# Effect of reference population size and available ancestor genotypes on imputation of Mexican Holstein genotypes[1]

**A. García-Ruiz,\*† F. J. Ruiz-Lopez,\*†[2] G. R. Wiggans,‡ C. P. Van Tassell,‡ and H. H. Montaldo§**
*Facultad de Estudios Superiores Cuautitlán, Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México, México, DF 04510, México
†Centro Nacional de Investigación en Fisiología y Mejoramiento Animal, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Ajuchitlán, Querétaro 76280, México
‡Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350
§Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México, México, DF 04510, México

## ABSTRACT

The effects of reference population size and the availability of information from genotyped ancestors on the accuracy of imputation of single nucleotide polymorphisms (SNP) were investigated for Mexican Holstein cattle. Three scenarios for reference population size were examined: (1) a local population of 2,011 genotyped Mexican Holsteins, (2) animals in scenario 1 plus 866 Holsteins in the US genotype database (GDB) with genotyped Mexican daughters, and (3) animals in scenario 1 and all US GDB Holsteins (338,073). Genotypes from 4 chip densities (2 low density, 1 mid density, and 1 high density) were imputed using findhap (version 3) to the 45,195 markers on the mid-density chip. Imputation success was determined by comparing the numbers of SNP with 1 or 2 alleles missing and the numbers of differently predicted SNP (conflicts) among the 3 scenarios. Imputation accuracy improved as chip density and numbers of genotyped ancestors increased, and the percentage of SNP with 1 missing allele was greater than that for 2 missing alleles for all scenarios. The largest numbers of conflicts were found between scenarios 1 and 3. The inclusion of information from direct ancestors (dam or sire) with US GDB genotypes in the imputation of Mexican Holstein genotypes increased imputation accuracy by 1 percentage point for low-density genotypes and by 0.5 percentage points for high-density genotypes, which was about half the gain found with information from all US GDB Holsteins. A larger reference population and the availability of genotyped ancestors improved imputation; animals with genotyped parents in a large reference population had higher imputation accuracy than those with no or few genotyped relatives in a small reference population. For small local populations, including genotypes from other related populations can aid in improving imputation accuracy.

**Key words:** imputation, Mexican Holstein, ancestor genotype, reference population

## INTRODUCTION

Genomic selection is a recent technology that has enhanced genetic improvement in dairy cattle but its success relies on numerous factors. The most important are methodology for calculating marker effects (Hayes et al., 2009; Daetwyler et al., 2010; VanRaden et al., 2010), marker panel density (Solberg et al., 2008; Hayes et al., 2009), trait heritability (Hayes et al., 2009), and number of genotyped animals in the reference population (Hayes et al., 2009). To improve the reliability of genomic predictions, increasing the number of genotyped animals is more important than using higher density panels (VanRaden et al., 2010). In developing countries such as Mexico, genotyping is still expensive and high-density marker panels are generally more expensive than low-density panels. Using low-density panels is an alternative that may lead to more genotyped animals.

Because genomic technology has evolved rapidly, the number and sets of markers to be used for genomic prediction have changed over time (Hayes et al., 2009). To combine information based on different markers or marker densities, missing markers for animals genotyped with lower density panels are imputed from genotype information of relatives or from haplotypes of animals genotyped with higher density panels (Druet et al., 2010; VanRaden et al., 2010). Druet et al. (2010) showed that imputation efficiency is higher when the size of the reference population, marker density, and expected proportion of the genome inherited from the reference population are increased. Bouwman et al. (2014)

reported that the number of genotyped relatives has a direct effect on imputation accuracy for animals without genotypes, and Khatkar et al. (2012) demonstrated the importance of considering genotyped relatives when imputing low-density genotypes, particularly when the sire's genotype was not included in the reference genotypes. Although different imputation methods with high accuracy have been implemented (Browning and Browning, 2011; VanRaden et al., 2011; Hickey et al., 2012), the choice of the optimal method depends on population structure (Johnston et al., 2011).

Using imputation as part of genomic selection reduces genotyping costs and increases both the size of the reference population and the number of markers for which effects are estimated, which increases the reliability of genetic predictions and consequently the expected genetic improvement. Reducing genotyping cost also makes the technology more accessible to breeders (Berry and Kearney, 2011; VanRaden et al., 2011). For dairy cattle, reliability of genomic predictions using imputation varies according to trait and according to population size and structure. Reliability improvements of approximately 2 percentage points have been reported in simulation studies that included a set of 500,000 SNP per imputed genotype compared with a 50,000-marker subset (VanRaden et al., 2011).

Canadian and US dairy cattle genetics are widely used around the world. Using genotypes from those animals to impute local genotypes would increase the size of the local reference population considerably and could increase the accuracy of imputing genotypes, thus improving genomic evaluations in many countries. Such an improvement should be possible for Mexico, because Mexican Holstein breeders have depended heavily on US and Canadian genetics. The primary objective of this study was to determine the effect on imputation of Mexican Holstein genotypes from increasing the size of the Mexican Holstein reference population by adding Holsteins from the US genotype database. The effect of relatedness of animals in the reference population on imputation also was evaluated.

## MATERIALS AND METHODS

### Scenarios

Three imputation scenarios were defined based on source and number of genotyped animals. For scenario 1, only genotypes of a local population of 2,011 genotyped Mexican Holsteins were included. For scenario 2, genotypes of animals in scenario 1 plus genotypes of 866 Holsteins in the US genotype database (**GDB**) with genotyped Mexican daughters were included. For scenario 3, genotypes of animals in scenario 1 and all

US GDB Holsteins available at the time of the study (338,073) were included.

### Data

*Genotypes.* The genotyped Mexican Holstein population (scenario 1) included 1,971 cows and 40 sires. For cows, 183 were genotyped with the Illumina BovineLD BeadChip v1.1 (**6K**; Illumina Inc., 2013), 277 with the GeneSeek Genomic Profiler BeadChip v1 (**9K**; Neogen Corp., 2013a), 686 with the Illumina BovineSNP50 BeadChip v2 (**50K**; Illumina Inc., 2011), and 825 with the GeneSeek Genomic Profiler HD BeadChip (**77K**; Neogen Corp., 2013b). All Mexican sires had 50K genotypes. Of the US GDB Holsteins, 839 bulls and 47 cows had genotyped daughters in Mexico (scenario 2). All US GDB dams of genotyped Mexican daughters had 50K genotypes; the US GDB sires included 533 US, 270 Canadian, and 22 European bulls with 50K genotypes and 10 US and 4 Canadian bulls with 77K genotypes. For scenario 3, 338,073 US GDB genotypes were included.

*Pedigrees.* Two different pedigree files were used in the analysis: 27,625 animals for scenarios 1 and 2; 938,662 animals for scenario 3.

### Imputation

Missing genotypes were predicted by combining population and pedigree haplotypes with findhap (version 3) software (VanRaden, 2015). The imputation goal for all scenarios was to fill in any missing genotypes for the 45,195 SNP from the 50K chip that were then being used in US genomic evaluations. Of those 45,195 SNP, 6,842 were included in 6K genotypes, 8,196 in 9K genotypes, and 28,048 in 77K genotypes, and a different number of SNP was imputed for each chip density: 38,353 for 6K genotypes, 36,999 for 9K genotypes, and 17,147 for 77K genotypes. For imputation, genotypes were first coded as 0 = BB, 1 = AB, 2 = AA, or 5 = both alleles unknown. Then, SNP genotypes for each scenario were imputed, and the results were compared using SAS software (version 9.3; SAS Institute Inc., Cary NC).

A SNP genotype cannot always be determined through imputation. If only 1 parental allele could be determined (i.e., B and unknown allele or A and unknown allele), then the SNP genotype was designated as M1. If neither parental contribution could be determined (i.e., both alleles missing), the SNP genotype was designated as M2. Coefficients of determination ($R^2$) calculated as squared correlations between estimated and true SNP genotypes (VanRaden et al., 2011), percentages of missing alleles (M1 and M2), and differenc-

es in imputed SNP (conflicts) between scenarios were used to determine the success of genomic imputation (imputation accuracy). Khatkar et al. (2012) reported that increasing the number of reference animals has a positive effect on imputation. Therefore, scenario 3 was expected to be the most accurate for imputation, and scenarios 1 and 2 were compared with scenario 3. Conflict frequency was calculated as the percentage of 50K SNP that were conflicts between scenarios.

### Ancestor Genotypes

Genotyped animals were grouped by category of genotyped ancestors: animals without parental genotypes; animals with dam or sire genotyped; and animals with dam, sire, and a grandsire genotyped. Animals with other combinations of genotyped ancestors were not considered because of small group sizes. Numbers of M1 and M2 SNP genotypes and SNP conflicts were compared using SAS within chip density by ancestor group. Only scenarios 2 and 3 were analyzed, because scenario 1 had too few genotyped ancestors to form groups.

### Comparison of Imputed and True Genotypes

To determine the accuracy of imputed genotypes from low-density chips (6K and 9K), 10 groups were formed by random sampling of the 686 cows without replacement; each group contained 10 cows with 50K genotypes. Group size was limited to 10 cows to avoid excessive reduction in imputation accuracy from removing many high-density genotypes that are critical for accurate SNP imputation; a group size of 10 cows was considered large enough that a single animal could not have undue influence. For each animal, only 6K or 9K SNP were included in the 50K genotype, and the genotype was filled through imputation and compared with the original 50K genotype. The same animal groups were used for both chip densities. The total numbers of missing SNP (**MM**), which was calculated as M1 + M2, and conflicts were used to compare true and imputed genotypes.

## RESULTS AND DISCUSSION

### Imputation Accuracy

Imputation accuracy was improved by increasing genotype chip density and the number of reference animals (Figure 1). Imputation $R^2$ using only local genotypes (scenario 1) were 96, 96, 99, and 99% when imputing from 6K, 9K, 50K, and 77K genotypes, respec-

tively. When information from US GDB Holsteins with genotyped Mexican daughters was added to the Mexican genotypes (scenario 2), imputation $R^2$ increased by almost 1 percentage point for 6K and 9K genotypes and by half a percentage point for 77K genotypes. When all US GDB Holstein data were included (scenario 3) and compared with the results for scenario 1, an increase of almost 3 percentage points was observed for 6K and 9K genotypes and 1 percentage point for 77K genotypes. As expected, no increase in imputation $R^2$ was found for 50K genotypes in any scenario because of the small number of SNP that actually were imputed. These results were consistent with those reported in other studies using the same (Wiggans et al., 2012) or different (Zhang and Druet, 2010; Johnston et al., 2011; Khatkar et al., 2012) imputation methods. All studies reported higher imputation accuracy with an increased number of reference animals.

### Missing Alleles

Percentages of 50K alleles that were missing decreased as more animals were added to the reference population for both M1 and M2 SNP genotypes (Table 1) regardless of chip density used for genotyping. Numbers of animals with missing alleles in $\geq 1$ SNP genotype followed that pattern only for 6K, 9K, and 77K M2 SNP genotypes. As expected, the correspond-
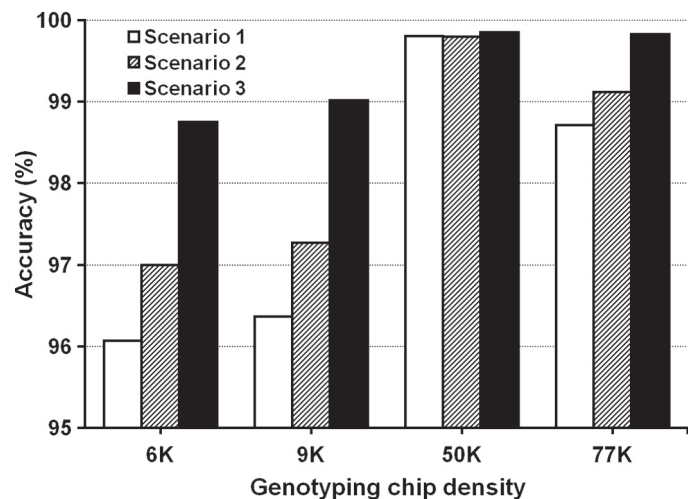


**Figure 1.** Accuracy of imputation to 45,195 (50K) markers for genotypes based on 4 genotyping chip densities (usable markers in common with 50K: 6K = 6,842; 9K = 8,196; 50K = 45,195; and 77K = 28,048) and 3 reference population scenarios (1 = 2,011 Mexican genotyped Holsteins; 2 = 2,011 Mexican genotyped Holsteins plus 886 related Holsteins in the US genotype database; and 3 = 2,011 Mexican genotyped Holsteins and 338,073 Holsteins in the US genotype database); accuracy was measured as the squared correlation between estimated and true genotypes and expressed as a percentage.

**Table 1.** Percentages of 45,195 (50K) genotype alleles that were missing[1] and numbers of animals with missing alleles by genotyping chip density and reference population scenario

| Genotyping chip density[2] | Scenario[3] | Animals (no.) | Missing 50K alleles (%) | | Animals with missing alleles at ≥1 SNP genotype (no.) | |
|---|---|---|---|---|---|---|
| | | | M1 | M2 | M1 | M2 |
| 6K | 1 | 183 | 4.75 | 0.92 | 183 | 183 |
| | 2 | 183 | 3.16 | 0.05 | 183 | 180 |
| | 3 | 183 | 0.38 | 0.01 | 183 | 26 |
| 9K | 1 | 277 | 6.03 | 0.94 | 277 | 277 |
| | 2 | 277 | 3.58 | 0.04 | 277 | 217 |
| | 3 | 277 | 0.22 | 0.02 | 277 | 25 |
| 50K | 1 | 726 | 0.17 | 0.86 | 726 | 726 |
| | 2 | 1,598 | 0.21 | 0.03 | 896 | 802 |
| | 3 | 1,598 | 0.01 | 0.01 | 690 | 196 |
| 77K | 1 | 825 | 4.36 | 1.13 | 820 | 820 |
| | 2 | 839 | 2.23 | 0.07 | 835 | 739 |
| | 3 | 839 | 0.13 | 0.02 | 815 | 153 |

[1]M1 = 1 allele missing; M2 = both alleles missing.

[2]Usable markers in common with 45,195 (50K): 6K = 6,842; 9K = 8,196; 77K = 28,048..

[3]Scenario 1 = 2,011 Mexican genotyped Holsteins; scenario 2 = 2,011 Mexican genotyped Holsteins plus 886 related Holsteins in the US genotype database; scenario 3 = 2,011 Mexican genotyped Holsteins and 338,073 Holsteins in the US genotype database.

ing M1 and M2 percentages by scenario were lower for 50K genotypes than for SNP genotypes from other chip densities, especially for M1 SNP. These results can be explained partly by the imputation process, which first divides each chromosome into segments, determines the haplotypes, and then matches genotypes with the haplotype list (VanRaden et al., 2010, 2011). When the original genotype has more markers, haplotype matching is more precise and imputation accuracy tends to be higher. In addition, the proportion of SNP to be imputed decreases as chip density increases.

The percentage of SNP genotypes designated as M1 was larger than that for M2 for all scenarios regardless of chip density, except for animals with 50K genotypes in scenario 1. The M1 percentage was largest for 9K genotypes followed by 6K and 77K genotypes for scenarios 1 and 2. For scenario 3, the M1 percentage was largest for 6K genotypes followed by 9K and 77K genotypes. The M2 percentage generally decreased as the number of SNP to be imputed increased for all scenarios.

### Conflicts

The number of detected conflicts increased as the number of genotyped animals or the number of imputed SNP (Table 2) increased. Conflict frequency never exceeded 3.8%, possibly because the 6K chip was the lowest density chip studied. Conflicts have a greater effect on evaluation reliability than do missing markers because missing markers are replaced by population

allele frequencies during the calculation of genomic evaluations (Wiggans et al., 2011).

### Ancestor Genotypes

For animals grouped by availability of genotyped ancestors, Table 3 shows differences between scenarios 2 and 3 for M1, M2, and conflict numbers by genotyping chip density and ancestor genotype group. Imputation problems for scenario 2 were more numerous with fewer genotyped ancestors, a pattern also reported by Johnston et al. (2011), who concluded that the number of correctly imputed genotypes increases as relatedness between genotyped ancestors and animals to be imputed increases. Other studies on genotype imputation from low to high density (e.g., Ma et al., 2013; Carvalheiro et al., 2014) reported only marginal benefits on imputation accuracy as the relatedness between imputed and reference animals increased. The benefit in this study may be caused partly by the imputation method used. In a comparison of 5 methods for imputing Swedish and Finnish Red Cattle marker genotypes, Ma et al. (2013) reported that IMPUTE2 (Howie et al., 2009) and Beagle (Browning and Browning, 2009) had the greatest imputation accuracy; however, the effect of relatedness on percentage of correctly imputed alleles was largest for findhap (version 2; VanRaden, 2015) and other imputation methods that used pedigree information. Carvalheiro et al. (2014) used Beagle as well as FImpute (Sargolzaei et al., 2011) to study Nelore cattle; although they found higher imputation accuracies with

**Table 2.** Numbers of imputed alleles for scenarios 1 and 2 that conflict with those imputed for scenario 3[1] and conflict frequencies by genotyping chip density

| Genotyping chip density[2] | Conflicts (no.) | | Conflict frequency[3] (%) | |
|---|---|---|---|---|
| | Scenario 1 | Scenario 2 | Scenario 1 | Scenario 2 |
| 6K | 1,718 | 325 | 3.80 | 0.72 |
| 9K | 1,705 | 506 | 3.77 | 1.12 |
| 50K | 28 | 16 | 0.06 | 0.04 |
| 77K | 746 | 689 | 1.65 | 1.52 |

[1]Scenario 1 = 2,011 Mexican genotyped Holsteins; scenario 2 = 2,011 Mexican genotyped Holsteins plus 886 related Holsteins in the US genotype database; scenario 3 = 2,011 Mexican genotyped Holsteins and 338,073 Holsteins in the US genotype database.

[2]Usable markers in common with 45,195 (50K): 6K = 6,842; 9K = 8,196; 77K = 28,048.

[3]Conflict frequency = 100(number of conflicts/45,195).

FImpute than with Beagle, they concluded that the optimal imputation method is unique to the genotyping chip and the population studied. When Khatkar et al. (2012) examined the effect of including sire genotype in the imputation process, a slightly lower allelic error rate was found when sire was included (2.61% with sire vs. 3.34% without sire). Similar findings were observed in this study, although the effect of including a parent was less for 50K and 77K genotypes than for 6K and 9K genotypes (Table 3).

In addition to reference population size and relatedness of genotypes, chip density also determined imputation success between scenarios 2 and 3. Within genotyped ancestor group, animals with 50K genotypes had the fewest imputation problems, and differences between scenarios 2 and 3 decreased slightly with more ancestors with 50K genotypes (Table 3). Because imputation improvement pertains only to SNP that are not called during genotyping and the 50K chip contains all SNP, the number of correctly filled SNP increased

slightly when genotypes for immediate family members were available.

The largest differences between scenarios 2 and 3 for imputation problems (17.2% of SNP) were found for animals with 6K genotypes and no parental genotypes (Table 3). However, comparison with the other genotyped ancestor groups was not possible because only animals with no parental genotypes had 6K genotypes. For animals with 9K genotypes, those with only sire or dam genotyped had more imputation problems than those with a grandsire also genotyped (14.8 vs. 3.3% of SNP). Animals with 77K genotypes had imputation problems for 7.8% of SNP for scenario 2 compared with scenario 3 when no parental genotypes were available, which decreased to 6.7% when a dam or sire genotype was available and to 1.6% when genotypes were available for dam, sire, and a grandsire. When the reference population is small, related genotyped animals with high-density genotypes should be included to increase imputation success.

**Table 3.** Differences in the numbers and percentages of missing alleles[1] and imputed allele conflicts for scenario 2 compared with scenario 3[2] by availability of genotyped ancestors and genotyping chip density

| Genotyped ancestors | Genotyping chip density[3] | Animals (no.) | Imputation difference | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | M1 | | M2 | | Conflicts | | All | |
| | | | No. | % | No. | % | No. | % | No. | % |
| Neither dam nor sire | 6K | 18 | 2,897 | 6.4 | 40 | 0.1 | 4,827 | 10.7 | 7,764 | 17.2 |
| | 50K | 12 | 159 | 0.4 | 20 | <0.1 | 168 | 0.4 | 347 | 0.8 |
| | 77K | 33 | 1,627 | 3.6 | 62 | 0.1 | 1,837 | 4.1 | 3,526 | 7.8 |
| Dam or sire | 9K | 23 | 2,896 | 6.4 | 20 | <0.1 | 3,762 | 8.3 | 6,678 | 14.8 |
| | 50K | 299 | 124 | 0.3 | 10 | <0.1 | 136 | 0.3 | 270 | 0.6 |
| | 77K | 156 | 1,444 | 3.2 | 34 | 0.1 | 1,560 | 3.5 | 3,038 | 6.8 |
| Dam, sire, and a grandsire | 9K | 20 | 584 | 1.3 | 6 | <0.1 | 898 | 2.0 | 1,488 | 3.3 |
| | 50K | 36 | 85 | 0.2 | 8 | <0.1 | 99 | 0.2 | 192 | 0.4 |
| | 77K | 47 | 583 | 1.3 | 11 | <0.1 | 705 | 1.6 | 1,299 | 2.9 |

[1]M1 = 1 allele missing; M2 = both alleles missing.

[2]Scenario 2 = 2,011 Mexican genotyped Holsteins plus 886 related Holsteins in the US genotype database; scenario 3 = 2,011 Mexican genotyped Holsteins and 338,073 Holsteins in the US genotype database.

[3]Usable markers in common with 45,195 (50K): 6K = 6,842; 9K = 8,196; 77K = 28,048.

**Table 4.** Numbers and percentages of missing markers (MM) and conflicts after SNP validation by genotyping chip density and reference population scenario

| Genotyping chip density[1] | Scenario[2] | MM | | Conflicts | |
|---|---|---|---|---|---|
| | | No. | % | No. | % |
| 6K | 1 | 2,965 | 6.6 | 1,277 | 2.8 |
| | 2 | 2,398 | 5.3 | 1,022 | 2.3 |
| 9K | 1 | 3,252 | 7.2 | 1,181 | 2.6 |
| | 2 | 2,641 | 5.8 | 936 | 2.1 |

[1]Usable markers in common with 45,195 (50K): 6K = 6,842; 9K = 8,196; 77K = 28,048.

[2]Scenario 1 = 2,011 Mexican genotyped Holsteins; scenario 2 = 2,011 Mexican genotyped Holsteins plus 886 related Holsteins in the US genotype database.

### *Comparison of Imputed and True Genotypes*

Imputed genotypes were validated only for scenarios 1 and 2, because only extremely small differences were found between scenarios 2 and 3 and the computational demands of the imputation process would have been large for scenario 3. Within scenario, more MM were found for 9K genotypes than for 6K genotypes (Table 4), whereas the number of conflicts was slightly higher for 6K genotypes. The percentage of markers with 1 or 2 missing alleles (Table 4) was higher for 9K genotypes (7.2% for scenario 1 and 5.8% for scenario 2) than for 6K genotypes (6.6 and 5.3%, respectively). However, 6K genotypes had a higher conflict rate (2.8% for scenario 1 and 2.3% for scenario 2) than did 9K genotypes (2.6 and 2.1%, respectively). Similar error rates (2.7%) were reported by Zhang and Druet (2010) when a panel of 6,000 markers was imputed to 45,836 markers using a reference group of 2,000 animals. Weigel et al. (2010) reported error rates of 3.6 to 5.8% for imputation from 8,680 markers using a reference group of 2,542 Jerseys, which was slightly higher than the 9K conflict rates in this study. Although imputation for 9K genotypes is expected to be slightly more accurate than for 6K genotypes (0.3 percentage points for all scenarios; Figure 1), the 6K and 9K genotyping chips are no longer available commercially.

### CONCLUSIONS

Imputation efficiency in Mexican Holstein cattle based on M1, M2, and conflicts was affected by the size of the reference population used for imputation, the SNP density of the chip used for the original genotype, and the availability of ancestor genotypes. The inclusion of information from direct ancestors with US GDB genotypes in the imputation of Mexican Holstein genotypes increased imputation accuracy by 1 percentage point for low-density (6K and 9K) genotypes and by 0.5 percentage points for high-density (77K) genotypes, which was about half the gain found with information

from all US GDB Holsteins. Numbers of MM markers and imputation conflicts decreased when all US GDB genotypes were included. A larger reference population and the availability of genotyped ancestors improved imputation; animals with genotyped parents in a large reference population had higher imputation accuracy than those with no or few genotyped relatives in a small reference population. For small local populations, including genotypes from other related populations can be an important tool for improving accuracy of imputation.

### 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. http://dx.doi.org/10.1017/S1751731111000309.

Bouwman, A. C., J. M. Hickey, M. P. L. Calus, and R. F. Veerkamp. 2014. Imputation of non-genotyped individuals based on genotyped relatives: Assessing the imputation accuracy of a real case scenario in dairy cattle. Genet. Sel. Evol. 46:6 http://dx.doi.org/10.1186/1297-9686-46-6.

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 individuals. Am. J. Hum. Genet. 84:210–223. http://dx.doi.org/10.1016/j.ajhg.2009.01.005.

Browning, B. L., and S. R. Browning. 2011. A fast, powerful method for detecting identity by descent. Am. J. Hum. Genet. 88:173–182. http://dx.doi.org/10.1016/j.ajhg.2011.01.010.

Carvalheiro, R., S. A. Boison, H. H. R. Neves, M. Sargolzaei, F. S. Schenkel, Y. T. Utsunomiya, A. M. Pérez O'Brien, J. Sölkner, J.

C. McEwan, C. P. Van Tassell, T. S. Sonstegard, and J. F. Garcia. 2014. Accuracy of genotype imputation in Nelore cattle. Genet. Sel. Evol. 46:69. http://dx.doi.org/10.1186/s12711-014-0069-1.

Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185:1021–1031. http://dx.doi.org/10.1534/genetics.110.116855.

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. http://dx.doi.org/10.3168/jds.2010-3255.

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. http://dx.doi.org/10.3168/jds.2008-1646.

Hickey, J. M., B. P. Kinghorn, B. Tier, J. H. J. van der Werf, and M. A. Cleveland. 2012. A phasing and imputation method for pedigreed populations that results in a single-stage genomic evaluation. Genet. Sel. Evol. 44:9 http://dx.doi.org/10.1186/1297-9686-44-9.

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. http://dx.doi.org/10.1371/journal.pgen.100529.

Illumina Inc. 2011. BovineSNP50 Genotyping BeadChip. Accessed Dec. 30, 2014. http://res.illumina.com/documents/products/datasheets/datasheet_bovine_snp5o.pdf.

Illumina Inc. 2013. BovineLD v1.1 Genotyping BeadChip. Accessed Dec. 30, 2014. http://res.illumina.com/documents/products/datasheets/datasheet_bovineld.pdf.

Johnston, J., G. Kistemaker, and P. G. Sullivan. 2011. Comparison of different imputation methods. Interbull Bull. 44:25–33.

Khatkar, M. S., G. Moser, B. J. Hayes, and H. W. Raadsma. 2012. Strategies and utility of imputed SNP genotypes for genomic analysis in dairy cattle. BMC Genomics 13:538 http://dx.doi.org/10.1186/1471-2164-13-538.

Ma, P., R. F. Brøndum, Q. Zhang, M. S. Lund, and G. Su. 2013. Comparison of different methods for imputing genome-wide marker genotypes in Swedish and Finnish Red Cattle. J. Dairy Sci. 96:4666–4677. http://dx.doi.org/10.3168/jds.2012-6316.

Neogen Corp. 2013a. GeneSeek Genomic Profiler for dairy cattle. Accessed Dec. 30, 2014. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP-LD_Dairy.pdf.

Neogen Corp. 2013b. GeneSeek Genomic Profiler HD for dairy cattle. Accessed Dec. 30, 2014. http://www.neogen.com/Agrigenomics/pdf/Slicks/GGP_HD_Dairy.pdf.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2011. FImpute—An efficient imputation algorithm for dairy cattle populations. J. Dairy Sci. 94(E-Suppl. 1):421. (Abstr.)

Solberg, T. R., A. K. Sonesson, J. A. Woolliams, and T. H. E. Meuwissen. 2008. Genomic selection using different marker types and densities. J. Anim. Sci. 86:2447–2454. http://dx.doi.org/10.2527/jas.2007-0010.

VanRaden, P. M. 2015. findhap.f90, Find haplotypes and impute genotypes using multiple chip sets and sequence data. Accessed Feb. 25, 2015. http://aipl.arsusda.gov/software/findhap/.

VanRaden, P. M., J. R. O'Connell, G. R. Wiggans, and K. A. Weigel. 2010. Combining different marker densities in genomic evaluation. Interbull Bull. 42:113–117.

VanRaden, P. M., J. R. O'Connell, G. R. Wiggans, and K. A. Weigel. 2011. Genomic evaluations with many more genotypes. Genet. Sel. Evol. 43:10. http://dx.doi.org/10.1186/1297-9686-43-10.

Weigel, K. A., G. de los Campos, A. I. Vazquez, G. J. M. Rosa, D. Gianola, and C. P. Van Tassell. 2010. Accuracy of direct genomic values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. J. Dairy Sci. 93:5423–5435. http://dx.doi.org/10.3168/jds.2010-3149.

Wiggans, G. R., T. A. Cooper, P. M. VanRaden, K. M. Olson, and M. E. Tooker. 2012. Use of the Illumina Bovine3K BeadChip in dairy genomic evaluation. J. Dairy Sci. 95:1552–1558. http://dx.doi.org/10.3168/jds.2011-4985.

Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, future. J. Dairy Sci. 94:3202–3211. http://dx.doi.org/10.3168/jds.2010-3866.

Zhang, Z., and T. Druet. 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. J. Dairy Sci. 93:5487–5494. http://dx.doi.org/10.3168/jds.2010-3501.