

Effect of Cows in the Reference Population: First Results in Swiss Brown Swiss

B. Bapst¹, C. Baes¹, F. R. Seefried¹, A. Bieber², H. Simianer³ and B. Gredler¹

¹Qualitas AG, Chamerstrasse 56, 6300 Zug, Switzerland

²Research Institute of Organic Agriculture, 5070 Frick, Switzerland

³Department of Animal Sciences, University of Göttingen, 37075 Göttingen, Germany

beat.bapst@qualitasag.ch

Abstract

The effect of including 1,236 cows in the current reference population of Swiss Brown Swiss cattle on the accuracy of genomic breeding values was studied. The accuracy of genomic predictions based on reference populations consisting of bulls only and bulls and cows was compared. The gain in accuracy was very small or not existing at all. Reasons could be that the number of cows was too small for a reference population consisting of 4,085 bulls and that the cows and bulls were too closely related. Nevertheless, genotyping cows and subsequent inclusion in the reference population should be accelerated.

Key words: genomic prediction, reference population, cows

Introduction

Accurate genomic predictions require large reference populations (RP) of bulls with reliable estimated breeding values (EBV) (Goddard and Hayes, 2009). For small dairy breeds the pool of eligible bulls for the RP is limited and the annual increase of the number of bulls to be included in the RP is moderate. Genotyping cows with inexpensive low density chips (e.g. Illumina BovineLD BeadChip) and subsequent imputation to a 50k SNP array to enlarge the RP could be a cost effective strategy (Gredler *et al.*, 2011). To date only a few countries have included cows in the RP (e.g. Pryce *et al.*, 2011; Dasonneville *et al.*, 2012). However, currently, international genotype sharing projects such as InterGenomics (2013) allow small populations, such as Brown Swiss dairy cattle, to improve the potential. The genotyped Swiss Brown Swiss population consists of 1,236 cows and 4,085 bulls. Currently, only bulls are included in the RP for genomic predictions. The objective of this study was to compare the accuracy of genomic prediction for milk yield (MKG) and somatic cell score (SCS) based on a RP containing bulls only (RP_B) and a RP containing bulls and cows (RP_BC).

Furthermore, we investigated the effect that relatedness between cows and bulls in the

RP_BC has on the accuracy of genomic predictions. The effect of production level of cows (elite or average) on the same accuracy in the same RP_BC was also analysed.

Material and Methods

In total, 4,085 and 3,891 bulls with EBV reliabilities $\geq 65\%$ for MKG and SCS, respectively, were included in the RP_B. A total of 1,236 and 1,162 cows with EBV reliabilities $\geq 50\%$ for MKG and SCS, respectively, were included in the RP_BC. All bulls and 480 cows were genotyped using the Illumina Bovine SNP50 Beadchip (50k) while 756 cows were genotyped using the Illumina BovineHD BeadChip (HD). Single nucleotide polymorphisms (SNP) present on the 50k chip but not on the HD chip were imputed according to the average allele frequency. SNP with a call rate $< 90\%$ were discarded from further analyses.

The prediction of genomic breeding values was carried out by applying BayesC ($\pi=0.95$) using the software package GenSel (Fernando and Garrick, 2009). SNP effects were estimated using deregressed EBV's for MKG and SCS as the response variables following the deregression procedure described by Garrick *et al.* (2009).

In a first step (step 1) direct genomic breeding values (DGV) were estimated based on either RP_B or RP_BC. Cows were grouped in specific sets according to their EBV reliabilities, which resulted in 4 different reference sets (Table 1 and 2). To test the accuracy of the DGV's, a forward prediction was carried out. The validation population (VP) consisted of 250 randomly sampled bulls born between 2005 and 2009. The accuracy of

estimated DGV was calculated as the correlation between the DGV's and the previously concealed deregressed EBV's ($r_{dgv,ebv}$) of bulls in the VP.

For taking into account that $r_{dgv,ebv}$ depends on the VP, a total of 120 replicates were executed for both RP_B and RP_BC within each scenario.

Table 1. Number of bulls and cows in the reference population for 4 scenarios for milk yield and somatic cell scores.

Scenario	Reference population	Milk yield			Somatic cell scores		
		Bulls n	Cows n	Total n	Bulls n	Cows n	Total n
All	Bulls reliability % ≥ 65	4,085	1,236	5,321	3,891	1,162	5,053
	Cows reliability % ≥ 50						
Cow55	Bulls reliability % ≥ 65	4,085	1,189	5,274	3,891	1,014	4,905
	Cows reliability % ≥ 55						
Cow60	Bulls reliability % ≥ 65	4,085	1,021	5,106	3,891	396	4,287
	Cows reliability % ≥ 60						
Cow65	Bulls reliability % ≥ 65	4,085	243	4,328	3,891	45	3,936
	Cows reliability % ≥ 65						

Table 2. Mean (Mean), minimum (Min) and maximum (Max) reliability of each scenario for milk yield and somatic cell scores of cows and bulls in the RP.

Scenario	Reference population	Milk yield			Somatic cell scores		
		Mean	Min	Max	Mean	Min	Max
All	Bulls reliability % ≥ 65	0.85	0.65	0.99	0.87	0.65	0.99
	Cows reliability % ≥ 50						
Cow55	Bulls reliability % ≥ 65	0.85	0.65	0.99	0.87	0.65	0.99
	Cows reliability % ≥ 55						
Cow60	Bulls reliability % ≥ 65	0.85	0.65	0.99	0.87	0.65	0.99
	Cows reliability % ≥ 60						
Cow65	Bulls reliability % ≥ 65	0.85	0.65	0.99	0.87	0.65	0.99
	Cows reliability % ≥ 65						

Secondly, the effect of the relationship between bulls and cows in RP_BC on the accuracy of genomic predictions for MKG was evaluated. A RP containing all cows from scenario Cow60 (Table 1) plus a bull subset of equal size (1,021) was used. Three levels of relatedness between the bulls and cows were considered: low, medium and high relatedness.

To construct the separate levels of relatedness, the genomic relationships between each bull and cow was calculated, after which the file was separated into low, medium and

high relatedness (sets), each containing 1,021 bull-cow "pairs". For each set, the DGV's for MKG were estimated, fitting the genomic relationship matrix as in VanRaden *et al.* (2009). As in step 1, the VP was sampled 120 times and the accuracy for each set was computed.

The third objective (step 3) was to explore the effect of the cows' production level (i.e. whether the cow is an elite or average cow) on the accuracy of genomic predictions. A sample of 264 cows was genotyped within the

framework of an internal program belonging to the breeding organization Braunvieh Schweiz. This framework awards the best 1% cows of the Swiss cow population an “elite” status based on EBVs for yield, SCS and conformation traits (Braunvieh Schweiz, 2011). The remaining 972 genotyped cows were part of the LowInputBreeds (2013) project, which investigates breeding strategies for low input and organic farmers. These cows were considered “average”, since their breeding values for yield traits represent the population mean. Therefore, scenario “All” (Table 1) from step 1 was carried out using all cows and only average cows corresponding to the LowInputBreeds project. Again, the accuracy of genomic predictions was assessed based on 120 replicates.

Results and Discussion

The main results for step 1 are shown in Figures 1 and 2. For MKG, the accuracy was slightly higher in all scenarios after cows were included in the RP. For SCS, no clear trend could be observed. For scenario Cow55 and Cow60, accuracies were slightly lower compared to using the RP_B (only bulls).

In Australia, 10,000 randomly chosen cows became part of the Holstein RP, which consists to date of 3,000 bulls. Pryce *et al.* (2013) observed an increase of the reliability of genomic prediction between 4 - 8% after adding cows, depending on the trait. Unfortunately no further studies are available that show the effect of including cows to the RP using real data.

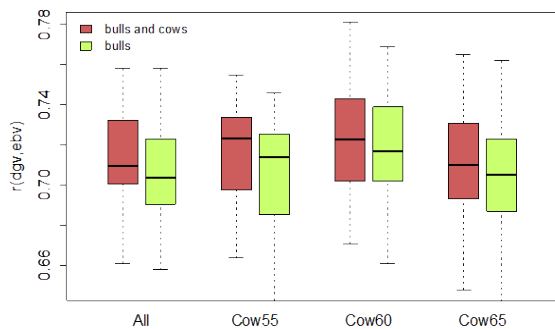


Figure 1. Accuracy of 120 replicates with and without cows in the RP for MKG for the 4 scenarios.

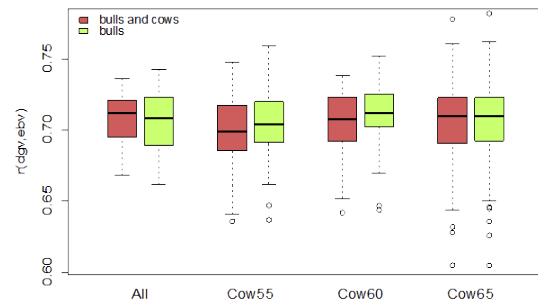


Figure 2. Accuracy of 120 replicates with and without cows in the RP for SCS for the 4 scenarios.

Different simulation studies (e.g. Thomasen, 2013) however show that the effect of having cows in the RP, especially for small breeds, could be favourable for an increased accuracy of prediction.

It generally seems that the inclusion of about 1,000 cows to a RP of 4,000 bulls is not enough to observe a significant increase in the accuracy of genomic prediction. The close genomic relatedness between bulls and cows in the RP could contribute to this problem.

Pszczola *et al.* (2013) showed that an increase in the accuracy of genomic prediction is achieved by minimizing relatedness between individuals within the RP and maximizing the relatedness between individuals in the RP and the VP.

Figure 3 shows that all cows included in this study are closely related. The cow set from the LowInputBreeds project includes many half sib families, contributing to this close relatedness. Hence, the cow set does not match the full genetic diversity shown by the bulls (Figure 3). This can also be observed for a subgroup of animals containing a high proportion of Original Braunvieh genes, which can be observed in Figure 3 as a small cluster on the left hand side of the figure.

The results from step 2 are shown in Figure 4. The accuracy of genomic prediction was highest when the relatedness between cows and bulls in the RP was low and decreased with increasing relatedness between cows and bulls. This confirms the findings of Pszczola *et al.* (2013) as previously described.

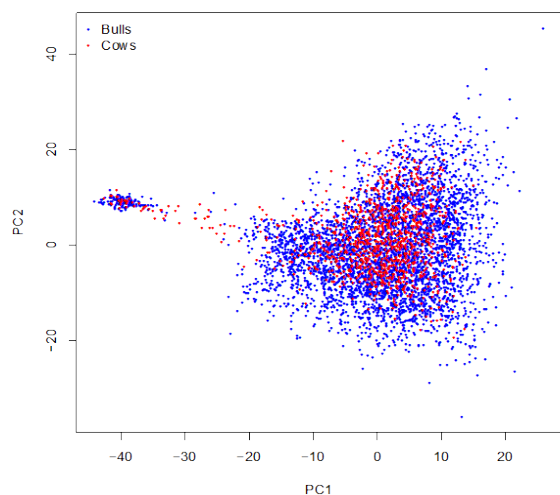


Figure 3. Structure of the RP for scenario 1 fragmented in bulls and cows.

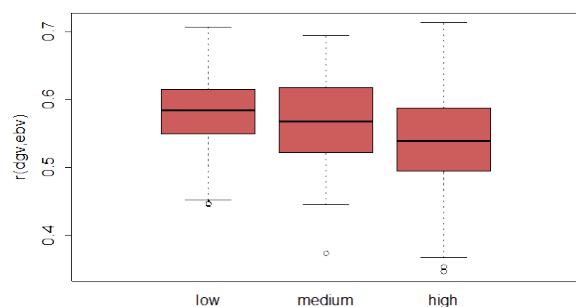


Figure 4. Accuracy for MKG depending on the relationship (low, medium, high) between the cow set and the bull set in the RP.

Evaluating the effect of cow production level on the accuracy of genomic prediction (step 3), did not show any significant difference in the accuracy for MKG and SCS when the “elite” cows were excluded from the RP. Dassonneville *et al.* (2012) previously showed that a bias in genomic evaluations can occur when a large number of elite cows are present in the RP, which was confirmed by Wiggans *et al.* (2011). This disputes our findings. However, the small number of “elite” cows available for this study could explain that no significant production level could be detected.

Conclusions

In this study, we could not detect a significant increase of accuracy of genomic prediction when varying the relatedness between bulls and cows in the RP. Neither can we conclude that the production level of the cow has a significant effect on the accuracy of genomic prediction. However, it is highly likely that a larger proportion of less related cows relative to bulls in the RP is required to detect the increase of accuracy previously described in literature. It has to be stressed that the findings of this study do not take away from the great potential that inclusion of cows in the RP has shown in literature. Therefore, the genotyping of cows should continue to be encouraged, stimulated by the continuous decrease of genotyping costs.

Acknowledgements

We thank the Association of Swiss Cattle Breeders (ASR) for financial support and Braunvieh Schweiz, Genotype pool Germany-Austria, Associazione Nazionale Allevatori Bovini della Razza Bruna and Beltsville Agricultural Research Centre for providing genotypes.

The authors gratefully acknowledge co-funding from the European Commission under LowInputBreeds, FP7 – project No KBBE 222 632 (The views expressed by the authors do not necessarily reflect the views of the European Commission, nor do they in any way anticipate the Commission’s future policy in this area.)

References

- Braunvieh Schweiz. 2011. Elite cows. Requirements for getting the award. Braunvieh Schweiz, Zug.
http://homepage.braunvieh.ch/xml_1/intern/et/de/application/d4/f519.cfm Accessed Aug. 15, 2013.

- Dassonneville, R., Baur, A., Fritz, S., Boichard, D. & Ducrocq, V. 2012. Inclusion of cow records in genomic evaluations and impact on bias due to preferential treatment. *Genet. Sel. Evol.* 44:40.
- Fernando, R.L. & Garrick, D.J. 2009. *GenSel-User manual for a portfolio of genomic selection related analyses*. Animal Breeding and Genetics, Iowa State University, Ames.
<http://www.biomedcentral.com/content/supplementary/1471-2105-12-186-s1.pdf>
Accessed Aug. 15, 2013.
- Garrick, J.D., Taylor, J.F. & Fernando, R.L. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* 41:55.
- Goddard, M.E. & Hayes, B.J. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet.* 10:6, 381-391.
- Gredler, B., Seefried, F.R., Schuler, U., Bapst, B., Schnyder, U. & Hickey, J.M. 2011. Imputation in Swiss Cattle Breeds. *Interbull Bulletin* 44, 8-11.
- Intergenomics. 2013. Project of the European Brown Swiss Federation for developing a common genomic evaluation program for the Brown Swiss breed at Interbull.
<http://www.brown-swiss.org/ScientificResults/genomics.html>
Accessed Aug. 15, 2013
- LowInputBreeds. 2013. Development of integrated livestock breeding and management strategies to improve animal health, product quality and performance in *European organic and 'low input' milk, meat and egg production*. EU FP7 - project No KBBE 222 632.
<http://www.lowinputbreeds.org/> Accessed Aug. 15, 2013
- Pryce, J.E., Blake, M., Davis, K.E. & Beggs, D.S. 2013. Challenges of health data recording – an Australian perspective. *ICAR 2013 Technical Meeting*. Aarhus, Denmark.
http://www.icar.org/Documents/Aarhus_2013/Manuscripts/Pryce.pdf Accessed Aug. 15, 2013.
- Thomasen, J.R. 2013. Genomic selection in small dairy cattle populations. *PhD Thesis*. Aarhus, Denmark.
https://pure.au.dk/portal/files/53600950/PhDthesis_J_rnRindThomasen.pdf Accessed Sept. 9, 2013.
- VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:11, 4414-4423.
- Wiggans, G.R., Cooper, T.A., VanRaden, P.M. & Cole, J.B. 2011. Technical note: Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. *J. Dairy Sci.* 94, 6188–6193.