Different Strategies for Genomic Prediction of Average Daily Weight Gain in Feedlot in Nellore Finishing Steers

*A.L. Somavilla**¶, L.C.A. Regitano†‖, F.B. Mokry‡, M.A. Mudadu†, R.R. Tullio†, M.L. Nascimento#, L.L. Coutinho#‖ and D.P. Munari*‖.
*Univ. Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, Brazil,
¶Fellowship Sao Paulo Research Foundation (FAPESP) n.2012/23702-8, n.2013/21644-3,
†Embrapa Pecuária Sudeste, ‡Univ. Federal de São Carlos, #Univ. de São Paulo/Esalq, ‖CNPq Fellowship.

ABSTRACT: Average daily weight gain data were recorded for 804 Nellore finishing steers during feedlot periods in Brazil between 2007 and 2009 and two experimental stations. These animals were genotyped with a 700k SNP panel, which yielded three other SNP subset strategies (TagSNPs, GGP HDi and GGP LDi). The genomic breeding values were estimated by the rr-BLUP approach and their accuracies were computed to compare the strategies. It was possible to observe some differences among the genetic variances assessed by each of the subsets. GGP HDi 90K and TagSNPs strategies estimated GEBVs with a wider distribution than SNPs from the whole dataset. This can be an evidence that using more informative markers in a small sample could result in better accuracies estimations. However, the differences in the variances of the GEBV suggest a necessity for a new parameterization of the model.
Keywords: beef cattle; *Bos indicus*; growth; SNPs

Introduction

In livestock production, desired animals are those with faster growth rates, which spend less time in pasture or feedlots, as such animals can be sent for processing on the market earlier and for less money than their slower maturing counterparts. In this context, the number of days to achieve mature weight (Machado, Aquino and Gonçalves (1999)) or average daily weight gain (ADG) (Sarmento, Pimenta Filho, Ribeiro et al. (2003)) can be used as a selection criteria. To minimize the disadvantages of increased inbreeding rates of the traditional BLUP methodology, studies on molecular markers, which might explain individuals genetic variability, are being conducted, and found that was possible to predict the genomic breeding value (GEBV) of individuals based on dense marker information (Meuwissen, Hayes and Goddard (2001)).

Genomic selection has several advantages when compared to traditional BLUP, one of which is the decrease in inbreeding rates (Daetwyler, Villanueva, Bijma et al. (2007)), since GEBV becomes less correlated among related individuals. In addition, over generations, the value of a specific individual has less influence in selection of its offspring and the genetic variability is maintained in the long term. Furthermore, it is possible to reduce generation intervals, as a consequence of more accurate GEBV predictions from young animals.

Genomic selection has been applied to dairy cattle production in the USA (VanRaden, Van Tassell, Wiggans et al. (2009)), and numerous studies have been conducted in order to identify appropriate methodologies to specific breeds and traits, which result in more accurate GEBVs. Based on the importance of Nellore (*Bos indicus*) cattle in Brazil, it is necessary to properly identify methodologies that allow genomic selection of animals with higher growth rates.

Materials and Methods

Samples. During the mating seasons of 2007 through 2009, 804 steers, offspring's of 34 Nellore bulls, were generated through fixed-time artificial insemination on five farms, where they are raised to around 21 months of age and later moved to either Embrapa Southeast Livestock (São Carlos, SP) or Embrapa Beef Cattle (Campo Grande, MS) for three seasons of feedlot experiment periods (2009/2010/2011).

Phenotype Data. Animals were fed corn or sorghum silage with at least 35% dry matter, and concentrate containing corn, soybean meal, soybean hull, cotton seed, limestone, mineral mixture, urea and monensin (Rumensin®), accounting around 13% crude protein and 71% total digestible nutrients. The diet was provided twice a day in which the feed offered (total mixture composed by silage + concentrate) were adjusted daily to ensure ad libitum consumption. The animals were weighed every 14 days without fasting, for an average period of 91 days.

Genotype Data. The Illumina BovineHD BeadChip, which contains more than 777k single nucleotide polymorphisms (SNP), was used for genotyping 780 steers and 34 bulls. The quality control criteria to keep markers were: SNP call rate > 98%, minor allele frequency (MAF) > 1%, and sample call rate > 90%. Unplaced SNPs, mitochondrial and sexual SNPs were discarded. The Beagle v.3.3.2 (Browning and Browning (2009)) software was used for phase inference and imputation of missing genotypes.

Statistical analyses. A linear regression analysis of live weight over time was performed. The intercept was used as the adjusted initial weight and the slope as the average daily gain during the feedlot period (ADG). A linear model was used to adjust the phenotype with estimates of

systematic effects (de los Campos, Hickey, Pong-Wong et al. (2013)). The model included fixed effects of contemporary groups (season + farm of birth + farm of feedlot), initial weight (IW), feedlot period (FP), and slaughter age (SLA). The adjusted phenotype was the overall mean plus the residual value. The rr-BLUP analysis (equivalent to GBLUP) implemented in GS3 software (Legarra, Ricard and Filangi (2012)) was performed to predict genomic breeding values (GEBVs) comparing four SNP sets strategies: 1) 427,813 SNPs (whole dataset) which passed quality control; 2) 219,572 TagSNPs selected from dataset 1 by using Tagger tool (Bakker, Yelensky, Pe'er et al. (2005)) implemented in Haploview (Barrett, Fry, Maller et al. (2005)); 3) 14,294 SNPs from GeneSeek Genomic Profiler (GGP) LDi 20K after quality control; 4) 61,017 SNPs from GGP HDi 90K. First, the variance components were estimated using VCE (option of GS3), which implements a MCMC analysis. The convergence analysis was performed using the BOA package (Smith (2007)). To allow for cross-validation, each dataset was randomly divided in four subsets, such that all animals were part of the validation (V) set at least once. The MCMC option was used to estimate the SNPs effects on training (T) subsets, and the PREDICT option to estimate phenotypes of animals in V subsets. The accuracy of each strategy was estimated by the Pearson correlation between adjusted phenotypes and GEBVs $r(y'_i, GEBV_i)$.

Results and Discussion

Descriptive statistics (Table 1) from ADG were close to those reported for Nellore cattle (Marques, Magnabosco, Lopes et al. (2013)), but using only pedigree information, they found higher genetic additive variance and heritability (0.55 ± 0.021) than the ones estimated in this study. Animals slaughter age were much higher on this study, almost double in some cases, and it might contribute to decrease the genetic variability of the trait among individuals. As observed in heritability estimate, it was not possible to assess the genetic variance using SNPs from GGP LDi 20K (h^2 = 0.0004). Previous analysis on GGP LDi 20K dataset (not shown) suggested that extent of linkage disequilibrium in Nellore breed is shorter when compared to Holstein cattle, so this could contribute for this result (Habier, Fernando and Garrick (2013)). For the other three approaches, heritability estimates were low (0.04, 0.13 and 0.14 for datasets 1, 2 and 4, respectively). Despite containing less SNPs, it seems that dataset 4 (GGP HDi 90K) was as informative as dataset 2 (TagSNPs) and more than dataset 1. This was expected because SNPs from GGP HDi 90K were chosen to represent variation among *Bos indicus* individuals. Disregarding the GGP LDi 20K set, in which the accuracy was not computed, the estimates from three other subsets were similar (Table 2), therefore it seems GGP HDi 90K has produced the same estimates with fewer markers available. Low estimates may be a consequence of small training population size and low heritability coefficients, since both affect prediction accuracy (Bolormaa, Pryce, Kemper et al. (2013)). The distribution of GEBVs was wider in the GGP HDi 90K subset (Figure 1C) than in TagSNPs (Figure 1B) and whole datasets (Figure 1A). When comparing the top 5 selected animals among SNPs subset strategies, there were 86.84, 34.21, and 28.94% agreement between subsets 1-2, 2-3 and 1-3, respectively, and 28.94% among all subsets. Subset 2 (TagSNPs) could assess the genetic variance and higher GEBVs estimation than subset 1, an indication that selecting the most representative SNPs might be a reasonable option.

Table 1. Descriptive statistics for average daily weight gain (ADG), initial weight (IW), feedlot period (FP) and slaughter age (SLA) in Nellore cattle

	Mean	SD	CV	Min	Max
ADG, kg	1.16	0.32	27.50	0.02	2.190
IW, kg	362.24	49.47	13.66	204.20	514.90
FP, days	91.73	19.99	21.80	48.00	119.00
SLA, days	739.78	47.06	6.34	634.00	862.00

SD = standard deviation, CV = coefficient of variation, Min = minimum, Max = maximum

Table 2. Accuracies for genomic prediction of average daily weight gain in Nellore cattle using four different density markers panel

Approach[§]	accuracy
1	0.19
2	0.18
3	-
4	0.20

[§]1 = SNPs after quality control, 2 = TagSNPs, , 3 = GGP LHi 20K, 4 = GGP HDi 90K.

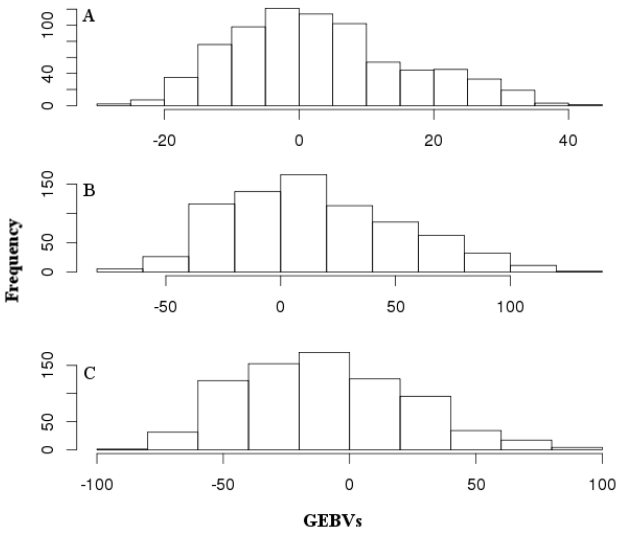

Figure 1. Distribution of GEBVs estimated by whole dataset (A), TagSNPs (B) and GGP HDi 90K (C)

Conclusion

The rr-BLUP is a feasible methodology to predict genomic breeding values for average daily weight gain in Nellore cattle. GGP HDi 90K subset yield the best estimates in this study, an evidence that using more informative markers in a small sample could result in better estimations. However, the differences in the variances of the GEBV suggest a necessity for a new parameterization of the model.

Literature Cited

Bakker, P. I. W., R. Yelensky, I. Pe'er et al. (2005). *Nat. Genet.* 37(11): 1217-1223.

Barrett, J. C., B. Fry, J. Maller et al. (2005). *Bioinformatics* 21(2): 263-265.

Bolormaa, S., J. Pryce, K. Kemper et al. (2013). *Genet. Sel. Evol.* 45(1): 43.

Browning, B. L. and S. R. Browning (2009). *Am. J. Hum. Genet.* 84(2): 210-223.

Daetwyler, H. D., B. Villanueva, P. Bijma et al. (2007). *J. Ani. Breed. Genet.* 124(6): 369-376.

de los Campos, G., J. Hickey, R. Pong-Wong et al. (2013). *Genetics* 193: 327 - 345.

Habier, D., R. L. Fernando and D. J. Garrick (2013). *Genetics* 194(3): 597-607.

Legarra, A., A. Ricard and O. Filangi (2012). GS3: Genomic Selection, Gibbs Sampling, Gauss-Seidel (and BayesCπ). http://snp.toulouse.inra.fr/~alegarra/.

Machado, P. F. A., L. H. Aquino and T. M. Gonçalves (1999). *Ciênc. Agrotec.* 23(1): 197-204.

Marques, E. G., C. U. Magnabosco, F. B. Lopes et al. (2013). *Biosci. J.* 29(1): 159-167.

Meuwissen, T., B. Hayes and M. Goddard (2001). *Genetics* 157: 1819 - 1829.

Sarmento, J. L. R., E. C. Pimenta Filho, M. N. Ribeiro et al. (2003). *Rev. Bras. Zootecn* 32: 325-330.

Smith, B. J. (2007). *J. Stat. Soft.* 21(11): 1-37.

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans et al. (2009). *J. Dairy Sci.* 92(1): 16-24.