### JOURNAL OF ANIMAL SCIENCE

The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

## Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens

C. Y. Chen, I. Misztal, I. Aguilar, S. Tsuruta, T. H. E. Meuwissen, S. E. Aggrey, T. Wing and W. M. Muir

*J Anim Sci* 2011.89:23-28. doi: 10.2527/jas.2010-3071 originally published online Oct 1, 2010;

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://jas.fass.org/cgi/content/full/89/1/23



www.asas.org

# Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens

C. Y. Chen,\*<sup>1,2</sup> I. Misztal,\* I. Aguilar,\*† S. Tsuruta,\* T. H. E. Meuwissen,‡ S. E. Aggrey,§ T. Wing,# and W. M. Muir||

\*Department of Animal and Dairy Science, University of Georgia, Athens 30602-2771; †Instituto Nacional de Investigación Agropecuaria, Las Brujas 90200, Uruguay; ‡Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, NO-1432 As, Norway; §Department of Poultry Science, University of Georgia, Athens 30602-2772; #Cobb-Vantress Inc., PO Box 1030, Siloam Springs, AR 72761-1030; and ||Department of Animal Science, Purdue University, West Lafayette, IN 47907-1151

**ABSTRACT:** Data of broiler chickens for 2 pure lines across 3 generations were used for genomic evaluation. A complete population (full data set; FDS) consisted of 183,784 and 164,246 broilers for the 2 lines. The genotyped subsets (SUB) consisted of 3,284 and 3,098 broilers with 57,636 SNP. Genotyped animals were preselected based on more than 20 traits with different index applied to each line. Three traits were analyzed: BW at 6 wk (BW6), ultrasound measurement of breast meat (BM), and leg score (LS) coded 1 = no and 2 = vesfor leg defect. Some phenotypes were missing for BM. The training population consisted of the first 2 generations including all animals in FDS or only genotyped animals in SUB. The validation data set contained only genotyped animals in the third generation. Genetic evaluations were performed using 3 approaches: 1) phenotypic BLUP, 2) extending BLUP methodologies to utilize pedigree and genomic information in a single step (ssGBLUP), and 3) Bayes A. Whereas BLUP and ssGBLUP utilized all phenotypic data, Bayes A could use only those of the genotyped subset. Heritabilities were 0.17 to 0.20 for BW6, 0.30 to 0.35 for BM, and 0.09 to 0.11 for LS. The average accuracies of the validation population with BLUP for BW6, BM, and LS were 0.46, 0.30, and <0 with SUB and 0.51, 0.34, and 0.28 with FDS. With ssGBLUP, those accuracies were 0.60, 0.34, and 0.06 with SUB and 0.61, 0.40, and 0.37 with FDS, respectively. With Bayes A, the accuracies were 0.60, 0.36, and 0.09 with SUB. With SUB, Bayes A and ssGBLUP had similar accuracies. For traits of high heritability, the accuracy of Bayes A/SUB and ssGBLUP/FDS were similar, and up to 50% better than BLUP/FDS. However, with low heritability, ssG-BLUP/FDS was 4 to 6 times more accurate than Bayes A/SUB and 50% better than BLUP/FDS. An optimal genomic evaluation would be multi-trait and involve all traits and records on which selection is based.

**Key words:** chicken, genetic evaluation, genomic prediction

© 2011 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2011. 89:23–28 doi:10.2527/jas.2010-3071

#### INTRODUCTION

Availability of large panels of SNP markers created an interest in genomic selection (Meuwissen et al., 2001) as a tool to increase the accuracy of prediction for young animals. These tools were first used in selection of dairy cattle (Hayes et al., 2009; VanRaden et al., 2009). A typical implementation in dairy cattle

Received April 9, 2010.

Accepted September 22, 2010.

is a multi-step procedure that requires 1) traditional evaluation with an animal model, 2) extraction of pseudo-observations, and 3) estimation of genomic effects for genotyped animals (Van Raden, 2008; Hayes et al., 2009; VanRaden et al., 2009). In step 3, various distributions of SNP effects may be postulated, but for most traits the assumption of equal variance of each marker yields accuracies similar to those obtained with different assumptions of distributions of SNP effects (e.g.; Bayes A; Hayes et al., 2009). Assumption of equal variances is equivalent to BLUP with a genomic relationship matrix (VanRaden, 2008).

The strategy in dairy cattle relies on many sires with increased accuracies. In species with less accuracy of

<sup>&</sup>lt;sup>1</sup>Current address: Newsham Choice Genetics, 701 Crown Industrial Ct., Chesterfield, MO 63005.

<sup>&</sup>lt;sup>2</sup>Corresponding author: cychen9@uga.edu

24 Chen et al.

Table 1. Description of phenotypic records in the complete data set FDS<sup>1</sup> for the 2 lines

	Line 1			Line 2		
$\mathrm{Item}^2$	Male	Female	Total	Male	Female	Total
BW6, 100 g						
No. of records	89,578	94,206	183,784	79,333	84,913	164,246
Mean	26.15	22.94	24.50	25.26	21.92	23.53
SD	3.14	2.43	3.22	3.08	2.29	3.17
$BM, cm^2$						
No. of records	7,163	33,751	40,914	7,033	33,543	40,576
Mean	45.60	42.22	42.81	44.81	40.32	41.09
SD	5.65	5.09	5.35	5.48	4.66	5.10
LS, 1 and 2						
No. of records	89,578	94,206	183,784	79,333	84,913	164,246
Mean	1.26	1.12	1.19	1.24	1.09	1.16
SD	0.44	0.33	0.39	0.42	0.28	0.37
No. of sires with records		_	281	_	_	308
No. of sires without records	_	_	287	_	_	273
No. of dams with records	_	_	2,447	_	_	2,398
No. of dams without records	_		2,151	_	_	2,010
Animals in pedigree <sup>3</sup>	_	_	186,222	_	_	166,529

<sup>&</sup>lt;sup>1</sup>Phenotypic records of ungenotyped and genotyped animals across 3 generations.

genotyped animals, pseudo-observations are harder to obtain, and steps 1 and 2 may be eliminated by performing step 3 directly on phenotypic records (Dekkers et al., 2009; González-Recio et al., 2009). However, it is not clear whether the increase of accuracy due to using the genomic information overcomes the loss of accuracy due to ignoring records on ungenotyped animals. A related issue is selection bias with multiple trait selection (Van Vleck, 1968; Pollak et al., 1984).

Misztal et al. (2009) proposed a single-step BLUP (ssGBLUP) for genomic evaluation where an additive relationship matrix is modified to incorporate the genomic information. The ssGBLUP was successfully applied for final scores of over 6 million Holsteins with accuracy superior to a multi-step procedure (Aguilar et al., 2010). The ssGBLUP is suitable for multiple-trait analyses.

The objective of this study was to compare BLUP, ssGBLUP, and a multi-step procedure on a large set of commercial data in broiler chickens using records on all or genotyped animals.

#### MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because the data were obtained from an existing database.

#### Data

Phenotypic data of broiler chickens from 2 pure lines across 3 generations (**G1**, **G2**, and **G3**) were provided by Cobb-Vantress Inc. (Siloam Springs, AR). Line 1 is a sire line, selected primarily for growth rate, meat

yield, feed conversion, and livability. Reproduction traits are also considered in the selection. Line 2 is a dam line, selected primarily for reproduction than production. Traits analyzed included BW at 6 wk (BW6, 100 g), ultrasound area of breast meat (BM, cm<sup>2</sup>), and leg angle (leg score; LS) coded 1 = acceptable and 2 =not acceptable. Phenotypic records of all animals (full data set; **FDS**) and only genotyped animals (**SUB**) for the 3 generations were analyzed. Genotyped animals in SUB were selected based on more than 20 traits with a different index applied to each line. Descriptions of phenotypic records for FDS are shown in Table 1. A total of 183,784 and 164,246 broilers were available in the initial complete data set for lines 1 and 2, respectively. The BW6 and LS were recorded for all animals, except for BM only 40,914 (line 1) and 40,576 (line 2) broilers were measured. Complete pedigrees were available for all animals. For FDS, numbers of animals in the pedigree (including sires and dams without records) were 186,222 for line 1 and 166,529 for line 2.

Genotypes for 57,636 SNP based on the SNP panel developed by Groenen et al. (2009) were determined in 3,284 (line 1) and 3,098 (line 2) broilers across 3 generations. Descriptions of phenotypic records for genotyped animals in SUB are shown in Table 2. All genotyped animals had records for BW6 and LS, but for BM there was a slightly reduced number of 3,099 and 2,993 animals for the respective lines. For the SUB data set, phenotypic information was limited to those animals genotyped, whereas the pedigree also included the parents of the first generation genotyped. For SUB, numbers of animals in the pedigree (including sires and dams without records) were 4,013 for line 1 and 3,722 for line 2.

<sup>&</sup>lt;sup>2</sup>BW6 = BW at 6 wk; BM = ultrasound measurement of breast meat; LS = leg score.

<sup>&</sup>lt;sup>3</sup>Numbers of animals with records and numbers of sires and dams without records in the pedigree.

Table 2. Description of phenotypic records of data set SUB<sup>1</sup> for the 2 lines

	Line 1				Line 2		
$\mathrm{Item}^2$	Male	Female	Total	Male	Female	Total	
BW6, 100 g	-						
No. of records	670	2,614	3,284	666	2,432	3,098	
Mean	29.25	24.03	25.09	26.77	22.43	23.36	
SD	2.43	1.95	2.94	2.42	1.78	2.63	
$BM, cm^2$							
No. of records	594	2,505	3,099	656	2,337	2,993	
Mean	46.38	42.04	42.87	45.27	39.81	41.00	
SD	5.30	5.16	5.47	5.31	4.37	5.12	
LS, 1 and 2							
No. of records	670	2,614	3,284	666	2,432	3,098	
Mean	1.18	1.04	1.07	1.30	1.08	1.12	
SD	0.39	0.20	0.25	0.46	0.26	0.33	
No. of sires with records	_	_	90	_	_	87	
No. of sires without records	_	_	220	_	_	186	
No. of dams with records	_	_	785	_	_	735	
No. of dams without records	_	_	509	_	_	438	
Animals in pedigree <sup>3</sup>	_	_	4,013		_	3,722	

<sup>&</sup>lt;sup>1</sup>Phenotypic records of genotyped animals across 3 generations.

The 2 data sets, FDS and SUB, were split into training and validation data sets for each line for the genetic prediction. Training data set consisted of records from G1 and G2. Validation data set contained 799 genotyped animals in G3, which also had phenotypic records. Numbers of animals in training and validation data sets for the 2 lines are presented in Table 3.

#### Models and Analyses

The single-trait model was

$$v = Xb + Zu + e$$
.

where y is the vector of observations for BW6, BM, and LS; b is the vector of fixed effects including contemporary group (house-hatch) and sex; **u** is the vector of random additive genetic effects, combining polygenic (breeding values based on pedigree) and genomic (breeding values based on genotypes) breeding values; X and Z are incidence matrices; e is the vector of random residuals. Based on the preliminary analyses, the maternal genetic effects were very small for all 3 traits, whereas the maternal permanent environmental effect was larger for BW6 but not for BM and LS. Therefore, the analysis for BW6 included the vector of random maternal permanent environmental effects (mp), and the corresponding incidence matrix (W) was added in the model. For FDS, levels of maternal permanent environmental effects were 4,598 for line 1 and 4,408 for line 2. For SUB, levels were 1,290 for line 1 and 1,162 for line 2. Contemporary groups were nested within generations. For FDS, there were 98 contemporary groups for each line. For SUB, there were 57 contemporary groups for each line.

For regular BLUP analysis, the (co)variance matrix was assumed to be

$$\operatorname{var} \begin{bmatrix} \mathbf{u} \\ \mathbf{mp} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \sigma_{\mathbf{u}}^2 & 0 & 0 \\ 0 & \mathbf{I}_{\mathbf{m}} \sigma_{\mathbf{mp}}^2 & 0 \\ 0 & 0 & \mathbf{I}_{\mathbf{n}} \sigma_{\mathbf{e}}^2 \end{bmatrix},$$

where **A** is a numerator relationship matrix of dimension corresponding to numbers of animals in the pedigree described above, **I** is the identity matrix of appropriate dimension (**m** or **n**), and  $\sigma_u^2$ ,  $\sigma_{mp}^2$ , and  $\sigma_e^2$  were additive, maternal permanent, and residual variances, respectively. In ssGBLUP, the **A** matrix was replaced by the **H** matrix with the following inverse (Aguilar et al., 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

**Table 3.** Numbers of animals in the training and validation data sets for the 2 lines

Item	Line 1	Line 2
Training data set <sup>1</sup>		
SUB	2,485	2,299
FDS	155,899	134,918
Validation data set <sup>2</sup>	799	799

<sup>&</sup>lt;sup>1</sup>Training data set in SUB consisted of phenotypes for genotyped animals in generations G1 and G2; training data set in FDS consisted of phenotypes for ungenotyped and genotyped animals in generations G1 and G2.

<sup>&</sup>lt;sup>2</sup>BW6 = BW at 6 wk; BM = ultrasound measurement of breast meat; LS = leg score.

<sup>&</sup>lt;sup>3</sup>Numbers of animals with records and numbers of sires and dams without records in the pedigree.

<sup>&</sup>lt;sup>2</sup>Validation data set contained genotyped animals in generation G3, which also had phenotypic records.

26 Chen et al.

**Table 4.** Estimates<sup>1</sup> of variance components and heritability using the complete data set (FDS)<sup>2</sup> with phenotypic BLUP for the 2 lines

Item <sup>3</sup>	$\sigma_{ m u}^2$	$\sigma_{ m mp}^2$	$\sigma_{ m e}^2$	$h^2$
Line 1				
BW6	1.03	0.40	3.69	0.20
$_{\mathrm{BM}}$	4.04	_	9.61	0.30
LS	0.02		0.13	0.11
Line 2				
BW6	0.85	0.32	3.83	0.17
$_{\mathrm{BM}}$	4.34		7.95	0.35
LS	0.01	_	0.12	0.09

 $<sup>^{1}\</sup>sigma_{u}^{2}$ ,  $\sigma_{mp}^{2}$ , and  $\sigma_{e}^{2}$  are additive, maternal permanent, and residual variances, respectively;  $h^{2}$  is the heritability.

where  $\mathbf{H}$  is a modified relationship matrix incorporating genomic information,  $\mathbf{A}_{22}^{-1}$  corresponds to the inverse of numerator relationship matrix for genotyped animals, and  $\mathbf{G}$  is a genomic relationship matrix. Matrix  $\mathbf{G}$  was created as in Aguilar et al. (2010) assuming equal allele frequencies. Using current allele frequencies (results not shown) did not affect rankings.

Estimates of variance components were obtained using single-trait models with the complete and subset data sets including records across 3 generations for each line. Estimates of variance components obtained with FDS were used for genetic evaluations. Genetic evaluations were performed by modified BLUP90IOD (Tsuruta et al., 2001; Misztal et al., 2002; Aguilar et al., 2010) using BLUP (no genomic information) and ssG-BLUP (with genomic information) for FDS and SUB. An extra analysis was done by Bayes A approach (Meuwissen et al., 2001) with SUB only. In addition, Bayes B with a range of  $\pi$  values (the proportion of SNP with null effects) was also compared, but the fit with  $\pi = 0$ , corresponding to Bayes A, fit the data best. Thus, Bayes B comparisons were not included in the results. Model predictive ability,  $r(\hat{u}, u + e)$ , was defined as the correlation between predicted breeding value (û) and the sum of true breeding value (u) and residual (e), using the formula shown by Legarra et al. (2008). Accuracy, correlation between predicted and true breeding values, was calculated as  $r(\hat{u}, u) = r(\hat{u}, u + e) / h$ , where h is the square root of heritability. Estimates of heritability obtained with FDS were used for calculating accuracies.

#### RESULTS AND DISCUSSION

Table 4 summarizes estimates of variance components using FDS with phenotypic BLUP for the 2 lines. Estimates of heritability for line 1 were 0.20, 0.30, and 0.11 for BW6, BM, and LS, respectively. For line 2 the estimates for the same traits were 0.17, 0.35, and 0.09. In general, the estimates for both lines were similar.

Table 5 summarizes estimates of variance components using SUB with phenotypic BLUP. For all traits, the estimates additive variances were somewhat smaller in SUB than those in FDS; whereas the estimates for the residual variances were almost 3 times smaller in SUB for BW6 and LS, the residual variance was larger in SUB for BM. The change in residual variances in SUB resulted in increased heritability for BW6 and decrease heritability for BM across lines, with mixed changes for LS.

The accuracies of prediction based on phenotypes only or on both phenotypes and genomic information for the 2 lines is shown in Table 6. Based on the definition of accuracy and amount of information for phenotypes and genotypes, greater accuracies of predictions for the validation generation should be expected with 1) FDS, 2) greater heritabilities, and 3) the genomic information (Muir, 2007; Hayes et al., 2009; VanRaden et al., 2009). Also, accuracy is a function of heritability and number of phenotypic records being used. Fewer records are required to achieve greater accuracy in the genomic evaluation for traits with greater heritability (Goddard, 2009; Hayes et al., 2009). With similar heritability and the amount of records being used, lines 1 and 2 would be expected to have similar accuracies for some traits. In the simulation study by Neuner et al. (2009), accuracy using a genotyped subset with marker information was similar to that using FDS data without marker information. For the genotyped subset, only 1 QTL effect was simulated with partial pedigree being used in their study. With relative larger amounts of SNP and FDS pedigree being applicable, accuracies using genomic information combined with phenotypes from the genotyped animals (SUB) would be substantially greater than phenotypic BLUP using the full data (FDS) without the genomic information for some traits with increased heritability.

As Table 6 shows, not all the expectations were realized. First, accuracies are less for BM than for BW6 despite its greater heritability. This is likely caused by

**Table 5.** Estimates<sup>1</sup> of variance components and heritability using data set SUB<sup>2</sup> with phenotypic BLUP for the 2 lines

Item <sup>3</sup>	$\sigma_{ m u}^2$	$\sigma_{ m mp}^2$	$\sigma_{ m e}^2$	$h^2$
Line 1				
BW6	0.56	0.34	1.37	0.25
$_{\mathrm{BM}}$	3.10		11.48	0.21
LS	0.005	_	0.05	0.09
Line 2				
BW6	0.64	0.59	1.47	0.24
$_{\mathrm{BM}}$	4.06		9.77	0.29
LS	0.01	_	0.06	0.20

 $<sup>^{1}\</sup>sigma_{u}^{2}$ ,  $\sigma_{mp}^{2}$ , and  $\sigma_{e}^{2}$  are additive, maternal permanent, and residual variances, respectively;  $h^{2}$  is the heritability.

<sup>&</sup>lt;sup>2</sup>Phenotypes for ungenotyped and genotyped animals across generations G1, G2, and G3.

 $<sup>^{3}</sup>$ BW6 = BW at 6 wk; BM = ultrasound measurement of breast meat; LS = leg score.

<sup>&</sup>lt;sup>2</sup>SUB consisted of phenotypes for genotyped animals in generations G1, G2, and G3.

 $<sup>^{3}</sup>$ BW6 = BW at 6 wk; BM = ultrasound measurement of breast meat; LS = leg score.

**Table 6.** Accuracy<sup>1</sup> based on no genomic or genomic information for the 2 lines with methods of BLUP, single-step BLUP (ssGBLUP), and 2-step Bayes A (if available)

Item <sup>2</sup>	No genomic	information		Genomic		
	BLUP		ssGBLUP		Bayes A	
	$\mathrm{SUB}^3$	$\mathrm{FDS}^4$	SUB	FDS	SUB	
Line 1						
BW6	0.46	0.51	0.60	0.61	0.60	
$_{\mathrm{BM}}$	0.30	0.34	0.34	0.40	0.36	
LS	<0	0.28	0.06	0.37	0.09	
Line 2						
BW6	0.39	0.24	0.50	0.44	0.47	
$_{\mathrm{BM}}$	0.27	0.33	0.45	0.51	0.51	
LS	0.24	0.43	0.15	0.73	0.11	

<sup>&</sup>lt;sup>1</sup>Accuracy defined as correlations between predicted and true breeding values. It was calculated as predictive ability divided by the square root of heritability. Estimates of heritability obtained with the full data set (FDS) were used for calculating accuracies.

incomplete data recording. Only about 22% animals in FDS and 90% in SUB had BM records. Also, for line 1, the heritability of BM in SUB was greater than for BW6. Additionally, a large range of accuracies were obtained for LS. A negative accuracy (-0.01) was found with phenotypic BLUP in SUB, whereas a greatest accuracy (0.73) was found with ssGBLUP in FDS despite its decreased heritability compared with that in BW6 and BM. Large accuracy for LS score could be an artifact of its binary nature, as calculation of accuracy assumes a linear trait. However, this should not influence ranking because the same assumption applied to all the methods and data sets.

For traditional BLUP, the use of FDS over SUB improved accuracy for all traits in line 1, whereas for line 2, BW6 decreased; the deterioration occurred regardless of the use of the genomic information. Efforts to refine the model to eliminate this anomaly were unsuccessful. Also, the pedigree is not an obvious problem because the accuracies for BM and LS followed expectation. A possible explanation for the decreased accuracy of BW6 in FDS of line 2 is selection bias resulting from multiple trait selection (Van Vleck, 1968; Pollak et al., 1984). Pollak et al. (1984) demonstrated that preselection caused upward bias for the worst animals and downward bias for the best animals. Before the current 3 generations, the animals were selected using multi-trait BLUP and for more traits than in the current genomic selection program. The current analysis is based on single traits; thus, bias from multi-trait selection has been introduced in the FDS.

There is a question whether the accuracies of the genomic prediction would have been greater with a methodology that accounted for nonequal distribution of SNP marker effects, such as Bayes A or Bayes B (Meuwissen et al., 2001). Table 6 also contains accuracies obtained with Bayes A. The results are similar to

ssGBLUP/SUB, with Bayes A slightly more accurate for BM and LS in line 1 and BM in line 2 but less accurate for BW6 and LS in line 2. Whereas greater performance of Bayes A can be attributed to major genes, its decreased performance is probably due to the increased number of parameters estimated. Also, some differences may be due to the Monte Carlo Markov chain sampling present in the Bayes A procedure.

Comparison of genomic information included vs. excluded, using the same number of phenotypes (i.e., BLUP/SUB vs. ssGBLUP/SUB or Bayes A), shows that inclusion of genomic information always increased accuracy except for LS in line 2, which could be an artifact due to its binary nature or a sampling error. Two important questions remain. First, does inclusion of all pedigree phenotypic information add a substantial increase in accuracy over just using phenotypes of animals genotyped? Comparison of ssGBLUP/SUB with ssGB-LUP/FDS shows that in most cases an improvement is seen and sometimes is dramatic, such as with LS where approximately a 500% relative improvement is seen, whereas for BM approximately only a 25% relative improvement resulted. The large increase in accuracy for LS could be an artifact because LS score had a binary nature but was evaluated as a linear trait. In contrast, for BW6 either none (line 1) or a decline (line 2) was seen. Again, the decline in BW6 with line 2 is assumed to be related to the multi-trait selection bias issue as discussed previously. Second, using all phenotypic information available in the pedigree, how much does the genomic information add? Comparison of BLUP/FDS with ssGBLUP/FDS shows that for all traits in both lines the relative improvement is between 17 and 83%, the greatest relative improvement for the trait with poor heritability, as predicted by Muir (2007).

Regarding a multi-step vs. single step approach to the problem, one can adapt a multi-step method as

<sup>&</sup>lt;sup>2</sup>BW6 = BW at 6 wk; BM = ultrasound measurement of breast meat; LS = leg score.

<sup>&</sup>lt;sup>3</sup>SUB consisted of phenotypes for genotyped animals.

<sup>&</sup>lt;sup>4</sup>FDS consisted of phenotypes for both ungenotyped and genotyped animals.

28 Chen et al.

used in dairy cattle (i.e., extract pseudo observations for genotyped animals based on a regular evaluation and then run genomic selection programs only for genotyped animals). However, such an approach may be complicated and possibly biased if average accuracy is poor and especially when genotyped animals become part of the training population. The other option is to use a single-step approach that can easily provide multiple-trait computations without extra effort. The ssGBLUP as applied here assumed equal distribution of SNP effects. If those distributions are markedly uneven due to large major genes, a genomic relationship matrix may be constructed with distributions obtained externally.

Abnormalities in accuracies found in this study could be due to faults in the model, although every effort was made to identify irregularities in the data and refine the model (e.g., checking the pedigree), using models with different combinations of fixed effects, and using different allele frequencies (0.5 and current). As can be seen in Table 6, problems in accuracy between FDS and SUB were the same regardless of whether the genomic information was included. Therefore, before undertaking expensive genotyping, it would be cost-effective to identify and correct the source of abnormalities such as unaccounted-for selection, specific editing, and additional effects influencing the trait. These efforts would use BLUP with the training and validation populations; such analyses are normally not undertaken with the regular evaluation.

#### Conclusions

Based on results in this study, 3 important conclusions are seen. First, one should always augment the genotyping data set with all phenotypes available in the pedigree and evaluate using an ssGBLUP approach. In this case, one benefits from an increased amount of information from the genotypes and from information obtained from the pedigree. Second, for traits with poor hereditability and especially for preselected traits, the accuracy of the evaluation using a genotyped subset would be poor, with or without the genomic information. In such cases the genomic information can still provide large increases of accuracy, but only if the complete population is used. Third, the most accurate evaluation would involve not only the complete populations but also with multiple-trait models and all traits on which the selection was practiced. Such evaluation is possible with the single-step methodology.

#### LITERATURE CITED

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743-752.
- Dekkers, J. C. M., C. Stricker, R. L. Fernando, D. J. Garrick, S. J. Lamont, N. P. O'Sullivan, J. E. Fulton, J. Arango, P. Settar, A. Kranis, J. McKay, K. A. Watson, A. Koerhuis, and R. Preisinger. 2009. Implementation of genomic selection in egg layer chickens. J. Anim. Sci. 87(E-Suppl. 2):iii. (Abstr.)
- Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica 136:245–257.
- González-Recio, O., D. Gianola, G. J. M. Rosa, K. A. Weigel, and A. Kranis. 2009. Genome-assisted prediction of a quantitative trait measured in parents and progeny: Application to food conversion rate in chickens. Genet. Sel. Evol. 41:3–13.
- Groenen, M. A. M., P. Wahlberg, M. Foglio, H. H. Cheng, H. J. Megens, R. P. M. A. Crooijmans, F. Besnier, M. Lathrop, W. M. Muir, G. K. Wong, I. Gut, and L. Andersson. 2009. A highdensity SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. Genome Res. 19:510–519.
- Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433–443.
- Legarra, A., C. Robert-Granie, E. Manfredi, and J. M. Elsen. 2008.
  Performance of genomic selection in mice. Genetics 180:611–618.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.
- Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J. Dairy Sci. 92:4648–4655.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Proc. 7th World Congr. Genet. Appl. Livest. Prod. Montpellier, France. Commun. No. 28–07.
- Muir, W. M. 2007. Comparison of genomic and traditional BLUPestimated breeding valueaccuracy and selection response under alternative trait and genomic parameters. J. Anim. Breed. Genet. 124:342–355.
- Neuner, S., C. Edel, R. Emmerling, G. Thaller, and K. U. Götz. 2009. Precision of genetic parameters and breeding values estimated in marker assisted BLUP genetic evaluation. Genet. Sel. Evol. 41:26.
- Pollak, E. J., J. van der Werf, and R. L. Quaas. 1984. Selection bias and multiple trait evaluation. J. Dairy Sci. 67:1590–1595.
- Tsuruta, S., I. Misztal, and I. Stranden. 2001. Use of the preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. J. Anim. Sci. 79:1166–1172.
- Van Vleck, L. D. 1968. Selection bias in estimation of the genetic correlation. Biometrics 24:951–962.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

This article cites 16 articles, 4 of which you can access for free at: http://jas.fass.org/cgi/content/full/89/1/23#BIBL References

Citations This article has been cited by 1 HighWire-hosted articles: http://jas.fass.org/cgi/content/full/89/1/23#otherarticles