Proceedings of the World Congress on Genetics Applied to Livestock Production, 11.474

Genomic prediction for dry matter intake in Nordic Holsteins

G. Su[1], D. G. M. Gordo[1], B. Li[1], G. P. Aamand[2], M. S. Lund[1], J. Lassen[1]

[1] Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark
[2] Nordic Cattle Genetic Evaluation, DK-8200 Aarhus N, Denmark
Guosheng.su@mbg.au.dk (Corresponding Author)

Summary

Genetic improvement of feed efficiency (FE) is an important issue since feed is the major cost in dairy farms. However, individual feed intake is difficult or costly to measure, which limits selection accuracy. Genomic selection is expected to be a promising tool to start selection for FE. The objective of this study is to investigate accuracy of genomic prediction for dry matter intake (DMI) in the case of a small reference data in Nordic Holsteins. Two data sets were used in this study. The first data set included 53,914 week-records of DMI records from 1,623 lactations of 827 cows. Among these cows, 330 individuals were genotyped with 54k chip. The other data set included 5,409 progeny-tested bulls with marker data of 54k chip and de-regressed proof of milk yield. Breeding value were predicted using a univariate and a bivariate (with milk yield as assistant trait) conventional BLUP model, as well as a univariate single-step model (ssGBLUP) and a bivariate ssGBLUP model.

Using single-trait analysis, correlation between EBV and corrected 305d DMI for genotyped animals increased from 0.287 to 0.393 in validation scenario without sibs in training data, and from 0.366 to 0.432 in scenario allowing sibs in training data, when moving conventional BLUP model to ssGBLUP model. The ssGBLUP also led to a slight improvement of EBV accuracy for non-genotyped animals. The bivariate models using milk yield information resulted in a considerable increase of model-based reliability for both genotyped and non-genotyped animals, but an increase of correlation between EBV and corrected 305d DMI only for genotyped animals. Although we used a small reference population, the results indicated that genomic prediction leads to better prediction accuracy than conventional BLUP method. To obtain a higher accuracy of EBV for DMI, more data are required.

Keywords: dairy cattle, dry matter intake, feed efficiency, genomic prediction, single-step GBLUP

Introduction

Feed is the greatest cost for dairy farmers, limiting the profitability of dairy production. Thus, improving feed efficiency (FE) through genetic selection is an important opportunity for the dairy industry. In Holsteins, heritability estimates of feed intake range from 0.12 to 0.53 (de Haas et al., 2015; Li et al., 2016) and therefore is likely to be changed via genetic improvement.
Recording FE is, however, complicated because these traits are difficult and costly to measure. This situation may result in a limited amount of information and, consequently, a low accuracy of estimated breeding values. Although the number of records for feed intake is low, the utilization of genomic selection is suggested as a promising tool to start selection for FE (Calus et al., 2013).

Previous studies have reported that genomic selection for feed intake is possible with an accuracy of around 0.4 (de Haas et al., 2012, 2015). It is important to highlight that most studies have validated the genomic breeding values (GEBV) in cows and growing heifers, and the reference animals are contemporaries of the validation animals. Therefore, attention has to be taken when using these results into accuracy of bull selection. A validation strategy more consistent with real life scenario is required for validation of prediction accuracy for feed intake.

The objective of this study is to investigate accuracy of genomic prediction for DMI in a small reference data in Nordic Holsteins, using single-step GBLUP model (Christensen and Lund, 2010).

Materials and Methods

Data

Two data sets were used in this study. The first data set included 53,914 week-records of DMI records from 1,623 lactations of 827 cows, covering calving year from 1991 to 2014 and birth year of cows from 1988 to 2013. Among these cows, 330 individuals were genotyped with 54k chip. The other data set included 5,409 progeny-tested bulls with marker data of 54k chip. De-regressed proof of milk yield in this data set were used as assistant trait to predict breeding value of DMI using a bivariate model. The raw data of DMI were edited by deleting the records out of the period from lactation week 1 to 44, the lactations with less than five week-records, and the lactations over the third parity. See details of raw DMI data in Li et al (2016).

Statistical models

To account for non-genetic effect accurately and to perform a validation procedure, corrected 305 d DMI (DMI_{305}) were derived from week-records using the following model,

$$DMI_w = \text{Herd\_Year\_Season} + \text{FeedingTrial} + \text{Parity\_ AgeGroup} + b_1 fwk + b_2(fwk)^2 + b_3 lwk + b_4(lwk)^2 + PE + b_{PE1}L1 + b_{PE2}L2 + A + b_{a1}L1 + b_{a2}L2 + \text{residual},$$

In the model, fwk = wk of days in milking/44, lwk=log(44/wk of days in milking), L1 and L2 were the first and second orders of Legendre polynomials. b_1-b_4 were fixed regression coefficients to describe lactation curve of DMI. PE, b_{PE1} and b_{PE2} were intercept and regression coefficients for random permanent effect. A , b_{a1}, and b_{a2} are intercept and regression coefficients for random additive genetic effect. To account for heterogeneous residual variances, different weights according to residual variances estimated from analysis on each 4 weeks were applied to residual variances in different lactation periods. The corrected DMI_{305} was calculated as the accumulation of additive genetic and permanent environment as well as residual effects over 305 d according to the parameters estimated from the above model.

Breeding values of DMI_{305} were predicted using a univariate and a bivariate (with milk yield as extra source of data information) conventional BLUP model, as well as a univariate ssGBLUP model and a bivariate ssGBLUP model. The basic model is

$$DMI_{305} = \mu + \text{additive genetic} + \text{residual}$$

In the model, the number of week-records was used as weight to account for heterogeneous residual variance of DMI_{305}.

When using ssGBLUP model, a weight of 0.20 was put on pedigree-based relationship matrix for building the joint marker-pedigree relationship matrix, and the genomic relationship matrix was adjusted to be in similar scale as pedigree relationship matrix as proposed by Christensen et al (2012). Estimation of variance components and prediction of breeding values were carried out using DMU package (Madsen and Jensen, 2013).

Validation

Accuracy of EBV for DMI_{305} using different models was assessed by a 5-fold cross-validation procedure. Two scenarios of validation were performed. In the first cross-validation scenario, the half-sib families in the whole DIM_{305} data were randomly divided into five subsets (family validation, VFAM). In the other scenario, the individuals in the whole DIM305 data were randomly divided into five subsets (random individual validation, VRAN). In each fold validation, one subset was used as test data and the other subsets as training data. The two scenarios differed in relationship between test and training animals. In VFAM, the test animals did not have sibs in the corresponding training data, thus a relatively distant relationship between test and training animals. Conversely, in VRAN, the test animals had sibs in training data set, indicating some relationship between test and training animals. Accuracy of EBV was measured by two definitions. One is the correlation between EBV and corrected DMI_{305}. Since the DMI data were collected from a long period between 1991 and 2014, both EBV and DMI_{305} were adjusted for year mean in calculation of the correlation to avoid that correlation was overestimated caused by genetic trend across years. The other measure is model-based reliability which was estimated according to prediction error variance derived from inverse of left-hand of mixed model equation.

Results and discussion

As shown in Table 1 and 2, accuracies of EBV for genotyped animals were higher than those for non-genotyped animals even when using conventional BLUP. This could be explained by the fact that genotyped animals were born in recent years during which more cows had DMI records, thus there were more data information for predicting breeding values. As expected, accuracies of EBV in validation scenario VRAN were higher than those in VFAM, because of the closer relationship between test and training animals.

Table 1. Correlation between corrected DIM_{305} and EBV from univariate (MT1) and bivariate (MT2) analyses using BLUP and ssGBLUP models.

Scenario	Type of animal	BLUP MT1	BLUP MT2	SSGBLUP MT1	SSGBLUP MT2
VFAM	All	0.269	0.287	0.318	0.320
	Genotyped	0.287	0.340	0.393	0.440
	Non-genotyped	0.258	0.257	0.277	0.258
VRAN	All	0.333	0.332	0.358	0.352
	Genotyped	0.366	0.387	0.432	0.447
	Non-genotyped	0.309	0.301	0.316	0.300

Although there were only 330 cows with genotypes, ssGBLUP led to higher accuracy of EBV than conventional BLUP. According to Table 1, the improvement of accuracy was not only for genotyped animals but also for non-genotyped animals in both scenarios, although there was a larger improvement for genotyped animals than for non-genotyped animals. The gain from genotype information in scenario VFAM was larger than in VRAN. This indicates that the benefit from genomic prediction is larger for animals with a more distant relationship with the relatives having phenotypic records. The scenario of VFAM could be consistent with the real life scenario in cattle breeding where young bulls were selected before the daughters and sibs have records. According to model-based reliability (Table 2), the improvement of EBV accuracy was only for genotyped animals, and the gains from genotype information were not different in the two validation scenarios.

Table 2. Model-based reliability of EBV from univariate (MT1) and bivariate (MT2) analyses using BLUP and ssGBLUP models

Scenario	Type of animal	BLUP MT1	BLUP MT2	SSGBLUP MT1	SSGBLUP MT2
VFAM	All	0.261	0.299	0.276	0.323
	Genotyped	0.328	0.365	0.379	0.440
	Non-genotyped	0.216	0.255	0.207	0.245
VRAN	All	0.323	0.343	0.333	0.365
	Genotyped	0.401	0.416	0.442	0.489
	Non-genotyped	0.271	0.271	0.260	0.382

The gains from bivariate models applying information of milk yield were not consistent between the two measures of accuracies. According to the correlation between the EBV and DMI_{305}, multiple-trait analysis increased accuracy of EBV only for the genotyped animals. For the based reliability model, however, the bivariate model using information of milk yield largely increased the reliability of EBV.

The above results confirm that genomic selection is a promising tool to selection for DMI. In order to obtain a high accuracy of EBV for DMI to satisfy practical breeding, however, more DMI data is required. A solution to overcome the costly-to-measure issue would be an international collaboration so that phenotypic and genotypic information from several populations could be used to predict genomic breeding values for DMI. Within this context, some international collaborative studies using feed intake in Holsteins have shown the possibility to increase the accuracy of genomic estimated breeding values for feed intake (de Haas et al., 2012, 2015). Using data on individual daily DMI of Holstein cows and heifers from 10 populations, de Haas et al. (2015) have shown that the accuracy of genomic prediction increases from 0.37 when performing within country evaluation to 0.44 when applying a joint evaluation. It should be noted that there are still some challenges to get large and well defined DMI data through international cooperation.

Conclusions

Even though we used a small reference population, genomic prediction using ssGBLUP improved the accuracy of EBV for DMI. Current multiple-trait model using information of correlated trait could slightly increase the prediction accuracy for DMI. To obtain a higher accuracy of EBV for

DMI, more data are required.

List of References

Calus, M. P. L., D. P. Berry, G. Banos, Y. de Haas, and R. F. Veerkamp, 2013. Genomic selection: the option for new robustness traits? Advances in Animal Biosciences, 4(03):618-625.

Christensen, O. F., and M. S. Lund, 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2.

Christensen, O. F., P. Madsen, B. Nielsen, T. Ostersen, and G. Su, 2012. Single-step methods for genomic evaluation in pigs. Animal 6:1565–1571.

de Haas, Y., M. P. L. Calus, R. F. Veerkamp, E. Wall, M. P. Coffey, H. D. Daetwyler, B. J. Hayes, and J. E. Pryce, 2012. Improved accuracy of genomic prediction for dry matter intake of dairy cattle from combined European and Australian data sets. Journal of dairy science, 95(10):6103-12.

de Haas, Y., J. Pryce, M. Calus, E. Wall, D. Berry, P. Løvendahl, N. Krattenmacher, F. Miglior, K. Weigel, D. Spurlock, K. Macdonald, B. Hulsegge, and R. Veerkamp, 2015. Genomic prediction of dry matter intake in dairy cattle from an international data set consisting of research herds in Europe, North America, and Australasia. Journal of Dairy Science, 98(9):6522-6534.

Li, B., W. Fikse, J. Lassen, M. Lidauer, P. Løvendahl, P. Mäntysaari, and B. Berglund, 2016. Genetic parameters for dry matter intake in primiparous Holstein, Nordic Red, and Jersey cows in the first half of lactation. Journal of Dairy Science, 99(9):7232 - 7239.

Madsen, P. and J. Jensen, 2013. A Package for Analysing Multivariate Mixed Models. ver. 6, rel. 5.2 http://dmu.agrsci.dk/DMU/Doc/Current/dmuv6_guide.5.2.pdf.