Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in *Bos taurus*, *Bos indicus*, and composite beef cattle[1]

S. Bolormaa,*[2] J. E. Pryce,* K. Kemper,† K. Savin,* B. J. Hayes,* W. Barendse,‡ Y. Zhang,§ C. M. Reich,* B. A. Mason,* R. J. Bunch,‡ B. E. Harrison,‡ A. Reverter,‡ R. M. Herd,# B. Tier,§ H.-U. Graser,§ and M. E. Goddard*†

*Victorian Department of Primary Industries, Bundoora, VIC 3083, Australia; †School of Land and Environment, University of Melbourne, Parkville, VIC 3010, Australia; ‡CSIRO Animal, Food and Health Sciences, Queensland Bioscience Precinct, St. Lucia, QLD 4067, Australia; §Animal Genetics and Breeding Unit, Armidale, NSW 2351, Australia; and #New South Wales Department of Primary Industries, Beef Industry Centre, Armidale, NSW 2351 Australia

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

ABSTRACT: The aim of this study was to assess the accuracy of genomic predictions for 19 traits including feed efficiency, growth, and carcass and meat quality traits in beef cattle. The 10,181 cattle in our study had real or imputed genotypes for 729,068 SNP although not all cattle were measured for all traits. Animals included *Bos taurus*, Brahman, composite, and crossbred animals. Genomic EBV (GEBV) were calculated using 2 methods of genomic prediction [BayesR and genomic BLUP (GBLUP)] either using a common training dataset for all breeds or using a training dataset comprising only animals of the same breed. Accuracies of GEBV were assessed using 5-fold cross-validation. The accuracy of genomic prediction varied by trait and by method. Traits with a large number of recorded and genotyped animals and with high heritability gave the greatest accuracy of GEBV. Using GBLUP, the average accuracy was 0.27 across traits and breeds, but the accuracies between breeds and between traits varied widely. When the training population was restricted to animals from the same breed as the validation population, GBLUP accuracies declined by an average of 0.04. The greatest decline in accuracy was found for the 4 composite breeds. The BayesR accuracies were greater by an average of 0.03 than GBLUP accuracies, particularly for traits with known genes of moderate to large effect mutations segregating. The accuracies of 0.43 to 0.48 for IGF-I traits were among the greatest in the study. Although accuracies are low compared with those observed in dairy cattle, genomic selection would still be beneficial for traits that are hard to improve by conventional selection, such as tenderness and residual feed intake. BayesR identified many of the same quantitative trait loci as a genomewide association study but appeared to map them more precisely. All traits appear to be highly polygenic with thousands of SNP independently associated with each trait.

Key words: accuracy of genomic estimated breeding value, BayesR, genomic best linear unbiased prediction, genomewide association study, multibreed, single breed

© 2013 American Society of Animal Science. All rights reserved.
J. Anim. Sci. 2013.91:3088–3104
doi:10.2527/jas2012-5827

INTRODUCTION

Genomic selection refers to selection decisions that are based on breeding values predicted using genomewide marker data such as SNP (Meuwissen et al., 2001). In beef cattle genomic predictions are attractive because many traits that affect profitability of beef production, such as feed conversion efficiency, have been hard to select for because they are expensive to measure or are measured only on relatives of breeding bulls (e.g., carcass and meat quality traits). Accurate genomic EBV (**GEBV**) would lead to greater genetic gain for these traits.

The calculation of GEBV depends on a reference population that has been measured for the trait and

[1]The authors acknowledge the funding from the former Cooperative Research Centre for Beef Genetic Technologies in Australia for this study and ARC project DPI093502. We thank to Jim Cook and Christian Girard (Animal Genetics and Breeding Unit, NSW, Australia) for providing the phenotypes.
[2]Corresponding author: bolormaa.sunduimijid@dpi.vic.gov.au
Received September 6, 2012.
Accepted April 3, 2013.

genotyped for the markers. The accuracy of GEBV on selection candidates depends on the size of this reference population and the extent of the linkage disequilibrium (**LD**) between SNP and QTL (Hayes et al., 2009; VanRaden et al., 2009). Assembling large enough reference populations for accurate GEBV prediction is a major challenge, especially where the phenotype is difficult or expensive to measure. A key question in calculating GEBV in beef cattle is whether a breed-specific reference population should be used for each breed requiring GEBV or whether a common reference population should be used for all breeds. Combining breeds into a common reference population increases its size, but LD exists only over shorter distances in a multibreed population and this tends to reduce the accuracy of GEBV and may counter the benefits of a larger reference population. In principle, the accuracy could decline as a result of merging breeds.

A multibreed reference population should be advantageous if the phase of LD between a SNP and a causal mutation or QTL is same in all breeds. This is only to be expected if chromosome segments containing the SNP and the QTL in different breeds have descended from a common ancestor without recombination. This situation will occur if the SNP and the QTL are very closely linked. Linkage disequilibrium phase is poorly conserved between *Bos taurus* breeds at distances typical of the inter-SNP distance in the 50K SNP chip (De Roos et al., 2008) but is well conserved at distances typical of the Illumina high density (**HD**) 800K SNP chip (www.illumina.com/applications/agriculture.ilmn). Therefore in this paper we use the Illumina HD SNP chip.

The statistical methods for genomic selection, or prediction, were introduced by Meuwissen et al. (2001). The term "BLUP" assumes that all SNP effects are drawn from the same normal distribution and hence all have small effects. The term "BayesB" assumes that some SNP have no effect on the trait whereas the effects of other SNP are assumed to follow a *t*-distribution. This allows some SNP to have a large effect. Here we compare BLUP with a new method, called BayesR (Erbe et al. 2012), which has similar properties to BayesB but is faster to compute than BayesB.

Mapping QTL using a genomewide association study (**GWAS**) can be done with the same data that are used as a reference population for genomic prediction. In fact, genomic prediction and GWAS are very similar; both estimate the effect of each SNP on the trait but the typical procedure for GWAS estimates the effect of 1 SNP at a time whereas genomic prediction estimates all SNP effects simultaneously.

In this paper we report GWAS and genomic prediction for 19 growth, feed efficiency [measured as residual feed intake (**RFI**)], and carcass and meat quality traits of beef cattle; we compare the accuracy of genomic prediction using BLUP and BayesR using either a common reference population or breed-specific reference populations.

MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because no new animals were handled in this experiment. The experiment was performed on trait records and DNA samples that had been collected previously.

Cattle Populations

The cattle were sourced from 9 different populations of 3 breed types; they include 4 different *Bos taurus* (**Bt**) breeds [1,743 Angus (**AN**), 223 Murray Grey (**MG**), 717 Shorthorn (**SS**), and 613 Hereford (**HH**)], 1 *Bos indicus* (**Bi**) breed [3,384 Brahman (**BB**) cattle], 3 composite (**Bt×Bi**) breeds [550 Belmont Red (**BR**), 598 Santa Gertrudis (**SG**), and 1,826 Tropical Composites (**TC**)], and 1 recent Brahman cross (**BX**) population (527 F_1 crosses of BB with Limousin, Charolais, AN, SS, and HH; Johnston et al., 2003; Barwick et al., 2009; Bolormaa et al., 2011).

Genotype Data

The SNP marker data used in this study was obtained from 5 different SNP panels: the Illumina HD Bovine SNP chip, comprising 777,963 (**800K**) SNP markers, the BovineSNP50K version 1 and version 2 BeadChip (Illumina, San Diego, CA), comprising 54,001 and 54,609 (**50K**) SNP, respectively, the IlluminaSNP7K panel, comprising 6,909 (**7K**) SNP, and the ParalleleSNP10K chip (Affymetrix, Santa Clara, CA), comprising 11,932 (**10K**) SNP. All SNP were mapped to the UMD 3.1 build of the bovine genome sequence assembled by the Center for Bioinformatics and Computational Biology at University of Maryland (www.cbcb.umd.edu/research/bos_taurus_assembly.shtml).

Stringent quality control procedures were applied to the SNP data of each platform. The SNP were excluded if the call rate per SNP [this is the proportion of SNP genotypes that have a **GC** (Illumina GenCall) score above 0.6] was less than 90% and they had duplicate map positions (2 SNP with the same position but with different names) or an extreme departure from Hardy-Weinberg equilibrium (e.g., SNP in autosomal chromosomes with both homozygous genotypes observed but no heterozygotes). Furthermore, if the call rate per individual was less than 90%, those animals were removed from the SNP data. The SNP data were

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Table 1. The number of genotypes used in this study by each breed within each platform[1]

Breed ID[2]	Breed type[3]	HD	50K	7K	10K	Breed definition
AN	Bt	133	1,016	282	312	Angus
BB	Bi	302	2,702	380		Brahman
BR	Bt×Bi		265	258		Belmont Red
BX	Bt×Bi		455	72		recent Brahman crosses
HH	Bt	30	437	146		Hereford
MG	Bt	19	97	107		Murray Gray
SG	Bt×Bi	45	213	340		Santa Gertrudis
SS	Bt	30	594	93		Shorthorn
TC	Bt×Bi	320	1,073	460		Tropical Composites

[1]HD = high density; 50K, 7K, and 10 K = approximately 50,000, 7,000, and 10,000 SNP markers (respectively) from commercial SNP chips

[2]ID = abbreviation used for breeds.

[3]Bt = *Bos taurus*; Bi = *Bos indicus*; Bt×Bi = composites.

Table 2. The accuracy of imputation (R^2) obtained from Beagle of the genotyped data[1]

	7K data		50K data		10K data	
Imputation/ breed[2]	7K to 50K	50K to 800K	50K to 50K	50K to 800K	3K to 50K	50K to 800K
ANMG[3]	0.89	0.94	0.98	0.95	0.88	0.96
BB	0.78	0.90	0.96	0.90		
BR	0.80	0.92	0.98	0.93		
BX			0.95	0.85		
HH	0.75	0.92	0.97	0.90		
SG	0.75	0.93	0.94	0.93		
SS	0.87	0.92	0.96	0.91		
TC	0.76	0.93	0.96	0.95		
Mean	0.80	0.92	0.97	0.90	0.88	0.96

[1]50K, 7K, 3K and 800K = approximately 50,000, 7,000, 3,000, and 800,000 SNP markers (, respectively) from commercial SNP chips

[2]Angus (AN), Brahman (BB), Belmont Red (BR), Hereford (HH), recent Brahman crosses (BX), Murray Grey (MG), Santa Gertrudis (SG), Shorthorn (SS), and Tropical Composites (TC).

[3]ANMG = genotypes of AN and MG animals were imputed together.

edited within breed group and within each platform and were subsequently combined.

After all the tests were applied, 729,068 SNP of the HD SNP chip were retained on 1,698 animals and the missing genotypes were filled using the BEAGLE program (Browning and Browning, 2011). The genotypes for each SNP were encoded in the top/top Illumina A/B format and then genotypes were reduced to 0, 1, and 2 copies of the B allele. The imputations of the 7K, 10K, and 50K SNP genotype data to 729,068 SNP were performed in 2 sequential stages: from 7K or 10K or 50K data to a common 50K data and then from the common 50K data to 800K data. Imputation was done within breed and using 30 iterations of BEAGLE. The HD genotypes of each breed type (212 Bt and 302 Bi) were used as a reference set to impute the 50K genotypes of each pure breed within the corresponding breed type. For the 4 composite breeds, all the HD genotypes were used as a reference set to impute the 50K genotypes of each breed. The number of genotypes for each platform is given in Table 1 and the mean R^2 values for the accuracy of imputation provided by BEAGLE are in Table 2. After imputation, an additional quality control step was applied based on comparing allele frequencies between SNP platforms to detect SNP with very different allele frequencies indicating incorrect conversion between platforms. In total, 10,181 animals that have a record for at least 1 trait and also had SNP genotypes were used in this study.

Population Structure

The relationships between the breeds used in this study were investigated in 4 different ways. This was done to ensure that the genomic relationship matrix (**GRM**), subsequently used for the genomic BLUP (**GBLUP**) analysis, was representative of the expected relationships within and across breeds. The GRM was constructed using genotypes of all 10,181 animals from different breeds by the method of Yang et al. (2010). This method standardizes the genotype calls (x_{ij}) that have mean = 0 and variance = 1 by the transformation

$$z = (x_{ij} - 2p_i)/[2p_i(1 - p_i)]^{1/2},$$

in which x_{ij} are genotypes of ith SNP of the jth individual scored as 0 or 2 for homozygotes and 1 for heterozygotes and p_i is allele frequency of ith SNP. The first method uses the GRM directly, in which the GRM estimates the relationship between animals relative to a base in which the allele frequency is the average allele frequency over the whole dataset. Thus, the average relationship between animals within a breed is twice the inbreeding (***F***) value of the breed since this base and the average relationship between animals from 2 different breeds is twice the inbreeding since the 2 breeds diverged relative to this base. The second method uses the first 2 principal components from the GRM to show the degree of differentiations between the 10,181 genotypes of the 9 breeds. The third method calculates pairwise F_{ST} coefficients, which are a measure of the genetic differentiation between breeds, using the 729,068 SNP markers by the formula of Hedrick (2005). These pairwise F_{ST} values estimate the average inbreeding in the two breeds relative to a base consisting of an F_2 cross. The F_{ST} values were converted to estimates of inbreeding (F) since the two breeds diverged from an ancestral population using this formula: $F = 2F_{ST}/(1 + F_{ST})$. These inbreeding values were used to draw a tree representing the divergence of the breeds since they shared a common ancestor. Finally, the inbreeding F from the GRM for each breed

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Table 3. Number of records, mean, SD, and heritability estimate (h^2) of each trait for the genotyped animals and their 5-generation ancestors

	No. records for the genotyped animals[1]										Animals with full records				
Trait[2]	AN	BB	BR	BX	HH	MG	SG	SS	TC	All	All	h^2	Mean	SD	Trait definition[3]
RFI	1,091	675	260	245	181	75	219	547	733	4,026	4,837	0.36	–1.5	2.1	Residual feed intake, kg
DFI[4]	639	675	260	245	181	75	219	89	733	3,116	3,732	0.47	12.1	2.2	Daily feed intake over RFI test period, kg
ADG[5]	639	122	260	245	181	75	219	89	106	1,936	2,245	0.38	1.4	0.4	ADG over RFI test period, kg
MMIDWT[6]	639	122	260		181	75	219	89		1,585	1,851	0.56	89.9	14.5	Metabolic midtest weight in the RFI test period, $kg^{0.73}$
LLPF	870	1,187	514	508	560	212	558	235	714	5,358	10,327	0.25	4.6	1.0	Peak force measured in longissimus lumborum muscle, kg
SP8	1,617	693	542		548	188	563	628		4,779	8,465	0.44	11.3	4.8	Exit scanned P8 fat depth, mm
SRIB	1,617	693	542		548	188	563	628		4,779	8,468	0.42	7.7	3.7	Exit scanned rib fat, mm
SEMA	1,507	668	521		537	188	528	590		4,539	8,112	0.17	66.9	11.0	Exit scanned eye muscle area, cm^2
CRBY	681	344	308	93	292	197	419	305	45	2,684	3,639	0.47	67.0	3.4	Carcass retail beef yield, %
CIMF	929	1,159	545	526	608	223	595	517	722	5,824	11,200	0.4	3.4	1.9	Percent intramuscular fat measured in Longissimus lumborum muscle, %
CMARB	463	1,187	446	329	459	105	431	109	699	4,228	9,018	0.23	0.7	0.8	Ausmeat marble score as assessed by MSA grader, score
CWT	968	1,281	550	527	613	223	598	518	732	6,010	11,819	0.52	278.0	56.3	Carcass weight, kg
PW_lwt	1,650	3,346	549	527	568	192	592	641	1,819	9,884	16,079	0.45	230.9	53.4	BW measured postweaning, kg
SF_lwt	1,698	1,248	549	525	569	205	592	704	733	6,823	12,739	0.43	354.6	72.3	BW measured at feedlot entry, kg
PW_hip	676	2,532	478		474	170	380	208	1,441	6,359	10,515	0.53	119.0	7.9	Hip height measured postweaning, cm
SF_hip	236	616	140		204	25	136	53	444	1,854	4,494	0.46	130.0	8.0	Hip height measured at feedlot entry, cm
HUMP		508							624	1,132	2,099	0.29	140.2	37.0	Hump height as assessed by MSA grader, mm
PWIGF	152	337						137	292	918	1,678	0.25	262.2	147.4	IGF-I concentration measured post weaning, ng/mL
EIGF		537							566	1,103	2,058	0.24	507.1	178.8	IGF-I concentration measured at feedlot entry, ng/mL

[1]AN = Angus; BB = Brahman; BR = Belmont Red; BX = recent Brahman crosses; HH = Hereford; MG = Murray Grey; SG = Santa Gertrudis; SS = Shorthorn; TC = Tropical Composites.

[2]Abbreviation used for traits.

[3]P8 = a point located at the intersection of a line drawn anterior to the Tuber ischii and another drawn ventrally from the spinus process of the third sacral vertebra; MSA = Meat standards Australia.

[4,5,6]Based on number of records that was received from different research groups.

were compared with the SNP heterozygosity. The mean observed heterozygosity ($\boldsymbol{HE_o}$) per breed was calculated as the average number of heterozygous SNP for each animal from a breed and also by comparing this value to the heterozygosity assuming Hardy-Weinberg equilibrium [$\boldsymbol{HE_{oHW}}$; i.e., the average of $2p_i(1 - p_i)$ across all SNP, in which p_i is an allele frequency of SNP i in a breed]. The heterozygosity predicted from the GRM ($\boldsymbol{HE_{pr}}$) is $0.38(1 - F)$, in which F is the inbreeding within each breed relative to the base population and 0.38 is the heterozygosity in the base population as calculated by

$$\left[\sum_j 2p_j\left(1-p_j\right)/N\right],$$

in which p_j is the allele frequency of jth SNP across all breeds and N is number of total SNP.

Phenotype Data

Phenotypes for 19 different traits were collated from 5 different sources. The trait definitions and number of

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

records for each trait (and per breed) are given in Table 3. The data sources included the Beef Co-operative Research Centre Phase I (**CRCI**), Phase II (**CRCII**), and Phase III (**CRCIII**), the Trangie selection lines, and the Durham Shorthorn group. Not all cattle were measured for all traits. For instance, cows that were recorded for lifetime reproduction were not measured for meat quality. The CRCI dataset contained 4,526 cattle that were measured for all the traits listed in Table 3 excluding hump height (**HUMP**) and IGF-I traits, using standard procedures described by Upton et al. (2001), Johnston et al. (2003), Reverter et al. (2003), and Robinson and Oddy (2004). The CRCII dataset consisted of 2,084 cows with BW and height data and 1,180 steers that were measured for the majority of traits listed on Table 3. A complete description of the design, methods, and analyses of carcass and meat quality assessments is given by Wolcott et al. (2009). The CRCIII dataset contained 1,111 bulls that were measured for BW and height data. Two BW were analyzed: one (**PW_lwt**) was taken shortly after weaning and the other (**SF_lwt**) was taken at feedlot entry for steers or at the end of their first wet season for heifers (full details of the protocols used for weighing the animals are given by Barwick et al., 2009).

The fourth data source was AN cattle (355 selection line cattle and 452 of their progeny) that were from the divergent RFI selection lines based at the Trangie Agricultural Research Centre, New South Wales, Australia (Arthur et al., 2001; Bolormaa et al., 2011). The fifth source was 473 Durham Shorthorn steers from the progeny test program conducted by Shorthorn Beef between 2000 and 2010 (for complete detail see www.durhamresearch.com.au/; Wolcott et al., 2010). Performance data for the Durham Shorthorns was recorded on all groups of calves from birth to slaughter including carcass traits and meat quality, and for some groups of animals feed intake and ADG was recorded at the Beef CRC Tullimba research feedlot in New South Wales, Australia (Johnston and Graser, 2010).

Statistical Analysis

Mixed models fitting fixed and random effects simultaneously were used for estimating heritabilities, associations with SNP, and genomic prediction using the GBLUP method. Variances of random effects were estimated in each case by REML. All the models used to analyze the traits consistently included dataset, breed, cohort, and sex as fixed effects (Table 3). Other fixed effects fitted to the traits varied by trait. The fixed effects were fitted as nested within a dataset. For carcass weight (**CWT**), postweaning weight (PW_lwt), feedlot entry IGF-I (**EIGF**) and HUMP, a maternal permanent environmental effect was fitted, as preliminary analyses (conducted at Animal Genetics and Breeding Unit, Armidale) revealed that it accounted for part of the variation. For the other traits, maternal effect did not account for a large proportion of the variation and was therefore not included. Further details of the models used in the analysis are reported by Johnston (2001), Johnston et al. (2003), Reverter et al. (2003), Robinson and Oddy (2004), Barwick et al. (2009), Wolcott et al. (2009), and Bolormaa et al. (2011).

The estimates of heritability were calculated based on all animals with phenotype data (even if they lacked SNP genotypes) and their 5-generation ancestors using mixed model: trait = mean + fixed effects + animal + error, with animal and error fitted as random effects, and additionally, dam effect was fitted as random effect for traits including CWT, PW_lwt, EIGF, and HUMP (Table 3). Raw means, SD, and heritability estimates are given in Table 3.

Genomewide Association Studies

The association between each SNP and each of the traits was assessed by a regression analysis using the ASReml software (Gilmour et al., 2009). The model used was the same as for estimating heritability, but SNP (SNP at *i*th chromosomal position) was additionally fitted as a covariate (Bolormaa et al., 2011). The GWAS were conducted 4 times for each trait, first using all available data and then within only Bt, Bi, or Bt×Bi animals. Using Bolormaa et al. (2011), the false discovery rate was calculated as

$$P(1- A/T)/[(A/T)(1-P)],$$

in which P is the P-value tested (e.g., 0.0001), A is the number of SNP that were significant at the P-value tested, and T is the total number of SNP tested (i.e. 729,068).

Genomic Prediction Methods

Genomic BLUP. Genomic EBV were calculated based on

$$\mathbf{y} = \mu + \mathbf{Xb} + \mathbf{g} + \mathbf{e}$$

in which **y** is the vector of observed phenotypic values of the animals, μ is the overall mean, **b** is vector of fixed effects, **X** is design matrix relating observations to the corresponding fixed effects, **e** is the vector of random errors, and **g** is a vector of breeding value with var(g) = $G\sigma_g^2$, in which σ_g^2 is genetic variance and G is the GRM. The criterion for a SNP to be included in the GRM was that the minor allele frequency had to be greater than 0.005. In the analysis of each trait, only the elements of the GRM corresponding to animals with phenotypes were used to reduce the computing time. For all GBLUP analyses,

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

the validation animals were included in the GRM but had unknown phenotypes in the calculation of GEBV. The estimates of GEBV were performed using ASReml software (Gilmour et al., 2009). In a second analysis of GBLUP within breed, the elements of the GRM between animals from different breeds were set to 0. This means that only animals of the same breed contributed to the predictions of GEBV of the validation animals.

BayesR. Another possible assumption about SNP effects is that many SNP have no effect on the trait because they are not in LD with any of the mutations that explain the variation in the traits. BayesR (Erbe et al., 2012) is an extension of BayesCπ (Habier et al. 2011), with the difference that SNP effects are assumed to come from a mixture of normal distributions. BayesCπ assumes that a proportion of SNP (π) have no effect while the rest (1 – π) have effects drawn from a normal distribution. The variance of each component of the mixture was fixed (0 or 0.01 or 0.1 or 1% of the genetic variance) but the number of SNP belonging to each component of the mixture was assumed to come from a multinomial distribution with proportions p_i (i = 1, 2, 3 or 4) in which the p_i are drawn from a Dirichlet distribution (a multivariate generalization of a beta distribution) with pseudo-counts of 1 for each component of the mixture. Thus, the prior assumed that the 4 components of the mixture are equally probable but with minimal prior knowledge of these probabilities. Gibbs sampling was used to sample from the posterior distributions of the parameters including SNP effects with 5,000 iterations of burn-in followed by 100,000 iterations. As the BayesR software used in this study does not allow a full model to be fitted, residuals were calculated by adjusting the phenotypes for the fixed effects (residuals = raw phenotype – mean – fixed effects) using ASReml (Gilmour et al., 2009). These residuals were then used as phenotypes in the analysis. The BayesR analysis fits the effect of the SNP and a residual polygenic effect. Using the estimated SNP effects, a vector of GEBV was calculated for the animals in the validation set GEBV $= \mathbf{Z}\hat{\mathbf{s}}$, in which $\mathbf{Z}$ is a matrix of the standardized genotypes of the animals, and $\hat{\mathbf{s}}$ is the vector of SNP effects predicted by BayesR. The SNP effects from each genomic prediction method were used to calculate GEBV for animals that were not part of the reference dataset (i.e., that were in validation populations; see below).

Validation Populations

For BayesR and GBLUP, the validation populations were same. Five-fold cross-validation was used by dividing the population (for each trait) into 5 parts of equal size such that each had a mixture of breeds and all of the progeny from randomly selected sires. Thus, the analysis was performed 5 times using each division of the data in turn as a validation group and the other 4 divisions as the reference population. The data were split into 5 sets by allocating all of the offspring of randomly selected sires to 1 of the 5 datasets. In this way no animal used for validation had paternal half sibs in the reference population.

The main aim of the study was to determine how well the GEBV predict the true breeding values (**TBV**) of individual animals. If the TBV of individuals were known, the accuracy of the GEBV would be the correlation between the GEBV and the TBV. In practice, the TBV are unknown, and the only data available are phenotypes, which are made up of the effect of the TBV and the environmental effect. Given this limitation, the accuracy of the GEBV was derived as follows.

For each validation population, the accuracy of genomic prediction was calculated as a correlation between GEBV and corrected phenotype within each breed. The correlation between the GEBV and the corrected phenotypes was divided by the square root of h^2, estimated by using the pedigree of the all animals recorded and their 5-generation ancestors. Thus, we report accuracy as the estimated correlation between the GEBV and the TBV. Accuracies are reported only when the number of records per breed is more than 200. When combining accuracies across breeds, the overall accuracy was the mean accuracy within breeds weighted by number of the records in each breed.

Comparisons of SNP Effects among Three Methods

The effects of each SNP on each trait were estimated by the GWAS and by the BayesR analysis and are implicit in the GBLUP analysis. The SNP effects of GBLUP were estimated using this formula: $\hat{\mathbf{s}} = \mathbf{Z}'G^{-1}\hat{g}\left(\sigma_s^2/\sigma_g^2\right)$, in which G is the GRM, $\hat{g}$ is the GEBV for each animal from the entire population, σ_s^2/σ_g^2 is the ratio of the additive genetic variance explained by each SNP (σ_a^2) to the genetic variance explained by all SNP (σ_g^2), and $\mathbf{Z}$ is the matrix of standardized genotypes. It is assumed that all SNP contribute equally to the genetic variance so σ_s^2/σ_g^2 was approximated as $1/N$ in which N is number of total SNP (i.e., 729,068).

RESULTS

Population Structure

Figure 1 shows the first 2 principal components of the GRM. The first principal component, represented on the horizontal axis, captures 87.8% of the total variation, clearly distinguishing the taurine and indicine breeds as well as their crossbreeds. The BB and taurine groups are

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

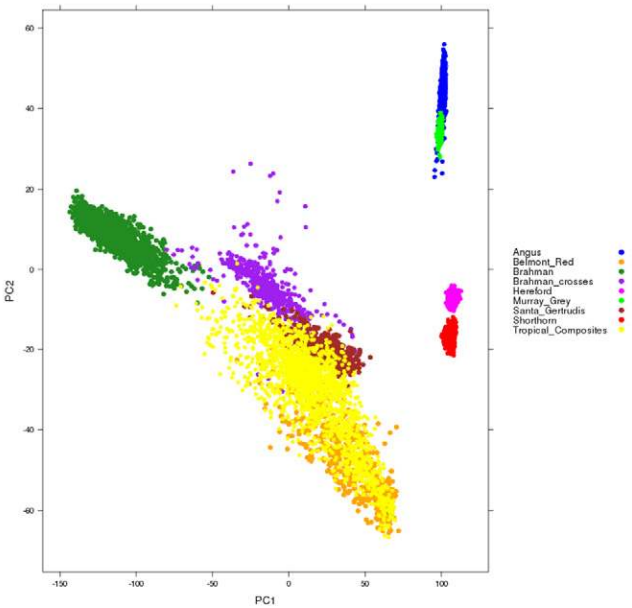

Figure 1. Principal component (PC) decompositions of genomic relationship matrix constructed from genotypes at 729,068 SNP markers for 9 different breeds. The first principal component (PC1) lies on the horizontal axis and the second principal component (PC2) on the vertical axis. See online version for figure in color.

at opposite extremes of this component and the crossbreed groups (BR and SG) appear in the middle. The second component explains 3.7% of the total variation and clearly distinguishes the taurine breeds (HH, SS, and AN). The MG overlaps the AN group as expected. Tropical composites and BR are separated from the taurine breeds in this dimension; perhaps because they have African Bt breeds in their ancestry (e.g., Africander).

The F_{ST} values between the breeds are shown in Table 4. They have been converted to estimates of the inbreeding because a common ancestor and used to draw a tree representing the divergence of the breeds (Fig. 2). Thus, the SNP estimate the inbreeding in BB because they shared a common ancestor with the taurine breeds to be 0.32. The composite breeds are not part of the tree because they have not evolved independently but resulted from crosses of BB and Bt breeds.

The GRM can also be used to estimate the inbreeding within breeds and the relationships among breeds. However the GRM uses a different base to that used to draw Fig. 2. The base for the GRM is a population with the allele frequency equal to the allele frequency over the 10,181 animals in the dataset. The average relationship between animals within a breed estimates twice the F of the breed since the base. These average F values are shown in Table 4. Similarly, the average relationship between animals of 2 different breeds estimates twice the inbreeding of an F_1 cross between the breeds relative to the base (Table 4). These values can be negative; for instance, the average relationship between AN and BB is –0.18 (Table 4). This indicates that a BB×AA F_1 is less inbred than a population with the average allele

Table 4. Pairwise F_{ST} (a measure of the genetic differentiation between breeds; upper off-diagonal elements) and inbreeding (F; lower off-diagonal and diagonal elements) estimated from the genomic relationship matrix[1]

Breed	AN	BB	BR	BX	HH	MG	SG	SS	TC
AN	0.22	0.19	0.07	0.08	0.07	0.03	0.08	0.06	0.07
BB	–0.18	0.21	0.10	0.05	0.21	0.20	0.09	0.20	0.07
BR	0.01	–0.06	0.08	0.03	0.07	0.08	0.04	0.07	0.01
BX	–0.03	0.02	–0.01	0.03	0.09	0.09	0.03	0.09	0.01
HH	0.12	–0.19	0.05	–0.01	0.34	0.08	0.09	0.08	0.08
MG	0.19	–0.17	0.02	–0.02	0.13	0.27	0.09	0.07	0.08
SG	0.00	–0.03	0.00	–0.01	0.01	0.00	0.09	0.07	0.02
SS	0.12	–0.18	0.03	–0.02	0.12	0.12	0.05	0.26	0.07
TC	–0.01	–0.02	0.03	0.00	0.01	0.00	0.00	0.02	0.02

[1]AN = Angus; BB = Brahman; BR = Belmont Red; BX = recent Brahman crosses; HH = Hereford; MG = Murray Grey; SG = Santa Gertrudis; SS = Shorthorn; TC = Tropical Composites.

frequency over the whole dataset. The inbreeding of an F_1 cross is an estimate of the inbreeding in the ancestral population from which the 2 breeds diverged. Therefore, the average relationships between breeds given by the GRM (Table 4) can be compared with the F_{ST} values also given in Table 4. For instance, the average inbreeding of AN and BB relative to their common ancestor is estimated to be (0.22 + 0.21)/2 – –0.18 = 0.395 from the GRM relationships and 0.32 [= 2 × 0.19/(1 + 0.19)] from the F_{ST} values in Table 4. All other breed pairs show a similar good agreement between these 2 estimates of inbreeding because a common ancestor.

Observed heterozygosity for each breed is shown in Table 5. There is good agreement between the observed heterozygosity and that expected under Hardy-Weinberg equilibrium except for the BX. This result is expected for the BX population because they are F_1 crosses and therefore their genotype frequencies are not always in Hardy-Weinberg equilibrium. Also, Table 5 shows predicted heterozygosity based on the inbreeding values implied by the GRM and the heterozygosity in the base population (0.38). Table 5 shows a good agreement of predicted and observed heterozygosity. This implies that the base population used for the GRM does not distort the relationships among the breeds too badly. However, there are some anomalies such as the high observed heterozygosity in HH. This may be due to the use of HH sequence data to discover the SNP on the SNP chip. Similarly the heterozygosity of BB would be greater than that of the taurine breeds if unbiased postweaning were used but the SNP on the chip were predominantly discovered in Bt data.

Genomewide Association Studies

Genomewide association studies, in which each SNP is tested separately for an association with the trait, were performed for all animals and in the 3 breed

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

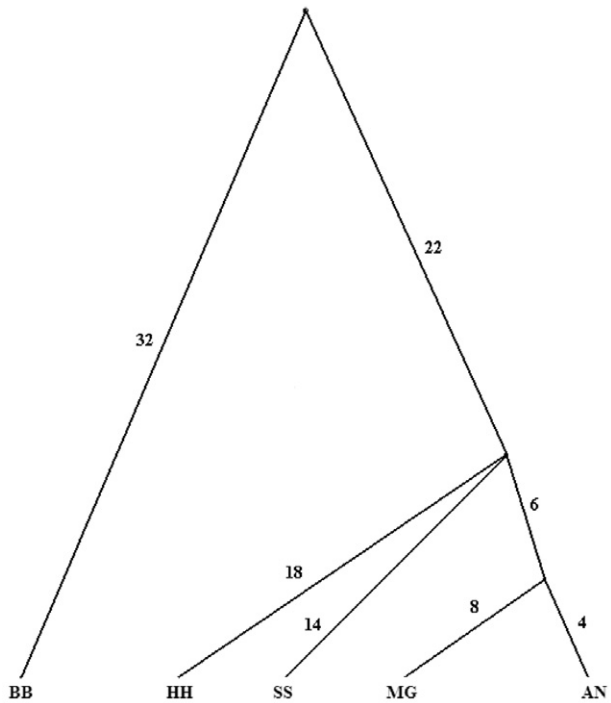

Figure 2. Evolutionary tree showing relationships between breeds based on estimates of inbreeding because breeds shared a common ancestor derived from F_{ST} statistics. The numbers beside each branch represent the inbreeding (inbreeding F in percent) along that branch. BB = Brahman; HH = Hereford; SS = Shorthorn; MG = Murray Grey; AN = Angus.

groups (Bt, Bi, and Bt×Bi; Table 6). For instance, 489 SNP were significant ($P < 0.0001$) for RFI in the joint analysis of all breed types. Because 729,068 SNP were tested, this corresponds to a false discovery rate (**FDR**) of 14% (Table 6). Table 7 shows the most highly significant ($P < 10^{-8}$) SNP. More than 1 SNP was found within narrow regions on chromosomes 5, 7, 14, and 29. In a number of cases those SNP have associations with more than 1 trait.

An alternative presentation of the results to the calculation of FDR is a quantile-quantile (**Q-Q**) plot shown for RFI in Fig. 3. The Q-Q line when all breed types are combined deviates most from expectations indicating that the FDR is lowest in the combined analysis. In the case of RFI, the FDR is lower in the analysis of the Bt data than in the analysis of the composite cattle. For the majority of the traits, FDR were less when all data were analyzed jointly than when separate analyses were performed within Bt, BB, and composite animals (Table 6). However, there was no consistent pattern for FDR when comparing the analyses of Bt, BB, and composite cattle.

Estimates of FDR varied between traits. They were low for BW, peak force (**LLPF**), and IGF-I traits, ranging from 2 to 9% (Table 6). A moderate FDR (12 to 30%) was found for the remaining of traits excluding scanned eye muscle area (**SEMA**). In Fig. 4A the number of significant SNP for each trait is plotted against the number of records for that trait (**T**) multiplied by the h^2 (***Th²***). On average,

Table 5. Average inbreeding estimated from genomic relationships and observed and predicted heterozygosities for each breed

Breed[1]	No. Animals	Diag[2]	Off-diag[2]	HE_{pr}[3]	HE_o[4]	HE_{oHW}[5]
AA	1,787	1.23	0.44	0.30	0.27	0.28
BB	3,471	1.24	0.43	0.30	0.26	0.27
BR	572	1.10	0.15	0.35	0.34	0.35
BX	527	0.94	0.05	0.37	0.41	0.37
HH	648	1.34	0.68	0.25	0.31	0.31
MG	223	1.28	0.55	0.28	0.26	0.26
SG	640	1.10	0.17	0.35	0.33	0.34
SS	742	1.28	0.52	0.28	0.25	0.26
TC	1,881	1.04	0.04	0.37	0.36	0.36

[1]AN = Angus; BB = Brahman; BR = Belmont Red; BX = recent Brahman crosses; HH = Hereford; MG = Murray Grey; SG = Santa Gertrudis; SS = Shorthorn; Tropical Composites

[2]Diag and Off-diag = average within breed diagonal and off-diagonal elements of genomic relationship matrix (GRM)

[3]HE_{pr} = heterozygosity predicted from the GRM

[4]HE_o = observed heterozygosity

[5]HE_{oHW} = heterozygosity when SNP are assumed to be in Hardy-Weinberg equilibrium.

Table 6. Number of SNP and their false discovery rates (FDR; %) at $P < 0.0001$ for single SNP regressions using 729,068 SNP regressed on phenotypes within each breed line and for all animals together[1]

	Number of SNP				FDR (%)			
Item[2]	Bt	Bi	Bt×Bi	All	Bt	Bi	Bt×Bi	All
RFI	508	148	276	489	13.6	46.8	25.1	14.2
DFI	368	118	178	303	18.8	58.7	38.9	22.8
ADG	72	286	324	311	96.2	24.2	21.4	22.3
MMIDWT	189	107	140	239	36.6	64.7	49.5	29.0
LLPF	254	149	692	2,012	27.3	46.5	10.0	3.4
SP8	281	232	304	664	24.6	29.8	22.8	10.4
SRIB	249	221	143	233	27.8	31.3	48.4	29.7
SEMA	154	193	165	122	45.0	35.9	42.0	56.8
CRBY	203	129	162	351	34.1	53.7	42.7	19.7
CIMF	213	396	434	1,189	32.5	17.5	15.9	5.8
CMARB	97	261	217	255	71.4	26.5	31.9	27.2
CWT	311	1,648	761	4,541	22.3	4.2	9.1	1.5
PW_lwt	206	1,028	1,348	1,822	33.6	6.7	5.1	3.8
SF_lwt	341	1,547	502	1,279	20.3	4.5	13.8	5.4
PW_hip	266	1,644	2,205	2,913	26.0	4.2	3.1	2.4
SF_hip	281	238	134	365	24.6	29.1	51.7	19.0
HUMP		171	271	328		40.5	25.5	21.1
PWIGF	79	1,064	124	609	87.7	6.5	55.8	11.4
EIGF		826	195	768		8.4	35.5	9.0
Mean	240	548	451	989	37.8	28.4	28.9	16.6

[1]Bt = *Bos taurus*; Bi = *Bos indicus*; Bt×Bi = composites.

[2]RFI = residual feed intake; DFI = daily feed intake; MMIDWT = metabolic mid-weight; LLPF = peak force; SP8 = scanned P8 fat depth; SRIB = scanned rib fat; SEMA = scanned eye muscle area; CRBY = retail beef yield; CIMF = intramuscular fat; CMARB = marble score; CWT = carcass weight; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; PW_hip = postweaning hip height; SF_hip = feedlot entry hip height; HUMP = hump height; PWIGF = postweaning IGF-I; EIGF = feedlot entry IGF-I.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Table 7. Highly significant SNP ($P < 10^{-8}$) across 19 traits studied[1]

SNP name	BTA	Position (bp)	Traits[2]	No. SNP[3]
BovineHD0100014190	1	50,495,850	CWT	1
BovineHD0500035405	5	42,641,456–51,093,247	PW_hip	10
BovineHD0500013789	5	47,731,856	PW_lwt, SF_lwt	1
BovineHD0500013895	5	48,069,099	PW_lwt	1
BovineHD0500014025	5	48,623,407	HUMP	1
BovineHD0500030518	5	106,417,996–110,712,255	PW_hip	3
BovineHD0600010840	6	39,431,268	PW_lwt	1
BovineHD0600010868	6	39,591,153	PW_hip	1
BovineHD0600010976	6	40,093,712	PW_lwt, PW_hip, SF_hip	1
BovineHD0600011539	6	42,411,772	PW_hip	1
BovineHD0600012396	6	4,5564,507	PW_hip, PW_lwt	1
BovineHD0600034321	6	62,735,880–6,8101,121	LLPF	2
ARS-BFGL-NGS-107035	7	93,007,435–9,3287,387	CRBY	2
BovineHD0700027195	7	93,079,409	CWT	1
BovineHD0700028765	7	98,540,675–107,723,475	LLPF	5
BTB-01182680	9	81,368,713	CIMF	1
BovineHD1000025368	10	89,027,305–94,456,158	LLPF	2
BovineHD1000027795	10	96,286,865	CIMF	1
BovineHD1300014692	13	51,444,238	CWT	1
BovineHD1400006275	14	21,784,075	CWT	1
BovineHD1400006558	14	22,751,455	CWT, PW_hip	1
BovineHD1400006612	14	22,913,491	SF_lwt	1
BovineHD1400006916	14	23,831,754	PWIGF, EIGF	1
BovineHD1400006992	14	24,114,365	PW_lwt	1
BovineHD1400007050	14	24,312,107	CWT, SF_lwt, PW_hip	1
BovineHD1400007153	14	24,621,142	RFI	1
BovineHD1400007259	14	25,015,640	PW_hip, PW_lwt, SF_lwt, CWT, EIGF	1
BovineHD1400007323	14	25,276,491	PWIGF	1
BovineHD1400007333	14	25,329,035	PW_hip, PW_lwt	1
BovineHD1400007333	14	25,329,035	CWT	1
BovineHD1400007333	14	25,329,035	SF_lwt	1
BovineHD1400007343	14	25,376,827	PWIGF, EIGF	1
BovineHD1400007584	14	26,326,039	PW_hip	1
BovineHD1400007658	14	26,587,761	EIGF	1
BovineHD1400007683	14	26,664,554	PW_lwt	1
BovineHD1400007684	14	26,666,557–35,583,587	CWT	7
BovineHD4100011372	14	27,208,716–27,337,201	SF_lwt	2
BovineHD1400007858	14	27,521,068–31,099,513	PW_hip	3
BovineHD1400008064	14	28,068,938–29,411,154	PW_lwt	2
BovineHD1400008387	14	28,957,602	PW_lwt, CWT	1
ARS-BFGL-NGS-30322	14	49,289,017	CIMF	1
BovineHD2000001543	20	4,873,556	SF_lwt	1
BovineHD2900011261	29	37,272,045– 38,728,128	LLPF	3
BovineHD2900011409	29	41,027,630–49,493,765	LLPF	9

[1]For regions where several highly significant SNP were observed, the most significant SNP is presented.

[2]Traits that were associated with significant SNP ($P < 10^{-8}$); CWT = carcass weight; PW_hip = postweaning hip height; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; HUMP = hump height; SF_hip = feedlot entry hip height; LLPF = peak force; CRBY = retail beef yield; CIMF = intramuscular fat; PWIGF =postweaning IGF-I; EIGF = feedlot entry IGF-I; RFI = residual feed intake.

[3]Number of SNP that were significant at $P < 10^{-8}$ within the specified region.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

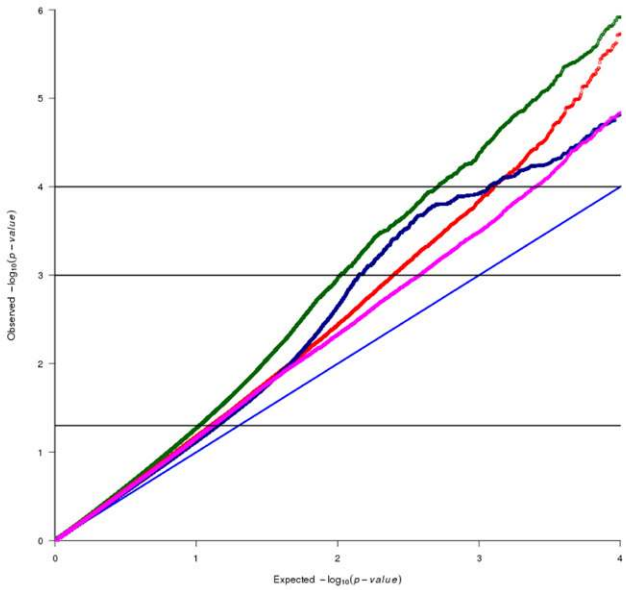

Figure 3. Quantile–quantile plot of *P*-values from single SNP genomewide association study of residual feed intake in *Bos taurus* (red), *Bos indicus* (navy blue), crossbreds (pink), and all breeds and crosses together (dark green). Observed and expected *P*-values would fall on the light blue line if there were no association. The top horizontal line is $P < 0.0001$, middle horizontal line is $P < 0.001$, and the bottom horizontal line is $P < 0.05$. See online version for figure in color.

traits with high values of (***Th*²**) have more significant SNP ($R^2 = 0.43$). However, there were some exceptions. For example, postweaning IGF-I (**PWIGF**) and EIGF, which were only recorded on 918 and 1,103 animals, respectively, have a large number of significant SNP. This is due in part to the large number of significant SNP surrounding the gene *PLAG1* (chromosome 14:25001906..25052394; – strand), which has a large effect on IGF concentration. On the other hand, scanned rib fat (**SRIB**) and scanned P8 fat depth (**SP8**), which are highly heritable and recorded for close to 5,000 animals, have a relatively low number of significant SNP (Fig. 4A).

False discovery rate also varied between breed type and trait combinations. For instance, FDR was less for growth traits in BB and Bt×Bi cross animals compared with Bt animals (Table 6). Such results might reflect mutations of larger effect segregating in particular breed types such as *PLAG1* segregating in BB and composite animals.

Genomic Predictions

Using GBLUP, the average correlation between the TBV and GEBV was 0.27 across traits and breeds (Table 8). The weighted (by numbers within breed)

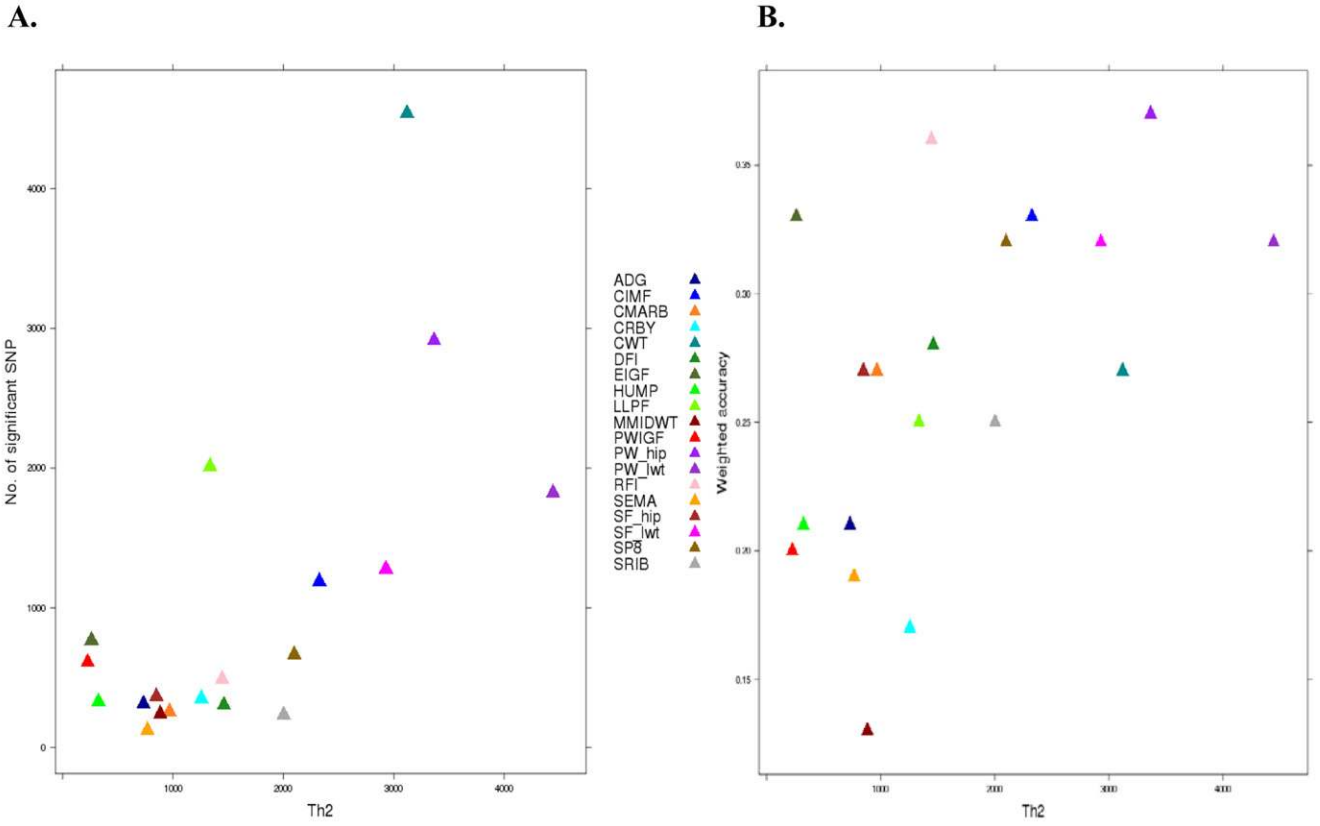

Figure 4. A) Relationship between number of significant SNP ($P < 0.0001$) and the number of records for particular trait (*T*) times heritability (h^2; Th^2). B) Relationship between the weighted genomic BLUP accuracy and the number of records (*T*) for particular trait times heritability (h^2; Th^2). CIMF = intramuscular fat; CMARB = marble score; CRBY = retail beef yield; CWT = carcass weight; DFI = daily feed intake; EIGF = feedlot entry IGF-I; HUMP = hump height; LLPF = peak force; MMIDWT = metabolic mid-weight; PWIGF = postweaning IGF-I; PW_hip = postweaning hip height; PW_lwt = postweaning BW; RFI = residual feed intake; SEMA = scanned eye muscle area; SF_hip = feedlot entry hip height; SF_lwt = feedlot entry BW; SP8 = scanned P8 fat depth; SRIB = scanned rib fat. See online version for figure in color.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Table 8. Average weighted accuracies of genomic EBV across breeds of the 5-fold cross-validation populations using genomic BLUP (GBLUP) method[1]

Trait[2]	AN	BB	BR	BX	HH	MG	SG	SS	TC	ALL (SD)
RFI	0.58	0.31	0.40	0.19			0.37	0.18	0.36	0.36 (0.14)
DFI	–0.02	0.32	0.37	0.24			0.36		0.32	0.28(0.15)
ADG	0.24		0.24	0.13			0.21			0.21 (0.05)
MMIDWT	0.25		0.04				0.10			0.13 (0.11)
LLPF	0.05	0.40	0.34	0.25	0.11	0.71	0.22	0.11	0.26	0.25 (0.20)
SP8	0.50	0.19	0.31		0.14		0.33	0.25		0.32 (0.13)
SRIB	0.42	0.22	0.27		0.07		0.25	0.09		0.25 (0.13)
SEMA	0.15	0.23	0.35		0.20		0.11	0.16		0.19 (0.08)
CRBY	0.26	0.23	–0.03		–0.04		0.21	0.22		0.17 (0.14)
CIMF	0.31	0.27	0.43	0.37	0.20	0.32	0.40	0.27	0.42	0.33 (0.08)
CMARB	0.21	0.19	0.56	0.12	0.14		0.16		0.45	0.27 (0.17)
CWT	0.16	0.28	0.33	0.19	0.32	0.39	0.29	0.19	0.38	0.27 (0.08)
PW_lwt	0.27	0.38	0.29	0.26	0.20		0.24	0.11	0.42	0.32 (0.10)
SF_lwt	0.42	0.31	0.30	0.21	0.26	0.39	0.27	0.19	0.41	0.32 (0.08)
PW_hip	0.20	0.44	0.32		0.23		0.24	0.17	0.47	0.37 (0.12)
SF_hip	0.23	0.28			0.32				0.26	0.27 (0.04)
HUMP		0.25							0.18	0.21 (0.05)
PWIGF		0.31							0.07	0.20 (0.17)
EIGF		0.43							0.24	0.33 (0.13)
Average across breed GBLUP (SD)	0.26 (0.15)	0.30 (0.08)	0.30 (0.14)	0.22 (0.08)	0.18 (0.10)	0.45 (0.17)	0.25 (0.09)	0.18 (0.06)	0.33 (0.12)	0.27 (0.07)
Average within breed GBLUP (SD)	0.26 (0.16)	0.27 (0.10)	0.20 (0.14)	0.13 (0.09)	0.17 (0.11)	0.36 (0.20)	0.16 (0.15)	0.18 (0.08)	0.24 (0.10)	0.23 (0.07)
% compared with within breed GBLUP	101	109	152	168	103	126	157	100	134	117

[1]Empty cells = not estimable or removed if the number of records fewer than 200 for the particular trait; SD = SD of accuracies across breeds and traits. AN = Angus; BB = Brahman; BR = Belmont Red; BX = recent Brahman crosses; HH = Hereford; MG = Murray Grey; SG = Santa Gertrudis; SS = Shorthorn; TC = Tropical Composites.

[2]RFI = residual feed intake; DFI = daily feed intake; MMIDWT = metabolic mid-weight; LLPF = peak force; SP8 = scanned P8 fat depth; SRIB = scanned rib fat; SEMA = scanned eye muscle area; CRBY = retail beef yield; CIMF = intramuscular fat; CMARB = marble score; CWT = carcass weight; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; PW_hip = postweaning hip height; SF_hip = feedlot entry hip height; HUMP = hump height; PWIGF = postweaning IGF-I; EIGF = feedlot entry IGF-I.

accuracies for GBLUP were 0.36 for RFI, between 0.21 and 0.37 for growth traits [PW_lwt, SF_lwt, postweaning hip height (**PW_hip**), feedlot entry hip height (**SF_hip**), and HUMP] and between 0.17 and 0.33 for carcass and meat quality traits [SP8, SRIB, retail beef yield (CRBY), intramuscular fat (**CIMF**), SEMA, CWT and marble score (**CMARB**)]. The lowest accuracies were estimated for metabolic mid-weight (**MMIDWT**), CRBY, and SEMA, which had the greatest FDR among the traits studied in the single SNP regression analysis. There was a tendency (R^2 = 0.30) for accuracy to increase as the number of phenotypes in the reference set multiplied by the heritability (Th^2) increased (Fig. 4B) in the same way as the number of significant SNP increased with Th^2 (Fig. 4A).

Comparing breeds, the greatest average accuracy of GEBV was found in TC followed by BR, BB, and AN ignoring MG (Table 8). Two Bt breeds (HH and SS) tended to have lower average accuracy than other breeds. The relationship between the average accuracy and mean number of records across traits are shown in Fig. 5. As seen in Fig. 5, TC and BB breeds had the greatest average accuracy as well as the largest mean number of records. The SS breed had the smallest number of records after the MG breed and the lowest accuracy of GEBV. The MG breed that is closely related to the AN breed had the greatest average accuracy, but the accuracy was averaged across only 4 traits (Table 8). On average, BayesR accuracies (Table 9) were greater than GBLUP accuracies by 0.03 (Table 8). The increase in accuracy from GBLUP to BayesR was greatest in the composite and BB breeds (Table 9 vs. Table 8) and for traits that had the most significant SNP in the GWAS (Fig. 6).

In Table 10 the posterior means for the number of SNP in each component of the mixture distribution of BayesR are presented (e.g., the number of SNP with effects drawn from normal distributions with variances of either 0, 0.01, 0.1, and 1% of the genetic variance). On average, 19 SNP, 271 SNP, and 4,029 SNP had an effect that explained 1, 0.1, and 0.01% of genetic variance, respectively (Table 10). The traits EIGF, PWIGF, and LLPF had the largest number of SNP explaining 1% of the genetic variance

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

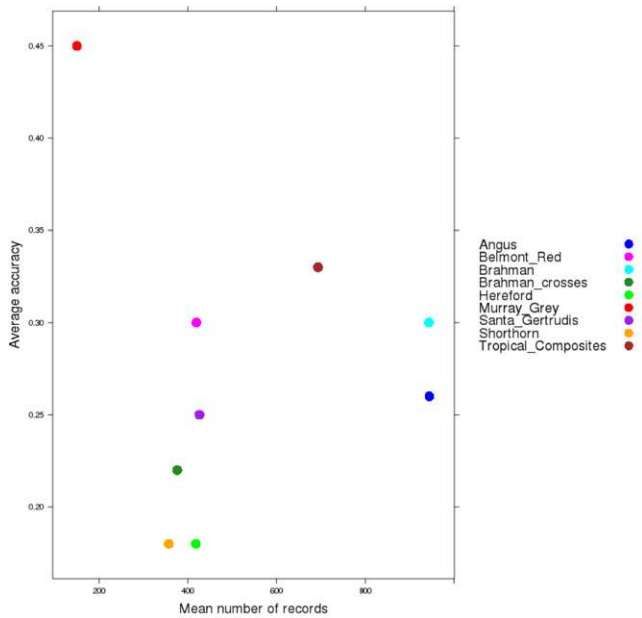

Figure 5. Relationship between the average genomic BLUP accuracy across traits and mean records across traits. See online version for figure in color.

(Table 10). The polygenic variance that is explained by the pedigree but not by the SNP varied from 6 to 70% of the total genetic variance (Table 10) and was largest for traits such as HUMP, ADG, CRBY, MMIDWT, and SRIB where few SNP were significant in the GWAS. We repeated the GBLUP analysis adding a polygenic effect to the model and estimated its variance, which varied widely from 0.01 to 0.55 (Table 10). Generally, the amount of variance explained by the polygenic effect for each trait was similar between BayesR and GBLUP except for some traits such as RFI and IGF-I traits. The accuracy of GBLUP GEBV was hardly changed by fitting the polygenic effect and so this model was not used for the main GBLUP analysis.

Breed Specific Compared with All Breed Training Populations

When the training population was restricted to animals from the same breed as the validation population (e.g., only BB animals used to predict Brahman GEBV) GBLUP accuracies dropped by an average of 0.04 (Table 8). The greatest drop in accuracy was found for the 4 composite breeds (e.g., the difference in accuracies for BR was 0.10). This shows that the composite breeds

Table 9. Average weighted accuracies of genomic EBV of the 5-fold cross-validation populations using BayesR method by breed and across breeds[1]

Trait[2]	AN	BB	BR	BX	HH	MG	SG	SS	TC	ALL (SD)
RFI	0.59	0.33	0.38	0.16			0.38	0.16	0.36	0.36 (0.15)
DFI	0.00	0.25	0.34	0.22			0.31		0.35	0.26 (0.13)
ADG	0.24		0.18	0.27			0.23			0.23 (0.04)
MMIDWT	0.27		0.07				0.15			0.17 (0.10)
LLPF	0.41	0.38	0.55	0.31	0.30	0.61	0.53	0.15	0.43	0.41 (0.14)
SP8	0.48	0.15	0.30		0.06		0.30	0.21		0.28 (0.15)
SRIB	0.38	0.22	0.31		0.00		0.27	0.06		0.23 (0.15)
SEMA	0.10	0.20	0.38		0.18		0.04	0.11		0.16 (0.12)
CRBY	0.35	0.17	0.04		0.01		0.21	0.24		0.20 (0.13)
CIMF	0.29	0.23	0.41	0.32	0.13	0.26	0.33	0.24	0.43	0.29 (0.09)
CMARB	0.10	0.21	0.53	0.08	0.17		0.15		0.43	0.25 (0.17)
CWT	0.18	0.31	0.42	0.26	0.32	0.42	0.35	0.17	0.42	0.30 (0.10)
PW_lwt	0.29	0.41	0.38	0.35	0.20		0.31	0.06	0.46	0.36 (0.13)
SF_lwt	0.42	0.37	0.39	0.27	0.32	0.39	0.33	0.24	0.44	0.36 (0.07)
PW_hip	0.25	0.48	0.40		0.26		0.25	0.15	0.51	0.41 (0.14)
SF_hip	0.24	0.29			0.31				0.30	0.29 (0.03)
HUMP		0.30							0.26	0.28 (0.03)
PWIGF		0.60							0.23	0.43 (0.26)
EIGF		0.56							0.40	0.48 (0.11)
Average (SD)	0.29 (0.15)	0.32 (0.13)	0.34 (0.14)	0.25 (0.08)	0.19 (0.12)	0.42 (0.14)	0.28 (0.11)	0.16 (0.07)	0.39 (0.08)	0.30 (0.09)

[1]Empty cells = not estimable or removed if the number of records fewer than 200 for the particular trait; SD = standard deviation of accuracies across breeds and traits. AN = Angus; BB = Brahman; BR = Belmont Red; BX = recent Brahman crosses; HH = Hereford; MG = Murray Grey; SG = Santa Gertrudis; SS = Shorthorn; TC = Tropical Composites.

[2]RFI = residual feed intake; DFI = daily feed intake; MMIDWT = metabolic mid-weight; LLPF = peak force; SP8 = scanned P8 fat depth; SRIB = scanned rib fat; SEMA = scanned eye muscle area; CRBY = retail beef yield; CIMF = intramuscular fat; CMARB = marble score; CWT = carcass weight; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; PW_hip = postweaning hip height; SF_hip = feedlot entry hip height; HUMP = hump height; PWIGF = postweaning IGF-I; EIGF = feedlot entry IGF-I.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

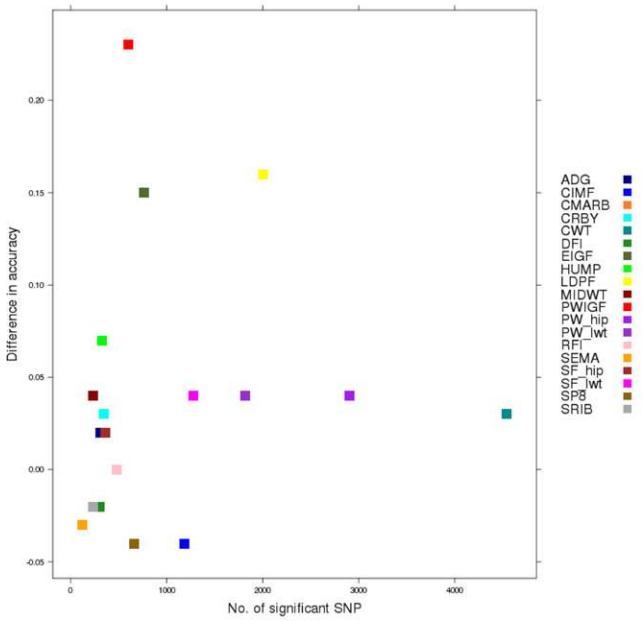

Figure 6. Relationship between number of significant SNP ($P < 0.0001$) and the difference in accuracy between BayesR and genomic BLUP. CIMF = intramuscular fat; CMARB = marble score; CRBY = retail beef yield; CWT = carcass weight; DFI = daily feed intake; EIGF = feedlot entry IGF-I; HUMP = hump height; LDPF = peak force; MIDWT = metabolic mid-weight; PWIGF = postweaning IGF-I; PW_hip = postweaning hip height; PW_lwt = postweaning BW; RFI = residual feed intake; SEMA = scanned eye muscle area; SF_hip = feedlot entry hip height; SF_lwt = feedlot entry BW; SP8 = scanned P8 fat depth; SRIB = scanned rib fat. See online version for figure in color.

gained more advantage than Bi and Bt breeds from using a training population that includes all breeds instead of a breed-specific training population.

Comparisons of Genomewide Association Studies, BLUP, and BayesR for Mapping QTL

The estimates of SNP effects were compared for BayesR and GBLUP (in which all SNP are fitted simultaneously) to those from the GWAS where SNP were fitted individually. The 729 SNP with the largest effects from the BayesR and GBLUP were frequently significant in the GWAS analysis (Table 11). For example, for RFI and LLPF, out of the 729 SNP with largest BayesR effects there were 590 (81%) and 394 (54%) SNP, respectively, that were also significant ($P < 0.001$) in the single SNP regression analysis. Similarly, for the same 2 traits, out of the 729 SNP with largest GBLUP effects there were 540 (74%) and 671 (92%) SNP, respectively, that were also significant ($P < 0.001$) in the single SNP regression analysis.

To judge to what extent the 2 genomic selection methods (BayesR and GBLUP) agree with each other, SNP were ranked according to their absolute effect on each trait. The proportion of the 729 largest BayesR effects that were also found among the largest 729 GBLUP SNP was then calculated. The greatest common proportion (0.54 to 0.63) of large SNP effects was for RFI, daily feed intake (**DFI**), SEMA, and CMARB. The lowest agreement was found for SF_hip and LLPF (0.39 and 0.33, respectively) where the largest BayesR effects were found near or within the regions of genes, such as *PLAG1*, Calpain (chromosome 29:44064420..44089990; + strand), and Calpastatin (chromosome 7:98524257..98581260; + strand), with known large effect mutations affecting the trait (Fig. 4B). When we used up to 7,290 (1% out of total SNP tested) the agreement between the 2 genomic methods gradually increased for all traits indicating that many SNP with small effects were captured by both genomic prediction methods.

Figure 7 shows examples of the typical difference between GWAS, BayesR, and BLUP estimates of SNP effects. The scale for BayesR or BLUP SNP effects presented on Fig. 7 is an arbitrary scale chosen so that we could display the results of all 3 methods on the same graph. In general, peaks occur at approximately the same chromosomal location. However, BLUP generates SNP solutions that do not vary widely between SNP. Genomewide association studies find many significant SNP in a region of the chromosome presumably because

Table 10. Posterior probabilities of SNP effect distributions in BayesR[1]

	Number of SNP in each component				Polygenic variance[2]	
Trait[3]	0	0.01%	0.1%	1%	BayesR	GBLUP[4]
RFI	721,264	7,498	296	6	0.11	0.01
DFI	723,619	5,224	214	8	0.19	0.16
ADG	726,876	1,876	293	15	0.47	0.35
MMIDWT	724,115	4,678	240	11	0.36	0.45
LLPF	727,358	1,419	254	36	0.27	0.26
SP8	723,095	5,384	578	6	0.12	0.26
SRIB	723,762	5,068	229	3	0.32	0.42
SEMA	724,762	3,972	319	8	0.22	0.05
CRBY	726,268	2,570	215	13	0.37	0.35
CIMF	727,127	1,506	428	5	0.24	0.18
CMARB	724,119	4,519	415	11	0.06	0.14
CWT	723,076	5,921	64	5	0.27	0.18
PW_lwt	723,765	5,214	76	11	0.24	0.29
SF_lwt	722,627	6,319	107	14	0.22	0.21
PW_hip	724,212	4,636	203	14	0.13	0.10
SF_hip	725,187	3,597	261	21	0.22	0.07
HUMP	728,106	771	80	17	0.70	0.55
PWIGF	724,729	3,774	416	72	0.17	0.02
EIGF	725,821	2,598	459	81	0.15	0.00
Mean	724,731	4,029	271	19	0.25	0.21

[1]0 or 0.01 or 0.1 or 1% of the total genetic variance.

[2]Polygenic variance as a proportion of the genetic variance for each trait using BayesR and GBLUP methods.

[3]RFI = residual feed intake; DFI = daily feed intake; MMIDWT = metabolic mid-weight; LLPF = peak force; SP8 = scanned P8 fat depth; SRIB = scanned rib fat; SEMA = scanned eye muscle area; CRBY = retail beef yield; CIMF = intramuscular fat; CMARB = marble score; CWT = carcass weight; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; PW_hip = postweaning hip height; SF_hip = feedlot entry hip height; HUMP = hump height; PWIGF = postweaning IGF-I; EIGF = feedlot entry IGF-I.

[4]GBLUP = genomic BLUP.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Table 11. The percentage of the 729 SNP with largest genomic BLUP (GBLUP) and BayesR effects that were also significant at different thresholds of *P*-values in the single SNP regression analyses

	GBLUP				BayesR			
Trait[1]	$P < 0.0001$	$P < 0.001$	$P < 0.01$	$P < 0.05$	$P < 0.0001$	$P < 0.001$	$P < 0.01$	$P < 0.05$
RFI	29	74	98	100	34	81	99	100
DFI	24	70	99	100	23	73	99	100
ADG	19	57	96	100	28	75	99	100
MMIDWT	18	55	94	100	25	76	99	100
LLPF	74	92	98	98	37	54	80	96
SP8	14	53	94	99	35	75	98	100
SRIB	16	58	98	100	21	71	99	100
SEMA	15	61	99	100	16	73	100	100
CRBY	25	68	99	100	27	67	95	100
CIMF	31	71	99	100	27	69	95	99
CMARB	27	83	100	100	25	79	99	100
CWT	45	72	95	100	32	61	86	98
PW_lwt	53	85	100	100	44	80	99	100
SF_lwt	40	77	99	100	25	70	97	100
PW_hip	57	83	99	100	38	71	97	99
SF_hip	18	64	97	100	23	61	96	99
HUMP	17	67	98	100	27	64	98	100
PWIGF	57	90	99	100	25	70	99	100
EIGF	65	90	100	100	30	73	99	100

[1]RFI = residual feed intake; DFI = daily feed intake; MMIDWT = metabolic mid-weight; LLPF = peak force; SP8 = scanned P8 fat depth; SRIB = scanned rib fat; SEMA = scanned eye muscle area; CRBY = retail beef yield; CIMF = intramuscular fat; CMARB = marble score; CWT = carcass weight; PW_lwt = postweaning BW; SF_lwt = feedlot entry BW; PW_hip = postweaning hip height; SF_hip = feedlot entry hip height; HUMP = hump height; PWIGF = postweaning IGF-I; EIGF = feedlot entry IGF-I.

all these SNP are in LD with the same QTL. BayesR tends to find small numbers of SNP very close together with much larger estimated effects than the surrounding SNP. Therefore, it is possible that BayesR maps QTL more precisely than the other methods. This is supported by the small (0.03) increase in GEBV accuracies for BayesR compared with the GBLUP.

DISCUSSION

The GRM is a good representation of the relationships between and within breeds for the animals included in our study. This is shown by the agreement with *F* values and the observed heterozygosities. In our analyses breed differences were always fitted as a fixed effect so the relationships between breeds were not used in the GWAS or calculation of GEBV. Therefore, the important consideration in using the GRM for genomic prediction is that it correctly reflects the amount of genetic variation within each breed. Erbe et al. (2012) calculated a GRM based on a large sample of Holsteins and a small sample of Jerseys and concluded that this distorted the within-breed genetic variance because the Jersey breed was further from the base (allele frequency for the whole dataset) than the Holstein breed. Therefore Jerseys appeared to have less genetic variance. Consequently Erbe et al. (2012) corrected the GRM to a better base. We tested for such a distortion by comparing the observed heterozygosity of SNP with that predicted from the GRM and found no major distortion. The largest anomalies, such as the high observed heterozygosity in Herefords and the less than expected heterozygosity in Brahman animals, are intrinsic to the SNP data used and so cannot be eliminated by a change to the GRM. Therefore we used the GRM as estimated for genomic prediction.

The accuracy of GEBV was affected by trait, size of the training population, and breed and, to a lesser extent, by the type of training population (within or across breed) and statistical method (GBLUP vs. BayesR). Traits that had a large training population and a high heritability and breeds with a large training population tended to give greater accuracies than average. This is expected from theory (Goddard, 2009) in which Th^2 is a critical parameter.

The use of a single multibreed reference population for all breeds either increased or did not decrease the accuracy of GEBV. There are also practical advantages of using a common reference population or, equivalently, a common prediction equation based on SNP genotypes for all breeds. Using a common training population leads to greater accuracy than breed-specific training populations especially for the Bi×Bt composite breeds. This is not surprising given that these breeds should share chromosome segments with both the Bt and Brahman cattle and hence benefit from their inclusion in the training population. Similarly, the reduced FDR in the combined analyses does not indicate that the same

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

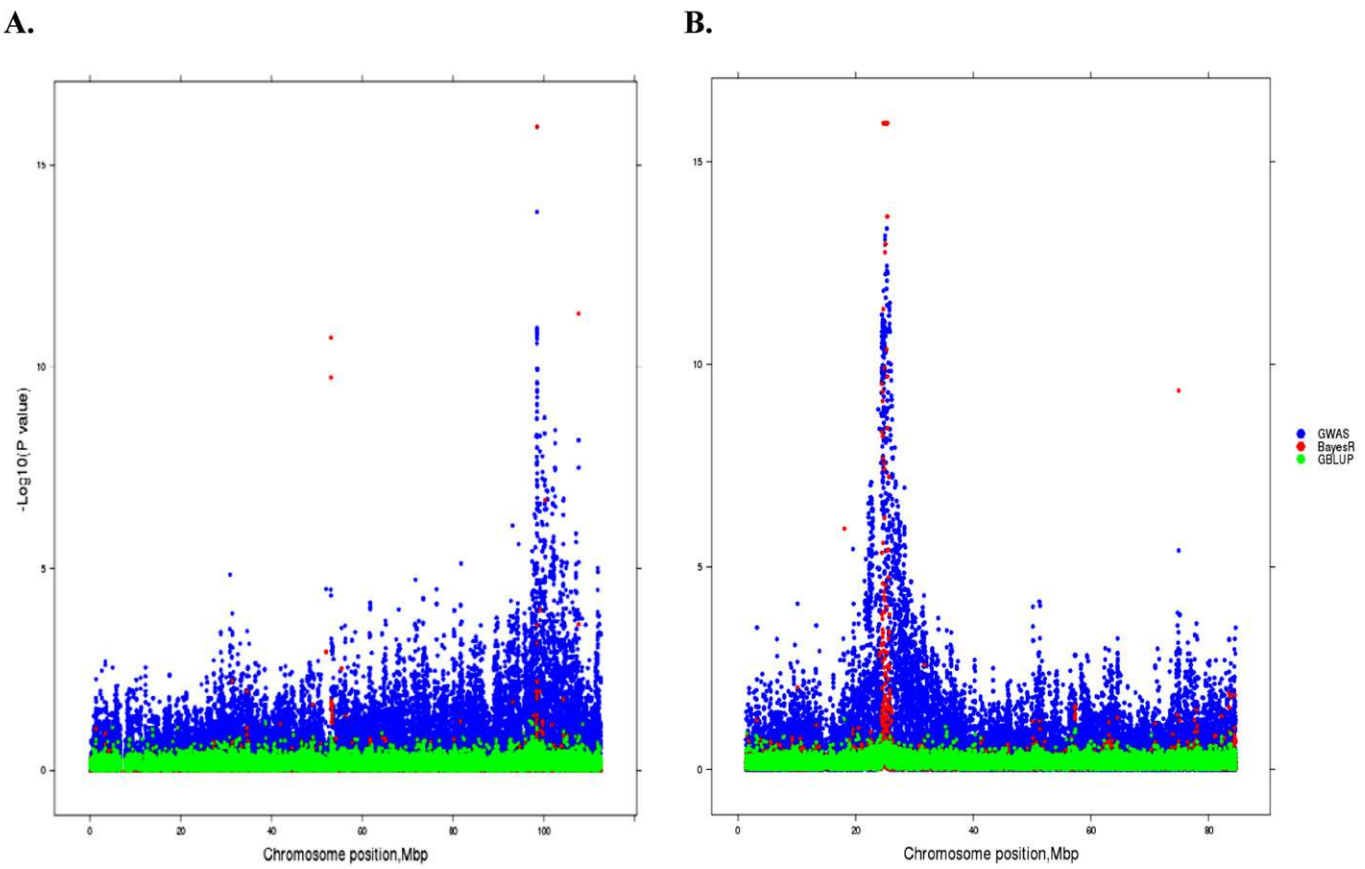

Figure 7. The $-\log_{10}$(*P*-values) of single SNP regressions (blue) and SNP effect predicted from genomic BLUP (GBLUP; green) and BayesR (red) on chromosome 7 for peak force (7A) and on chromosome 14 for postweaning IGF-I (7B). The absolute value effects for BayesR or GBLUP were scaled arbitrarily to fit genomewide association study (GWAS). See online version for figure in color.

QTL segregate in all breed types; it might be that some QTL segregate in both Bt and composite animals and other QTL segregate in Brahman and composite animals.

BayesR gave greater accuracy of GEBV than GBLUP and the advantage was greatest (up to 0.23) for traits with mutations of moderate effect segregating and a high number of significant SNP. For instance, LLPF is known to be affected by polymorphisms in Calpain 1 and Calpastatin (Barendse, 2002, 2007; Page et al., 2002; White et al., 2005; Johnston and Graser, 2010). Highly significant SNP in these genes were indeed found in the GWAS. Similarly, a large number of significant SNP was found associated with IGF concentration in Brahman animals and the accuracy of GEBV increased from 0.25 with GBLUP to 0.41 with BayesR. *PLAG1* is 1 gene with a known large effect mutation for IGF concentration that is segregating in Australian Brahman cattle (Hawken et al., 2011). Precorrection for fixed effects (e.g., breed) could cause some bias in the estimation of SNP effects with BayesR. However, we do not believe that this bias is serious and despite any bias BayesR slightly outperforms GBLUP.

The accuracy of GEBV reported here are less than those typically found in dairy cattle (Hayes et al., 2009; VanRaden et al., 2009; Erbe et al., 2012). This is probably because the value of Th^2 is much greater within Holstein than within any of the breeds represented here. For instance, we have 1,650 Angus measured for PW_lwt with a heritability of 0.45 (Th^2 = 742) whereas VanRaden et al. (2009) had 10,000 Holsteins with daughter averages for milk yield (h^2 = 0.8; Th^2 = 8,000). Hence, it is not surprising that the accuracy in Holstein is 0.7 compared with 0.27 for our Angus. Interestingly, the accuracy of 0.42 reported by Pryce et al. (2012) for the accuracy of RFI in growing Holstein heifer calves is very similar to what we report here, perhaps because as the reference set in Pryce et al. (2012) consisted of heifers with their own phenotypes rather than bulls with daughter averages. The low accuracies of GEBV limit the benefit from genomic selection of beef cattle. However, the GEBV are still of value for some traits that are difficult to improve by traditional selection such as RFI (accuracy = 0.36) and LLFP (accuracy = 0.41). The results and theory strongly suggest that the accuracies of all traits will increase if the training population is increased in size.

There are few other reports on accuracy of genomic predictions in beef cattle (e.g., Kizilkaya et al., 2010; Toosi et al., 2010; Garrick, 2011; Weber et al., 2012). The simulation study of Toosi et al. (2010) found that a multibreed training population was poor at predicting breeding values in a breed that is not included in the

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

training data. However, if the validation breed was a part of the multibreed training set, then the accuracies were similar to within-breed prediction (Toosi et al., 2010). In our findings, multibreed prediction generally gave more accurate genomic prediction than single-breed prediction. This may be due to use of high density markers and more data per breed for the cross-validation. Our findings are supported by a more recent report by Weber et al. (2012) in which the prediction equations trained in multibreed populations were more accurate (especially for Angus and Hereford subpopulations) than the prediction equations trained in a single breed. Using the 50K SNP data, Weber et al. (2012) derived genomic predictions by training and cross-validating using 2 U.S. beef populations of diverse breed composition for selected growth and carcass traits. The authors also reported that the accuracies of GEBV between populations were variable, ranging between 0.11 to 0.50 for growth traits and –0.02 to 0.40 for carcass traits when the prediction equations were trained in multibreed populations, whereas for prediction equations trained in a single breed the accuracies were between 0.05 to 0.53 for growth traits and –0.01 to 0.28 for carcass traits. The within-Angus accuracy of GEBV using 384 SNP selected for their associations with traits reported by Garrick et al. (2009) and Garrick (2011) were 0.59 for marbling, 0.32 for backfat, 0.58 for rib eye, 0.44 for carcass weight, and 0.35 for yearling weight. These estimates were much greater than our accuracies for those traits.

The Q-Q plot deviates from expectation even at low values of $-\log_{10}$ (P-value). This is sometimes interpreted as a sign of uncontrolled population structure. However, this is not likely in this case as we fitted a breed and a polygenic effect in the model. This type of deviation from expectation is also the expected pattern for a trait with many loci causing genetic variation (Yang et al., 2010) and this is the likely interpretation in this case. This interpretation is supported by the thousands of SNP that the BayesR analysis fits, each accounting for approximately 0.0001 of the genetic variance, and by the finding that the accuracy of GEBV calculated with GBLUP is almost as high as those calculated with BayesR. Nevertheless some SNP show highly significant ($P < 10^{-8}$) associations with 1 or more traits (Table 7).

The GWAS, BLUP, and BayesR tend to identify the same SNP as associated with a trait. The prior distribution of SNP effects assumed by BLUP forces all estimated SNP effects to be small. Consequently, we expect that the effect of a single QTL will be predicted by a linear combination of many SNP. The observed results are in agreement with this expectation in that many SNP have small effects and no SNP have large effects. Despite this the largest SNP effects often correspond with the most significant SNP from the GWAS. For traits with some mutations of larger effect (e.g., explaining >1% of the genetic variance), it seems likely that BayesR places more weight on SNP close to the causal mutation than does GBLUP and consequently makes more accurate predictions of breeding values.

Conclusion

Traits with a large number of recorded and genotyped animals and with high heritability gave the greatest accuracy of GEBV. Using a common training population gave greater accuracy than using breed-specific training populations (on average 0.04) especially for composites between *Bos indicus* and *Bos taurus* breeds. BayesR gave greater accuracy of GEBV (up to 0.23) than GBLUP for traits with a high number of significant SNP indicating some polymorphisms of moderate effect on the trait. The average accuracies across traits using BayesR and GBLUP were 0.30 and 0.27, respectively, but they varied widely between breeds and between traits. All traits appear to be highly polygenic with thousands of SNP independently associated with each trait.

LITERATURE CITED

Arthur, P. F., J. A. Archer, R. M. Herd, and G. J. Melville. 2001. Response to selection for net feed intake in beef cattle. In: Proc. Assoc. Advmt. Anim. Breed. Genet. Queenstown. New Zealand. 14:135–138.

Barendse, W. J. 2002. DNA markers for meat tenderness. Patent Application WO02064820. U.S. Patent 7625698-B2.

Barendse, W., B. E. Harrison, R. J. Hawken, D. M. Ferguson, J. M. Thompson, M. B. Thomas, and R. J. Bunch. 2007. Epistasis between calpain 1 and its inhibitor calpastatin within breeds of cattle. Genetics 176:2601–2610.

Barwick, S. A., M. L. Wolcott, D. J. Johnston, H. M. Burrow, and M. T. Sullivan. 2009. Genetics of steer daily feed intake and residual feed intake in tropical beef genotypes and relationships among intake, body composition, growth and other post-weaning measures. Anim. Prod. Sci. 49:351–366.

Bolormaa, S., B. J. Hayes, K. Savin, R. Hawkin, W. Barendse, D. Johnston, P. Arthur, R. Herd, and M. E. Goddard. 2011. Genome wide association studies for feedlot and growth traits in cattle. J. Anim. Sci. 89:1684–1697.

Browning, B. L. and S. R. Browning. 2011. A fast, powerful method for detecting identity by descent. Am. J. Hum. Genet. 88:173-182.

De Roos, A. P. W., B. J. Hayes, R. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein Friesian, Jersey and Angus cattle. Genetics 179:1503–1512.

Erbe, M., B. J. Hayes, L. K. Matukumalli, S. Goswami, P. J. Bowman, M. Reich, B. A. Mason, and M. E. Goddard. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J. Dairy Sci. 95:4114–4129.

Garrick, D. J. 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. Genet. Sel. Evol. 43:17.

Garrick, D. J., J. F. Taylor, and R. L. Fernando. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genet. Sel. Evol. 41:55.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN International Ltd., Hemel Hempstead, UK.

Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica (The Hague) 136:245–257.

Habier, H., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. BMC Bioinf. 12:186.

Hawken, R. J., Y. D. Zhang, M. R. S. Fortes, E. Collis, W. C. Barris, N. J. Corbet, P. J. Williams, G. Fordyce, R. G. Holroyd, J. R. W. Walkley, W. Barendse, D. J. Johnston, K. C. Prayaga, B. Tier, A. Reverter, and S. A. Lehnert. 2011. Genome-wide association studies of female reproduction in tropically adapted beef cattle. J. Anim. Sci. 90:1398–1410.

Hayes, B. J., P. J. Bowman, A. C. Chamberlain, and M. E. Goddard. 2009. Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433–443.

Hedrick, P. W. 2005. Genetics of populations. 3rd ed. page 489, Jones and Barlett Publishers, Inc. Sudbury, Massachusetts, USA.

Johnston, D. J. 2001. Selecting for marbling and its relationship with other important economic traits. What impact does it have? In: Proceedings of Marbling Symposium, Coffs Harbour, Australia. p. 88–93.

Johnston, D. J., and H.-U. Graser. 2010. Estimated gene frequencies of GeneSTAR markers and their size of effects on meat tenderness, marbling and feed efficiency in temperate and tropical beef cattle breeds across a range of production systems. J. Anim. Sci. 88:1917–1935.

Johnston, D. J., A. Reverter, D. M. Ferguson, J. M. Thompson, and H. M. Burrow. 2003. Genetic and phenotypic characterisation of animal, carcass and meat quality traits for temperate and tropically adapted beef breeds. 3. Meat quality traits. Aust. J. Agric. Res. 54:135–147.

Kizilkaya, K., R. Fernando, and D. Garrick. 2010. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. J. Anim. Sci. 88:544–551.

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.

Page, B. T., E. Casas, M. P. Heaton, N. G. Cullen, D. L. Hyndman, C. A. Morris, A. M. Crawford, T. L. Wheeler, M. Koohmaraie, J. W. Keele, and T. P. L. Smith. 2002. Evaluation of single nucleotide polymorphisms in CAPN1 for associations with meat tenderness in cattle. J. Anim. Sci. 80:3077–3085.

Pryce, J. E., J. Arias, P. J. Bowman, S. R. Davis, K. A. Macdonald, G. C. Waghorn, W. J. Wales, Y. J. Williams, R. J. Spelman, and B. J. Hayes. 2012. Accuracy of genomic predictions of residual feed intake and 250 day bodyweight in growing heifers using 625,000 SNP markers. J. Dairy Sci. 95:2108–2119.

Reverter, A., D. J. Johnston, D. Perry, M. E. Goddard, and H. M. Burrow. 2003. Genetic and phenotypic characterisation of animal, carcass and meat quality traits for temperate and tropically adapted beef breeds. 2. Abattoir carcass traits. Aust. J. Agric. Res. 54:119–134.

Robinson, D. L., and V. H. Oddy. 2004. Genetic parameters for feed efficiency, fatness, muscle area and feeding behaviour of feedlot finished beef cattle. Livest. Prod. Sci. 90:255–270.

Toosi, A., R. Fernando, and J. Dekkers. 2010. Genomic selection in admixed and crossbred populations. J. Anim. Sci. 88:32–46.

Upton, W. H., H. M. Burrow, A. Dundon, D. R. Robinson, and E. B. Farrell. 2001. Cooperative Research Centre breeding program design, measurements and database-methods that underpin CRC research results. Aust. J. Exp. Agric. 41:943–952.

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.

Weber, K. L., R. M. Thallman, J. W. Keele, W. M. Snelling, G. L. Bennett, T. P. Smith, T. G. McDaneld, M. F. Allan, A. L. Van Eenennaam, and L. A. Kuehn. 2012. Accuracy of genomic breeding values in multibreed beef cattle populations derived from deregressed breeding values and phenotypes. J. Anim. Sci. 90:4177–4190.

White, S. N., E. Casas, T. L. Wheeler, S. D. Shackelford, M. Koohmaraie, D. G. Riley, C. C. Chase Jr., D. D. Johnson, J. W. Keele, and T. P. L. Smith. 2005. A new single nucleotide polymorphism in CAPN1 extends the current tenderness marker test to include cattle of *Bos indicus*, *Bos taurus*, and crossbred descent. J. Anim. Sci. 83:2001–2008.

Wolcott, M. L., D. J. Johnston, S. A. Barwick, C. L. Iker, J. M. Thompson, and H. M. Burrow. 2009. Genetics of meat quality and carcass traits and the impact of tenderstretching in two tropical beef genotypes. Anim. Prod. Sci. 49:383–398.

Wolcott, M. L., H.-U. Graser, and D. J. Johnston. 2010. Effects of early weaning on growth, feed efficiency and carcass traits in Shorthorn cattle. Anim. Prod. Sci. 50:315–321.

Yang, J., B. Benyamin, N. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 42:565–569.

Downloaded from https://academic.oup.com/jas/article/91/7/3088/4717036 by guest on 20 August 2022