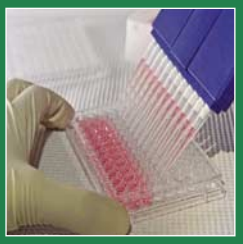
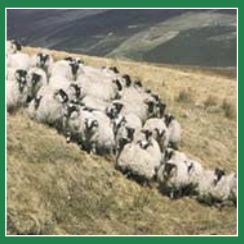
ISSN 2040-4700

APRIL 2012

VOLUME 3 PART 1



Advances in Animal Biosciences

Healthy Food from Healthy Animals

Proceedings of the British Society of Animal Science
and the Association of Veterinary Teaching
and Research Work

advances in animal biosciences

CAMBRIDGE UNIVERSITY PRESS

131

Comparison between BayesC and GBLUP methods to estimate genomic breeding values in threshold phenotypes

Masoud Shirali[1,2], Seyed Reza Miraei-Ashtiani[1], Abbas Pakdel[1], Chris Haley[2,3], Pau Navarro[3], Ricardo Pong-Wong[2]
[1]Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Islamic Republic of Iran, [2]Division of Genetics and Genomics, The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, UK, [3]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK *Email: masoud.shirali@gmail.com*

Introduction Accurate estimation of breeding values is one of the main aims in breeding programs. There are two main approaches in genomic selection for estimation of breeding values. The first approach assumes that all single-nucleotide polymorphisms (SNPs) have effects on the trait variance such as GBLUP and the second approach assumes that just some SNPs contribute to the trait variance such as BayesC. The objective of present study was to compare the accuracy in estimating genomic estimated breeding values (GEBVs) using two diverse methods, GBLUP and BayesC in threshold simulated traits.

Material and methods A genome of four chromosomes each of 100 cM length was simulated with 4000 SNPs equally spaced and 40 or 200 QTL randomly positioned on the genome and their effects sampled from a normal distribution. Linkage disequilibrium between the SNPs was generated by allowing a population of 50 mating pairs to undergo 50 generations of random mating. Thereafter, the population was expanded to 1000 per generation and 5 extra more generations were created. Given the QTL effects and the individuals' genotypes, breeding values were calculated and rescaled to have a variance of 0.3. Phenotypes were, then, simulated by adding an environmental effect sampled from a normal distribution with variance equal to 0.7 (i.e. h^2 =0.3). This continuous phenotype (CON) was, then used to calculate three threshold traits: TH1, TH2, TH3 by setting a threshold value and assigning individuals to be a "case" if their phenotype was above the threshold, or "control" otherwise. The threshold points for TH1, TH2, TH3 were at the mean, mean + SD and mean + 2SD, representing an incidence rate of approximately 50%, 16% and 2% respectively. A study was then carried out to compare the GBLUP and the BayesC method in term of the accuracy of their GEBVs. A set of 500, 1000 or 3000 phenotyped individuals (training population) was used to calculate the GEBV of 1000 unphenotyped individuals from last generation (testing population). The correlation between estimated and true genomic breeding values was used as measure of accuracy. T-test was used to compare BayesC and GBLUP estimations for all cases. A total of 10 replicates were used in the simulation.

Results The average accuracies for the different methods are shown in Table 1. As expected the accuracy of all methods and traits increased with the size of the training population. Additionally, the accuracy was greater for the continuous trait and decreased in the threshold traits when the incidence rate decreased. Decreasing the incidence rate from 50% (TH1) to 2% (TH3) resulted in approximately a loss of 30-40% in the GEBV accuracy. Compared with the results from the continuous trait, the accuracies when assuming an incidence rate of 2% were around 40-50% lower.
For the scenario assuming 40 QTL, in general the accuracy of the BayesC method was better than or at least the same as those obtained with GBLUP. On the other hand when assuming 200 QTL, the slight advantage of BayesC disappeared with GBLUP having slightly better accuracy for some scenarios. This diminishing advantage of BayesC over GBLUP has being previously reported when comparing them in scenarios considering continuous traits (Daetwyler *et al*., 2010).

Table 1 Average accuracies for BayesC and GBLUP for different numbers of QTL and training population sizes

# QTL	Trait	Training population					
		500		1000		3000	
		BC	GB	BC	GB	BC	GB
40	TH1	0.648	0.624	0.725	0.719	0.846	0.802
	TH2	0.557	0.543	0.652	0.624	0.784	0.748
	TH3	0.363	0.361	0.471	0.470	0.578	0.565
	CON	0.736	0.731	0.813	0.797	0.892	0.858
200	TH1	0.650	0.650	0.739	0.720	0.797	0.796
	TH2	0.605	0.607	0.667	0.675	0.734	0.747
	TH3	0.414	0.418	0.463	0.484	0.534	0.574
	CON	0.743	0.740	0.802	0799	0.859	0.852

BC: BayesC, GB: GBLUP

Conclusions The results from this study showed the advantage of genomic selection methods in threshold traits. Even for the least favourable scenario assuming an incidence rate of 2%, the accuracy of those methods were at least half those observed when assuming a continuous trait. Comparing both methods, BayesC tended to show slightly better accuracy when the trait was controlled by 40 QTL, but they disappeared when the number of QTL affecting the trait rose to 200. For the later situation, GBLUP had better accuracy for some scenarios.

Acknowledgements The first author acknowledges financial support from Iranian Ministry of Science, Research and Technology to support study at the Roslin institute.

References

Daetwyler H D, Pong-Wong R, Villanueva B, Woolliams J A 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185, 1021-31.