Imputation of genotypes with low-density chips and its effect on reliability of direct genomic values in Dutch Holstein cattle

H. A. Mulder,^{*1} M. P. L. Calus,^{*} T. Druet,[†] and C. Schrooten[‡]

^{*}Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, PO Box 65, 8200 AB Lelystad, the Netherlands

[†]Unit of Animal Genomics, Faculty of Veterinary Medicine and Centre for Biomedical Integrative Genoproteomics, University of Liège, B-4000, Liège, Belgium

[‡]CRV, PO Box 454, 6800 AL Arnhem, the Netherlands

ABSTRACT

Genomic selection using 50,000 single nucleotide polymorphism (50k SNP) chips has been implemented in many dairy cattle breeding programs. Cheap, low-density chips make genotyping of a larger number of animals cost effective. A commonly proposed strategy is to impute low-density genotypes up to 50,000 genotypes before predicting direct genomic values (DGV). The objectives of this study were to investigate the accuracy of imputation for animals genotyped with a low-density chip and to investigate the effect of imputation on reliability of DGV. Low-density chips contained 384, 3,000, or 6,000 SNP. The SNP were selected based either on the highest minor allele frequency in a bin or the middle SNP in a bin, and DAGPHASE, CHROMIBD, and multivariate BLUP were used for imputation. Genotypes of 9,378 animals were used, from which approximately 2,350 animals had deregressed proofs. Bayesian stochastic search variable selection was used for estimating SNP effects of the 50k chip. Imputation accuracies and imputation error rates were poor for low-density chips with 384 SNP. Imputation accuracies were higher with 3,000 and 6,000 SNP. Performance of DAGPHASE and CHROMIBD was very similar and much better than that of multivariate BLUP for both imputation accuracy and reliability of DGV. With 3,000 SNP and using CHROMIBD or DAGPHASE for imputation, 84 to 90% of the increase in DGV reliability using the 50k chip, compared with a pedigree index, was obtained. With multivariate BLUP, the increase in reliability was only 40%. With 384 SNP, the reliability of DGV was lower than for a pedigree index, whereas with 6,000 SNP, about 93% of the increase in reliability of DGV based on the 50k chip was obtained when using DAGPHASE for imputation. Using genotype probabilities to predict gene content increased imputation accuracy and the reliability of DGV and is

therefore recommended for applications of imputation for genomic prediction. A deterministic equation was derived to predict accuracy of DGV based on imputation accuracy, which fitted closely with the observed relationship. The deterministic equation can be used to evaluate the effect of differences in imputation accuracy on accuracy and reliability of DGV.

Key words: imputation, single nucleotide polymorphism, low-density chip, genomic selection

INTRODUCTION

In the last couple of years, genotyping of bulls and cows with 50,000 SNP (50k) chips has become common practice in dairy cattle breeding programs applying genomic selection (e.g., Hayes et al., 2009; Calus, 2010). With genomic selection, thousands of SNP marker effects are estimated using a reference population of animals with genotypes and phenotypes (Meuwissen et al., 2001). With these estimated marker effects, direct genomic values (DGV) can be calculated for any genotyped animal, and selection of animals can be performed on genomic EBV, which is basically an index of the DGV and the pedigree index (VanRaden et al., 2009). Still, the costs for genotyping an animal using a 50k chip are such that routine genotyping of commercial dairy cows is hampered. Recently, Habier et al. (2009) suggested the use of low-density SNP chips that are cheaper, which makes it feasible to genotype a larger number of animals.

Direct calculation of DGV based on a low-density chip results in lower accuracy (Weigel et al., 2009; Moser et al., 2010) and requires separate estimation of SNP effects for both the low-density and the 50k chip. However, with the use of a low-density chip, the missing SNP (those that are on the 50k chip) can be predicted or imputed based on a reference panel that is genotyped for the 50k chip. Using imputed 50k SNP genotypes is expected to give greater accuracy for DGV than using only the subset of SNP on the low-density chip (Weigel et al., 2010a). Different imputation methods are available and used in human

Received April 27, 2011.

Accepted October 25, 2011.

¹Corresponding author: han.mulder@wur.nl

genetics to increase power in genome-wide association studies by combining data of individuals genotyped with different SNP chips (Browning, 2008; Li et al., 2009; Marchini and Howie, 2010). These methods rely mostly on population-wide linkage disequilibrium (LD) information, which is feasible if dense marker chips are used. Weigel et al. (2010b) used fastPhase (Scheet and Stephens, 2006) and IMPUTE (Howie et al., 2009) to impute from a low-density chip to the 50k chip and reported that between 73.3 and 93.5% of the genotypes were correctly imputed in Jersey cattle when 5 or 10% of the genotypes were available. In animal populations, however, linkage information based on extended pedigrees and genotyped ancestors can be used to infer haplotypes and to impute SNP (e.g., Habier et al., 2009). Calus et al. (2011) used simulated data to compare the imputation accuracy of multivariate BLUP on gene content, as proposed by Mulder et al. (2010b), with fastPhase and Beagle (Browning and Browning, 2007) and found that Beagle generally performed best, but that the multivariate BLUP method performed better at low marker density. Calus et al. (2011), Druet et al. (2010), and Zhang and Druet (2010) showed that information of genotyped ancestors is particularly important for accuracy of imputation when LD levels are low. Druet and Georges (2010) developed a hidden Markov model that combines both linkage and LD information for imputation and implemented this method in the DAGPHASE software. This method was also used by the EuroGenomics consortium to impute SNP from the standard Illumina BovineSNP50 BeadChip (Illumina, San Diego, CA) to 2 customized Illumina 50k chips and vice versa, with an overlap of 10,954 and 19,784, respectively, with the standard Illumina BovineSNP50 BeadChip (Druet et al., 2010). They reported error rates of 0.5% and imputation accuracies of 0.98 when using 2,000 animals in the reference population; that is, animals that were genotyped for 2 virtual chips of approximately 27,500 SNP derived from the SNP chips. Zhang and Druet (2010) compared DAGPHASE and CHROMIBD (Druet and Farnir, 2011). The CHROMIBD program uses linkage information to estimate identity-by-descent (IBD) probabilities between chromosomes of an animal of interest and chromosomes of the closest genotyped ancestors in all branches of the pedigree. Both methods were tested on a subset of data used in this study with a varying number of markers on the low-density chip. CHROMIBD was only more efficient than DAGPHASE at lower marker densities when several and recent ancestors were genotyped. Allelic imputation error rates were reported to be between 2.0 and 11.7% for chips varying in size between 6,000 markers and 384 (Zhang and Druet, 2010). The DAGPHASE and CHROMIBD methods were not compared

with the multivariate BLUP method until now. So far, studies for a low-density chip selected either equally spaced SNP (Habier et al., 2009) or almost equally spaced SNP with the highest minor allele frequency (MAF; Weigel et al., 2010a; Zhang and Druet, 2010). Single nucleotide polymorphisms with a higher MAF have greater information content and therefore are expected to provide more information for imputation. It is, however, unknown how large the benefit is when SNP are selected on the highest MAF, in addition to being equally spaced.

The effect of imputation of the low-density chip to a high-density chip on accuracy of DGV was first studied by Habier et al. (2009) using simulation. They showed that the accuracy of DGV based on a combination of high-density SNP chip data and imputed low-density SNP chip data was about 4.5% lower than when using the high-density SNP chip in all cases. Weigel et al. (2010a) were the first to use real data to investigate the effect of imputation of the low-density chip to the 50k chip on the accuracy of DGV. They reported that DGV based on a low-density chip with 3,000 SNP and with imputation of other SNP yielded 95% of the accuracy of DGV based on the 50k chip in Jersey cattle. They used the IMPUTE software for imputation, which was the best method in their earlier study (Weigel et al., 2010b).

Two issues were not addressed in previous studies that may affect the accuracy or reliability of DGV based on imputed SNP. The first issue is that some imputation packages provide both the most likely genotype as well as genotype probabilities, which can be used to calculate predicted gene contents or "dosage." The use of predicted gene contents may be preferable for calculation of DGV, as it accounts for uncertainty due to imputation. The second issue is that there are numerous measures of imputation accuracy and different names exist for the same parameter. Roughly speaking, the measures can be divided into 2 groups: (1) proportion of correctly imputed alleles or genotypes (e.g., Weigel et al., 2010b) or incorrectly imputed alleles (Druet et al., 2010; Zhang and Druet, 2010) or genotypes, and (2) (squared) correlation between true and imputed genotype (e.g., Huang et al., 2009; Mulder et al., 2010b). The correlation measure may be preferable because it is not dependent on minor allele frequency (Browning and Browning, 2009). However, the relationship between imputation accuracy and accuracy of DGV is unknown.

The first aim of this study, therefore, was to investigate the accuracy of imputation for animals genotyped with a low-density chip, based on 3 different marker densities, 2 types of marker selection, and using 3 different methods (DAGPHASE, CHROMIBD, and multivariate BLUP) for imputation. The second aim was to

investigate the effect of imputation on the reliability of DGV as well as the effect of using the most likely genotype or the predicted gene content. The third aim was to derive an empirical relationship between imputation accuracy and accuracy of DGV.

MATERIALS AND METHODS

Data

Genotype Data. The initial data set contained 9,849 bulls and cows, predominantly Holsteins, genotyped with a customized Illumina 50k SNP chip (Charlier et al., 2008). In constructing this data set, genotypes were removed if they did not match the genotype of parent(s) or progeny. Single nucleotide polymorphisms with MAF <0.0025, with genotypes in less than 90% of the animals, as well as SNP that deviated from Hardy-Weinberg equilibrium (difference between observed and expected frequency of heterozygotes >0.15; Wiggans et al., 2009) were removed. Only SNP on BTA 1 to 29 were used. Finally, 39,759 SNP were used in the analyses. All included individuals were Black and White and Red and White Holstein animals and were born after January 1, 1990. This resulted in 9,378 animals genotyped and used in the analyses. For these genotyped animals, their whole known pedigree comprised 19,718 animals.

SNP Selection for Low-Density Chip. Three different numbers of SNP for the low-density chip were considered: 384, 3,000, and 6,000. Those numbers were based on the size of currently commercially available SNP chips. The SNP were selected in 2 ways after dividing the whole genome into 384, 3,000, or 6,000 equally sized segments. Within each segment, the SNP closest to the physical midpoint of the segment was selected (hereafter **Mid**) or the SNP with the highest MAF was selected (hereafter **MAF**). For SNP selected based on MAF, adjacent selected SNP were restricted to have an r^2 value <0.2 whenever possible. Using these 3 densities and the 2 types of SNP selection, 6 different low-density chips were constructed in total.

Animal Selection for Low-Density Chip. The individuals selected for the low-density chip were animals without genotyped offspring and with known

pedigrees. In addition, traceability was used as a selection criterion. Traceability is a simple measure that summarizes how many and which types of ancestors are genotyped (see Appendix). Traceability was also used in Druet et al. (2010) and Zhang and Druet (2010). Here, animals were selected with a traceability of at least 0.40. Animals without offspring were selected, because in practice most interest is in screening young animals (without offspring) with the low-density chip. In total, 8,145 animals were candidates to be genotyped for the low-density chip.

Two batches of animals genotyped with the low-density chip were made, in which half of the animals were selected for each traceability (*T*) class ($T = 1.0$, $0.95 \leq T < 1.0$, $0.75 \leq T < 0.95$, $0.55 \leq T < 0.75$, and $0.4 \leq T < 0.55$). Numbers of considered candidates for the low-density chip per traceability class are given in Table 1. If the animal was selected in the first batch for the low-density chip, it was on the 50k chip in the second batch and vice versa. Genotyped animals not selected for the low-density chip were always considered to be on the 50k chip. Therefore, in the first (second) batch, there were 5,304 (5,307) animals genotyped for the 50k chip, and 4,074 (4,071) animals were genotyped for each of the considered low-density chips.

Imputation Methods

Three methods were used for imputation: localized haplotype clustering model using DAGPHASE and Beagle software (Druet and Georges, 2010), genotyped ancestor hidden Markov model using CHROMIBD and Beagle software (Druet and Farnir, 2011), and multivariate BLUP using MiXBLUP (Mulder et al., 2010a,b). To limit computational requirements, CHROMIBD was only run on the scenario MAF with 3,000 SNP; DAGPHASE and multivariate BLUP were run on all scenarios.

Localized Haplotype Clustering Using DAGPHASE and Beagle. In the first step, haplotypes of all animals were reconstructed as far as possible based on linkage information using LinkPHASE (Druet and Georges, 2010). In the second step, haplotypes of 50k genotyped animals were further reconstructed using LD

Table 1. Number of animals per traceability class that were considered as candidates for low-density chip genotyping, proportion of Red and White animals, and average birth year per traceability class

Traceability class	Number of animals for low-density chip genotyping	Proportion Red and White	Average birth year
1.0	1,863	0.207	2008
0.95–1.00	2,839	0.218	2005
0.75–0.95	2,843	0.210	2001
0.55–0.75	468	0.280	1995
0.4–0.55	132	0.242	1992

information by iterating 20 times between Beagle 2.1.3. (Browning and Browning, 2007) and DAGPHASE (Druet and Georges, 2010). In each iteration, Beagle produced a directed acyclic graph (**DAG**). This DAG summarized all founder haplotypes that were present in the population. In the final step, linkage information from LinkPHASE and LD information from Beagle/DAGPHASE were integrated to reconstruct the haplotypes and impute the missing genotypes on the low-density chip. In this case, Beagle was run with scale parameter 2.0 and shift parameter 0.1. For more details on the method, see Druet and Georges (2010). Both LinkPHASE and DAGPHASE are members of PHASEBOOK, a collection of imputation software packages (Druet and Georges, 2010). In all scenarios, the most likely genotype was imputed, whereas for reasons of comparison the predicted gene content using the probability for each of the possible genotypes was also calculated for the scenario MAF with 3,000 SNP, as follows:

$$\hat{gc} = \sum_{i=1}^4 gc_i \times P_i,$$

where $\hat{gc}$ is the predicted gene content, gc_i is the gene content of genotype i ; that is, 0, 1, 1, and 2 (i.e., for genotypes 11, 12, 21 and 22, respectively, when one allele is coded as 1 and the other allele as 2), and P_i is the probability for genotype i . In the results section, results based on the most likely genotype are referred to as DAGPHASE and the results based on genotype probabilities as DAGPHASEgc.

Genotyped Ancestor Hidden Markov Model Using CHROMIBD and Beagle. The genotyped ancestor hidden Markov model involved several steps. The first 2 steps were identical to those of the localized haplotype clustering model; only the final imputation step was different. In this step, CHROMIBD calculated IBD probabilities between a chromosome and all the genotyped chromosomes of its parents (Druet and Farnir, 2011). For every branch in the pedigree, only the closest genotyped parent was considered. These IBD probabilities were converted to predicted gene contents by multiplying the parental genotype by the IBD probability:

$$\hat{gc} = \sum_i^2 \sum_j^n (IBD_{ij} \times s_{ij}) - 2,$$

where i is the paternal or maternal haplotype of the animal, j is the parental haplotype number, n is the number of possible parental haplotypes, IBD_{ij} is the IBD probability that the paternal (maternal) haplotype

originates from the parental haplotype j , and s_{ij} is the SNP allele (coded as 1 or 2) for the parental haplotype j of animal haplotype i ; 2 is subtracted to obtain a value of $\hat{gc}$ between 0 and 2. Results of the method are referred to as CHROMIBD.

Multivariate BLUP. The missing SNP were predicted using multivariate BLUP on gene contents following Mulder et al. (2010b) and Calus et al. (2011), being 0, 1, 1, and 2 for genotypes for genotypes 11, 12, 21, and 22, respectively. The multivariate BLUP method is an extension of the BLUP on gene content method as presented by Gengler et al. (2007, 2008), using information of flanking markers as correlated traits, as well as the additive genetic relationship matrix to predict the missing marker genotype. In this case, the nearest left and right marker on the low-density chip were used as correlated traits using the correlation between phased genotypes on the imputed and predicted loci. The squared value of this correlation is equal to the r^2 value that is used as a measure of LD (Hill and Robertson, 1968). The chosen variance components corresponded to a heritability of 0.99 for gene content (Mulder et al., 2010b). The multivariate BLUP method is not sensitive when using slightly lower values of heritability (Calus et al., 2011). Genetic groups were not included, because the animals used were all Holsteins, assuming no differences in allele frequencies in the base generation. The predicted gene contents were calculated as mean plus EBV. The mixed model equations were solved using MiXBLUP (Mulder et al., 2010a).

Evaluation of Imputation. Imputation accuracy was calculated as the correlation between gene contents of imputed and true genotypes. To obtain insight into individual variability in imputation accuracy, imputation accuracies per animal were calculated as the correlation between predicted and true gene contents across all loci on a chromosome and then averaged across all chromosomes. Bias in the variance of the predicted gene contents was assessed by regression of true on predicted gene contents. Mean squared error of prediction (**MSEP**) was calculated as the mean of the squared errors of prediction. Allelic imputation error rates were calculated as the average number of incorrectly imputed alleles when the method yielded discrete alleles. When results were expressed as gene content, allelic imputation error rates were calculated as the average deviation of the true and predicted gene contents divided by 2, because the number of gene contents is equal to half the number of alleles. Imputation accuracies, regressions, MSEP, and allelic imputation error rates were first calculated per SNP across both batches of animals considered genotyped with low-density chips, and then averaged across SNP per chromosome and then across chromosomes.

Prediction of DGV

The DGV and SNP effects were estimated for 10 traits using the genotypes from the 50k chip and phenotypes of animals in the reference population. The traits considered were protein yield, milk yield, udder depth score, overall feet and legs score, overall udder score, final conformation score, longevity, SCS, calving interval, and calving ease. Phenotypes used were deregressed proofs (**DRP**) from the national evaluation with a minimum effective daughter contribution of 10. The reference population was based on the first batch using only animals with phenotypes and genotypes with the 50k chip. The number of reference animals was 2,568 for calving ease, and varied between 2,339 (calving interval) and 2,386 (SCS, longevity) for the other traits. Prediction of DGV was performed with the method described by Calus et al. (2008); that is, a Bayesian method based on stochastic search variable selection (George and McCulloch, 1993). The model was

$$y_i = \mu + a_i + \sum_{j=1}^{nloc} \sum_{k=1}^2 SNP_{ijk} + e_i,$$

where y_i is the phenotypic record for animal i , μ is the overall mean, a_i is the random polygenic effect of animal i , SNP_{ijk} is a random effect for allele k at locus j of animal i , and e_i is a random residual for animal i . The SNP effects, denoted as SNP_{ijk} , were estimated as $q_{ijk} \times v_k$, following Meuwissen and Goddard (2004), where q_{ijk} is the size of the effect of allele k at locus j , v_k is a scaling factor in the direction vector for locus k , and $nloc$ is the number of SNP loci. More details on the prior distributions and the fully conditional distributions can be found in Meuwissen and Goddard (2004). All parameters were estimated with a Markov chain Monte Carlo method using Gibbs sampling with residual updating.

Estimated marker effects used in further analyses were averages of posterior marker effects of 4 Gibbs chains with 25,000 iterations each, of which the first 5,000 iterations were discarded as burn-in. The validation animals were animals with phenotypes in batch 1 that were considered genotyped with the low-density chip (i.e., using imputed genotypes for the other SNP). The number of validation animals varied between 1,382 and 1,391. Increase in reliability due to using SNP information was measured as the squared correlation between genomic and daughter-based EBV minus the squared correlation between pedigree and daughter-based EBV divided by the average reliability of DRP, following the Interbull validation test for genomic evaluation (Interbull, 2010):

$$\Delta R^2 = (R_{DGV,DRP}^2 - R_{PED,DRP}^2) / R_{DRP}^2, \quad [1]$$

where ΔR^2 is the increase in reliability, $R_{DGV,DRP}^2$ is the squared correlation between DGV and DRP, $R_{PED,DRP}^2$ is the squared correlation between pedigree index and DRP, and R_{DRP}^2 is the average reliability of the DRP of the validation bulls. For DAGPHASE, DGV were calculated for all scenarios, whereas for DAGPHASE with predicted gene content, multivariate BLUP, and CHROMIBD, DGV were only calculated for MAF3000. In addition, correlations between the DGV based on the 50k chip and the low-density chips were calculated per traceability class to investigate the effect of traceability on the reliability of DGV with imputed SNP.

Relationship Between Imputation Accuracy and Accuracy of DGV

The relationship between imputation accuracy and accuracy of DGV (i.e., the square root of the reliability) was derived as follows. According to Meuwissen (2009) the accuracy of DGV may be split up into a component due to prediction based on SNP and a part that could be predicted by traditional BLUP or pedigree index; that is,

$$r_{DGV} = r_{ped} + (1 - r_{ped})\rho, \quad [2]$$

where r_{DGV} (r_{ped}) is the accuracy of the DGV (pedigree index with traditional BLUP) and ρ is the proportion of inaccuracy in the pedigree index that could be explained by SNP. Equation [2] gives the upper bound of the accuracy of DGV when the animal is genotyped for the 50k chip. Equation [2] was modified to predict r_{DGV} when some markers are imputed and some are on a low-density chip; derivation details can be found in the Appendix:

$$r_{DGV,imp} = r_{ped} + (1 - r_{ped}) \frac{\rho}{(1 - r_{imp,ped})} [p(1 - r_{imp,ped}) + (1 - p)(r_{imp} - r_{imp,ped})], \quad [3]$$

where $r_{DGV,imp}$ is the accuracy of the DGV based on a low-density chip where most SNP are imputed, ρ is the proportion of accuracy of DGV that is not explained by r_{ped} when using a 50k chip, r_{imp} is the imputation accuracy of imputed SNP, $r_{imp,ped}$ is the imputation accuracy based on pedigree alone, not due to using markers on the low-density chip (set at 0.68, the average in this data set), and p is the proportion of SNP on the low-density chip relative to the high-density chip. Equation [3] shows that the accuracy of DGV increases linearly

Table 2. Accuracy of imputed SNP genotypes for MAF and Mid scenarios¹ for different numbers of SNP and different traceability classes using different methods of imputation

Method	Scenario	SNP (no.)	Traceability					
			Overall ²	1.0	0.95–1.0	0.75–0.95	0.55–0.75	0.4–0.55
Multivariate BLUP	MAF	384	0.647	0.710	0.644	0.620	0.562	0.494
		3,000	0.705	0.757	0.703	0.682	0.632	0.576
		6,000	0.732	0.779	0.723	0.712	0.668	0.619
	Mid	384	0.649	0.712	0.646	0.622	0.562	0.493
		3,000	0.696	0.751	0.695	0.695	0.620	0.564
		6,000	0.725	0.774	0.723	0.704	0.658	0.608
DAGPHASE	MAF	384	0.578	0.809	0.543	0.518	0.362	0.160
		3,000	0.884	0.964	0.897	0.861	0.717	0.566
		6,000	0.923	0.978	0.940	0.908	0.771	0.614
	Mid	384	0.561	0.782	0.525	0.505	0.351	0.143
		3,000	0.881	0.963	0.894	0.858	0.714	0.562
		6,000	0.921	0.977	0.938	0.905	0.768	0.611
DAGPHASEgc ³	MAF	3,000	0.919	0.974	0.929	0.905	0.800	0.660
CHROMIBD	MAF	3,000	0.926	0.980	0.952	0.910	0.757	0.567

¹The SNP on the low-density chip were selected when being closest to the physical midpoint of a bin (Mid), or with the highest minor allele frequency within a bin (MAF).

²Across all animals on low-density chip.

³Gene content predicted using DAGPHASE based on genotype probabilities.

with the imputation accuracy. Equation [3] was applied across scenarios and methods.

RESULTS

Accuracy of Imputation

Number of SNP on Chip and SNP Selection. Tables 2, 3, and 4 show the accuracy, allelic imputation error rate, and regression, respectively, per scenario for each method and traceability class. Increasing the number of SNP on the low-density chip increased imputation accuracy and decreased allelic imputation error rates, especially from 384 to 3,000 SNP. Selection of SNP on the low-density chip based on MAF or Mid had very small effect on accuracy and allelic imputation error rate. Regression coefficients were in general lower than unity, indicating that the variance of the imputed SNP genotypes was biased upward. Increasing the number of SNP on the low-density chip resulted in higher regression coefficients closer to unity. Regression coefficients were virtually the same for MAF and Mid (results of Mid are therefore not shown). The MSEP values showed similar trends as the allelic imputation error rates, accuracy, and regression coefficients (results not shown). It can be concluded that at least 3,000 SNP should be placed on a low-density chip. Selecting SNP based on MAF, in addition to being equally spaced, barely improved imputation accuracy.

Traceability. With higher traceability, accuracy of imputation increased and allelic imputation error rates and MSEP decreased. The regression coefficients were generally closer to unity when traceability increased.

For animals in which both sire and dam were genotyped (traceability = 1.0), accuracy was approaching unity and allelic imputation error rates were very low when using at least 3,000 SNP and DAGPHASE for imputation. For animals having the sire, maternal grandsire, and some more distant ancestors genotyped (traceability >0.75), accuracy of imputation was still reasonably high (≥ 0.86 for DAGPHASE and 3,000 SNP), but lower than for animals that had both sire and dam genotyped. For animals with only the sire and some distant ancestors genotyped (traceability <0.75), the accuracy of imputation was quite low (≥ 0.71 for DAGPHASE and 3,000 SNP) and allelic imputation error rates were very high, especially when using only 384 SNP on the low-density chip. It can be concluded that imputation accuracy is high for animals that have at least sire and maternal grandsire genotyped.

Method of Imputation. In most scenarios, DAGPHASE yielded higher imputation accuracies and lower allelic imputation error rates than multivariate BLUP, except for the scenario with 384 SNP and traceability <1.0. In addition, multivariate BLUP yielded greater accuracy for animals with traceability between 0.4 and 0.55. Regression coefficients were close to unity with multivariate BLUP, but deviated more from unity with DAGPHASE. Both DAGPHASEgc and CHROMIBD were tested on the scenario MAF3000 and yielded higher accuracies than DAGPHASE with the most likely genotype. As expected, allelic imputation error rates were not lower, except for the lower traceability classes. CHROMIBD yielded the highest accuracy and lowest error rates when traceability was >0.75. Regression coefficients of CHROMIBD and DAGPHASEgc were

Table 3. Allelic imputation error rate of imputed SNP genotypes for MAF and Mid¹ scenario for different numbers of SNP for different traceability classes using different methods of imputation

Method	Scenario	SNP (no.)	Traceability					
			Overall ²	1.0	0.95–1.0	0.75–0.95	0.55–0.75	0.4–0.55
Multivariate BLUP	MAF	384	16.3	14.4	16.4	17.0	18.3	18.9
		3,000	14.6	13.0	14.6	15.2	16.4	16.9
		6,000	13.5	12.0	13.5	14.0	15.1	15.7
	Mid	384	16.3	14.4	16.4	17.0	18.3	19.0
		3,000	14.9	13.3	14.9	15.5	16.8	17.4
		6,000	13.8	12.3	13.8	14.4	15.5	16.0
DAGPHASE	MAF	384	13.57	5.93	14.65	15.45	21.26	30.14
		3,000	3.78	0.98	3.16	4.40	10.69	19.11
		6,000	2.56	0.52	1.78	2.88	9.08	17.97
	Mid	384	14.06	6.64	15.12	15.82	21.67	30.77
		3,000	3.88	1.00	3.27	4.52	10.83	19.19
		6,000	2.61	0.54	1.83	2.96	9.16	18.04
DAGPHASEgc ³	MAF	3,000	4.25	1.27	4.04	5.06	9.45	15.02
CHROMIBD	MAF	3,000	3.81	1.05	2.59	4.84	12.03	17.56

¹The SNP on the low-density chip were selected when being closest to the physical midpoint of a bin (Mid), or with the highest minor allele frequency within a bin (MAF).

²Across all animals on low-density chip.

³Gene content predicted using DAGPHASE based on genotype probabilities.

closer to unity than DAGPHASE with the most likely genotype. It can be concluded that DAGPHASE and CHROMIBD gave much better imputation accuracy than multivariate BLUP, when at least 3,000 SNP were placed on a low-density chip.

Animal-Specific Accuracies. Individual imputation accuracies may depend heavily on the traceability of the animal. To explore this, individual imputation accuracies averaged across chromosomes (MAF 3,000 scenario) were plotted against traceability in Figure 1. Results showed an increase in the range of accuracies with decreasing traceability. At traceability levels <0.91, 2 distributions of imputation accuracies were observed when using DAGPHASE. Splitting animals into groups with no parents or one or more parents genotyped (Figure 1) revealed that almost all low imputation accuracies were for animals without genotyped parents.

Within this group, no clear differences in imputation accuracy were found between animals with 0, 1, or 2 grandparents genotyped (results not shown). Imputation accuracies obtained with multivariate BLUP did not fall into 2 distributions. The above results indicate that traceability as an indicator for imputation accuracy should be used with caution.

Reliabilities of DGV

Table 5 shows the increase in reliability of DGV relative to a pedigree index when using DAGPHASE for imputation. Surprisingly, with 384 SNP on a low-density chip, the reliability of the DGV was often lower than the reliability of the pedigree index. With 3,000 SNP, approximately 87% of the increase in reliability with a 50k chip could be obtained. With 6,000 SNP, 93%

Table 4. Regression of true on imputed SNP genotypes for MAF scenarios¹ with different numbers of SNP for different traceability classes using different methods of imputation

Method	SNP (no.)	Traceability					
		Overall ²	1.0	0.95–1.0	0.75–0.95	0.55–0.75	0.4–0.55
Multivariate BLUP	384	0.966	0.970	0.966	0.962	0.942	0.895
	3,000	0.983	0.984	0.983	0.980	0.964	0.927
	6,000	0.984	0.984	0.983	0.981	0.968	0.936
DAGPHASE	384	0.575	0.805	0.543	0.517	0.338	0.126
	3,000	0.866	0.961	0.887	0.842	0.625	0.414
	6,000	0.904	0.976	0.931	0.889	0.671	0.450
DAGPHASEgc ³	3,000	0.982	0.991	1.000	0.981	0.831	1.012
CHROMIBD	3,000	0.978	0.993	0.988	0.981	0.932	0.861

¹The SNP on the low-density chip were selected with the highest minor allele frequency within a bin (MAF).

²Across all animals on low-density chip.

³Gene content predicted using DAGPHASE based on genotype probabilities.

Figure 1. Individual imputation accuracies [obtained with multivariate BLUP or DAGPHASE using the most likely genotypes (DAGPHASE) or predicted gene content (DAGPHASEgc)] averaged across chromosomes for individual animals, plotted against their traceability (minor allele frequency, 3,000 SNP scenario).

of the increase in reliability with the 50k chip could be achieved. Differences between MAF and Mid were small and not significant ($P = 0.44$, paired Student t -test). Interestingly, Mid gave slightly higher reliability than MAF, whereas the imputation accuracy of Mid was slightly lower than for MAF. The effect of imputation on reliability of DGV was different for different traits. Thus, the low-density chip needs to contain at least 3,000 SNP to obtain a higher reliability of DGV compared with a pedigree index.

Table 6 shows the increase in reliability of DGV relative to a pedigree index when using DAGPHASE, CHROMIBD, or multivariate BLUP for imputation for the scenario with 3,000 SNP selected on MAF. DAGPHASE and CHROMIBD yielded higher reliability of

DGV than multivariate BLUP, as expected based on differences in imputation accuracy. Imputation with multivariate BLUP yielded, for all traits, increases in reliability relative to a pedigree index, but much smaller increases in reliability than DAGPHASE and CHROMIBD. CHROMIBD yielded slightly lower reliabilities than DAGPHASE and DAGPHASEgc gave higher reliabilities than DAGPHASE with the most likely genotype. We conclude that DAGPHASEgc is the best method for imputation with respect to reliability of DGV.

Table 7 shows the correlations between DGV based on a low-density chip with imputation to the 50k chip and DGV based on the 50k chip for different traceability classes. The group with traceability <0.55 was too

Table 5. Reliability ($R^2 \times 100$) of the pedigree index (Ped) and increase in reliability ($\Delta R^2 \times 100$) of the direct genomic values compared with the pedigree index¹ when using a 50k chip (50,000 SNP) or a low-density chip with 384, 3,000, or 6,000 SNP selected on the highest minor allele frequency (MAF) in a bin or the middle SNP (Mid) in a bin using DAGPHASE with the most likely genotype for imputation

Trait	R^2 Ped	ΔR^2 50K chip	ΔR^2					
			MAF			Mid		
			384	3,000	6,000	384	3,000	6,000
Protein yield	15.6	23.3	−2.2	19.7	20.7	0.2	19.6	21.4
Milk yield	31.6	19.5	−10.3	14.6	17.0	−9.4	14.3	18.0
Udder depth	13.0	16.4	−3.6	13.6	14.3	2.2	13.9	14.4
Feet and legs	22.1	16.8	6.8	16.9	15.8	8.1	14.1	16.0
Udder	20.4	12.0	−3.1	10.2	10.9	−4.3	11.5	11.7
Final score	20.3	13.9	3.4	11.9	13.0	−0.7	11.7	13.5
Longevity	18.7	11.4	−4.2	9.9	10.6	0.9	10.2	10.9
SCS	28.6	21.3	−1.6	18.4	20.0	−2.2	18.4	19.0
Calving interval	28.9	16.4	2.8	15.3	15.9	2.0	17.6	17.4
Calving ease	28.8	11.6	−1.1	10.2	11.2	−6.9	12.3	10.6
Average	22.8	16.3	−1.3	14.1	14.9	−1.0	14.4	15.3

¹See equation [1].

small and therefore not included in the table. Trends were very similar to those for imputation accuracy (Table 2) and the correlation increased with higher traceability. With 384 SNP on the low-density chip, correlations were quite low, indicating that the accuracy of DGV based on imputed SNP is low. With 3,000 and 6,000 SNP, correlations were above 0.96 when traceability was at least 0.75 and when CHROMIBD and DAGPHASE were used for imputation, indicating that animal reranking due to imputation was small. For animals with lower traceability (i.e., fewer or more distant genotyped ancestors), the correlation was somewhat lower. With multivariate BLUP, correlations between the DGV were about 0.9 for all traceability classes. We can conclude that correlations between DGV based on a 50k and on a low-density chip with imputation to 50k are high; this is especially true for the higher traceability classes, indicating that re-ranking of animals between both sets of DGV is small.

Relationship Between Imputation Accuracy and Accuracy of DGV

Figure 2 shows the observed and deterministically predicted accuracy of DGV based on equation (3) for protein yield (panel A) and for the average of all traits (panel B), respectively, as a function of imputation accuracy. Predicted accuracies showed good correspondence with observed values, especially for the average of all traits (R^2 , the squared correlation between observed and predicted accuracy, was 0.96 and 0.97 for protein yield and the average of all traits, respectively). Especially for protein yield, the predicted accuracy was lower than the observed accuracy. Bias (predicted – observed values) was −0.039 and −0.025 for protein yield and the average of all traits, respectively). The observed relationship between imputation accuracy and accuracy of DGV was approximately linear, in concordance with equation [3]. Bias in accuracy could be reduced by us-

Table 6. Increase in reliability of the direct genomic values ($\Delta R^2 \times 100$) in comparison to the pedigree index ($R^2 \times 100$)¹ when using a 50k chip (50,000 SNP) or a low-density chip with 3,000 SNP selected on the highest minor allele frequency in a bin using DAGPHASE (DP), DAGPHASE with predicted gene content (DPgc), CHROMIBD (CI), or multivariate BLUP (MB) for imputation

Trait	R^2 Ped	ΔR^2				
		50K	DP	DPgc	CI	MB
Protein yield	15.6	23.3	19.7	20.3	18.2	9.9
Milk yield	31.6	19.5	14.6	14.6	14.6	7.6
Udder depth	13.0	16.4	13.6	14.5	13.6	6.5
Feet and legs	22.1	16.8	16.9	15.6	14.9	5.2
Udder	20.4	12.0	10.2	11.5	10.2	6.0
Final score	20.3	13.9	11.9	12.3	11.2	5.8
Longevity	18.7	11.4	9.9	10.6	9.2	4.4
SCS	28.6	21.3	18.4	19.0	19.6	7.9
Calving interval	28.9	16.4	15.3	16.5	15.5	7.8
Calving ease	28.8	11.6	10.2	10.8	9.9	4.0
Average	22.8	16.3	14.1	14.6	13.7	6.5

¹See equation [1].

Table 7. Correlation between direct genomic values (DGV) based on a low-density chip with imputation to the 50k chip (50,000 SNP) and DGV based on the 50K chip for different imputation methods and traceability classes (average across 10 traits)

Imputation method ¹	Traceability				
	All	0.55–0.75	0.75–0.95	0.95–1.0	1.0
DP_MAF_0384	0.827	0.726	0.739	0.742	0.906
DP_MAF_3000	0.977	0.949	0.963	0.967	0.988
DP_MAF_6000	0.987	0.972	0.981	0.982	0.993
DP_Mid_0384	0.811	0.649	0.709	0.721	0.896
DP_Mid_3000	0.975	0.948	0.961	0.965	0.988
DP_Mid_6000	0.987	0.979	0.981	0.982	0.993
CI_MAF_3000	0.985	0.932	0.971	0.983	0.992
MB_MAF_3000	0.908	0.895	0.890	0.901	0.916
DPgc_MAF_3000	0.984	0.968	0.975	0.976	0.991

¹DP = DAGPHASE, DPgc = DAGPHASE with predicted gene content, CI = CHROMIBD, MB = multivariate BLUP, MAF = SNP selected within a bin on highest minor allele frequency, Mid = SNP selected closest to the midpoint of the bin, gc = gene content.

ing a lower accuracy of imputation based on pedigree only ($r_{imp,ped}$); for example, 0.57 instead of 0.68 gave the lowest bias. For other individual traits, the agreement between predicted and observed accuracies of DGV was good; predicted accuracies were generally slightly lower than observed accuracies. We conclude that equation [3] can be used to predict the accuracies of DGV using a low-density chip with imputation to a high-density chip to understand the relationship between the accuracy of imputation and the accuracy of DGV.

DISCUSSION

Imputation Accuracy

The first aim of this study was to investigate the accuracy of imputation for animals genotyped with a low-density chip, based on 3 different marker densities and

2 types of marker selection using 3 imputation methods (DAGPHASE, CHROMIBD, and multivariate BLUP). Results were in general very similar to those of Zhang and Druet (2010), who used partly the same data and the DAGPHASE and CHROMIBD software for imputation. They obtained slightly lower error rates than in this study, especially in the lower traceability classes. In both studies, imputation error rates were lower than in Weigel et al. (2010b), who used fastPhase and IMPUTE. Zhang and Druet (2010) extensively discussed both methods and compared them to other approaches. Calus et al. (2011) showed that multivariate BLUP performs in most cases worse than Beagle and fastPhase, except at lower marker densities, which was also observed in this study. Multivariate BLUP, however, can be performed with any BLUP software package such as MiXBLUP (Mulder et al., 2010a). To make results more

Figure 2. Observed (obs) and deterministically (det) predicted accuracy of direct genomic values (DGV) (see equation [3]) for protein yield (A) and the average of all traits (B) as a function of imputation accuracy across different scenarios (minor allele frequency, 384, 3,000, and 6,000 SNP) and different methods of imputation (DAGPHASE, multivariate BLUP, and CHROMIBD).

Table 8. Imputation accuracy for different traceability classes and computation time for chromosome 1 (BTA1) using a low-density chip with 3,000 SNP selected on the highest minor allele frequency in a bin using DAGPHASE, CHROMIBD, and Beagle

Method	Sc ¹	Sh ²	Accuracy					Time (h)
			1	0.95–1.0	0.75–0.95	0.55–0.75	0.4–0.55	
DAGPHASEgc ³	2	0.1	0.983	0.949	0.921	0.823	0.662	20.9
DAGPHASEgc	4	0.2	0.975	0.915	0.884	0.784	0.619	4.6
DAGPHASE	2	0.1	0.945	0.903	0.862	0.741	0.580	22.5
DAGPHASE	4	0.2	0.883	0.840	0.793	0.666	0.518	4.8
CHROMIBD	2	0.1	0.985	0.964	0.917	0.768	0.566	19.2
Beagle	1	0	0.964	0.956	0.933	0.917	0.880	106.2
Beagle_DP ⁴	1	0	0.964	0.945	0.907	0.774	0.605	~170 ⁵
Beagle_DPGc ⁴	1	0	0.967	0.957	0.930	0.824	0.669	~170 ⁵

¹Scale parameter used for Beagle.
²Shift parameter used in Beagle.
³Gene content predicted using DAGPHASE based on genotype probabilities.
⁴Beagle run combined with one DAGPHASE run to integrate linkage information.
⁵Run on another computer due to higher memory requirements; time is approximated.

comparable to the literature (e.g., Berry and Kearney, 2011), DAGPHASE and CHROMIBD were compared with Beagle on imputation accuracy for chromosome 1. Results in Table 8 showed that Beagle had slightly lower imputation accuracies than DAGPHASEgc and CHROMIBD for animals with both parents genotyped, but substantially higher imputation accuracies for animals with more distantly related genotyped ancestors. This shows that use of linkage information improves accuracy of imputation, but that linkage disequilibrium information is better modeled by Beagle. Beagle uses all haplotypes to estimate DAG and better scale and shift parameters and optimally weights different haplotypes at each iteration. In terms of computing time, DAGPHASE and CHROMIBD outperformed Beagle (Table 8). Running Beagle on the whole genome is very time consuming, but once Beagle has made a DAG, the same DAG can be used again for imputing some new animals, reducing computing time dramatically. This can be done with DAGPHASE as well.

In this study, the number of SNP on the low-density chip varied. Results indicate that at least 3,000 markers should be placed on the low-density chip to obtain reasonable imputation accuracies and therefore also sufficiently high accuracies of the DGV. Also, Zhang and Druet (2010) and Weigel et al. (2010a,b) concluded that 3,000 SNP are necessary on a low-density chip. The method of SNP selection hardly affected imputation accuracy, although slightly higher accuracies were obtained when SNP were selected on MAF, because the information content increases when the MAF increases.

Traceability

Imputation accuracy is dependent on the number of genotyped ancestors. If animals have only their sire genotyped and no other ancestors, imputation accuracy

is poor and the predictive ability of the DGV is low as a result. In this study, as in Druet et al. (2010) and Zhang and Druet (2010), traceability was used as a measure of how many ancestors are genotyped and how distant these ancestors are. Traceability can be interpreted as the probability that a given haplotype of an animal is also in the reference data used for imputation (Druet et al., 2010). Based on Figure 1, it can be concluded that traceability as a predictor of imputation accuracy should be used with caution. Results of DAGPHASE are sensitive to whether parents are genotyped or not, which resulted in 2 distributions of accuracy: one for animals where at least 1 parent was genotyped and one for animals with no parents genotyped, similar to using the maximum additive genetic relationship in Habier et al. (2010). This indicates that not only is the value of the traceability important, but also the number of generations between contributing ancestors and the imputed animal. Another factor that potentially affects imputation accuracy is the number of meiotic events between the animal and the ancestor that has the same haplotype in the reference data, because this affects the probability that one or more recombinations took place in the haplotype. Accounting for recombination in the traceability measure, that is, $T_c = \sum \left(\frac{1}{2} \times (1 - c) \right)^n$, where T_c is the corrected traceability and c is the recombination probability, could not explain, however, the bimodal distribution of animal specific accuracies in Figure 1. Therefore, it is suggested to use both traceability and the number of parents genotyped as predictors of imputation accuracy.

Animal-Specific Accuracy

The animal-specific imputation accuracies allow the identification of variation in imputation accuracy.

Animal-specific accuracies, however, are biased upward for various reasons. The most obvious reason is that allele frequencies affect the level of the true and predicted gene contents and, as such, bias the animal-specific accuracies. The bias can be removed by subtracting the mean gene content from the true and predicted gene contents. Bias is greater when the imputation accuracy is lower but is constant given the imputation accuracy. Therefore, the bias is not expected to strongly affect the distribution of animal-specific imputation accuracies. Furthermore, due to using linkage information and pedigree information for imputation, imputation results of flanking markers are correlated, which inflates the animal-specific accuracy, but this is not expected to affect the distribution of animal-specific imputation accuracies between animals. Therefore, animal accuracy is a helpful tool in identifying animals with low imputation accuracy.

Reliability of DGV

The second aim of the study was to investigate the effect of imputation on the reliability of DGV. The effect of imputation on the reliability was detrimental in the case of a low-density chip with 384 SNP when DAGPHASE was used to impute the most likely genotype. It seems that due to the magnitude of the imputation errors, the noise added is greater than the benefit from using SNP to predict breeding values. Also, Weigel et al. (2010a) reported that the accuracy of DGV was lower when using imputed genotypes in the case of a chip with 366 SNP than when estimating SNP effects and predicting DGV based only on those 366 SNP. The decrease in reliability of DGV with imputation in this study was very similar to that in Weigel et al. (2010a).

Although allelic imputation error rates were very small, especially for the higher traceability classes, the effect on the accuracy of DGV was quite substantial. Huang et al. (2009) reported that the power in genome-wide association studies decreases substantially when using imputed genotypes, even at low allelic imputation error rates. Lin et al. (2010) concluded that allelic imputation error rates or the proportion of correct genotypes are poor measures of imputation accuracy because of their dependence on minor allele frequency, which was not, however, observed in our data (results not shown). The imputation accuracy, calculated as the correlation between true and imputed genotypes does not depend on minor allele frequency. Furthermore, imputation accuracy as well as the reliability of DGV was higher when using predicted gene contents instead of the most likely genotype, which was not reflected, however, in the allelic imputation error rates. In addition, we showed that the relationship between imputation accuracy and the

accuracy of DGV is approximately linear and can be approximated by equation [3], a modification of equation [3] by Meuwissen (2009). Therefore, it is suggested that imputation accuracy is a more suitable measure of precision of imputation than allelic imputation error rates, and therefore a better predictor of the accuracy and reliability of DGV based on imputed genotypes than allelic imputation error rates.

Results suggest that using genotype probabilities to calculate gene contents or “dosage” is likely beneficial for the accuracy of imputation as well as the reliability of DGV. By using the predicted gene content, the uncertainty of imputation is taken into account and as a result, the gene content is regressed toward the mean genotype. Therefore, the DGV are also regressed toward the mean. Only those SNP with high imputation accuracy will have a positive effect on the reliability of DGV, whereas those with a low accuracy will not add much to the reliability in addition to the pedigree information.

CONCLUSIONS

This study shows that it is feasible to obtain high reliability of DGV for animals genotyped with a low-density chip combined with imputation to a 50k chip. Low-density chips should contain at least 3,000 SNP. Selection of SNP on a low-density chip using either the middle SNP or the SNP with the highest MAF in equally spaced bins has a negligible effect on imputation accuracy. We recommend the use of advanced imputation software for imputation, such as Beagle, DAGPHASE, or CHROMIBD. Moreover, we recommend using predicted gene contents rather than the most likely genotype to account for uncertainty of imputation leading to higher imputation accuracy and reliability of DGV than using the most likely genotype. The effect of differences in imputation accuracy on reliability of DGV can be evaluated using the deterministic equation derived in this study.

ACKNOWLEDGMENTS

CRV BV (Arnhem, the Netherlands) is acknowledged for financial support to H. A. M., M. P. L. C., and T. D. and for making the genotype data available for this study. T. D. is research associate from the Fonds de la Recherche Scientifique (FNRS). Further financial support to T. D. was provided by grants of the Service Public de Wallonie.

REFERENCES

- Berry, D. P., and J. F. Kearney. 2011. Imputation of genotypes from low- to high-density genotyping platforms and implications for genomic selection. *Animal* 5:1162–1169.

- Browning, B. L., and S. R. Browning. 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated animals. *Am. J. Hum. Genet.* 84:210–223.
- Browning, S. R. 2008. Missing data imputation and haplotype phase inference for genome-wide association studies. *Hum. Genet.* 124:439–450.
- Browning, S. R., and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81:1084–1097.
- Calus, M. P. L. 2010. Genomic breeding value prediction: Methods and procedures. *Animal* 4:157–164.
- Calus, M. P. L., T. H. E. Meuwissen, A. P. W. De Roos, and R. F. Veerkamp. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178:553–561.
- Calus, M. P. L., R. F. Veerkamp, and H. A. Mulder. 2011. Imputation of missing SNP genotypes using a multivariate mixed model framework. *J. Anim. Sci.* 89:2042–2049.
- Charlier, C., W. Coppieters, F. Rollin, D. Desmecht, J. S. Agerholm, N. Cambisano, E. Carta, S. Dardano, M. Dive, C. Fasquelle, J. C. Frennet, R. Hanset, X. Hubin, C. Jorgensen, L. Karim, M. Kent, K. Harvey, B. R. Pearce, P. Simon, N. Tama, H. Nie, S. Vandeputte, S. Lien, M. Longeri, M. Fredholm, R. J. Harvey, and M. Georges. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat. Genet.* 40:449–454.
- Druet, T., and F. P. Farnir. 2011. Modeling of identity-by-descent processes along a chromosome between haplotypes and their genotyped ancestors. *Genetics* 188:409–419.
- Druet, T., and M. Georges. 2010. A hidden Markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. *Genetics* 184:789–798.
- Druet, T., C. Schrooten, and A. P. W. De Roos. 2010. Imputation of genotypes from different single nucleotide polymorphism panels in dairy cattle. *J. Dairy Sci.* 93:5443–5454.
- Gengler, N., S. Abras, C. Verkenne, S. Vanderick, M. Szydlowski, and R. Renaville. 2008. Accuracy of prediction of gene content in large animal populations and its use for candidate gene detection and genetic evaluation. *J. Dairy Sci.* 91:1652–1659.
- Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal* 1:21–27.
- George, E. I., and R. E. McCulloch. 1993. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* 88:881–889.
- Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2009. Genomic selection using low-density marker panels. *Genetics* 182:343–353.
- Habier, D., J. Tetens, F.-R. Seefried, P. Lichtner, and G. Thaller. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* 42:5.
- 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.
- Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* 38:226–231.
- Howie, B. N., P. Donnelly, and J. Marchini. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5:e1000529.
- Huang, L., C. Wang, and N. A. Rosenberg. 2009. The relationship between imputation error and statistical power in genetic association studies in diverse populations. *Am. J. Hum. Genet.* 85:692–698.
- Interbull. 2010. Interbull validation test for genomic evaluations—GEBV test. Accessed Mar. 29, 2011. http://www.interbull.org/images/stories/GEBV_validationtest_June2010.pdf.
- Li, Y., C. Willer, S. Sanna, and G. Abecasis. 2009. Genotype imputation. *Annu. Rev. Genomics Hum. Genet.* 10:387–406.
- Lin, P., S. M. Hartz, Z. H. Zhang, S. F. Saccone, J. Wang, J. A. Tischfield, H. J. Edenberg, J. R. Kramer, A. M. Goate, L. J. Bierut, and J. P. Rice. 2010. A new statistic to evaluate imputation reliability. *PLoS ONE* 5:e9697.
- Marchini, J., and B. Howie. 2010. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* 11:499–511.
- Meuwissen, T. H. E. 2009. Accuracy of breeding values of “unrelated” individuals predicted by dense SNP genotyping. *Genet. Sel. Evol.* 41:35.
- Meuwissen, T. H. E., and M. E. Goddard. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multi-trait data. *Genet. Sel. Evol.* 2004:261–279.
- 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.
- Moser, G., M. S. Khatkar, B. Hayes, and H. Raadsma. 2010. Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP. *Genet. Sel. Evol.* 42:37.
- Mulder, H. A., M. Lidauer, I. Strandén, E. A. Mantysaari, M. H. Pool, and R. F. Veerkamp. 2010a. MiXBLUP manual. Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Lelystad, the Netherlands.
- Mulder, H. A., T. H. E. Meuwissen, M. P. L. Calus, and R. F. Veerkamp. 2010b. The effect of missing marker genotypes on the accuracy of gene-assisted breeding value estimation: A comparison of methods. *Animal* 4:9–19.
- Pszczola, M., H. A. Mulder, and M. P. L. Calus. 2011. Effect of enlarging the reference population by (un)genotyped animals on the accuracy. *J. Dairy Sci.* 94:431–441.
- Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78:629–644.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggins, 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.
- Weigel, K. A., G. de los Campos, O. Gonzalez-Recio, H. Naya, X. L. Wu, N. Long, G. J. M. Rosa, and D. Gianola. 2009. Predictive ability of direct genomic values for lifetime net merit of Holstein sires using selected subsets of single nucleotide polymorphism markers. *J. Dairy Sci.* 92:5248–5257.
- Weigel, K. A., G. De los Campos, A. I. Vazquez, G. J. M. Rosa, D. Gianola, and C. P. Van Tassell. 2010a. Accuracy of direct genomic values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. *J. Dairy Sci.* 93:5423–5435.
- Weigel, K. A., C. P. Van Tassell, J. R. O’Connell, P. M. VanRaden, and G. R. Wiggins. 2010b. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using reference panels and population-based imputation algorithms. *J. Dairy Sci.* 93:2229–2238.
- Wiggins, G. R., T. S. Sonstegard, P. M. Vanraden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, F. S. Schenkel, and C. P. Van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92:3431–3436.
- Zhang, Z., and T. Druet. 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. *J. Dairy Sci.* 93:5487–5494.

APPENDIX

Traceability

Traceability was calculated as $T = \Sigma a$; where a is the additive genetic relationship between the animal of interest and the closest genotyped ancestor in a branch of the pedigree (calculated as $\frac{1}{2}^n$, where n is the number of generations between genotyped ancestor and animal), assuming no inbreeding. The summation was

across pedigree branches and stopped in a branch when a genotyped ancestor was found. In the cases that both parents were genotyped, summation stopped after one generation of pedigree, but in the case that only males were genotyped, the maternal branches were traced back to genotyped sires. Genotyped ancestors were traced back at maximum 7 generations in the pedigree.

When calculations are performed using the a -values from the numerator relationship matrix that includes inbreeding, this results in T values >1 for some animals due to inbreeding. Because of difficulty in interpretation when including inbreeding, we choose to calculate T based on a values that did not account for inbreeding.

Relationship Between Imputation Accuracy and Accuracy of DGV

Equation [2] (Meuwissen, 2009) was modified to predict r_{DGV} when some markers are imputed and some are on a low-density chip. If an animal is not genotyped at all and the genotypes were imputed based on genotyped ancestors, the accuracy of imputed genotypes could be predicted by selection index theory, assuming a heritability of 1.0 for the genotype. For an animal with both parents genotyped, this accuracy is $\sqrt{1/2}$. For the “average” genotyped animal in our data, the expected accuracy was 0.68. If an animal is not genotyped and its genotypes are imputed based on genotyped ancestors, the accuracy of DGV based on imputed genotypes is expected to be equal to the accuracy of a pedigree index ($r_{DGV,imp} = r_{ped}$); that is, the lower bound of accuracy of DGV. This expectation is supported by simulations of Pszczola et al. (2011), who found that the accuracy of DGV for animals not genotyped but

using imputed genotypes was approximately equal to accuracy of BLUP EBV. For animals genotyped on the low-density chip, one can differentiate between the SNP that are on the low-density chip and the ones that are imputed. Therefore, the weighted average of the additional imputation accuracy, relative to a scenario with only pedigree information, is $p(1 - r_{imp,ped}) + (1 - p)(r_{imp} - r_{imp,ped})$, where 1 is the imputation accuracy of SNP genotyped on the low-density chip, r_{imp} is the imputation accuracy of imputed SNP, $r_{imp,ped}$ is the imputation accuracy based on pedigree alone not due to using markers on the low-density chip (set at 0.68, the average in this data set), and p is the proportion of SNP on the low-density chip relative to the high-density chip. When combining these upper (i.e., equation [2]) and lower bounds of accuracy (i.e., $r_{DGV,imp} = r_{ped}$) and the weighted imputation accuracy of SNP for animals genotyped with a low-density chip [i.e., $p(1 - r_{imp,ped}) + (1 - p)(r_{imp} - r_{imp,ped})$], the accuracy of DGV based on a low-density chip and with imputation to a 50k chip may be approximated as

$$r_{DGV,imp} = r_{ped} + (1 - r_{ped}) \frac{\rho}{(1 - r_{imp,ped})} [p(1 - r_{imp,ped}) + (1 - p)(r_{imp} - r_{imp,ped})], \quad [3]$$

where $r_{DGV,imp}$ is the accuracy of the DGV based on a low-density chip where most SNP are imputed and ρ is the proportion of accuracy of DGV that is not explained by r_{ped} when using a 50k chip. When all markers are on the chip ($p = 1$), equation [3] reduces to equation [2]. When an animal is not genotyped ($p = 0$ and $r_{imp} = r_{imp,ped}$), equation [3] reduces to $r_{DGV,imp} = r_{ped}$.