^2 is the contemporary group variance for HP; p_2^2 is the contemporary group variance for DO, and $p_1 p_2$ is the contemporary group covariance of HP and DO.

Finally, the random residuals effect $\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \sim N(0, I \otimes E)$ in which $E = \begin{bmatrix} e_1^2 & e_1 e_2 \\ e_1 e_2 & e_2^2 \end{bmatrix}$

E is the residual variance-covariance matrix, e_1^2 is the residual variance for HP, e_2^2 is the residual variance for DO, and $e_1 e_2$ is the residual covariance of HP and DO.

The estimation of phenotypic variance explained by all SNPs, the so-called heritability, was calculated as $h^2 = \dfrac{\sigma^2}{\sigma^2 + e^2}$

h^2 is the heritability, σ^2 genetic variance for a given trait, and e^2 residual variance.

The genetic correlation (σ_g) was obtained as $\sigma_g = \dfrac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 \sigma_1^2}}$

The '-mlma' option within GCTA was used to perform univariate GWAS for Heifer Pregnancy, Days Open, and Days to Conception, using the same univariate model described for genomic REML analysis, except each SNP is independently fit as a fixed effect.

2.2.4. **Analysis with Genome-wide Efficient Mixed-Model Association (GEMMA) using pre-corrected phenotypes**

GEMMA was also used for genome-wide association analysis, due to the ability of GEMMA to use linear mixed models to control the population structure and relatedness [67] coupled with increased statistical power from multi-trait models [68]. However, there had been previously difficulty getting GEMMA to run with large numbers of fixed effects. Consequently, the analyses in GEMMA were performed using pre-corrected phenotypes from GCTA. The

adjusted phenotypes were the sum of individual estimated breeding values and corresponding residual. Because the phenotypes were pre-corrected, the model fit in GEMMA contained the phenotypes and genomic relationship matrix, and its mathematical representation is as follows:

$$Y = \mu + Zu + e \ (equation. 3)$$

$$h^2 = \frac{\sigma^2}{\sigma^2 + e^2} \ Equation. 4$$

And the multivariate equation is as,

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \ (Equation\ 5)$$

Where μ is the mean, and other letters and distribution are as described in previous equations.

2.2.5. Variance Components and Genomic predictions with Gibbs sampling

Due to the binary nature of heifer pregnancy, the ideal way of estimating variance and heritability would be to consider the liability scale [69]. The THRGIBBS1F90 [70, 71] is one of the BLUP90 programs [72] that can handle both continuous and binary variables; thus, it has been used to generate Gibbs samples, and in turn, those samples were used by the POSTGIBBS1F90 program to estimate the posterior variances and breeding values. The threshold model (THRGIBBS1F90) was run for 1,000,000 MCMC samples and samples were saved at every 200 iterations for HP and 50 for DO and DC. The initial burn-in was 100,000 samples. Because the BLUPF90 suite of programs accepts both genotyped and non-genotyped individuals, all heifers with phenotype information were used for variance component estimation and genomic predictions. For every analysis, the sample and pedigree sizes (3 generations) changed as the individuals with missing phenotypes were not included in the study. The HP univariate and multivariate models had the same sample size of 18,039 heifers, and the pedigree contained 47,050

individuals. The univariate DO model had 17,938 heifers and 46,919 individuals within the pedigree. On the other hand, the DC model among others, had the smallest sample and pedigree size of 14,231 and 40,473 individuals.

The number of individuals within the contemporary group has increased due to the expanded sample size. Hence, the contemporary group was fit as a fixed effect (rather than random as in the GCTA models). Quality control and relationship matrices were performed using the PREGSF90 program[73]. Analyses used the inverse of the H matrix (H^{-1}), which was built as a blended matrix of genomic (G) and pedigree relationship matrices as described in Aguilar et al. [74–76]. In addition, the genomic relationship matrix was produced following the method described by VanRaden [77].

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \textbf{\textit{(Equation. 6)}}$$

The equations described in the GCTA analysis were used to estimate the variance component and breeding values, except that in the models solved using THRGIBBS1F90, the contemporary group had a fixed effect. Consequently, there was no distribution of contemporary group random effect. The model accuracy and individuals' reliability were calculated using Linear Regression methods [78, 79] to test model bias. Selection decisions are typically made regarding the youngest generation, and to avoid the data leakage problem, the individuals within the contemporary groups were kept together when forming validation subsets. As a result, the youngest 15% of individuals were grouped into three validation sets. The first 5% youngest had a missing phenotype for the first validation. The second youngest 5% group had a missing record for the second validation and the same for the third group. The model's accuracy was calculated by comparing the estimated

breeding values of each partial set with the estimated breeding values from the complete set of individuals, as described by Bermann et al.[79].

2.2.6. De-regressed breeding values for Genome-Wide Association (GWAS)

SNP1101 program suite [80] has been used to perform de-regressed genome-wide association studies (GWASs).

The estimated breeding values divided by their reliabilities from the Gibbs sampling models were used as new phenotypes to perform GWAS, weighted by $\frac{1}{Reliability} - 1$ [80]. The individual's reliability was calculated using $1 - \frac{PEV}{(1+F_i)\sigma_a^2}$. PEV is the prediction error variance, F_i is the individual's breeding value coefficient, and σ_a^2 is the additive variance. The genomic relationship matrix was built by VanRaden's method [77].

2.2.7. Nearest Gene within 10Kb

After performing GWAS, the BovineMine v1.6 search tool [81] was used to find genes within 10kb of genome-wide significant SNPs for each trait separately, to ensure a strong linkage disequilibrium between the casual gene and associated variant [82].

2.3. Results

2.3.1. Variance Component, heritability, and Breeding values estimation

2.3.1.1. Univariate GCTA and GEMMA models

The heifer pregnancy rate among the individuals analyzed only in GCTA and GEMMA was 87.1%, while among all individuals was 81.3%. GCTA and GEMMA used linear mixed models during variance estimation, and both analyses were different in that pre-corrected

phenotypes were used in GEMMA throughout this study. Table 3 reports the genetic variance and heritability from univariate GCTA and GEMMA.

As expected, the reported heritability from GEMMA is higher than that of GCTA. That is because the property of continuous variables has more information than binary variables. Even though there is a small increase in heritability from GCTA to GEMMA, the difference is not significant. The increase in heritability might be associated with fewer noises introduced in GEMMA than GCTA during pre-correcting phenotypes.

Table 3 Estimates of genetic variance σ_g^2, residual variance σ_e^2, and heritability h^2 from univariate GCTA and GEMMA.

Trait	GCTA			GEMMA		
Variance	σ_g^2	σ_e^2	h^2	σ_g^2	σ_e^2	h^2
HP	0.0129 ± 0.0024	0.092 ± 0.0031	0.123	0.0132	0.0872	0.131
DC	100.641 ± 23.138	836.79 ± 29.740	0.107	105.212	831.17	0.112
DO	135.586 ± 29.615	1382.27 ± 39.700	0.100	134.589	1190.45	0.102

Among all three traits, Heifer Pregnancy (HP) is showing a high heritability compared to Days Open (DO) and Days to Conception (DC).

The HP heritability obtained from the linear model might not be reliable because of the binary nature of HP, which violates the assumptions of normal distribution for the linear models. For that reason, the obtained heritability was adjusted on the liability scale. As a result, the adjusted

HP heritability in GCTA was 0.33. GEMMA's liability scale was not needed because the used variables were continuous.

2.3.1.2. Bivariate GCTA and GEMMA

The bivariate model fit heifer pregnancy and days open. HP and DO model resulted in a correlation of -0.61 ± 0.095 and -0.63 ± 0.020 for GCTA and GEMMA, respectively. The variance estimates and heritabilities were not different from the univariate model. The bivariate model of DC and HP could not converge, probably because only pregnant heifers had values for days to conception. Not only was there a convergence problem, but it would not even be easy for results interpretation of the model that includes DC due to the missing records on non-pregnant heifers.

2.3.2. Gibbs Sampling Models variance components and genomic predictions

2.3.2.1. Univariate and bivariate outcomes

The threshold univariate model used 18,039 heifers with records on pregnancy; some of those did not have DC or DO records; thus, the sample size for the later traits was smaller than HP. The fixed effects from the Gibbs sampling model were smaller than the effect from linear models, as can be seen in table 4. However, there was not much difference between heritability estimated from previous linear models (GCTA & GEMMA) and the estimates from the thresholds model. Days to conception only showed a drastic change (0.107 to 0.0749). Table 5 shows the univariate outcomes from posterior distributions while figure 4 shows the convergence of all univariate models and the corresponding genetic variance and heritability. The reported values are the means from posteriors. The HP and DO bivariate model yielded a correlation of 0.86, considering that the HP solutions were on the liability scale. Even though a direct relationship between HP and DC was not estimated, Figure 3 shows a positive correlation between estimated breeding values of DO

and DC, thus, an indirect relation between HP and DC. Furthermore, the heritabilities from

bivariate were the same as univariate models.

Table 4. Heifer's age and season length estimated effect from linear and Gibbs sampling models on Heifer Pregnancy, Days to Conception, and Days Open.

Traits	Heifer Pregnancy		Days to Conception		Days Open	
Models	Linear	Gibbs	Linear	Gibbs	Linear	Gibbs
Heifer's Age	0.0003 ± 0.0002	-0.004 ± 0.00024	-0.178 ± 0.0189	-0.0164 ± 0.005	-0.176 ± 0.0217	-0.0204 ± 0.00438
Season Length	0.0076 ± 0.0018	-0.00166 ± 0.000565	0.0643 ± 0.0168	-0.0798 ± 0.0120	0.110 ± 0.0195	0.141 ± 0.0136
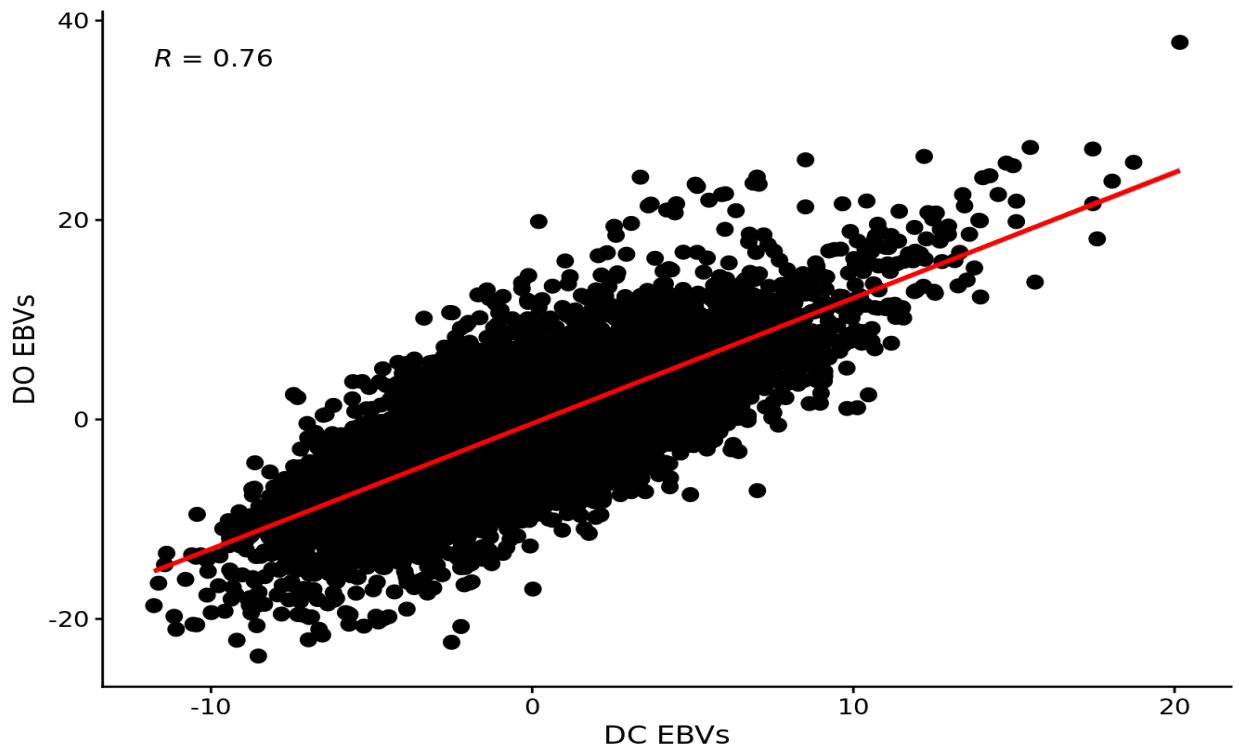
Figure 3 Correlation between estimated days open (DO) and days to conception (DC) breeding values (EBVs)

Table 5 The number of samples, mean, standard deviation, high posterior density (HPD), autocorrelation, and heritability (h2) estimates from the posterior distribution of the univariate threshold model.

Trait	samples	Posterior Mean		HPD		Autocorrelation		h^2
		Genetic	Residual	Genetic	Residual	Genetic	Residual	
HP	4500	0.117 ± 0.048	0.861 ± 0.0160	0.032; 0.21	0.833; 0.893	0.232	0.013	0.119
DC	18000	76.7 ± 18.84	947.0 ± 22.46	41.7; 113.8	903.6; 991.6	0.258	0.125	0.0749
DO	18000	170.61 ± 31.57	1495.0 ± 33.83	112.40; 233.6	1432; 1563.0	0.129	0.07	0.102

(a) HP

76.701

0.075

(b) DC
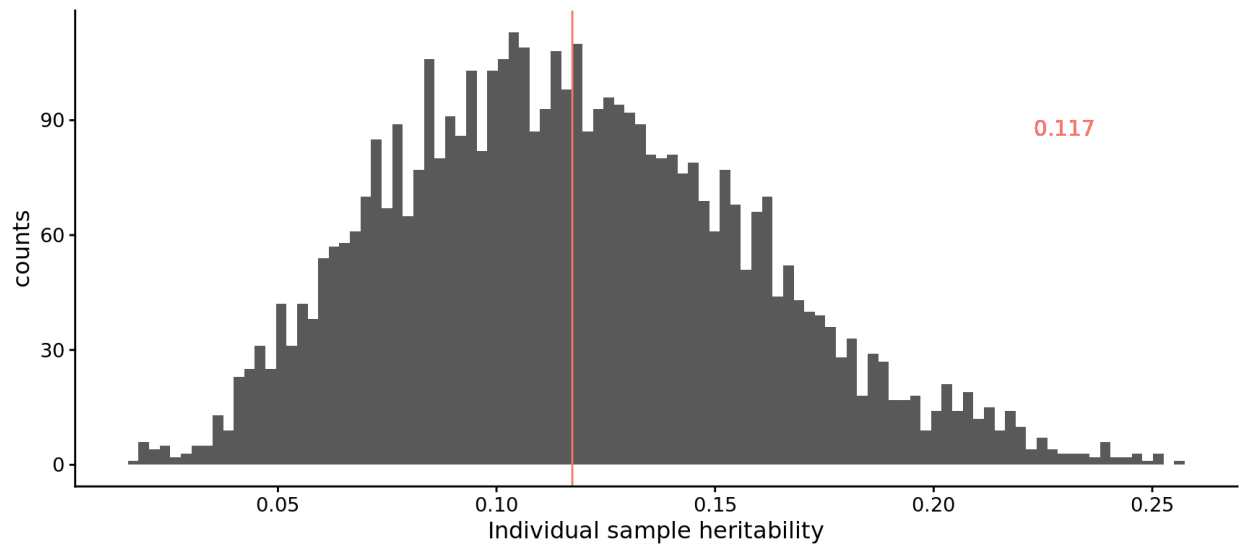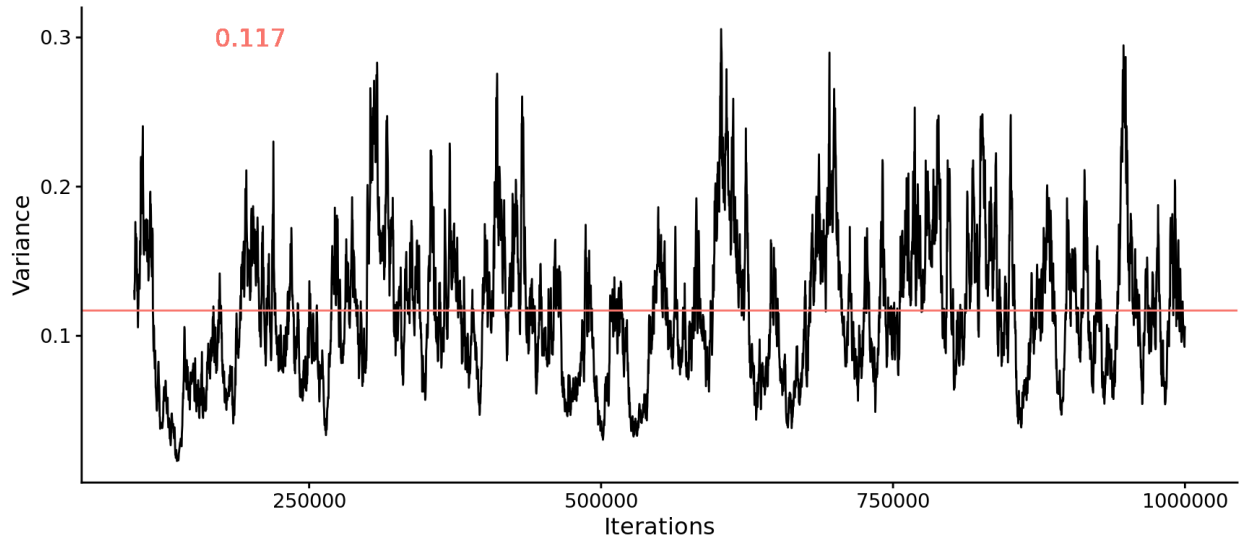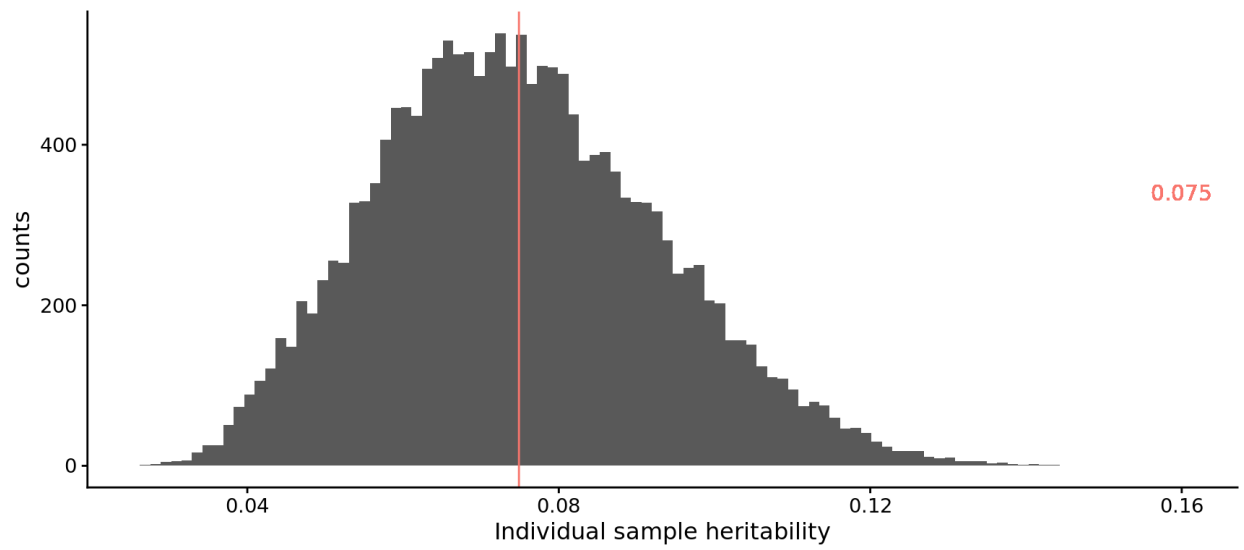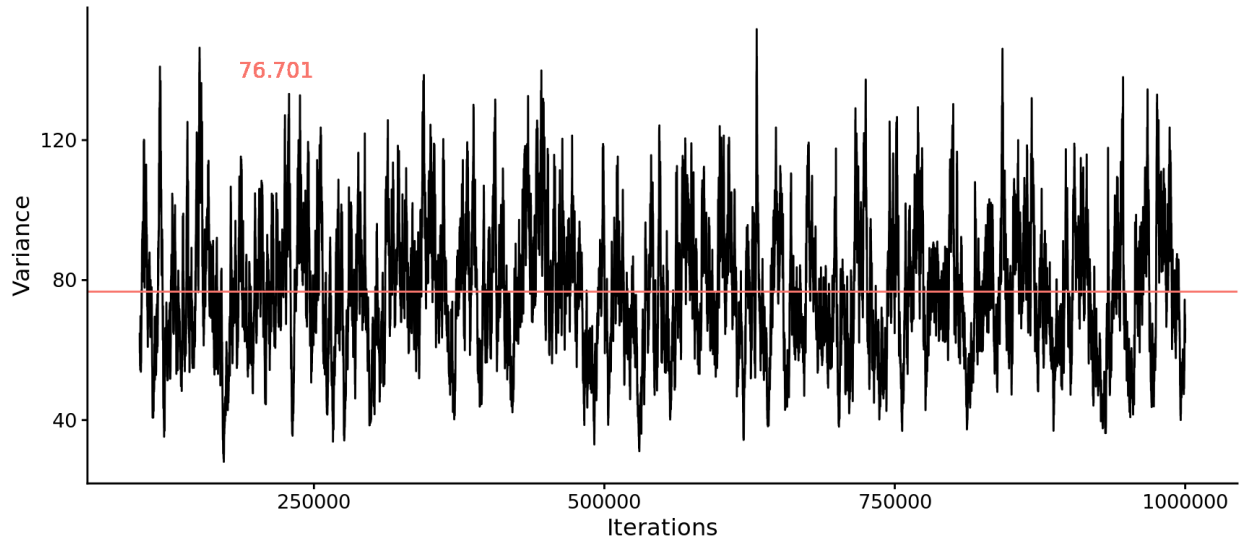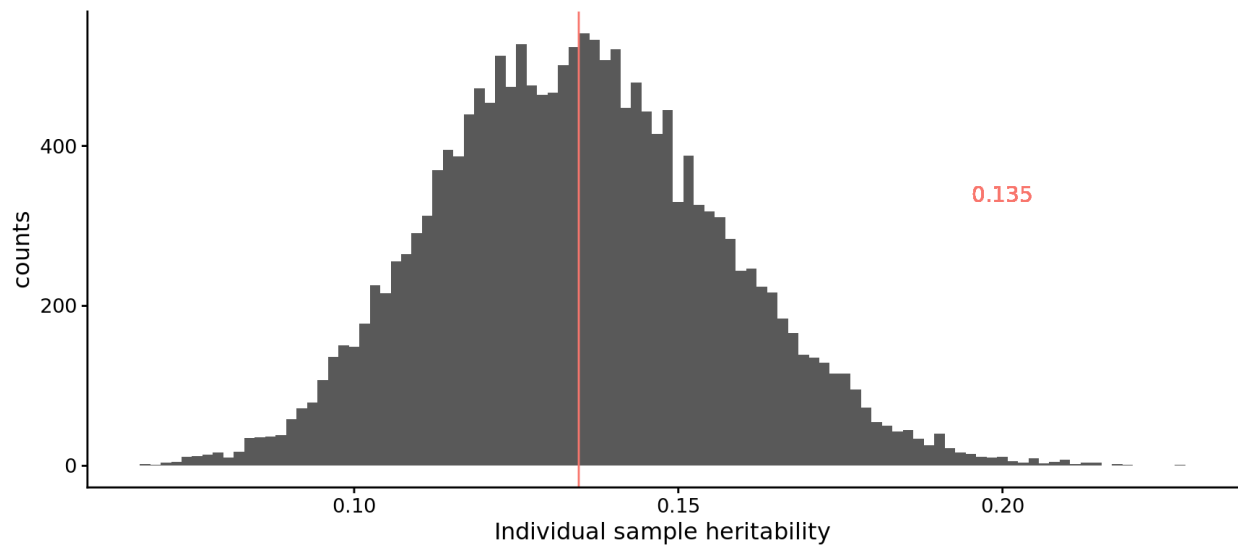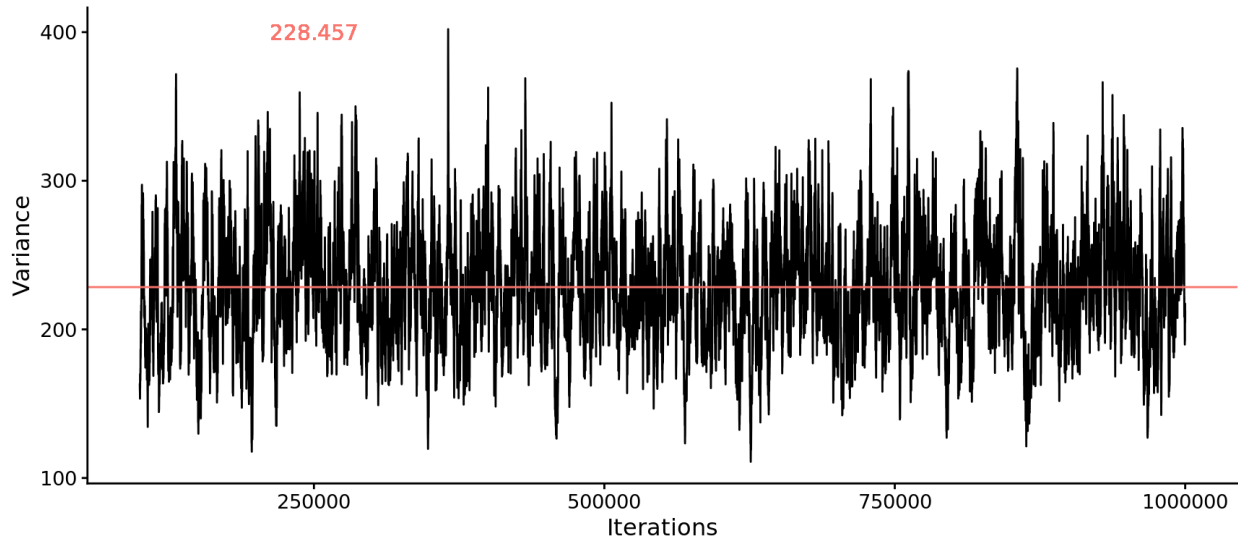
Figure 4 Trace and histogram plots for (a) Heifer pregnancy, (b) days to conception, (c) days open that show the convergence in threshold models using Gibb's sampling. The red line and their corresponding value show the mean posterior variance and heritability for trace plots and histograms.

2.3.2.2. Accuracy of the model

Three sets of the youngest heifers were selected as the validation sets to calculate the model accuracy. The accuracy formula was derived by Legarra et al. [78]. $Acc = \sqrt{\frac{cov(EBV_f, EBV_p)}{(1-F)\sigma_a^2}}$ where

Acc is the accuracy, cov is the covariance between the estimated full (EBV_f) and partial (EBV_p)

26

model breeding values, F is the mean inbreeding value of the individuals in the validation set, and σ_a^2 is the additive genetic variance from the full model. Among all the three traits, the model that fit days to conception resulted in higher accuracy of 0.38 ± 0.02, while days open and heifer pregnancy accuracies were 0.33 ± 0.03 and 0.25 ± 0.05, respectively. Figure 5 shows the correlation among EBVs in full and partial models for HP, DC, and DO in all three validation sets used to calculate the accuracy. Overall, there was an underestimation of the estimated breeding values for partial set given that the slopes are less than 1.

(a) HP

(b) DC

(c) DO
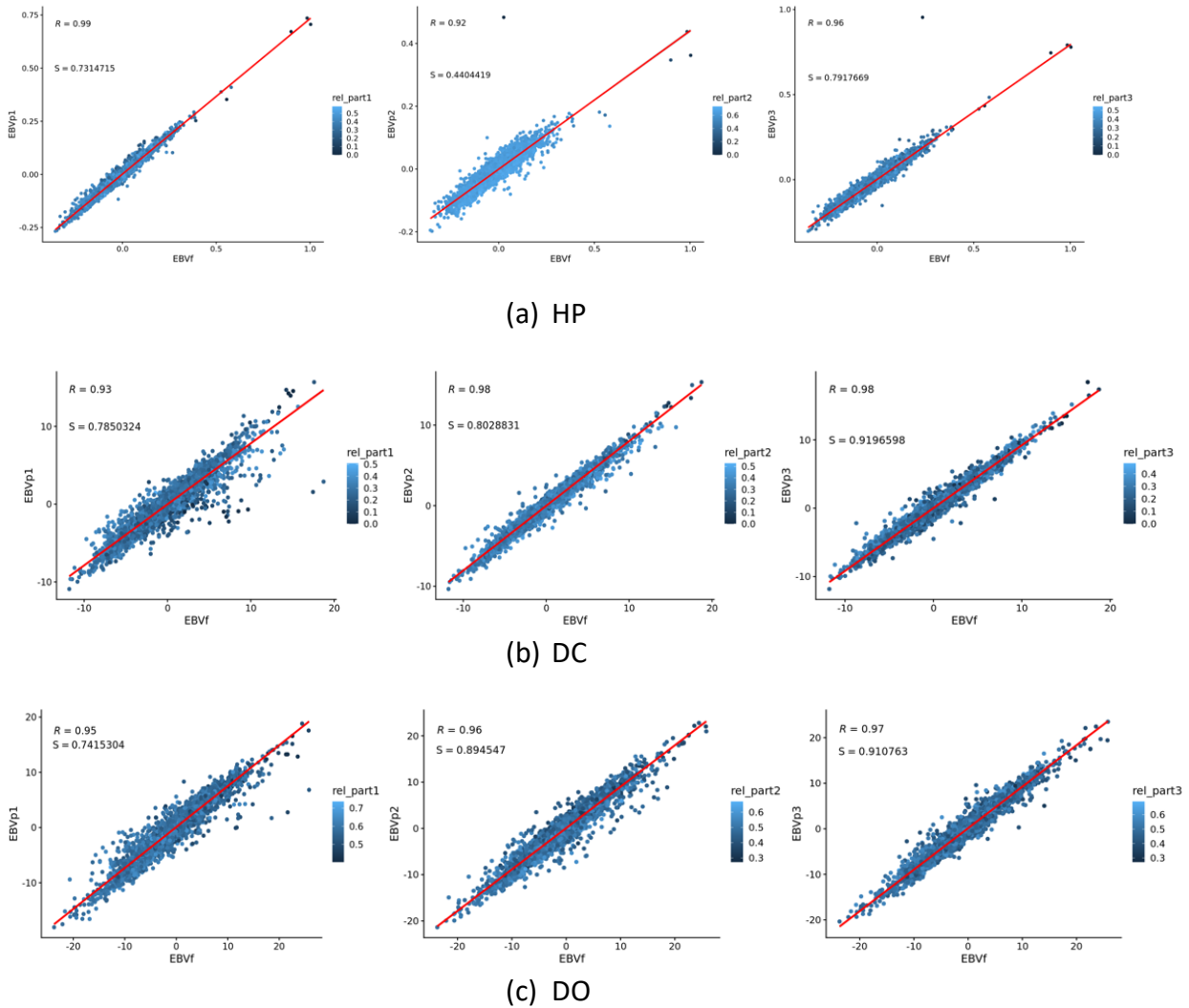
Figure 5 The relationship between full and partial breeding values in the validation set of (a) heifer pregnancy, (b) days to conception, and (c) days open.

The rel_part is the reliability in the partial validation set with the corresponding number (1, 2, and 3). The R is the Pearson correlation coefficient, S is the slope which is expected to be 1 is case there is no dispersion, and the blue line indicates regression.

2.3.3. Genome-wide association (GWAS) for SNP effect

2.3.3.1. Univariate GCTA and GEMMA

There was a lack of power in the GCTA univariate GWAS for all three traits as none of the SNPs reached the standard genome-wide significance level (5.0e-08) for common variants. On the other side, 62 SNPs found on chromosomes 2, 3, 4, 17, 18, 23, 25, and 26 reached the suggestive genome significance threshold level (p-value = 1.0e-05) in both GCTA and GEMMA univariate models. Among those, 47 SNPs were significant at the False Discovery Rate (FDR) less than 0.05, distributed on chromosomes 2, 3, and 23. Also, 12 SNPs on chromosomes 1, 2, 6, 13, and 24 for DC and 10 SNPs on chromosomes 8, 10, 12, and 22 for DO, reached suggestive significance levels (Manhattan and QQ plots in Appendix 2.6.3.). There was a weaker SNP association with DC and DO than HP. Neither DC nor DO had a significant SNP at the FDR < 0.05. Overall, there was an increase in power from GCTA to GEMMA (Figure 6). Apart from the SNPs that reached the suggestive significance level in GCTA, GEMMA showed additional 5 SNPs to reach the suggestive level, and 2 SNPs on chromosome 23 were significant at the genome-wide cutoff level (p-value < 5e-08), and 60 SNPs were significant at the 0.05 FDR.

(a)

(b)
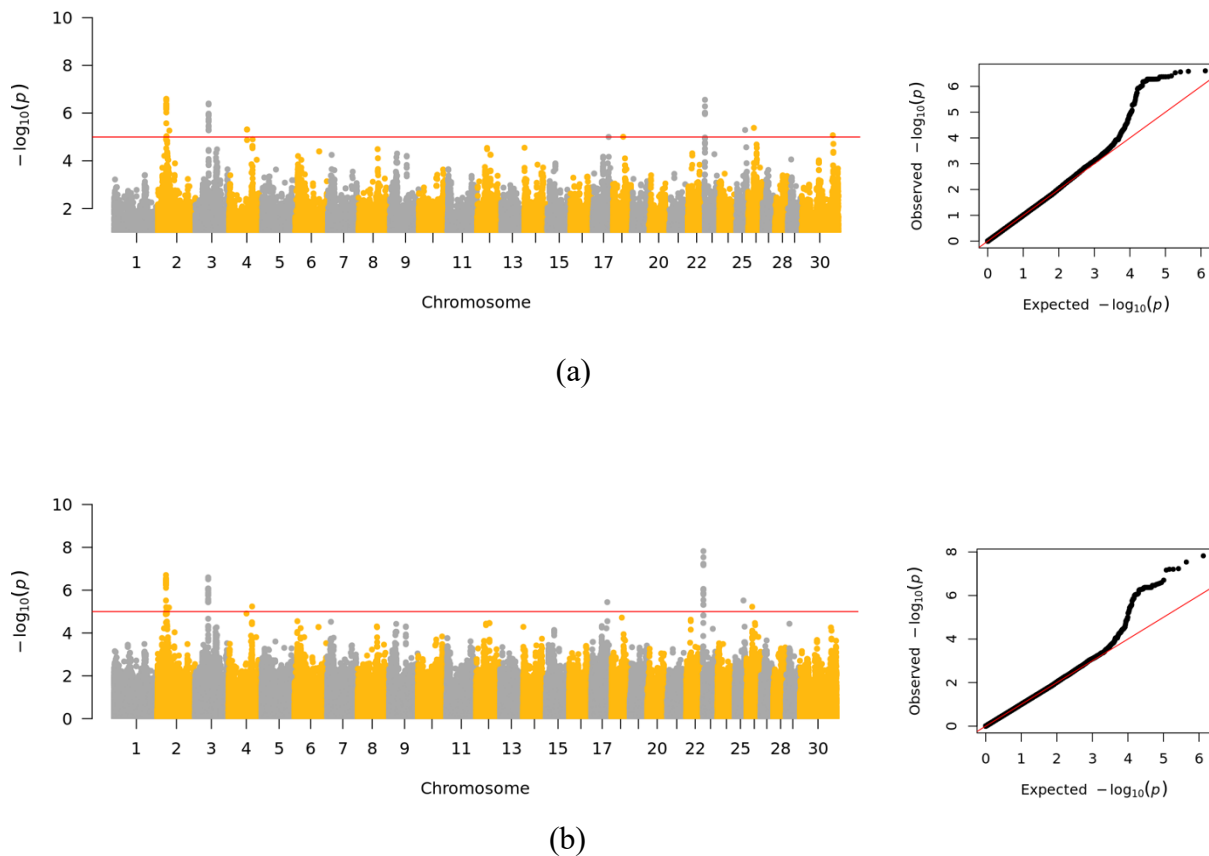
Figure 6 Manhattan plot and QQ plot for (a) GCTA and (b) GEMMA univariate GWAS in HP. The red line in Manhattan plot shows a cutoff level at p-value of 1e-05.

There was an increase in power from GCTA to GEMMA, and the deviation from red line in QQ plot show an association between SNPs and the traits of interest (HP).

2.3.3.2. **Bivariate GEMMA**

As described by [83], assessing the SNP association using a bivariate model is advantageous when two or multiple traits are believed to be correlated or affected by the same genes. During the bivariate GWAS analysis in GEMMA, both HP and DO were analyzed, given that those traits are genetically correlated. Among 662,044 SNPs, 59 reached the significance level (p-value = 1.0e-05) and 38 SNPs were significant at the FDR less than 0.05. In addition to the chromosomes from

29

the univariate models, there was a significant SNP on one more chromosome (BTA1) from the bivariate (Figure 7).
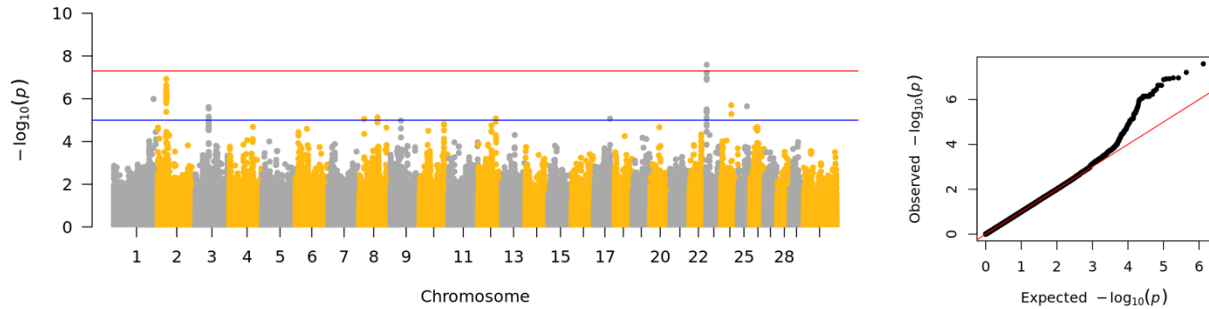
Figure 7 Manhattan plot and QQ plot for GEMMA bivariate GWAS in HP and DO. The red line shows the genome-wide significant level (p-value = 5e-08) while the blue line shows the suggestive significant level (p-value = 1e-05).

2.3.3.3. De-regressed Estimated Breeding Value (EBV) GWAS

Before performing the de-regressed EBV GWAS, the reliability was calculated from the Gibbs sampling models (Table 6). The same problem of having no statistical power for significant SNPs was observed for de-regressed EBVs. Only two SNPs reached the genome-wide significance level (p-value = 5e-08) for HP, while there were no significant SNPs in association with DC and DO (Figure 8). At the suggestive significant level, 58 SNPs and 2 SNPs were found to be associated with HP and DC, respectively. The Manhattan plots below show the significant SNPs for each trait and the corresponding chromosome loci.

Table 6 Reliabilities for heifer pregnancy (HP), days to conception, and days open on genotypes heifer (4097 individuals). SD is the standard deviation.

Trait	minimum	mean	maximum	SD
HP	0	0.1931	0.422	0.072
DC	0	0.223	0.493	0.072
DO	0.0792	0.419	0.667	0.075
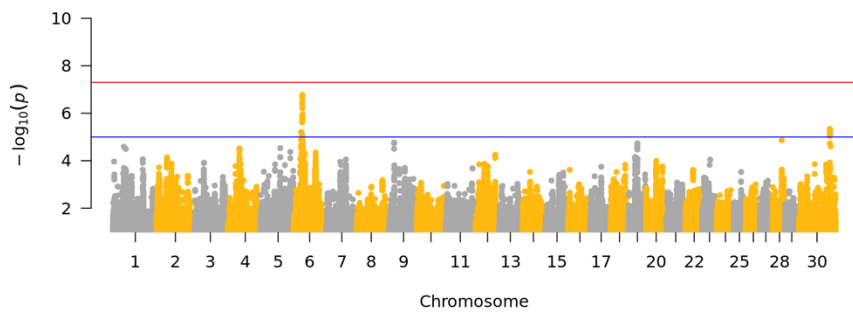
(a) HP
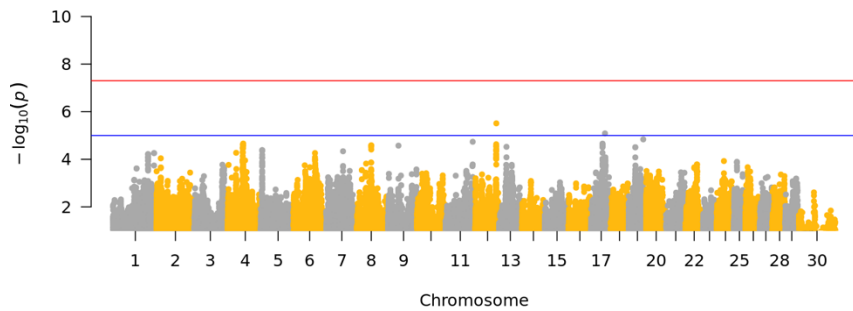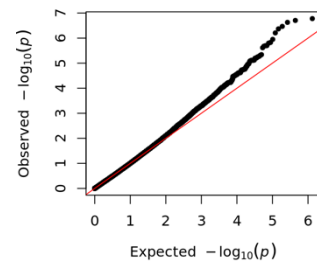
(b) DC
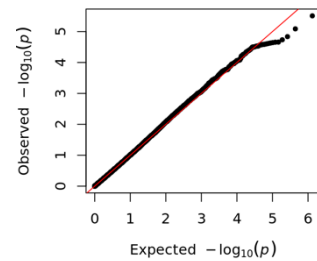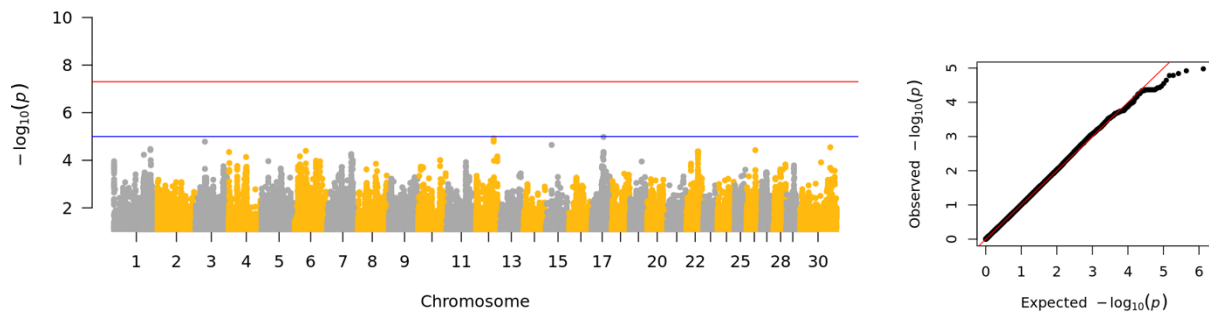
(c) DO

Figure 8 Manhattan QQ plots for (a) Heifer Pregnancy, (b) Days to Conception, and (c) Days Open based on de-regressed breeding values in the SNP1101 program.

The red line in the Manhattan plots shows the Genome-wide significant threshold level adjusted (p-value = 5e-08), and the blue line shows a suggestive significant level (p-value = 1e-05). The HP quantile-quantile (QQ) plot shows that there is an association (deviation from the red line) between a few SNPs and the traits while the plots for DC and DO reveal the lack of power

2.3.4. **Nearest genes within 10kb and Gene set enrichment analysis.**

Using Bovinemine v.16 with ARS-UCD1.2 genome assembly, there were six unique genes for HP and DO linear bivariate and 4 and 2 unique genes for HP and DC, respectively, from de-regressed EBV GWAS. There were no genes for DO as none of the SNPs reached at least the suggestive genome-wide significant level. g:Profiler [84] assisted in gene set enrichment analysis. At the Benjamin-Hochberg cutoff level of 0.05, the bivariate GEMMA resulted in 8 gene set terms. On the other hand, HP and DC de-regressed EBV GWAS analysis yielded 75 and 5 gene terms at the same cutoff level, respectively. GRID2 and ZMIZ1 were the most predominant genes in many terms, and they had been identified in other studies to affect fertility in different species (discussed later).

2.4. Discussion

In this study, the heifer pregnancy rate was 87.1%, which falls in the range of 75-95%, as reported in previous studies [23, 85]. Thus, we proceeded with the analysis of heifer pregnancy, knowing the pregnancy rate was in the normal range. Maintaining a high pregnancy rate is critical because problems associated with this trait are the leading cause of culling in the beef industry [86]. Although fertility traits are pivotal in the beef and dairy industry [20, 87–89], they are notorious for their low heritability, making it hard to make genetic progress [86, 90].

The fixed effects were small, and the negative sign shows that the increase in heifer's age will result in reduced days open and days to conception; however, the increase in season length will increase both days open and days to conception. The effect of both heifer's age and season length was relatively smaller than that of the previous two traits (DO & DC), and they were close to zero. The reduced effect from the linear model to Gibbs sampling suggests that the effect of both heifer's age and season length on the analyzed traits might be negligible when considering a very large sample size.

The current study confirms the low heritability nature of fertility traits. Even though the scaled heritability from the linear model was high (0.33), the threshold model (designed to handle binary data) did not confirm that heritability. Hence, the current study assumes algorithm problems during the adjustment and does not consider 0.33 heritability. The linear and threshold models resulted in HP heritability between 0.119 to 0.131 (Tables 3 and 5). This heritability is comparable to the 0.12 ± 0.01 HP heritability reported by Boldt et .al [23] in Red Angus. Various HP heritability estimates have been obtained across the breeds for the past years. For instance, Toelle and Robison reported the heifer pregnancy rate heritability of 0.06 among the Hereford breed [91]. Within the same breed, Evans et al. found the heritability of 0.138 ± 0.08 [21]. Also, in Angus

33

heifers, the pregnancy heritability was estimated to be between 0.13 and 0.27 [85, 92], and in Brangus, 0.07 [51]. Overall, many breeds show a low estimate of heritability. In their study, Martínez-Velázquez et al. combined nine breeds in one analysis, resulting in the heritability of 0.14 ± 0.03 [93]. Although the *Bos taurus* breeds show a low heifer pregnancy heritability, the inverse seems true for *Bos indicus* breeds. For example, the Nellore breeds showed a high HP heritability estimated to be between 0.37 to 0.57 [94–96]. It is not surprising that the heritability among *Bos taurus* and *Bos indicus* is different because these subspecies have distinct physiological differences that might explain such disparity.

Apart from heifer pregnancy, this study also considered two more fertility traits, days to conception and days open, which are novel in the beef industry. The linear models' genomic heritability estimates (without pedigree information) were 0.112 and 0.102 for DC and DO, respectively. In contrast, the combination of pedigree and genomic information models resulted in 0.0749 and 0.102 for DC and DO, respectively. The cause of the decline in DC heritability is not understood yet. However, it might be because of the increase in the ratio of pregnant heifers over non-pregnant, implying the bias in the previous model with a small sample size. Not many works have been done about DO and DC in the beef industry. In the study done in Spain, the reported heritability of DO in beef cattle ranged from 0.09 to 0.197 [36]. That range is not far from the obtained heritabilities in the current study, though the definition of days open (days from one calving to the next conception) does not quietly match the present study. In the dairy industry, the exact definition of DO - days from calving to next conception - has been used, and a couple of studies reported the heritability between 0.02 to 0.06 [44, 97–102]. The interval from first to successful AI in the study of black Japanese cows by Setiaji and Oikawa is close to DC in this study, and the obtained heritability was 0.028 [99]. The reported heritabilities in the current study

seem to be higher than those reported about the dairy industry. That might be predicated on the worsening of reproductive traits of the dairy cows that happened while focusing on improving the production traits (milk yield) [16].

The genetic correlation between DO and HP from the linear model was -0.61. The threshold model resulted in a 0.85 correlation, and the positive direction was because the solutions for HP were on the liability scale. The linear bivariate model for HP and DC failed to converge. While the threshold model could ignore missing values of DC and fit the bivariate model with only pregnant heifers, the results from that model would be biased due to the lack of not analyzing data on non-pregnant heifers. Consequently, the scaled model was not used to predict the correlation between HP and DC to avoid that bias. To the author's knowledge, there has not been any known study that evaluated the genetic correlation between DO and HP among beef cattle. However, some studies assessed the correlation with other fertility traits. The correlation between pregnancy rate and days open in Japanese black cows was -0.92 [99]. The correlation differences with the current study may be caused by how traits are defined and the unlikeness of the breeds. A negative correlation between DO and HP suggests that some common genes affect both traits differently. Consequently, it is possible to improve one trait while meticulously selecting the other. That is, while selecting on lower days open, there is a high probability of increasing the heifer pregnancy.

In this study, the model accuracies were 0.38, 0.33, and 0.25 for DC, DO, and HP, respectively. Despite its high heritability compared to the other two traits, the HP model resulted in lower accuracy than DC and DO models. That was expected as the quantitative variables contain more information than binary variables [103]. Therefore, selecting days to conception or days open might lead to more rapid genetic improvement than selection on heifer pregnancy.

Most previous studies on beef fertility used traditional genetic approaches (prediction based on pedigree) for variance and heritability estimation, and almost none considered the model validation. It was easy to overlook the fertility traits during selection decisions based on the results that showed they are lowly heritable. The current study agrees with the low heritability status of HP, DO, and DC, but the achieved accuracies are the best indicators that genomic selection can elaborate on these traits. There is a hope that with sophisticated genomic tools and statistical methods, animal producers should start to invest in improving fertility traits, expecting that there will be some improvement while applying genomic selection.

The bivariate linear mixed model of HP and DO gave 6 unique genes classified into 8 gene ontology terms (Table 7). Those genes are DLGAP1 FBXO21, LOC112445875, SCN3A, SLC37A1, and SLC38A1. According to the GWAS catalog [104], most of those genes (DLGAP1, FBXO21, SCN3A, and SLC38A11) are related to body mass index traits in humans; but it is not clear about the role they play in HP and DO or any other fertility traits in general.

Table 7 Gene terms at Benjamin-Hochberg correction p-value < 0.05 in association with HP and DO.

Source	Term name	Intersections	P-value
GO:BP	Hexose phosphate transport	SLC37A1	0.0224
GO:BP	Glucose-6-phosphate transport	SLC37A1	0.0224
GO:BP	Phosphate ion transmembrane transport	SLC37A1	0.0358
GO:BP	Phosphate ion transport	SLC37A1	0.0391
GO:BP	Inorganic ion transmembrane transport	SCN3A, SLC37A1	0.0391

GO:BP	Regulation of postsynaptic neurotransmitter receptor activity	DLGAP1	0.0391
KEGG	Glutamatergic synapse	DLGAP1	0.0394
KEGG	Taste transduction	SCN3A	0.0394

From de-regressed breeding values GWAS, there were 4 and 2 unique genes with 75 and 5 gene set terms (Appendix 2.6.4. table 4 and 5) for HP and DC, respectively, at the Benjamin-Hochberg p-value ≤ 0.05. Both genes identified from DC are indicated in the human GWAS catalog to be associated with body mass index [104]. If that is the case in this study, the body condition score effect on the fertility traits should be assessed in future research. Overall, most of the gene set terms for HP are related to the central nervous system, and it has been speculated that they may play a role in the secretion of some reproductive hormones. Moreover, GRID2 and ZMIZ1 genes have been identified in the previous studies to have an influence on reproductive traits among different species. For instance, Vohra et al. found GRID2 among the genes associated with the postpartum interval in Indian buffalo [105]. In addition, in human clinical studies, the intragenic deletion in GRID2 has been associated with a neurological disorder with developmental delay as one of the major symptoms [106]. Although it can be argued that the deletion in GRID2 resulted in a bigger effect that might not be seen when considering a small allele change (the case of SNPs), the fact that the function of this gene is not fully understood can be the reason to study in depth of its effect on fertility as it might be affecting the pubertal behavior.

Another group of researchers from China identified ZMIZ1 as one of the genes that were strongly associated with the interval from calving to first insemination among Chinese Holstein [107]. Not only was ZMIZ1 found to enrich the transcription activity of androgen receptors [108],

but May-Panloup et al. found that its less expression level was linked with a diminished ovarian reserve which is an indicator of subfertility [109]. Therefore, it would be important to conduct a validation study for this gene while assessing its effect on pregnancy.

Contrary to the expectation, a drastic reduction in power for the SNP effect was observed; thus, very few genes were associated with the analyzed traits while using de-regressed EBV for GWAS. One possible reason for that might be the small reliability of the estimated EBVs, and it implies that an enormous sample size with phenotypic and genotypic information is needed to conduct analyses on fertility that will result in high reliability, hence, the gain of power.

2.5. Conclusion

The availability of sophisticated computational methods and the improvement in DNA sequencing increase prediction accuracy in hard-to-measure reproductive traits. The current study estimated the genetic parameters for heifer pregnancy, days to conception, and days open which are the fertility traits that are easy to measure and contain enough information for genomic studies. Both linear and threshold models result in almost the same heritability for binary traits, indicating the robustness of the developed linear model tools that successfully handle binary data. Like the previous studies, the heritability of fertility was relatively low. Even though the multivariate between HP and DC did not converge, the high correlation between DO and HP shows that indirect selection for one of these traits can be used to improve the overall reproductive performance in Red Angus.

Furthermore, despite the variation in DC heritability, this variable yielded higher accuracy than the rest. Given its quantitative nature, future studies should assess the feasibility of using DC in multi-trait models that fit binary and continuous variables. Some genes were associated with HP, DC, and DO, but no known biological processes relate those genes to fertility. Further studies

are needed to confirm the influence of the obtained genes and their corresponding gene terms on heifer reproduction. While the current study considered only one breed, future studies should try multibreed analysis with a large sample size to confirm the effect of obtained genes (GRID2 & ZMIZ1) on beef heifer fertility and the feasibility of DC and DO, which are considered novel traits in the industry.

2.6.Reference:

1. Meirelles F v., Rosa AJM, Lôbo RB, Garcia JM, Smith LC, Duarte FAM. Is the American Zebu really Bos indicus? Genetics and Molecular Biology. 1999;22:543–6.
2. Ajmone-Marsan P, Garcia JF, Lenstra JA. On the origin of cattle: How aurochs became cattle and colonized the world. Evolutionary Anthropology. 2010;19:148–57.
3. Willham RL. GENETIC IMPROVEMENT OF BEEF CATTLE IN THE U N I T E D STATES : CATTLE , PEOPLE A N D THEIR INTERACTION 1 ' 2. 1982.
4. Garrick DJ, Golden BL. Producing and using genetic evaluations in the United States beef industry of today. J Anim Sci. 2009;87 14 Suppl.
5. Garrick DJ. The nature, scope and impact of genomic prediction in beef cattle in the United States. Genetics Selection Evolution. 2011;43:1–11.
6. Enns RM, Nicoll GB. Genetic change results from selection on an economic breeding objective in beef cattle. Journal of Animal Science. 2008;86:3348–57.
7. Marshall DM. Breed differences and genetic parameters for body composition traits in beef cattle. J Anim Sci. 1994;72:2745–55.
8. Bertrand JK, Green RD, Herring WO, Moser DW. Genetic evaluation for beef carcass traits. Journal of Animal Science. 2001;79 E-Suppl:E190.
9. Rolfe KM, Snelling WM, Nielsen MK, Freetly HC, Ferrell CL, Jenkins TG. Genetic and phenotypic parameter estimates for feed intake and other traits in growing beef cattle, and opportunities for selection. Journal of Animal Science. 2011;89:3452–9.
10. Golden BL, Garrick DJ, Newman S, Enns RM, Collins F, Zealand N. ECONOMICALLY RELEVANT TRAITS A FRAMEWORK FOR THE NEXT. Proceedings of the 32nd Research Symposium and Annual Meeting of the Beef Improvement Federation. 2000;:2–13.
11. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and characterization of a high-density SNP genotyping assay for cattle. PLoS ONE. 2009;4:e5350.
12. Martínez R, Bejarano D, Gómez Y, Dasoneville R, Jiménez A, Even G, et al. Genome-wide association study for birth, weaning and yearling weight in Colombian Brahman cattle. Genetics and Molecular Biology. 2017;40:453–9.
13. Grigoletto L, Ferraz JBS, Oliveira HR, Eler JP, Bussiman FO, Abreu Silva BC, et al. Genetic Architecture of Carcass and Meat Quality Traits in Montana Tropical® Composite Beef Cattle. Frontiers in Genetics. 2020;11:123.
14. Berry DP, Garcia JF, Garrick DJ. Development and implementation of genomic predictions in beef cattle. Animal Frontiers. 2016;6:32–8.
15. González-Recio O, Pérez-Cabal MA, Alenda R. Economic Value of Female Fertility and its Relationship with Profit in Spanish Dairy Cattle. Journal of Dairy Science. 2004;87:3053–61.
16. Lucy M. Reproductive loss in high-producing dairy cattle: where will it end? J Dairy Sci. 2001;84:1277–93.
17. Wiltbank JN. Research Needs in Beef Cattle Reproduction. Journal of Animal Science. 1970;31:755–62.
18. EPD Trends - Red Angus. https://redangus.org/genetics/epd-trends/#epd-trends-summer-2022. Accessed 16 Mar 2022.
19. Rowan T, Durbin H, Seabury C, Schnabel R, Decker J. Powerful detection of polygenic selection and evidence of environmental adaptation in US beef cattle. BioRxiv. 2020. https://doi.org/10.1101/2020.03.11.988121.

20. Cammack KM, Thomas MG, Enns RM. Reproductive Traits and Their Heritabilities in Beef Cattle. The Professional Animal Scientist. 2009;25:517–28.

21. Evans JL, Golden BL, Bourdon RM, Long KL. Additive Genetic Relationships Between Heifer Pregnancy and Scrotal Circumference in Hereford Cattle. J Anim Sci. 1999;77:2621–8.

22. Snelling WM, Golden BL, Bourdon RM. Within-herd genetic analyses of stayability of beef females. J Anim Sci. 1995;73:993–1001.

23. Boldt RJ, Speidel SE, Thomas MG, Enns RM. Genetic parameters for fertility and production traits in Red Angus cattle. Journal of Animal Science. 2018;96:4100–11.

24. Roberts AJ, Gomes Da Silva A, Summers AF, Geary TW, Funston RN. Developmental and reproductive characteristics of beef heifers classified by pubertal status at time of first breeding. Journal of Animal Science. 2017;95:5629–36.

25. Hersom M, Thrift T, Yelich J. Culling and Replacement Rate in the Beef Cow Herd. 2018;:1–4.

26. González-Recio O, Pérez-Cabal MA, Alenda R. Economic value of female fertility and its relationship with profit in Spanish Dairy Cattle. Journal of Dairy Science. 2004;87:3053–61.

27. What is the Economic Impact of Infertility in Beef Cattle? https://www.thebeefsite.com/articles/1698/what-is-the-economic-impact-of-infertility-in-beef-cattle/. Accessed 19 Mar 2022.

28. Mercadante VRG, Dias NW, Timlin CL, Pancini S. 375 Economic Consequences of Pregnancy Loss in Beef Cattle. Journal of Animal Science. 2020;98 Supplement_4:124–124.

29. Burns BM, Fordyce G, Holroyd RG. A review of factors that impact on the capacity of beef cattle females to conceive, maintain a pregnancy and wean a calf—Implications for reproductive efficiency in northern Australia. Animal Reproduction Science. 2010;122:1–22.

30. Snelling WM, Golden BL, Bourdon RM, Snellingl WM. Within-Herd Genetic Analyses of Stayability of Beef Females. 2013.

31. Doyle SP, Golden BL, Green RD, Brinks JS. Additive genetic parameter estimates for heifer pregnancy and subsequent reproduction in Angus females. Journal of Animal Science. 2000;78:2091–8.

32. van Melis MH, Eler JP, Rosa GJM, Ferraz JBS, Figueiredo LGG, Mattos EC, et al. Additive Genetic Relationships Between Heifer Pregnancy and Scrotal Circumference in Hereford Cattle. Journal of Animal Science. 2010;88:3809–13.

33. Fennewald DJ, Weaber RL, Lamberson WR. Genotype by environment interaction for stayability of red angus in the United States. Journal of Animal Science. 2018;96:422–9.

34. Minick Bormann J, Wilson DE. Calving day and age at first calving in angus heifers. Journal of Animal Science. 2010;88:1947–56.

35. Gutiérrez JP, Goyache F, Fernández I, Alvarez I, Royo LJ. Genetic relationships among calving ease, calving interval, birth weight and weaning weight in the Asturiana de los Valles beef cattle breed. Journal of Animal Science. 2007;85:69–75.

36. Goyache F, Gutiérrez JP, Fernández I, Royo LJ, Álvarez I. Genetic analysis of days open in beef cattle. Livestock Production Science. 2005;93:283–9.

37. Martínez-Velázquez G, Gregory KE, Bennett GL, van Vleck LD. Genetic relationships between scrotal circumference and female reproductive traits. Journal of Animal Science. 2003;81:395–401.

38. Speidel SE, Buckley BA, Boldt RJ, Enns RM, Lee J, Spangler ML, et al. Genome-wide association study of stayability and heifer pregnancy in red angus cattle. Journal of Animal Science. 2018;96:846–53.

39. Stegemiller MR, Murdoch GK, Rowan TN, Davenport KM, Becker GM, Hall JB, et al. Genome-wide association analyses of fertility traits in beef heifers. Genes (Basel). 2021;12:1–15.

40. MacNeil MD, Nkrumah JD, Woodward BW, Northcutt SL. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. J Anim Sci. 2010;88:517–22.

41. Vanek JK, Watts MJ, Brester GW. Carcass quality and genetic selection in the beef industry. Journal of Agricultural and Resource Economics. 2008;33:349–63.

42. Golden BL, Garrick DJ, Benyshek LL. Milestones in beef cattle genetic evaluation. J Anim Sci. 2009;87 14 Suppl.

43. Cushman RA, Allan MF, Kuehn LA. Characterization of biological types of cattle: Indicator traits of fertility in beef cows. Revista Brasileira de Zootecnia. 2008;37 SPECIALISSUE:116–21.

44. VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HO, Kuhn MT, et al. Development of a national genetic evaluation for cow fertility. Journal of Dairy Science. 2004;87:2285–92.

45. Naya H, Peñagaricano F, Urioste JI. Modelling female fertility traits in beef cattle using linear and non-linear models. Journal of Animal Breeding and Genetics. 2017;134:202–12.

46. Jamrozik J, McGrath S, Kemp RA, Miller SP. Genetic analysis of female fertility traits in Canadian Simmentals. Livestock Science. 2012;150:302–9.

47. Moorey SE, Biase FH. Beef heifer fertility: importance of management practices and technological advancements. J Anim Sci Biotechnol. 2020;11:97.

48. Perry GA. Factors affecting puberty in replacement beef heifers. Theriogenology. 2016;86:373–8.

49. Utter SD, Houghto PL, Corah LR, Simms DD, Spir MF, Butine MD. FACTORS INFLUENCING FIRST-SERVICE CONCEPTION AND OVERALL PREGNANCY RATES IN COMMERCIAL BEEF HEIFERS 1.

50. Philipsson J. Genetic aspects of female fertility in dairy cattle. Livestock Production Science. 1981;8:307–19.

51. Fortes MRS, Snelling WM, Reverter A, Nagaraj SH, Lehnert SA, Hawken RJ, et al. Gene network analyses of first service conception in Brangus heifers: Use of genome and trait associations, hypothalamic-transcriptome information, and transcription factors1. Journal of Animal Science. 2012;90:2894–906.

52. Toghiani S, Hay E, Sumreddee P, Geary TW, Rekaya R, Roberts AJ. Genomic prediction of continuous and binary fertility traits of females in a composite beef cattle breed. Journal of Animal Science. 2017;95:4787.

53. Akanno EC, Chen L, Abo-Ismail MK, Crowley JJ, Wang Z, Li C, et al. Genomic prediction of breed composition and heterosis effects in Angus, Charolais, and Hereford crosses using 50K genotypes. https://doi.org/101139/cjas-2016-0124. 2017;97:431–8.

54. Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL, et al. Heritability and bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. Journal of Animal Science. 2013;91:605–12.

55. Snelling WM, Cushman RA, Fortes MRS, Reverter A, Bennett GL, Keele JW, et al. Physiology and endocrinology symposium: How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. Journal of Animal Science. 2012;90:1152–65.

56. Rowan TN, Hoff JL, Crum TE, Taylor JF, Schnabel RD, Decker JE. A multi-breed reference panel and additional rare variants maximize imputation accuracy in cattle. Genetics Selection Evolution. 2019;51:1–16.

57. Eler JP, Bignardi AB, Ferraz JBS, Santana JL. Genetic relationships among traits related to reproduction and growth of Nelore females. Theriogenology. 2014;82:708–14.

58. Torell R. Back to Basics: Gestation length of the beef cow vs. dystocia. Angus Beef Bulletin. 2009; February:84.

59. Larson RL, White BJ, Laflin S. Beef Heifer Development. Veterinary Clinics of North America - Food Animal Practice. 2016;32:285–302.

60. Dow JS, Moore JD, Bailey CM, Foote WD. ONSET OF PUBERTY IN HEIFERS OF DIVERSE BEEF BREEDS AND CROSSES 1.

61. Moorey SE, Biase FH. Beef heifer fertility: Importance of management practices and technological advancements. Journal of Animal Science and Biotechnology. 2020;11.

62. Byerley DJ, Staigmiller RB, Berardinelli JG, Short RE. Pregnancy rates of beef heifers bred either on puberal or third estrus. J Anim Sci. 1987;65:645–50.

63. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9:1–9.

64. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics. 2007;81:559–75.

65. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. American Journal of Human Genetics. 2011;88:76–82.

66. Ugarte E, Alenda R, Carabaño MJ. Fixed or Random Contemporary Groups in Genetic Evaluations. Journal of Dairy Science. 1992;75:269–78.

67. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics. 2012;44:821–4.

68. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nature Methods 2014 11:4. 2014;11:407–9.

69. S Falconer BD. The inheritance of liability to certain diseases, estimated from the incidence among relatives. 1966;29:61.

70. Masuda Y. Variance Component Estimation. 2019;40:291–305. https://masuday.github.io/blupf90_tutorial/vc_gs.html. Accessed 14 Feb 2022.

71. Tsuruta S, Misztal I. THRGIBBS1F90 for estimation of variance components with threshold and linear models. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. 2006;:27–31.

72. Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z. Manual for BLUPF90 family of programs. Athens, USA. 2018;:142.

73. readme.pregsf90 [BLUPF90]. http://nce.ads.uga.edu/wiki/doku.php?id=readme.pregsf90. Accessed 14 Feb 2022.

74. Aguilar I, Misztal I, Legarra A, Tsuruta S. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. Journal of Animal Breeding and Genetics. 2011;128:422–8.

75. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science. 2010;93:743–52.

76. Legarra A, Aguilar I, Misztal I. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science. 2009;92:4656–63.

77. VanRaden P. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.

78. Legarra A, Reverter A. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method 01 Mathematical Sciences 0104 Statistics. Genetics Selection Evolution. 2018;50.

79. Bermann M, Legarra A, Hollifield MK, Masuda Y, Lourenco D, Misztal I. Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: An application in chicken mortality. Journal of Animal Breeding and Genetics. 2021;138:4–13.

80. Sargolzaei M. SNP1101 User Guide. 2014;0 March.

81. Elsik CG, Unni DR, Tayal A, Diesh CM, Hagen DE. P1038 BovineMine: A bovine genome data mining warehouse. Journal of Animal Science. 2016;94 suppl_4:33–33.

82. Gibbs RA, Taylor JF, van Tassell CP, Barendse W, Eversole KA, Gill CA, et al. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. Science (1979). 2009;324:528–32.

83. Stephens M. A Unified Framework for Association Analysis with Multiple Related Phenotypes) A Unified Framework for Association Analysis with Multiple Related Phenotypes. PLoS ONE. 2013;8:65245.

84. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research. 2019;47:W191–8.

85. Bormann JM, Totir LR, Kachman SD, Fernando RL, Wilson DE. Pregnancy rate and first-service conception rate in Angus heifers. Journal of Animal Science. 2006;84:2022–5.

86. Morris CA, Wilson JA, Bennett GL, Cullen NG, Hickey SM, Hunter JC. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. New Zealand Journal of Agricultural Research. 2000;43:83–91.

87. Day ML, Nogueira GP. Management of age at puberty in beef heifers to optimize efficiency of beef production. Animal Frontiers. 2013;3:6–11.

88. Weigel KA, Rekaya R. Genetic parameters for reproductive traits of Holstein cattle in California and Minnesota. Journal of Dairy Science. 2000;83:1072–80.

89. Larracharte A, Espasandin AC, Urioste JI. Longevity and reproductive traits in Angus cattle: Genetic parameters, predicted and realized genetic change. Livestock Science. 2021;250.

90. Foote RH. Inheritance of Fertility—Facts, Opinions, and Speculations. Journal of Dairy Science. 1970;53:936–44.

91. Toelle VD, Robison OW. Estimates of genetic correlations between testicular measurements and female reproductive traits in cattle. J Anim Sci. 1985;60:89–100.

92. Doyle SP, Golden BL, Green RD, Brinks JS. Additive genetic parameter estimates for heifer pregnancy and subsequent reproduction in Angus females. Journal of Animal Science. 2000;78:2091–8.

93. Martínez-Velázquez G, Gregory KE, Bennett GL, van Vleck LD. Genetic relationships between scrotal circumference and female reproductive traits. Journal of Animal Science. 2003;81:395–401.

94. Eler JP, Bignardi AB, Ferraz JBS, Santana JL. Genetic relationships among traits related to reproduction and growth of Nelore females. Theriogenology. 2014;82:708–14.

95. van Melis MH, Eler JP, Rosa GJM, Ferraz JBS, Figueiredo LGG, Mattos EC, et al. Additive genetic relationships between scrotal circumference, heifer pregnancy, and stayability in Nellore cattle. Journal of Animal Science. 2010;88:3809–13.

96. Eler JP, v. Silva II JA, Evans JL, Ferraz JBS, Dias F, Golden BL. Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle. Journal of Animal Science. 2004;82:2519–27.

97. Oseni S, Misztal I, Tsuruta S, Rekaya R. Genetic components of days open under heat stress. Journal of Dairy Science. 2004;87:3022–8.

98. Liu Z, Jaitner J, Reinhardt F, Pasman E, Rensing S, Reents R. Genetic evaluation of fertility traits of dairy cattle using a multiple-trait animal model. Journal of Dairy Science. 2008;91:4333–43.

99. Setiaji A, Oikawa T. Genetic parameters of reproductive traits from artificial insemination records of Japanese Black cows. Livestock Science. 2019;229:85–9.

100. López De Maturana E, Legarra A, Varona L, Ugarte E. Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data. Journal of Dairy Science. 2007;90:2012–24.

101. Schulman NF, Sahana G, Lund MS, Viitala SM, Vilkki JH. Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. Genetics Selection Evolution. 2008;40:195–214.

102. Oseni S, Tsuruta S, Misztal I, Rekaya R. Genetic Parameters for Days Open and Pregnancy Rates in US Holsteins Using Different Editing Criteria. Journal of Dairy Science. 2004;87:4327–33.

103. Kizilkaya K, Fernando RL, Garrick DJ. Reduction in accuracy of genomic prediction for ordered categorical data compared to continuous observations. Genetics Selection Evolution. 2014;46:1–13.

104. Buniello A, Macarthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research. 2019;47:D1005–12.

105. Vohra V, Chhotaray S, Gowane G, Alex R, Mukherjee A, Verma A, et al. Genome-Wide Association Studies in Indian Buffalo Revealed Genomic Regions for Lactation and Fertility. Frontiers in Genetics. 2021;12.

106. Veerapandiyan A, Enner S, Thulasi V, Ming X. A Rare Syndrome of GRID2 Deletion in 2 Siblings. Child Neurol Open. 2017;4:2329048X17726168.

107. Liu A, Wang Y, Sahana G, Zhang Q, Liu L, Lund MS, et al. Genome-wide Association Studies for Female Fertility Traits in Chinese and Nordic Holsteins. Scientific Reports 2017 7:1. 2017;7:1–12.

108. Li X, Zhu C, Tu WH, Yang N, Qin H, Sun Z. ZMIZ1 Preferably Enhances the Transcriptional Activity of Androgen Receptor with Short Polyglutamine Tract. PLOS ONE. 2011;6:e25040.

109. May-Panloup P, Ferr-Lhtellier V, Morinire C, Marcaillou C, Lemerle S, Malinge MC, et al. Molecular characterization of corona radiata cells from patients with diminished ovarian reserve using microarray and microfluidic-based gene expression profiling. Human Reproduction. 2012;27:829–43.
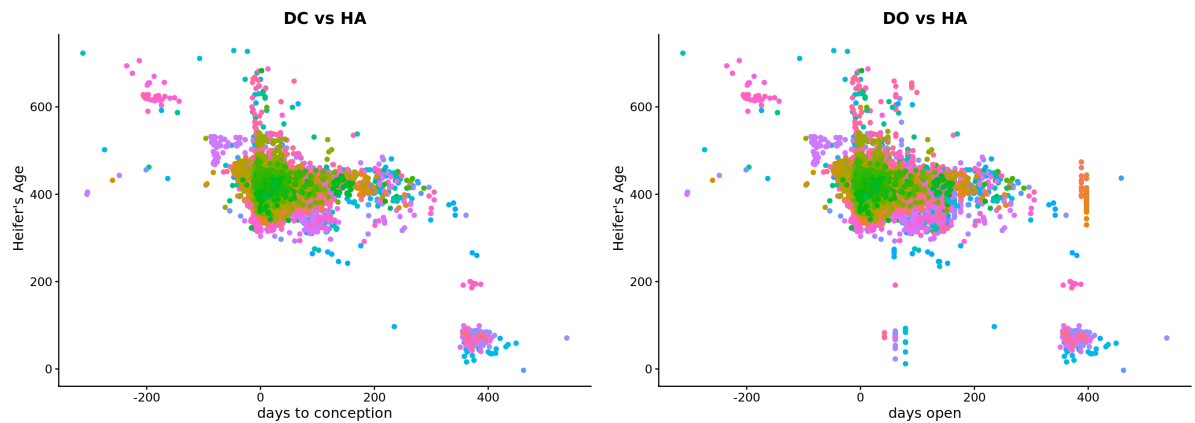
2.7. Appendix

2.7.1. Figures before the quality control
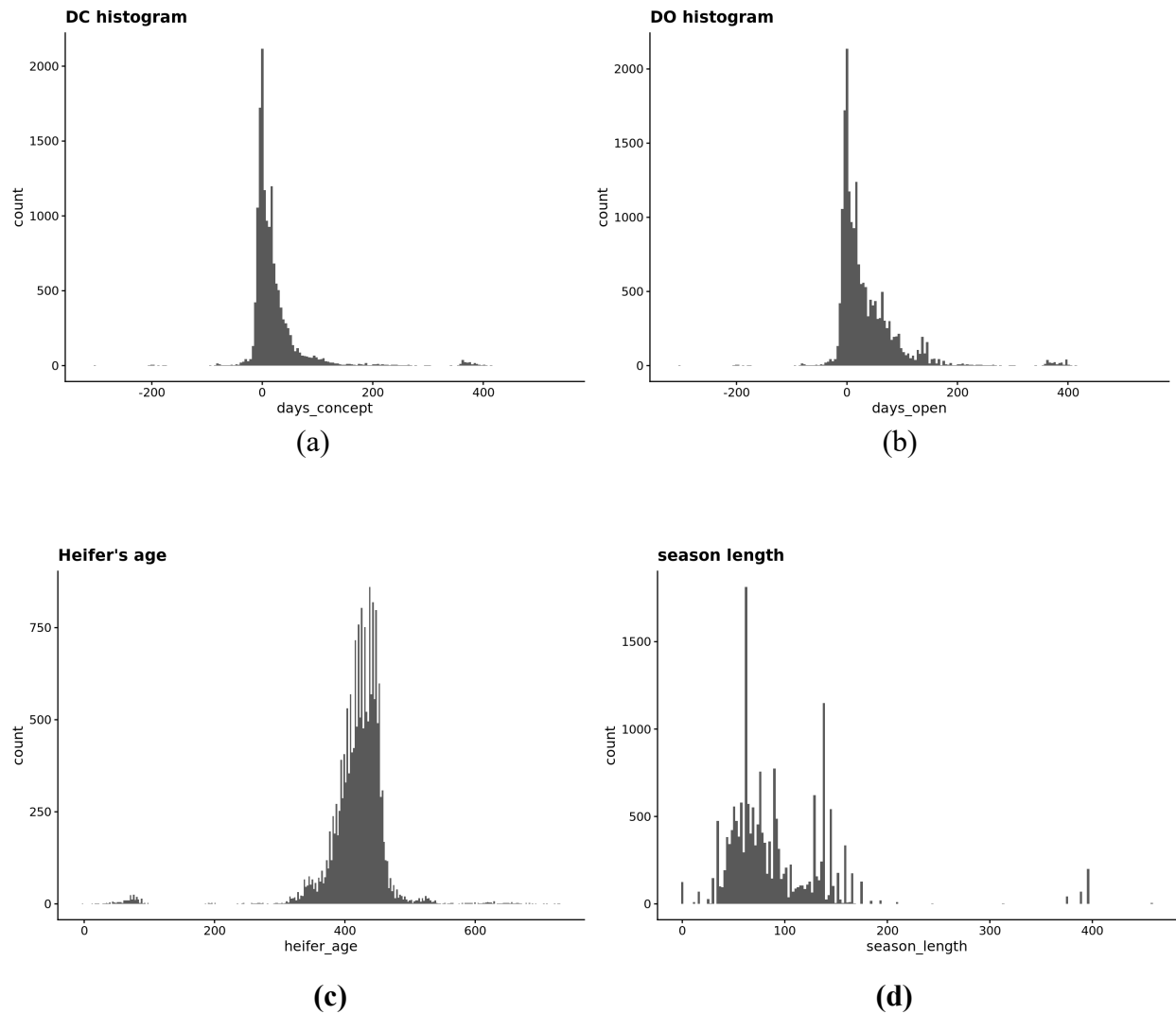
Figure 1. The histogram of (a) days to conception, (b) days open, (c) age, and season length before quality control.

For days interval (DC & DO), individuals less -50 and above 200 days seem to be outliers.

For age, many individuals are within 300 and 500 days old, while season length with many
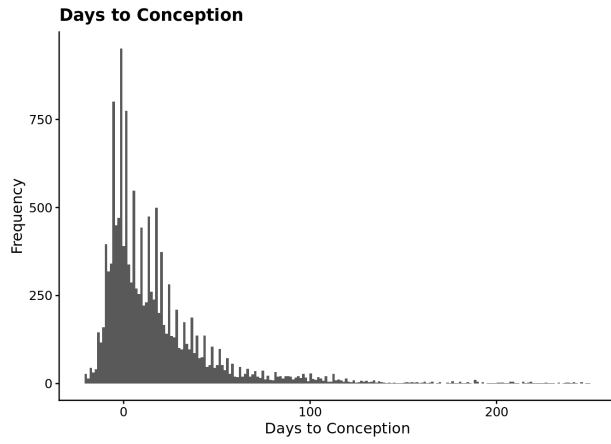
individuals ranges 0 to 200 days.

Figure 2. Scatter plots relating (a) days open and (b) days to conception with heifer's age after quality control. Note that colors indicate birthdate.

2.7.2. Figures after data quality control

The quality control was done based on the literature, and the distribution of raw data.

- First, changed the -9 days interval (DO&DC) to -9.1

- At the same time, removed heifers with more than 365 and less than -30 days and less than 200 days of age (Based on the distribution and assuming errors in data collection)

- Afterward, more than 250 and less than -21 DO and DC to missing

- Finally, less than 300 and more than 550 days of age, as well as more than 360 days of season length as missing.
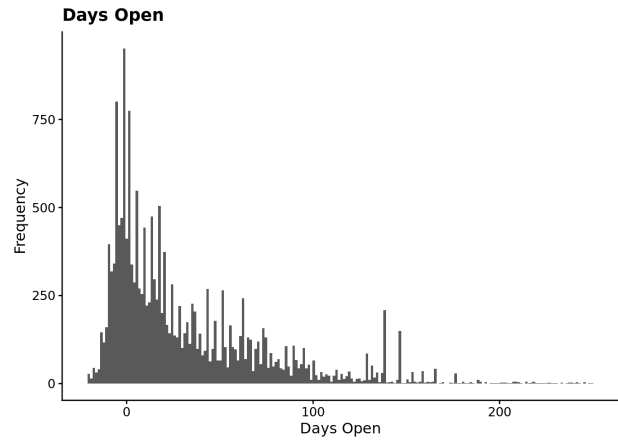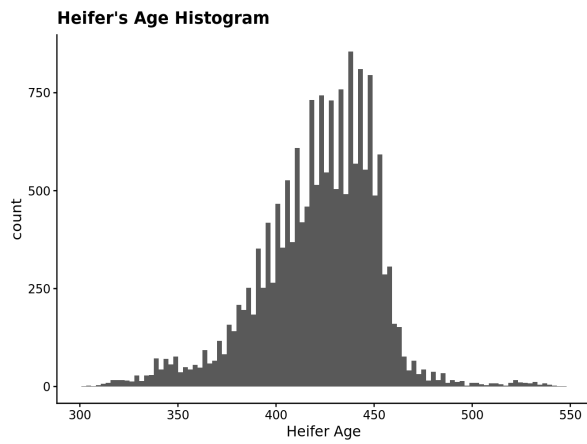
Figure 3. Distribution of (a) days to conception, (b) days open, (c) heifer's age, and (d) season length after quality control

Figure 4. Scatter plots relating (a) days open and (b) days to conception with heifer's age after quality control.

2.7.3. Days to conception and days open for GCTA and GEMMA Manhattan and QQ plots
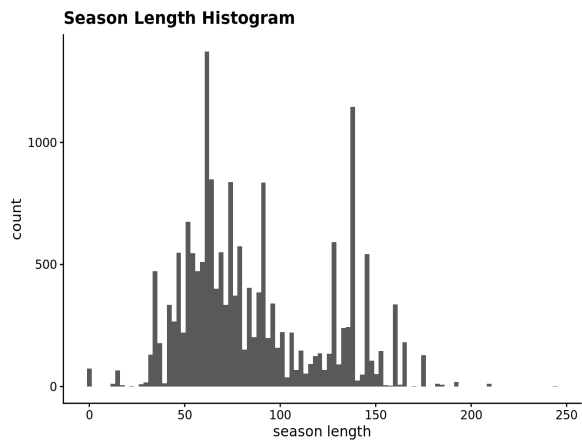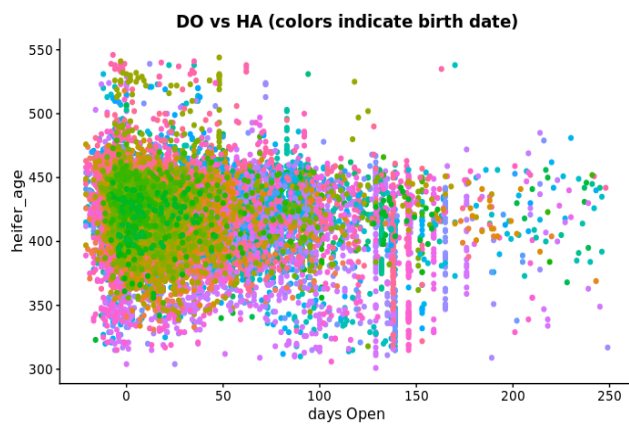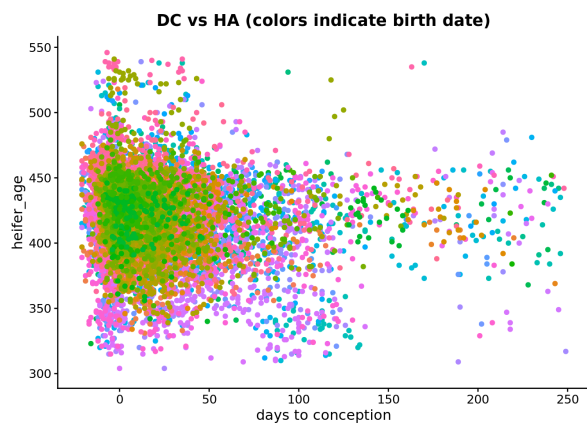

(a)

(b)

(c)

(d)
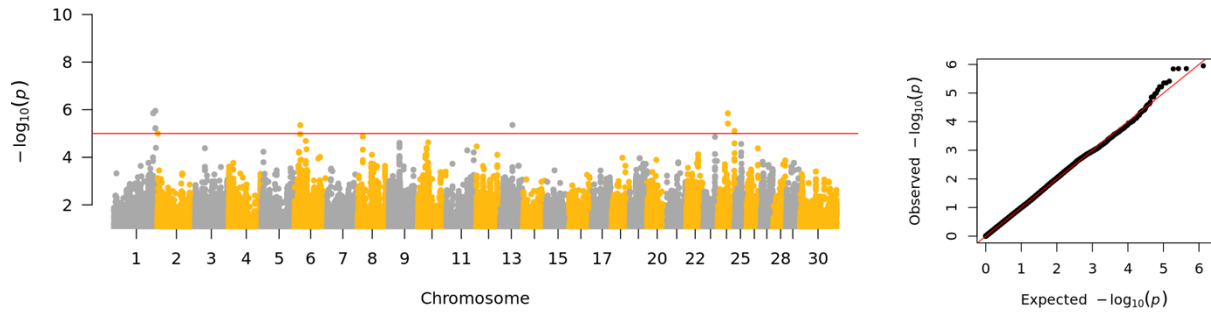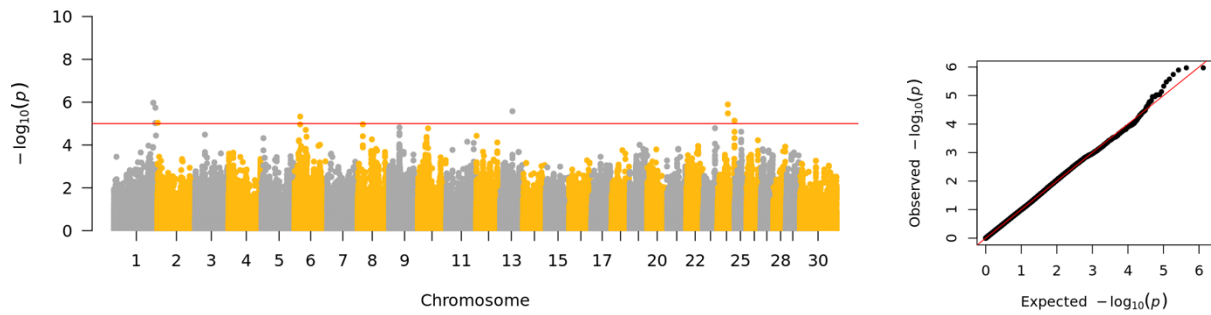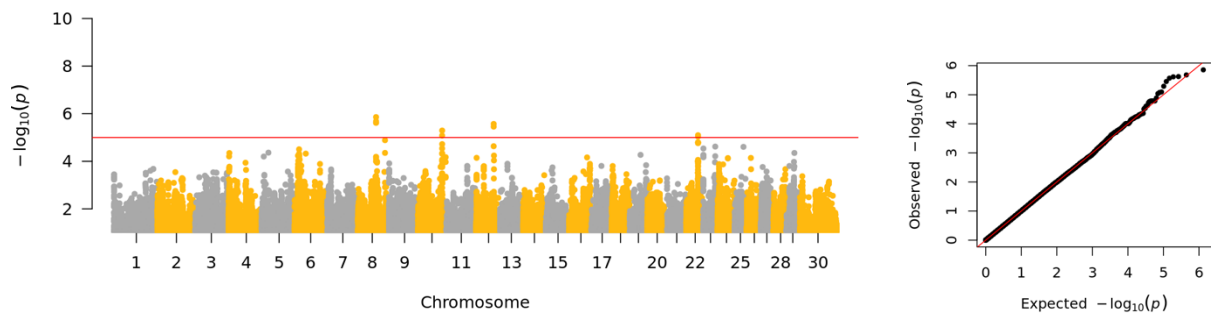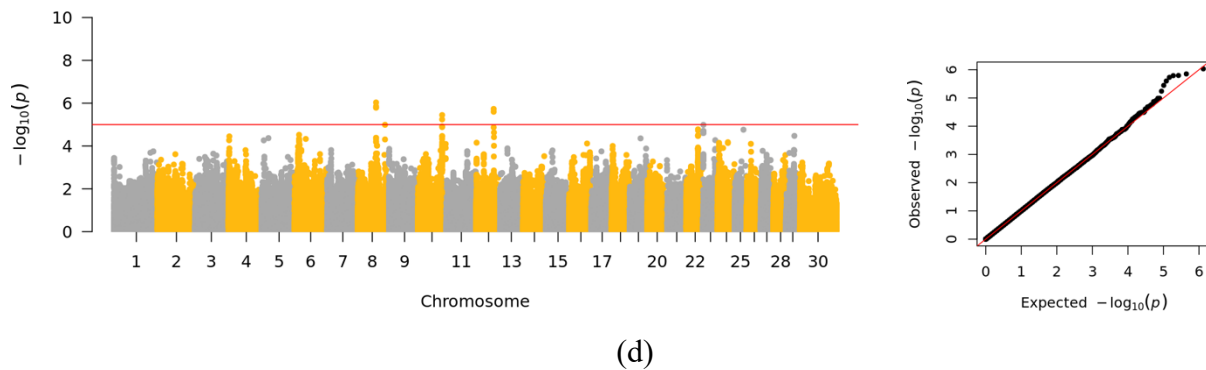
Figure 5. Manhattan and QQ plots for days to conception from GCTA (a) and GEMMA, and Manhattan and QQ plots for Days Open (DO) from GCTA (c) and GEMMA(d).

There was a weak association between a few SNPs and the both traits (DC - Days to conception and DO - Days Open). Also, there was an increase in power from GCTA to GEMMA, probably because of less noise introduced in GEMMA and most of the variation was due to genetics as estimated breeding values were used to generate pseudo-phenotypes.

2.7.4. Figures that show traits relationship among estimated breeding values, as well as the breeding values from linear models and Gibbs sampling.

(a)   (b)   (c)
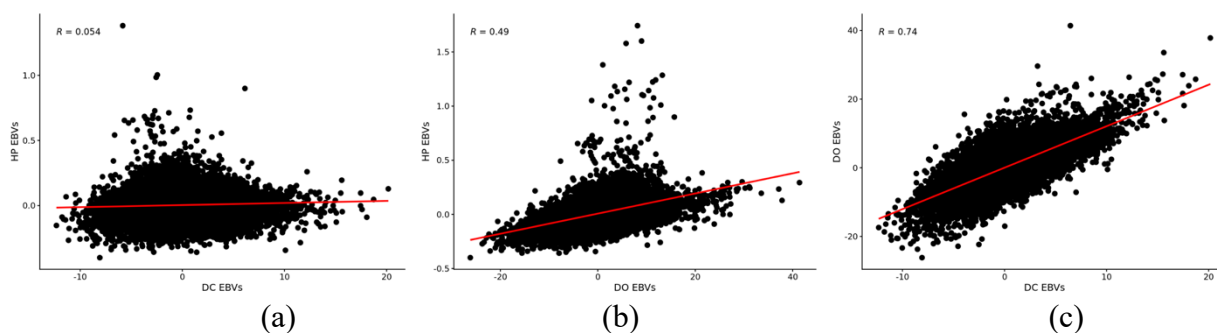
Figure 6. scatter plots that show the relationship between traits based on the estimated breeding values (a) heifer pregnancy and days to conception, (b) heifer pregnancy and days open, and (c) days open and days to conception.

The red lines show the regression line while R is the correlation between breeding values.
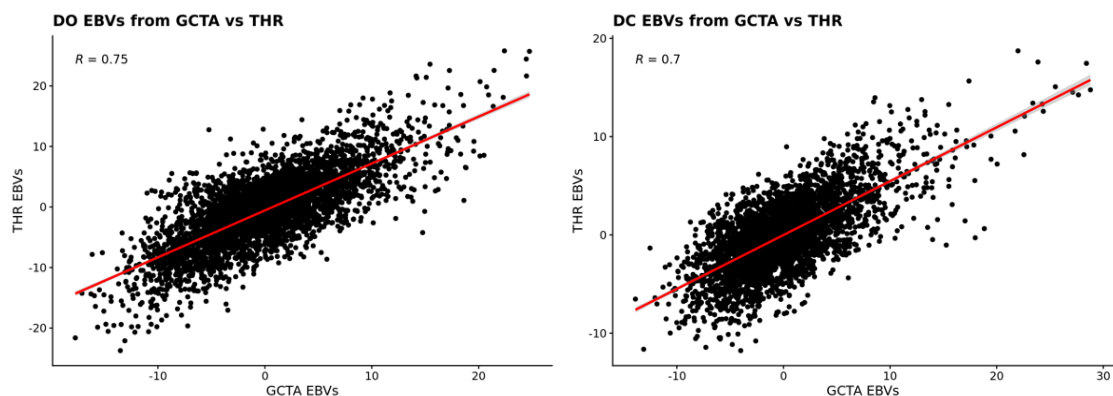
Figure 7. EBVs relationship from Gibbs sampling (THR) and linear model (GCTA) for days open (a) and days to conception (b). The red lines show the regression line while R is the correlation between breeding values.

2.7.5. Gene-Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) term enrichment from genes associated from linear and Gibbs sampling models

Table 1. Heifer pregnancy GO and KEGG terms from linear models (GCTA & GEMMA)

HP GO and KEGG terms from linear models			
Source	Term name	Intersection	P-value
GO:BP	regulation of metanephros size	PAX2	0.03
KEGG	Mucin type O-glycan biosynthesis	GALNT13	0.026
KEGG	Other types of O-glycan biosynthesis	GALNT13	0.026
KEGG	Taste transduction	SCN3A	0.031

Table 2. Days to conception GO and KEGG terms from linear models (GCTA & GEMMA)

DC GO and KEGG terms from linear models			
source	Term name	Intersection	P-value
GO:BP	Negative regulation of peptidase activity	SERPINB12, SERPINB13	0.0183
GO:BP	Negative regulation of endopeptidase activity	SERPINB12, SERPINB13	0.0183
GO:BP	Hexose phosphate transport	SLC37A1	0.0183
GO:BP	Glucose-6-phosphate transport	SLC37A1	0.0183

GO:BP	negative regulation of keratinocyte apoptotic process	SERPINB13	0.0183
GO:BP	Regulation of keratinocyte apoptotic process	SERPINB13	0.0183
GO:BP	Keratinocyte apoptotic process	SERPINB13	0.0183
GO:BP	negative regulation of proteolysis	SERPINB12, SERPINB13	0.0183
GO:BP	Negative regulation of hydrolase activity	SERPINB12, SERPINB13	0.0183
GO:BP	regulation of endopeptidase activity	SERPINB12, SERPINB13	0.0216
GO:BP	Phosphate ion transmembrane transport	SLC37A1	0.0224
GO:BP	Regulation of peptidase activity	SERPINB12, SERPINB13	0.0224
GO:BP	Regulation of postsynaptic neurotransmitter receptor activity	DLGAP1	0.0311
GO:BP	Phosphate ion transport	SLC37A1	0.0344
GO:BP	Negative regulation of catalytic activity	SERPINB12, SERPINB13	0.0344
GO:BP	Regulation of proteolysis	SERPINB12, SERPINB13	0.0386
GO:BP	Regulation of molecular function	DLGAP1, SERPINB12, SERPINB13	0.0407
GO:BP	Negative regulation of epithelial cell apoptotic process	SERPINB13	0.0407
GO:BP	Negative regulation of protein metabolic process	SERPINB12, SERPINB13	0.0476
GO:BP	Negative regulation of molecular function	SERPINB12, SERPINB13	0.0476
GO:BP	Regulation of neurotransmitter receptor activity	DLGAP1	0.0476
GO:BP	Negative regulation of cellular protein metabolic process	SERPINB12, SERPINB13	0.0476
GO:BP	Regulation of epithelial cell apoptotic process	SERPINB13	0.0476
GO:BP	Regulation of hydrolase activity	SERPINB12, SERPINB13	0.0476
KEGG	Glutamatergic synapse	DLGAP1	0.0264

Table 3. Days to conception KEGG terms from linear models (GCTA & GEMMA)

DO KEGG terms from linear models			
Source	Term name	Intersection	p-value
KEGG	Signaling pathways regulating pluripotency of stem cells	ESRRB	0.0341

Table 4. Heifer pregnancy GO and KEGG terms using de-regressed EBV (SNP1101)

source	Term name	Intersections	P-value
HP GO and KEEG terms using de-regressed EBV from threshold model			
GO:BP	Pyramidal neuron development	ZMIZ1	0.0285
GO:BP	Pyramidal neuron differentiation	ZMIZ1	0.0285
GO:BP	Pyramidal neuron migration to cerebral cortex	ZMIZ1	0.0285
GO:BP	Central nervous system neuron differentiation	GRID2, ZMIZ1	0.0285
GO:BP	Cerebellar granule cell differentiation	GRID2	0.0285
GO:BP	Regulation of postsynaptic specialization assembly	GRID2	0.0285
GO:BP	Regulation of postsynaptic density assembly	GRID2	0.0285
GO:BP	Cerebellar granular layer formation	GRID2	0.0285
GO:BP	Cerebellar granular layer morphogenesis	GRID2	0.0285
GO:BP	Radial glia-guided pyramidal neuron migration	ZMIZ1	0.0285
GO:BP	Cerebellar granular layer development	GRID2	0.0285
GO:BP	Positive regulation of long-term synaptic depression	GRID2	0.0285
GO:BP	Prepulse inhibition	GRID2	0.0285
GO:BP	Vitellogenesis	ZMIZ1	0.0285
GO:BP	Regulation of excitatory synapse assembly	GRID2	0.0285
GO:BP	Regulation of postsynaptic density organization	GRID2	0.0285

GO:BP	Regulation of long-term synaptic depression	GRID2	0.0285
GO:BP	Postsynaptic density assembly	GRID2	0.0285
GO:BP	Postsynaptic specialization assembly	GRID2	0.0299
GO:BP	Brain development	GRID2, ZMIZ1	0.0299
GO:BP	Cytoplasm organization	ZMIZ1	0.0299
GO:BP	Head development	GRID2, ZMIZ1	0.032
GO:BP	Long-term synaptic depression	GRID2	0.033
GO:BP	Telencephalon glial cell migration	ZMIZ1	0.033
GO:BP	Postsynaptic density organization	GRID2	0.033
GO:BP	Regulation of protein sumoylation	ZMIZ1	0.033
GO:BP	Ionotropic glutamate receptor signaling pathway	GRID2	0.033
GO:BP	Postsynaptic specialization organization	GRID2	0.033
GO:BP	Cell-cell adhesion	GRID2, ZMIZ1	0.033
GO:BP	Cerebral cortex radial glia-guided migration	ZMIZ1	0.033
GO:BP	Postsynapse assembly	GRID2	0.033
GO:BP	Cerebellar cortex formation	GRID2	0.033
GO:BP	Cell differentiation in hindbrain	GRID2	0.033
GO:BP	Excitatory synapse assembly	GRID2	0.033

GO:BP	Central nervous system development	GRID2, ZMIZ1	0.033
GO:BP	Ligand-gated ion channel signaling pathway	GRID2	0.033
GO:BP	Startle response	GRID2	0.033
GO:BP	Forebrain neuron development	ZMIZ1	0.033
GO:BP	neuron development	GRID2, ZMIZ1	0.0422
GO:BP	Negative regulation of synaptic transmission	GRID2	0.0422
GO:BP	Positive regulation of fibroblast proliferation	ZMIZ1	0.0422
GO:BP	Androgen receptor signaling pathway	ZMIZ1	0.0422
GO:BP	Hindbrain morphogenesis	GRID2	0.0422
GO:BP	Positive regulation of Notch signaling pathway	ZMIZ1	0.0422
GO:BP	Cerebellar cortex morphogenesis	GRID2	0.0422
GO:BP	Cerebral cortex radially oriented cell migration	ZMIZ1	0.0422
GO:BP	Forebrain neuron differentiation	ZMIZ1	0.0422
GO:BP	Glial cell migration	ZMIZ1	0.0422
GO:BP	Cerebellum morphogenesis	GRID2	0.0422
GO:BP	Heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules	GRID2	0.0422
GO:BP	Cerebral cortex cell migration	ZMIZ1	0.0428
GO:BP	Forebrain generation of neurons	ZMIZ1	0.0428

GO:BP	Glutamate receptor signaling pathway	GRID2	0.0439
GO:BP	Positive regulation of developmental process	GRID2, ZMIZ1	0.0439
GO:BP	Protein sumoylation	ZMIZ1	0.0439
GO:BP	Cerebellar cortex development	GRID2	0.0439
GO:BP	Positive regulation of synapse assembly	GRID2	0.0439
GO:BP	Regulation of multicellular organismal development	GRID2, ZMIZ1	0.0445
GO:BP	Positive regulation of multicellular organismal process	GRID2, ZMIZ1	0.046
GO:BP	Artery morphogenesis	ZMIZ1	0.046
GO:BP	Regulation of postsynapse organization	GRID2	0.046
GO:BP	Neuron differentiation	GRID2, ZMIZ1	0.046
GO:BP	Telencephalon cell migration	ZMIZ1	0.046
GO:BP	Forebrain cell migration	ZMIZ1	0.0472
GO:BP	Cell adhesion	GRID2, ZMIZ1	0.0472
GO:BP	Fibroblast proliferation	ZMIZ1	0.0472
GO:BP	Central nervous system neuron development	ZMIZ1	0.0472
GO:BP	Biological adhesion	GRID2, ZMIZ1	0.0472
GO:BP	Synaptic transmission, glutamatergic	GRID2	0.0472
GO:BP	Regulation of fibroblast proliferation	ZMIZ1	0.0472

GO:BP	Vasculogenesis	ZMIZ1	0.0481
GO:BP	Excitatory postsynaptic potential	GRID2	0.0491
GO:BP	Generation of neurons	GRID2, ZMIZ1	0.0491
KEGG	Glycosaminoglycan biosynthesis - heparan sulfate / heparin	HS3ST3A1	0.0227
KEGG	Long-term depression	GRID2	0.0281

Table 5. Days to conception GO and KEGG terms using de-regressed EBV (SNP1101)

DC GO & KEGG terms from de-regressed GWAS			
source	Term name	intersections	p-value
GO:BP	Activation of cysteine-type endopeptidase activity involved in apoptotic process by cytochrome c	DIABLO	0.0278
GO:BP	Poly-N-acetyllactosamine metabolic process	B3GNT4	0.0278
GO:BP	Poly-N-acetyllactosamine biosynthetic process	B3GNT4	0.0278
KEGG	Glycosphingolipid biosynthesis - lacto and neolacto series	B3GNT4	0.0209
KEGG	Apoptosis - multiple species	DIABLO	0.0209