Livestock Science ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Optimizing selection strategies of genomic selection in swine breeding program based on a dataset simulated ☆

Guoqing Tang [a,*], Junfeng Liu [a], Jia Xue [a], Ruifei Yang [a], Tianfei Liu [a], Zhiyao Zeng [a], Anan Jiang [a], Yanzhi Jiang [a], Mingzhou Li [a], Li Zhu [a], Lin Bai [a], Surong Shuai [a], Zhiquan Wang [b], Xuewei Li [a]

[a] College of Animal Science and Technology, Sichuan Agricultural University, Ya'an 625014, China
[b] Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada T6G2P5

ARTICLE INFO

Keywords:
Accuracy
Genomic selection
Genomic enhanced breeding value
Optimization
Pig

ABSTRACT

The conventional BLUP uses the phenotype and pedigree information to predict the estimated breeding values (EBV) of individuals in genetic evaluation. However, to obtain the phenotypes of interest in swine breeding program requires relatively a long time until the completion of performance testing, which produces an extra breeding cost for the culled pigs in swine industry. An alternative solution is to pre-select a predefined proportion of young replacement piglets for future performance testing through genomic selection, which could reduce the number of testing animals entering into performance testing program and hence reduce the breeding costs associated with the tests. In this study, four strategies of genomic selection applications in swine breeding program were compared through simulation to investigate the potential benefits in the different strategies in swine breeding program. For comparison purpose, the conventional BLUP selection was simulated as strategy 1 to form the benchmark basis for comparisons. Strategy 2 was an extreme case for applying genomic selection, where newborn piglets were selected directly based on their genomic enhanced breeding value (GEBV) only without further performance testing. Strategy 3 was to pre-select piglets, based on their GEBV, for entry to performance testing in early stage, and then the breeding stock were selected ultimately based on their EBVs predicted by BLUP method when the phenotypic records were available. Similar to strategy 3, strategies 4 and 5 also used the GEBVs to pre-select replacement piglets in an early stage; however, the breeding stocks in strategies 4 and 5 were selected based on the breeding values obtained using the bi-variable model and the conventional index method to combine GEBV and EBV information, respectively, when individuals had both GEBV and phenotypes available after the performance testing. Comparing these strategies, strategy 4 resulted in the highest accuracy in first three generations and achieved the best cumulative selection response in the last generation, followed by strategies 1, 3, 5, and 2. The proportion of pre-selection of boars and sows in the early stage affected the efficiency of genomic selection substantially.

1. Introduction

Genomic selection (GS) has become a very interesting field of research in livestock breeding in recent years and has attracted considerable attention from animal geneticists.

☆ This paper is part of the special issue entitled: Genomics Applied to Livestock Production, Guest Edited by Jose Bento Sterman Ferraz.

* Corresponding author. Tel.: +86 0835 2225215; fax: +86 0835 2886080.
E-mail address: tyq003@163.com (G. Tang).

http://dx.doi.org/10.1016/j.livsci.2014.04.023
1871-1413/

Research and applications of GS are growing rapidly worldwide. GS has been defined as selection of animals for breeding based on estimated breeding values calculated from the joint effects of genetic markers covering the entire genome (Goddard and Hayes, 2009) using a dense marker panel, so that all genes are expected to be in linkage disequilibrium with at least one of the markers on the panel. In a sense, GS is marker-assisted selection on a genome wide scale (Meuwissen, 2007). In 2009, only a few countries computed genomic enhanced breeding values (GEBV) for the Holstein dairy cattle breed, however, in 2011, data from 5 dairy breeds and 13 countries were provided for validation of national genomic evaluations at the Interbull Centre (Swedish University of Agricultural Sciences, Uppsala, Sweden; Loberg et al., 2011).

GS differs from the traditional Pedigree-BLUP (Henderson, 1975) breeding strategy. Implementation of genomic selection conceptually proceeds in two steps: (1) estimation of the effects of chromosome segments in a reference population and (2) prediction of GEBVs for animals in the selecting population (Goddard and Hayes, 2009). This second step is straightforward: to predict GEBVs of animals based on genotypes without phenotypes. In other words, the newborn piglets can be selected directly using their GEBVs and the biggest advantage of GS is to select early.

For traditional breeding strategy, the estimated breeding value (EBV) is calculated using the observed phenotypic record and pedigree information from the BLUP evaluation. However, the observed phenotype of production traits can only be measured when it is expressed in the animals' lifecycle, which may need a long time to process in swine breeding program. For example, the typical performance records can only be measured in a swine performance testing program when the body weight of pigs reach the range of 85–115 kg, which arose an extra breeding cost for those piglets culled in future. In addition, the market weights of breeding piglets are 15–50 kg in Chinese market, which cause the piglets culled in performance testing program to be difficult to sell since they are too large in weight.

Genomic selection has been adopted and implemented already in dairy cattle in recent years and young bulls have been selected directly using their GEBVs without progeny testing program in dairy breeding companies of several countries (Goddard and Hayes, 2009). However, this strategy has high risk to the practical breeding activities because the selection accuracy of GS will decrease gradually in next generations due to the recombination of markers and QTLs. A better strategy is applying GS to pre-select young replacement bulls using GEBVs in dairy cattle breeding, and then the young bulls selected will be used in a progeny testing program (Patry and Ducrocq, 2011a, 2011b; Patry et al., 2013). This strategy will reduce the number of young bulls entry into the progeny testing program, and thus reduce the breeding cost of dairy cattle. In addition, the bulls preselected will have both GEBV and phenotypic information of the traits of interest after the progeny test. The two data information can then be combined to further improve the selection accuracy and selection response in the dairy genetic improvement program.

Therefore, the objective of this paper is to investigate and compare four strategies of genomic selection application in swine breeding program based on their selection accuracy and expected selection responses.

2. Materials and methods

2.1. The swine data simulated

To obtain a realistic distribution of QTL effects and gene frequencies, a population of 4000 individuals was simulated in the initial generation with an equal number (2000) of male and female animals. The random mating was implemented in the system for many generations to allow the population to evolve until it reached the equilibrium between mutation and random genetic drift due to finite population size. For simplicity, natural and artificial selection in the past were negligible.

To construct the realistic Porcine60k (60,000) SNP panel, the porcine genome in this paper was assumed to consist of 19 chromosomes of 100 cMs each to mimic the porcine genome total of 19 chromosomes, and each with 3158 SNP markers. The position of each marker in a chromosome was determined by random sampling. The number of marker allele in the first generation was assumed to be 2 and with an equal (0.5) allele frequency. In this study, the backfat to 100 kg (B100) was simulated as the objective trait for selection and it was assumed to be influenced by 285 QTLs over the entire genome because some literature reported a quantitative trait was controlled generally by 200–2000 QTLs (such as http://www.aps.uoguelph.ca/~lrs/LRSsite/gws01.pdf). The position of each QTL in a chromosome was determined randomly. Similar to marker data, the initial QTL allele in the initial generation was also assumed to be 2 and with an equal allele frequency of 0.5. The effects of QTLs were assumed to follow a gamma distribution with shape parameter $\beta=0.4$ (Hayes and Goddard, 2001). Mutations occurred randomly at marker and QTL with mutation rates assumed to be 2.5×10^{-3} and 2.5×10^{-5} per locus per generation (Meuwissen et al., 2001). Every mutation at a marker locus resulted in a new unique marker allele, which made the mutated marker loci to become multiallelic loci. However, for meeting the biallelic SNP characteristics, these multiallelic markers were re-coded to SNP markers with the highest frequency in original marker coded to 1, while other alleles were coded to 2.

To reach a mutation-drift balance, populations were simulated stochastically for 1000 generations with 2000 individuals (1000 males and 1000 females). After these 1000 generations, 20 males and 200 females were sampled randomly from the population of 1000th generation to construct the initial founders of reference population. This differs from the implementation by Meuwissen et al. (2001) who increase the population size from 100 to 2000 after 1000 generations because the data simulated in this study was multiparous pigs. A total of 200 sows resulted in an average of 2000 piglets, and the progeny number of each sow was sampled from a normal distribution with the mean of 10 and the variance of 2.25. The numbers of sire and dam were kept as constant (20/200) in

Please cite this article as: Tang, G., et al., Optimizing selection strategies of genomic selection in swine breeding program based on a dataset simulated. Livestock Science (2014), http://dx.doi.org/10.1016/j.livsci.2014.04.023