Received: 19 July 2022 | Accepted: 31 January 2023

DOI: 10.1111/jbg.12763

Journal of Animal Breeding and Genetics WILEY

ORIGINAL ARTICLE

Genomic prediction of service sire effect on female reproductive performance in Holstein cattle: A comparison between different methods, validation population and marker densities

Rui Shi[1,2,3] | **Ziwei Chen**[1] | **Guosheng Su**[4] | **Hanpeng Luo**[1] | **Lin Liu**[5] | **Gang Guo**[6] | **Yachun Wang**[1]

[1]Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing, China

[2]Wageningen University & Research Animal Breeding and Genomics, Wageningen, the Netherlands

[3]Animal Production Systems Group, Wageningen University & Research, Wageningen, the Netherlands

[4]Center for Quantitative Genetics and Genomics, Aarhus University, Tjele, Denmark

[5]Beijing Dairy Cattle Center, Beijing, China

[6]Beijing Sunlon Livestock Development Co. Ltd, Beijing, China

Correspondence
Yachun Wang, Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China.
Email: wangyachun@cau.edu.cn

Funding information
Beijing Sanyuan Breeding Technology Ltd. Co. funded project (SYZYZ20190005); Earmarked fund for China Agriculture Research System (CARS-36); China Scholarship Council (No. 201913043); Hainan University; National Agricultural Genetic Improvement Program (2130135); Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62); National Agriculture Key Science & Technology Project (NK20221201)

Abstract

Reproductive traits of dairy cattle are bound to the actual efficiency of farm operation, which therefore show great economic importance. Among them, some traits were deemed to be simultaneously affected by service sire and mating cow. Service sires are proved to play an important role in reproduction process of cows. However, limited study explored the genetic effect of service sire (GESS), let alone the genomic prediction of this effect. In the present study, 2244 genotyped bulls together with phenotypic records were used to predict the GESS on conception rate, 56-day non-return rate, calving ease, stillbirth and gestation length. The feasibilities of multi-step genomic best linear unbiased predictor (msGBLUP) and single-step genomic best linear unbiased predictor (ssGBLUP) were investigated under different scenarios, that is, different marker densities and validation population. The predictive accuracies and unbiasedness for GESS ranged from 0.159 to 0.647 and from 0.202 to 2.018, respectively, when validated on young bulls, while the accuracies and unbiasedness ranged from 0.409 to 0.802 and 0.333 to 1.146 when validated on random split data sets. It is feasible to predict GESS on reproductive traits by using a linear mixed model and genomic data, and high-density marker panel had limited contribution to the prediction. This research investigated the potential factors that influence the genomic prediction of GESS on reproductive traits and indicated the possibility of genomic selection on GESS, both in ideal and practical circumstances.

KEYWORDS
dairy cattle, genomic prediction, reproduction, service sire

1 | INTRODUCTION

Reproductive performance, which is exceedingly economic important, has been widely emphasized in dairy industry for recent decades. Several studies have investigated genetic effects of cows/dams (GED), which are constituted in many breeding schemes of dairy cattle (Berry et al., 2014; Cole & VanRaden, 2018), while the genetic effect of service sire (GESS) is not yet completely explored, let alone the genomic analysis. Service sires are proved to influence the fertilization process, early stage of embryo development and consequently the herd reproduction (Jaton et al., 2017; Kropp et al., 2014). For insemination traits, GESS represents the specific ability of service sire to impregnate mating cow; while for calving traits, GESS measures the development of calves from a particular service sire during pregnant process. In contemporary dairy production system, artificial insemination was applied to the majority of cows, generating enormous amount of mating records for service sires. Taylor et al. (2018) indicated that these records cannot be well adopted in the genetic evaluation models of the US dairy industry due to little variation in male spermatozoa after bull studs' use. However, genetic parameters of GESS were successfully obtained by using the data of Chinese dairy farms, which provides fundamental resources for genomic selection (GS) scheme of GESS in China (Chen et al., 2021).

In the last decade, GS was widely applied to both bulls and cows in dairy cattle breeding, and several methods were generated, in which multi-step genomic best linear unbiased predictor (msGBLUP) and single-step genomic best linear unbiased predictor (ssGBLUP) are frequently used. A routine msGBLUP procedure consists of (1) deriving estimated breeding values (EBV) and genetic parameters based solely on phenotypic and pedigree information, (2) calculation of pseudo phenotypes such as de-regressed proof (DRP) using results of previous step and (3) obtaining direct genomic values (DGV) for genotyped animals and blending them with EBV (Doormaal et al., 2009). To avoid the potential information loss and double counting during msGBLUP steps, Misztal et al. (2009), Aguilar et al. (2010) and Christensen and Lund (2010) developed a methodology that simultaneously incorporated pedigree, phonotypic and genomic information into a single step to obtain genomically enhanced breeding value (GEBV) of individual animals. For traits with lower heritabilities, the unbiasedness and accuracy of prediction can be improved by using ssGBLUP method (Guarini et al., 2018).

In commercial livestock populations, which are intensively selected and have higher level of linkage disequilibrium (LD), the important quantitative trait loci (QTL) can be accurately tracked with fewer markers. In this case, the benefits of using high-density markers in GS are limited (VanRaden, 2020). However, it may be a different case when it comes to Chinese Holstein population. The bulls/semen were mostly imported from a wide range of countries, which means the service sires may be from different strains with various genetic architecture. In addition, reproductive traits are not yet included in the selection scheme of dairy cows in China. Therefore, high-density markers are likely to increase genomic prediction accuracy because of potential stronger LD with causal genes across the whole genome. Meanwhile, previous study found limited SNPs and overlapping genes were observed between GED and GESS (Chen et al., 2021), which indicates the causative mutations for GESS may be different from GED across the whole genome. Thus, a comparison on prediction accuracy of different marker densities will help to find the optimum predicting scenario of GESS. Previous study has proved that the relatedness between reference and validation individuals would impact the accuracy of prediction, which determines the outcomes of different validation strategies (Daetwyler et al., 2013).

In China, systematic genomic evaluations are currently solely applied to production traits, and there is still a lack of studies utilizing the GS procedure for reproductive traits in Chinese Holstein populations. Additionally, to our knowledge, no study so far has investigated the feasibility of GS for GESS. In this article, we compared the power of genomic predictions for GESS of various reproductive traits in Chinese Holstein cattle under different scenarios, using multi-step and single-step GBLUP methods, and discussed the prediction power of GESS.

2 | MATERIALS AND METHODS

2.1 | Phenotypic data

Reproductive events in 39 farms (Sunlon Livestock Development Co. Ltd., Beijing, China) from 1987 to 2020 in Beijing, China, were extracted from the AfiFarm software (AfiFarm, www.afimilk.com.cn) and used in the current study. Five reproductive traits were derived, which were conception rate (CR), 56-day non-return rate (NRR56), calving ease (CE), stillbirth (SB) and gestation

length (GL). A descriptive summary for each trait is shown in Table 1. The model to estimate variance components and calculation of heritabilities for GESS can be found in previous publication (Chen et al., 2021).

The phenotypic data in this analysis comprised 1952 service bulls and 163,818 cows. The final pedigree consisted of 503,118 cows and 151,273 bulls born between 1957 and 2020, which was provided and traced (at least three generations back) by the Beijing Dairy Cattle Center (BDCC, Beijing, China). The pedigree included cows that have phenotypic records, genotyped bulls and non-genotyped bulls used in ssGBLUP method.

2.2 | Genotypic data

A total of 2244 bulls were genotyped with the Illumina BovineSNP50 BeadChip (Illumina, Inc., San Diego, CA, United States). To compare the predictions of different panel density, these genotypes were imputed to 150 K using the BEAGLE v5.1 software (Browning et al., 2018) with a reference population consisting of 3119 cows and 81 bulls (Illumina 150 K Bovine Beadchip). The imputation accuracy was tested by masking and validating random markers (Chen et al., 2021). Only SNPs with imputation accuracy greater than 90% were kept for further analyses. Thereafter, both low-density (50 K) SNP panel and high-density (150 K) SNP panel were filtered by the thresholds below: minor allele frequency higher than 0.01 and fulfil Hardy–Weinberg equilibrium ($P > 10–7$). After editing, 44,825 markers in low-density panel and 110,304 markers in high-density panel located in the autosomes and pseudo-autosomal regions of the X chromosome were retained for further analysis.

2.3 | Calculation of pseudo phenotypes

For some prediction models, DRPs of service sire were used as dependent variables. For validation of genomic prediction, DRPs were used for benchmarking. The DRPs were estimated by following formulas (de Oliveira et al., 2018; VanRaden et al., 2009).

$$\mathrm{DRP} = \mathrm{PA} + \frac{\mathrm{EBV} - \mathrm{PA}}{R}$$

where EBV and PA are the estimated breeding value and parental average of GESS derived from the pedigree-based linear mixed model (see next section), respectively, and R is the de-regression factor computed as follows:

$$R = \frac{\mathrm{ERC}_{\mathrm{Ind-PA}}}{\mathrm{ERC}_{\mathrm{Ind-PA}} + \mathrm{ERC}_{\mathrm{PA}} + 1}$$

where $\mathrm{ERC}_{\mathrm{Ind-PA}}$ is the effective record contributions without PA information, $\mathrm{ERC}_{\mathrm{PA}}$ is the effective record contributions of PA. These parameters were computed as follows (VanRaden & Wiggans, 1991):

$$\mathrm{ERC}_{\mathrm{Ind-PA}} = \lambda \frac{\mathrm{REL}_{\mathrm{EBV}}}{1 - \mathrm{REL}_{\mathrm{EBV}}} - \mathrm{ERC}_{\mathrm{PA}}$$

$$\mathrm{ERC}_{\mathrm{PA}} = \lambda \frac{\mathrm{REL}_{\mathrm{PA}}}{1 - \mathrm{REL}_{\mathrm{PA}}}$$

where $\mathrm{REL}_{\mathrm{EBV}}$ and $\mathrm{REL}_{\mathrm{PA}}$ were the reliabilities of EBV and PA, respectively; λ was calculated by: $1 - h^2 / h^2$, and h^2 is the heritability of GESS in corresponding trait. Consequently, individuals with reliabilities of EBVs over 0.1 were retained for genomic prediction.

2.4 | Statistical models and prediction approaches

The pedigree-based linear mixed model was used to estimate variance components and breeding values using all original phenotypic data. The EBVs and their reliability from this univariate model were used to derive DRPs for further analysis. The model is:

TABLE 1 Descriptive statistics and heritabilities for target traits in this study[a]

Trait	Mean	SD	Minimum	Maximum	N	h^2_{ss}
CR	0.43	0.49	0	1	837,655	0.020
NRR56	0.50	0.50	0	1	857,821	0.024
CE	1.06	0.27	1	3	259,042	0.004
SB	1.07	0.25	0	1	273,367	0.078
GL	278.36	6.18	260	302	258,611	0.102

[a]CR, Conception rate (0 or 1); NRR56, 56 days' non-return rate (0 or 1); CE, Calving ease (1, 2 or 3); SB, stillbirth (0 or 1); GL, Gestation length (day); SD, standard deviation; N, number of records; h^2_{ss}, heritabilities of GESS for each trait.

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_1 u_f} + \boldsymbol{Z_2 u_m} + \boldsymbol{W_1 p_m} + \boldsymbol{W_2 p_f} + \boldsymbol{Z_3 h} + \boldsymbol{e} \quad (1)$$

where $\boldsymbol{y}$ represents the vector of insemination, pregnancy and calving events for 5 reproductive traits; $\boldsymbol{b}$ is the vector of fixed effects (i.e., AI technician, parity, semen type and number of insemination sequences for CR and NRR56 and calf sex, parity and group of calf size for CE, SB, and GL); $\boldsymbol{u_f}$ is a vector of additive genetic effects for GED; $\boldsymbol{u_m}$ is a vector of additive genetic effects for GESS; $\boldsymbol{p_m}$ and $\boldsymbol{p_f}$ are the vectors of the random permanent environment effects caused by the service sire and dam, respectively; $\boldsymbol{h}$ is the vector of the random herd-year-month effects; $\boldsymbol{e}$ is the vector of the random residual effects; $\boldsymbol{X}$, $\boldsymbol{Z_1}$, $\boldsymbol{Z_2}$, $\boldsymbol{W_1}$, $\boldsymbol{W_2}$ and $\boldsymbol{Z_3}$ represent the corresponding incidence matrices. We assumed that:

$$\begin{pmatrix} \boldsymbol{u_m} \\ \boldsymbol{u_f} \\ \boldsymbol{p_m} \\ \boldsymbol{p_f} \\ \boldsymbol{h} \\ \boldsymbol{e} \end{pmatrix} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{V})$$

with:

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{A} \otimes \begin{pmatrix} \sigma_m^2 & \sigma_{m,f} \\ \sigma_{m,f} & \sigma_f^2 \end{pmatrix} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I_1}\,\sigma_{pm}^2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I_2}\,\sigma_{pf}^2 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I_3}\,\sigma_h^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I_4}\,\sigma_e^2 \end{pmatrix}$$

where σ_m^2 and σ_f^2 are the additive genetic variances of service sire and dam, respectively; $\sigma_{m,f}$ is the genetic covariance of service sire and dam; σ_{pm}^2 and σ_{pf}^2 are the permanent environmental variances of service sire and dam, respectively; σ_h^2 is the herd-year-month variance; σ_e^2 is the residual variance; $\otimes$ is the Kronecker product function; $\boldsymbol{A}$ is the pedigree-based additive genetic relationship matrix among the animals; and $\boldsymbol{I_1}$ to $\boldsymbol{I_4}$ are identity matrices. Other detailed information and model assumptions can be found in previous publication (Chen et al., 2021).

Three genomic prediction approaches, which are ssGBLUP, msGBLUP$_G$ and msGBLUP$_H$, were used in the present study. For ssGBLUP approach, the expression of model is the same as **(1)**, except for the relationship matrix. The relationship matrix is a genotype-pedigree combined relationship matrix ($\boldsymbol{H}$, presented below), instead of $\boldsymbol{A}$ matrix. In this study, the response variable $\boldsymbol{y}$ in ssGBLUP was a vector of insemination or calving events for cows.

For msGBLUP, the model is:

$$\boldsymbol{y} = \boldsymbol{1}\mu + \boldsymbol{Wa} + \boldsymbol{e} \quad (2)$$

where $\boldsymbol{y}$ is the vector of DRPs of genotyped bulls; μ is the overall mean; $\boldsymbol{a}$ is the vector of DGVs for GESS captured by SNP markers; $\boldsymbol{e}$ is the vector of random residuals; $\boldsymbol{1}$ is a vector of all ones; and $\boldsymbol{W}$ is the corresponding incidence matrices. The residual vector e was normally distributed: $\boldsymbol{e} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$. For msGBLUP$_G$, we assume that $\boldsymbol{a} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{G}\sigma_a^2)$, where the $\boldsymbol{G}$ matrix only included genotyped bulls. For msGBLUP$_H$, $\boldsymbol{a} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{H}\sigma_a^2)$ where the $\boldsymbol{H}$ matrix further included also ungenotyped individuals and thus the model provided DGVs of ungenotyped individuals too. The σ_a^2 denotes additive genetic variance, which is GESS in the current study. Matrix $\boldsymbol{H}$ combined pedigree and genomic information by the following formula (Aguilar et al., 2010; Christensen & Lund, 2010; Misztal et al., 2009):

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{G_\omega} & \boldsymbol{G_\omega A_{22}^{-1} A_{21}} \\ \boldsymbol{A_{12} A_{22}^{-1} G_\omega} & \boldsymbol{A_{12} A_{22}^{-1} G_\omega A_{22}^{-1} A_{21} + A_{11} - A_{12} A_{22}^{-1} A_{21}} \end{bmatrix}$$

where $\boldsymbol{A_{22}}$ is the sub-matrix of the $\boldsymbol{A}$ for genotyped animals, built using all individuals in the pedigree; $\boldsymbol{A_{11}}$ is the sub-matrix of $\boldsymbol{A}$ matrix for non-genotyped animals, $\boldsymbol{A_{12}}$ and $\boldsymbol{A_{21}}$ are the sub-matrices of $\boldsymbol{A}$ for relationships between genotyped animals and non-genotyped animals; and $\boldsymbol{G_\omega} = (\boldsymbol{1} - \boldsymbol{\omega})\boldsymbol{G} + \boldsymbol{\omega A_{22}}$, where ω is a weight. In this study, $\omega = 0.20$ was used, which means that 20% of total genetic variance is due to the polygenic effects that were not described by the SNP markers. This value was chosen by a prior test within the same population, in which $\omega = 0.20$ generally generated best predictive abilities (results not shown). The $\boldsymbol{G}$ matrix, which represents the genomic relationship between individuals, can be derived directly from SNP data (VanRaden, 2008). The inverse of $\boldsymbol{H}$ is described as:

$$\boldsymbol{H^{-1}} = \boldsymbol{A^{-1}} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G_\omega^{-1} - A_{22}^{-1}} \end{bmatrix}$$

The AI-REML algorithm implemented in the DMUAI module of the DMU software (Madsen et al., 2010) was used to estimate variance components. These estimates were then set as true variance components in the subsequent prediction implemented by the DMU4 module. The $\boldsymbol{G}$ and $\boldsymbol{G_\omega^{-1}}$ were built using the G-matrix programs (Su & Madsen, 2011), while $\boldsymbol{H^{-1}}$ was constructed directly by DMU software with pedigree and $\boldsymbol{G_\omega^{-1}}$.

2.5 | Validation

Genomic predictions were validated using two approaches. The first approach was validation on young bulls, in which younger bulls (January 1, 2015 onwards for CR, NRR56, SB and GL; January 1, 2014 onwards for CE) were used as validation population and the other as reference population. The different birth year threshold was chosen to make a ratio close to 4 to 1 between the animals in reference population and validation population. Two bulls lost birth date, but they were recorded as relatives of other animals in the pedigree, and thus, they were both classified as young bull validation. The detailed description of reference and validation populations for each trait are listed in Table 2. The birth year thresholds and animals in validation population of ssGBLUP were the same as those of msGBLUP. As for reference population in ssGBLUP, old non-genotyped bulls with phenotypic records were included to be consistent with actual breeding practice (Table 3).

The second approach was a fivefold random cross-validation in which all genotyped animals were randomly divided into 5 subsets; each subset was used as validation population for once. The fivefold random cross-validation was implemented with 10 replicates for each trait. For ssGBLUP, non-genotyped bulls (Table 3) were always categorized as reference population. Table 4 summarized the random subsets in this validation approach.

The accuracy of predictions was evaluated as the correlation between DRP (from the full dataset) and DGV (from msGBLUP approaches) or GEBV (from ssGBLUP) of the individuals in validation population. Regression coefficients (b) of validation bulls' DRP on DGV (or GEBV) were used to measure the unbiasedness of predictions (Su, Brøndum, et al., 2012), which reached the most optimal status when they were close or equal to 1. For the fivefold random cross-validation, the accuracy and regression were calculated for each replicate. Thus, 50 accuracies and regression coefficients were obtained for each trait. The average and standard deviation of these 50 values for each trait were then calculated and reported.

3 | RESULTS AND DISCUSSION

This study investigated genomic prediction for GESS of reproductive traits in Chinese Holstein population using $msGBLUP_G$, $msGBLUP_H$ and ssGBLUP methods with different densities of markers.

For most reproductive traits, random cross-validation generated more accurate and unbiased prediction results than validation on young bulls for GESS (Tables 6 and 7). Overall, predictions for GL were the most accurate (with accuracy from 0.452 to 0.802) and unbiased (with b from 0.662 to 1.264), regardless of validation population, prediction methods and panel density. For the other traits, accuracies were below 0.500 when validated based on young bulls. For NRR56 and CE, the accuracies of random validation were much higher (from 0.515 to 0.660) than validation of young bulls. Most of these results are in general consistent with those for lowly heritable traits (Guarini et al., 2018; Rezende et al., 2019).

For other traits, the optimal prediction varied in different scenarios. Besides, NRR56 was not properly estimated when validated based on young bulls; thus, the corresponding accuracies and unbiasedness are not shown in the tables. This may be due to small number of validation animals and low accuracy of DRP. For fivefold random cross-validation, reasonable results were generated for all traits. Supplementary file S1 listed all the detailed statistical information of cross-validation.

TABLE 2 Structure of reference and validation data sets divided by birth year in msGBLUP.

	Validation population			Reference population					
Trait[a]	Birth year	N[c]	REL[d]	Birth year	N	REL	Ratio[b]	N	REL
CR	2015–2018	396 (2)	0.200	1993–2014	1668	0.302	4.21:1	2064	0.282
NRR56	2015–2018	395 (2)	0.197	1993–2014	1697	0.304	4.30:1	2092	0.284
CE	2014–2017	403 (2)	0.163	1994–2013	966	0.229	2.40:1	1369	0.209
SB	2015–2018	427 (2)	0.214	1984–2014	1745	0.343	4.09:1	2172	0.318
GL	2015–2018	410 (2)	0.217	1984–2014	1807	0.351	4.41:1	2217	0.326

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.

[b]Ratio between the number of animals in reference and validation population.

[c]Number of individuals.

[d]Mean reliability of EBVs for corresponding data set.

TABLE 3 Structure of reference data sets divided by birth year in ssGBLUP.

	Reference population							All[b]		
		N[c]			REL[d]					
Trait[a]	Birth year	P only[f]	G only[g]	P + G[h]	P only	G only	P + G	Ratio[e]	N	REL
CR	1965–2014	740	1162	506	0.279	0.234	0.464	6.08:1	2804	0.283
NRR56	1965–2014	748	1193	504	0.291	0.243	0.498	6.19:1	2840	0.294
CE	1965–2013	455	595	371	0.108	0.085	0.183	3.53:1	1824	0.128
SB	1965–2014	754	1243	502	0.445	0.406	0.610	5.85:1	2926	0.423
GL	1965–2014	743	1303	504	0.475	0.444	0.599	6.22:1	2960	0.446

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.
[b]Animals in reference population (this table) and validation population (Table 2).
[c]Number of individuals.
[d]Mean reliability of EBVs of corresponding data sets.
[e]Ratio between the animals in reference population and validation population (Table 2).
[f]Bulls that only have phenotypes.
[g]Bulls that only have genotypes.
[h]Bulls with both phenotype and genotype.

TABLE 4 Structure of randomly divided data sets.

	Genotyped animals					Non-genotyped animals[b]
Trait[a]	Group 1	Group 2	Group 3	Group 4	Group 5	
CR	412	412	412	412	416	740
NRR56	418	418	418	418	420	748
CE	273	273	273	273	277	455
SB	434	434	434	434	436	754
GL	443	443	443	443	445	743

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.
[b]Non-genotyped animals were always categorized as reference population in ssGBLUP.

TABLE 5 Accuracies and unbiasedness of estimates of validation population in young bull validation and random cross-validation predicted by msGBLUP$_G$ using 150 and 50 K panels.

	150 K				50 K			
	Young bull validation		Random cross-validation		Young bull validation		Random cross-validation	
Trait[a]	Accuracy	b[b]	Accuracy	b	Accuracy	b	Accuracy	b
CR	0.212	0.302	0.411 ± 0.058	0.946 ± 0.130	0.159	0.202	0.409 ± 0.069	0.837 ± 0.136
NRR56	/	/	0.551 ± 0.034	0.939 ± 0.086	/	/	0.535 ± 0.046	0.903 ± 0.092
CE	0.263	1.062	0.604 ± 0.074	1.020 ± 0.126	0.282	1.210	0.609 ± 0.076	0.993 ± 0.131
SB	0.415	1.089	0.426 ± 0.056	0.994 ± 0.128	0.414	0.918	0.438 ± 0.048	0.885 ± 0.116
GL	0.574	1.197	0.738 ± 0.027	1.007 ± 0.049	0.568	1.162	0.739 ± 0.027	0.998 ± 0.042

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.
[b]b: unbiasedness.

3.1 | Predictive ability validated with different types of validation populations

In this study, younger bull validation and random cross-validation sets were both used to assess predictive abilities for each trait. In general, random cross-validation generated higher prediction accuracies for each trait when msGBLUP$_G$ was implemented, compared to young bull validation (Table 5). Meanwhile, the unbiasedness generated by random cross-validation strategy was closer to 1 compared to those in young bull validation, regardless of panel density. Similar pattern was observed in results

of msGBLUP$_H$ (Table 6) and ssGBLUP (Table 7). The results in present study are consistent with the findings of Bolormaa et al. (2017) and Campos et al. (2018), who reported higher accuracies and better unbiasedness using random-split strategy in msGBLUPG method.

It is widely agreed that validation on young bulls is more consistent with real-life selection scenario, where the selection candidates do not have phenotypic information from sibs and offspring. However, given limited size of genotyped animals in this study, the accuracies of young bull validation may be less reliable, because of large uncertainty due to small size of validation population. In addition, validation on bulls based on correlation between GEBV and corrected phenotypes might underestimate accuracy of genomic prediction, if the bulls were selected for the traits of interest or for the traits correlated with the trait of interest. This is because correlation coefficient calculated from a sample of directionally selected data is much lower than that from a random sample.

Random cross-validation strategy can generate stable predictive results by greatly reducing the sampling error. Higher accuracies from this strategy are expected due to the facts that the validation cows were a random sample, the combined validation data were large (all animals were used for validation), and there was a closer relationship between validation and reference animals compared with validation on young bulls. However, in the random cross-validation, the relationship between validation and reference animals could be closer than the relationship between selection candidates and reference animals in practical breeding program. Thus, the result provided a potential upper limit of true predictive ability. The results of the current study indicated that the real predictive ability of GESS for each reproductive trait would be relatively higher than the results of young bull validation but lower than those of random cross-validation.

3.2 | Predictive ability with different marker density

For each animal in this study, both imputed high-density marker panel (150 K) and low-density marker panel (50 K) were used to predict genomic breeding values. In general, the performance metrics fluctuated across different panels, especially in young bull validation. The statistics of random cross-validation were the average values of multiple subsets, which explained the less variation between different panels. When the validation population was predicted by msGBLUP$_G$ (Table 5), higher accuracies and more unbiased results were always obtained by a high-density panel than those by a low-density panel in insemination traits (CR and NRR56). The greatest improvement of accuracy that caused by the increased marker density was observed in CR (from 0.159 to 0.212). However, in calving traits (CE, SB and GL), predictions were more accurate when a low-density panel was used, except for the comparable results of SB and GL in young bull validation (150 K vs. 50 K: 0.415 vs. 0.414 in SB; and 0.574 vs. 0.568 in GL). In young bull validation or random cross-validation, the b values, to a certain extent, were closer to 1 for all the traits (especially for CR) when 150 K panel was used. For msGBLUP$_H$ (Table 6), only CR and GL in young bull validation, together with NRR56 in random cross-validation benefitted from higher marker density (from 0.292 to 0.321 for CR, from 0.636 to 0.647 for GL and from 0.636 to 0.651 for NRR56). The unbiasedness of msGBLUP$_H$ was relatively comparable for each trait across the panels. Compared with other methods, the variation of accuracies caused by marker density was smaller in ssGBLUP method (Table 7). Lower density contributed the largest improvement to the accuracy of SB in young bull validation (from 0.299 to 0.314), while causing the largest decline in CE of young bull validation (from 0.428 to 0.417). By using this method, b values generated by 50 K panel outperformed those generated by 150 K in most cases.

A number of markers to include in genomic predictions have dramatically increased since GS was initially implemented, while animal breeders usually opted for low-density makers to reduce cost. Several researches have explored the benefits of utilizing denser marker panels in GBLUP method but found limited evidence supporting the practical use of high-density panel (Erbe et al., 2012; Ertl et al., 2014; Su, Brøndum, et al., 2012; VanRaden et al., 2017).

In this study, denser marker panel do improve the prediction power of CR when young bull validation was used (no matter which method was implemented). However, limited improvement or decline was observed for other traits, which were in line with the findings of previous research. In the present study, the average pair-wise LD was 0.122 and 0.183 for 50 and 150 K marker panel, respectively, calculated from marker data of the genotyped animals. These values were lower than that (0.209 for 54 K panel in Holstein population) reported by Su, Brøndum, et al. (2012); Su, Madsen, et al. (2012), which indicated that high-density markers were expected to capture QTL effect better and lead to better predictions in this population (VanRaden, 2020).

There are some potential reasons that could explain the small or negative gain from denser markers. Firstly, the markers of 50 K panel may explain almost the whole genetic variation of GESS in the population. Therefore, using 150 K panel may add little information and also extra noises in prediction models and thus lead to lower prediction ability. Secondly, the 150 K panel in this study

TABLE 6 Accuracies and unbiasedness of estimates of validation population in young bull validation and cross-validation predicted by msGBLUP$_H$ using 150 and 50 K panels.

	150 K				50 K			
	Young bull validation		Random cross-validation		Young bull validation		Random cross-validation	
Trait[a]	Accuracy	b[b]	Accuracy	b	Accuracy	b	Accuracy	b
CR	0.321	0.463	0.502 ± 0.064	1.146 ± 0.140	0.292	0.421	0.503 ± 0.072	1.131 ± 0.154
NRR56	/	/	0.651 ± 0.033	1.101 ± 0.081	/	/	0.636 ± 0.043	1.084 ± 0.091
CE	0.297	1.279	0.656 ± 0.077	1.117 ± 0.116	0.301	1.445	0.660 ± 0.075	1.117 ± 0.134
SB	0.429	1.079	0.465 ± 0.059	1.112 ± 0.134	0.439	1.074	0.487 ± 0.045	1.127 ± 0.122
GL	0.647	1.264	0.802 ± 0.021	1.078 ± 0.043	0.636	1.239	0.802 ± 0.022	1.080 ± 0.036

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.

[b]b: unbiasedness.

TABLE 7 Accuracies and unbiasedness of estimates of validation population in young bull validation and cross-validation predicted by ssGBLUP using 150 and 50 K panels.

	150 K				50 K			
	Young bull validation		Random cross-validation		Young bull validation		Random cross-validation	
Trait[a]	Accuracy	b[b]	Accuracy	Unbiasedness	Accuracy	b	Accuracy	b
CR	0.445	0.376	0.408 ± 0.056	0.635 ± 0.081	0.434	0.380	0.410 ± 0.057	0.640 ± 0.086
NRR56	/	/	0.515 ± 0.051	0.652 ± 0.073	/	/	0.518 ± 0.050	0.655 ± 0.074
CE	0.428	1.809	0.572 ± 0.080	0.752 ± 0.158	0.417	2.018	0.575 ± 0.081	0.763 ± 0.156
SB	0.299	0.324	0.285 ± 0.057	0.333 ± 0.064	0.314	0.337	0.284 ± 0.058	0.335 ± 0.067
GL	0.452	0.815	0.662 ± 0.026	0.796 ± 0.050	0.458	0.819	0.662 ± 0.026	0.797 ± 0.052

[a]CE, Calving ease; CR, Conception rate; GL, Gestation length; NRR56, 56 days' non-return rate; SB, Stillbirth.

[b]b: unbiasedness.

was imputed from the 50 K panel, and the imputation error may contribute to the limited gain of high-density markers (Ertl et al., 2014). In this study, most of the reference animals for imputation were cows, but a large proportion of them are not close relatives of imputed bulls, which may also influence the imputation accuracy (Xu et al., 2021) and compromise the potential gain of 150 K panel.

3.3 | Predictive abilities with different approaches

In msGBLUP$_G$ method, a highest accuracy of 0.738 and an unbiasedness of 1.007 were obtained for GL. In addition, the prediction results seemed to be less biased in calving traits (CE, SB and GL) than in insemination traits (CR and NRR56). For instance, the largest deviation of the unbiasedness of calving traits from 1 was 0.210, while that of insemination traits was 0.798. In msGBLUP$_H$ method, almost all the predictions of validation population were carried out with higher accuracy and comparable bias compared to that in msGBLUP$_G$ method (Table 6). These two approaches adopted the same reference sets and statistical model; thus, the mere difference between them was their genetic matrices. Table 7 showed corresponding statistics of ssGBLUP method. Compared with other methods, the results of ssGBLUP in the current study were quite different. For example, in young bull validation, the accuracies ranged from 0.159 to 0.739 in msGBLUP$_G$, from 0.292 to 0.802 in msGBLUP$_H$ while those in ssGBLUP were less varied, ranging from 0.284 to 0.662. Almost all the traits obtained lower accuracies in random cross-validation predicted by ssGBLUP method, compared to msGBLUP$_G$ and msGBLUP$_H$. Meanwhile, the results seemed to be less biased when it came to msGBLUP$_G$ and msGBLUP$_H$ (0.837–1.146), whereas high variation was detected by ssGBLUP (0.333–0.797). When elder bulls were used as reference population, more biased results were obtained in all the traits by ssGBLUP. Interestingly, the discrepancies across different validation strategies and marker panels for all the traits were relatively small when ssGBLUP was used.

Single-step method has been implemented in real breeding population of different species to predict the genomic breeding values for various traits, since the methodology of constructing $\boldsymbol{H}$ was proposed. msGBLUP$_H$ and ssGBLUP have been widely used in previous studies. Koivula et al. (2012) and Su, Brøndum, et al. (2012); Su, Madsen, et al. (2012) concluded that GEBVs from msGBLUP$_H$ were comparable or better than DGVs obtained from msGBLUP$_G$ for Nordic Red Dairy bulls. ssGBLUP was used widely in recent research, such as research for piglet mortality (Guo et al., 2015), carcass traits in beef cattle (Croué & Ducrocq, 2017) and functional traits of Holstein cattle (Guarini et al., 2018). The feasibility of msGBLUP$_H$ was also extended to the production traits of Chinese Holstein cows (Li et al., 2014). However, genomic prediction for lowly heritable traits in Chinese Holstein population, such as reproductive traits, has not been reported yet let alone the single-step methods. This study established the first platform for the GS of reproductive traits for Chinese Holstein cattle.

Accuracies of msGBLUP$_H$ were more promising compared to those of msGBLUP$_G$ for all the GESS of reproductive traits in this study, which were inconsistent with previous research. In msGBLUP$_H$, increased b values were observed for all the traits compared with msGBLUP$_G$ (Table 6), indicating a slight deflation of GEBV. Additionally, in young bull validation, the predictions were more biased for CE and GL; when using msGBLUP$_H$ method, this was conflicted with the results of other studies (Guarini et al., 2018; Ma et al., 2015). One reason for higher accuracy in msGBLUP$_H$ than msGBLUP$_G$ may be due to an appropriate weight on pedigree information. This might account for the variation that was not accounted for by markers.

In different scenarios of ssGBLUP, the b values were more deviated from 1 compared with msGBLUP$_G$ and msGBLUP$_H$, which may be due to the elder (born before 1985) non-genotyped bulls in reference population (Table 3). These senior bulls were highly selected, whose information may lead to biased genomic prediction. Therefore, based on the results of current study, msGBLUP$_H$ is expected to be the optimal approach for predicting GESS of reproductive traits in the current small population.

Several researchers have reported that inclusion of genotyped cow in reference population would reduce bias of ssGBLUP method (Guarini et al., 2018; Su et al., 2016). However, the aim of present study was to evaluate the feasibility of predicting GESS. Inclusion of genotyped cows may not add additional value to the prediction. Given limited genotyped bulls, msGBLUP$_H$ could be a good method for GS of GESS.

4 | CONCLUSIONS

The present study reported promising results regarding the genomic prediction for genetic effect of service sire on cow reproductive traits in China Holstein cattle. In several scenarios, improved predictive ability was observed using high-density maker panel, but the gain was relatively small. Collectively, this research illustrated the impact of methodology to perform GS on GESS of reproductive traits and provide the first insight into the feasibility of genomic prediction for lowly heritable traits in Chinese Holstein population.

AUTHOR CONTRIBUTIONS

ZC, RS, GS and YW have made substantial contributions to conception and design. ZC, RS, GS and YW have made substantial contributions to analysis and interpretation of data. ZC and RS have been involved in drafting the manuscript. ZC, RS and GS have been involved in revising the manuscript critically for important intellectual content. HL, LL and GG have made substantial contributions to acquisition of data. All authors have read and approved the final manuscript.

ACKNOWLEDGEMENTS

We would like to thank CAU cattle genetics & breeding research team members who participated in sample collection and provided helpful hints in discussion.

FUNDING INFORMATION

This work was supported by China Agriculture Research System (CARS-36), National Agricultural Genetic Improvement Program (2130135), Beijing Sanyuan Breeding Technology Ltd. Co. funded project (SYZYZ20190005), the Program for Changjiang Scholar and Innovation Research Team in University (IRT_15R62), China Scholarship Council (No. 201913043) and Hainan University. The funders had no role in the study design, data collection and analysis, preparation of the manuscript or in the decision to publish the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Rui Shi https://orcid.org/0000-0002-0047-1262

REFERENCES

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, *93*, 743–752.

Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal*, *8*, 105–121. https://doi.org/10.1017/S1751731114000743

Bolormaa, S., Brown, D. J., Swan, A. A., van der Werf, J. H. J., Hayes, B. J., & Daetwyler, H. D. (2017). Genomic predictionof reproduction traits for Merino sheep. *Anim. Genet.*, *48*, 338–348. https://doi.org/10.1111/age.12541

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *American Journal of Human Genetics*, *103*, 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Campos, G. S., Reimann, F. A., Cardoso, L. L., Ferreira, C. E. R., Junqueira, V. S., Schmidt, P. I., Braccini Neto, J., Yokoo, M. J. I., Sollero, B. P., Boligon, A. A., & Cardoso, F. F. (2018). Genomic prediction using different estimationmethodology, blending and cross-validation techniques for growth traits andvisual scores in Hereford and Braford cattle. *J. Anim. Sci.*, https://doi.org/10.1093/jas/sky175

Chen, Z., Brito, L. F., Luo, H., Shi, R., Chang, Y., Liu, L., Guo, G., & Wang, Y. (2021). Genetic and genomic analyses of service sire effect on female reproductive traits in Holstein cattle. *Frontiers in Genetics*, *12*, 713575. https://doi.org/10.3389/fgene.2021.713575

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics, Selection, Evolution*, *42*, 2. https://doi.org/10.1186/1297-9686-42-2

Cole, J. B., & VanRaden, P. M. (2018). Symposium review: Possibilities in an age of genomics: The future of selection indices. *Journal of Dairy Science*, *101*, 3686–3701. https://doi.org/10.3168/jds.2017-13335

Croué, I., & Ducrocq, V. (2017). Genomic and single-step evaluations of carcass traits of young bulls in dual-purpose cattle. *Journal of Animal Breeding and Genetics*, *134*, 300–307. https://doi.org/10.1111/jbg.12261

Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de los Campos, G., & Hickey, J. M. (2013). Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics*, *193*, 347–365. https://doi.org/10.1534/genetics.112.147983

de Oliveira, H. R., Silva, F. F., Brito, L. F., Guarini, A. R., Jamrozik, J., & Schenkel, F. S. (2018). Comparing deregression methods for genomic prediction of test-day traits in dairy cattle. *Journal of Animal Breeding and Genetics*, *135*, 97–106. https://doi.org/10.1111/jbg.12317

Doormaal, B., Kistemaker, G., Sullivan, P., Sargolzaei, M., & Schenkel, F. (2009). Canadian implementation of genomic evaluations. *Interbull Bulletin*, *40*, 214–218.

Erbe, M., Hayes, B. J., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., Mason, B. A., & Goddard, M. E. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, *95*, 4114–4129. https://doi.org/10.3168/jds.2011-5019

Ertl, J., Edel, C., Emmerling, R., Pausch, H., Fries, R., & Götz, K.-U. (2014). On the limited increase in validation reliability using high-density genotypes in genomic best linear unbiased prediction: Observations from Fleckvieh cattle. *Journal of Dairy Science*, *97*, 487–496. https://doi.org/10.3168/jds.2013-6855

Guarini, A. R., Lourenco, D. A. L., Brito, L. F., Sargolzaei, M., Baes, C. F., Miglior, F., Misztal, I., & Schenkel, F. S. (2018). Comparison of genomic predictions for lowly heritable traits using multi-step and single-step genomic best linear unbiased predictor in Holstein cattle. *Journal of Dairy Science*, *101*, 8076–8086. https://doi.org/10.3168/jds.2017-14193

Guo, X., Christensen, O. F., Ostersen, T., Wang, Y., Lund, M. S., & Su, G. (2015). Improving genetic evaluation of litter size and piglet mortality for both genotyped and nongenotyped individuals using a single-step method1. *Journal of Animal Science*, *93*, 503–512. https://doi.org/10.2527/jas.2014-8331

Jaton, C., Schenkel, F. S., Malchiodi, F., Sargolzaei, M., Price, C. A., Baes, C., & Miglior, F. (2017). Genetic analysis for quality of frozen embryos produced by Holstein cattle donors in Canada. *Journal of Dairy Science*, *100*, 7320–7329. https://doi.org/10.3168/jds.2017-12851

Koivula, M., Strandén, I., Su, G., & Mäntysaari, E. A. (2012). Different methods to calculate genomic predictions—Comparisons of BLUP at the single nucleotide polymorphism level (SNP-BLUP), BLUP at the individual level (G-BLUP), and the one-step approach (H-BLUP). *Journal of Dairy Science*, *95*, 4065–4073. https://doi.org/10.3168/jds.2011-4874

Kropp, J., Peñagaricano, F., Salih, S. M., & Khatib, H. (2014). Invited review: Genetic contributions underlying the development of preimplantation bovine embryos. *Journal of Dairy Science*, *97*, 1187–1201. https://doi.org/10.3168/jds.2013-7244

Li, X., Wang, S., Huang, J., Li, L., Zhang, Q., & Ding, X. (2014). Improving the accuracy of genomic prediction in Chinese Holstein cattle by using one-step blending. *Genetics, Selection, Evolution*, *46*, 66. https://doi.org/10.1186/s12711-014-0066-4

Ma, P., Lund, M. S., Nielsen, U. S., Aamand, G. P., & Su, G. (2015). Single-step genomic model improved reliability and reduced the bias of genomic predictions in Danish Jersey. *Journal of Dairy Science*, *98*, 9026–9034. https://doi.org/10.3168/jds.2015-9703

Madsen, P., Sørensen, P., Su, G., Damgaard, L. H., Thomsen, H., & Labouriau, R. (2010). DMU—A package for analyzing Multivariate mixed models. In: Proc. 9th World Congr. Genet.Appl. Livest. Prod., Leipzig, Germany. Paper 732.

Misztal, I., Legarra, A., & Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*, *92*, 4648–4655. https://doi.org/10.3168/jds.2009-2064

Rezende, F. M., Nani, J. P., & Peñagaricano, F. (2019). Genomic prediction of bull fertility in US Jersey dairy cattle. *Journal of Dairy Science*, *102*, 3230–3240. https://doi.org/10.3168/jds.2018-15810

Su, G., Brøndum, R. F., Ma, P., Guldbrandtsen, B., Aamand, G. P., & Lund, M. S. (2012). Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and red dairy cattle populations. *Journal of Dairy Science*, *95*, 4657–4665. https://doi.org/10.3168/jds.2012-5379

Su, G., Ma, P., Nielsen, U. S., Aamand, G. P., Wiggans, G., Guldbrandtsen, B., & Lund, M. S. (2016). Sharing reference data and including cows in the reference population improve genomic predictions in Danish Jersey. *Animal*, *10*, 1067–1075. https://doi.org/10.1017/S1751731115001792

Su, G., & Madsen, P. (2011). User's Guide for Gmatrix: A program for computing Genomic relationship matrix. https://dmu.ghpc.au.dk/dmu/Gmatrix/Doc/Previous/Gmatrix-User-Guide.pdf

Su, G., Madsen, P., Nielsen, U. S., Mäntysaari, E. A., Aamand, G. P., Christensen, O. F., & Lund, M. S. (2012). Genomic prediction for Nordic red cattle using one-step and selection index blending. *Journal of Dairy Science*, *95*, 909–917. https://doi.org/10.3168/jds.2011-4804

Taylor, J. F., Schnabel, R. D., & Sutovsky, P. (2018). Review: Genomics of bull fertility. *Animal*, *12*, s172–s183. https://doi.org/10.1017/S1751731118000599

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M. (2020). Symposium review: How to implement genomic selection. *Journal of Dairy Science*, *103*, 5291–5301. https://doi.org/10.3168/jds.2019-17684

VanRaden, P. M., Tooker, M. E., O'Connell, J. R., Cole, J. B., & Bickhart, D. M. (2017). Selecting sequence variants to improve genomic predictions for dairy cattle. *Genetics, Selection, Evolution*, *49*, 32. https://doi.org/10.1186/s12711-017-0307-4

VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., & Schenkel, F. S. (2009). Invited review: Reliability of genomic predictions for north American Holstein bulls. *Journal of Dairy Science*, *92*, 16–24. https://doi.org/10.3168/jds.2008-1514

VanRaden, P. M., & Wiggans, G. R. (1991). Derivation, calculation, and use of National Animal Model Information. *Journal of Dairy Science*, *74*, 2737–2746. https://doi.org/10.3168/jds.S0022-0302(91)78453-1

Xu, L., Niu, Q., Chen, Y., Wang, Z., Xu, L., Li, H., Xu, L., Gao, X., Zhang, L., Gao, H., Cai, W., Zhu, B., & Li, J. (2021). Validation of the prediction accuracy for 13 traits in Chinese Simmental beef cattle using a preselected low-density SNP panel. *Animals*, *11*, 1890. https://doi.org/10.3390/ani11071890

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Shi, R., Chen, Z., Su, G., Luo, H., Liu, L., Guo, G., & Wang, Y. (2023). Genomic prediction of service sire effect on female reproductive performance in Holstein cattle: A comparison between different methods, validation population and marker densities. *Journal of Animal Breeding and Genetics*, *140*, 355–365. https://doi.org/10.1111/jbg.12763