J. Dairy Sci. 96:668–678
http://dx.doi.org/10.3168/jds.2012-5702
© American Dairy Science Association®, 2013.

Genomic imputation and evaluation using high-density Holstein genotypes

P. M. VanRaden,*[1] D. J. Null,* M. Sargolzaei,† G. R. Wiggans,* M. E. Tooker,* J. B. Cole,* T. S. Sonstegard,‡ E. E. Connor,‡ M. Winters,§ J. B. C. H. M. van Kaam,# A. Valentini,‖ B. J. Van Doormaal,¶ M. A. Faust,** and G. A. Doak††

*Animal Improvement Programs Laboratory, Agricultural Research Service, US Department of Agriculture (USDA), Beltsville, MD 20705-2350
†Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, N1G 2W1, Canada
‡Bovine Functional Genomics Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350
§DairyCo, Agriculture and Horticulture Development Board, Kenilworth, Warwickshire, CV8 2TL, United Kingdom
#Associazone Nazionale Allevatori Frisona Italiana, 26100, Cremona, Italy
‖Dipartimento per la Innovazione nei Sistemi Biologici, Agroalimentari e Forestali (DIBAF), Universita della Tuscia via de Iellis, 01100 Viterbo, Italy
¶Canadian Dairy Network, Guelph, Ontario, N1K 1E5, Canada
**ABS Global, DeForest, WI 53532
††National Association of Animal Breeders, Columbia, MO 65205

ABSTRACT

Genomic evaluations for 161,341 Holsteins were computed by using 311,725 of 777,962 markers on the Illumina BovineHD Genotyping BeadChip (HD). Initial edits with 1,741 HD genotypes from 5 breeds revealed that 636,967 markers were usable but that half were redundant. Holstein genotypes were from 1,510 animals with HD markers, 82,358 animals with 45,187 (50K) markers, 1,797 animals with 8,031 (8K) markers, 20,177 animals with 6,836 (6K) markers, 52,270 animals with 2,683 (3K) markers, and 3,229 nongenotyped dams (0K) with >90% of haplotypes imputable because they had 4 or more genotyped progeny. The Holstein HD genotypes were from 1,142 US, Canadian, British, and Italian sires, 196 other sires, 138 cows in a US Department of Agriculture research herd (Beltsville, MD), and 34 other females. Percentages of correctly imputed genotypes were tested by applying the programs findhap and FImpute to a simulated chromosome for an earlier population that had only 1,112 animals with HD genotypes and none with 8K genotypes. For each chip, 1% of the genotypes were missing and 0.02% were incorrect initially. After imputation of missing markers with findhap, percentages of genotypes correct were 99.9% from HD, 99.0% from 50K, 94.6% from 6K, 90.5% from 3K, and 93.5% from 0K. With FImpute, 99.96% were correct from HD, 99.3% from 50K, 94.7% from 6K, 91.1% from 3K, and 95.1% from 0K genotypes. Accuracy for the 3K and 6K genotypes further improved by approximately 2 percentage points if imputed first to 50K and then to HD instead of imputing all genotypes directly to HD. Evaluations were tested by using imputed actual genotypes and August 2008 phenotypes to predict deregressed evaluations of US bulls proven after August 2008. For 28 traits tested, the estimated genomic reliability averaged 61.1% when using 311,725 markers vs. 60.7% when using 45,187 markers vs. 29.6% from the traditional parent average. Squared correlations with future data were slightly greater for 16 traits and slightly less for 12 with HD than with 50K evaluations. The observed 0.4 percentage point average increase in reliability was less favorable than the 0.9 expected from simulation but was similar to actual gains from other HD studies. The largest HD and 50K marker effects were often located at very similar positions. The single-breed evaluation tested here and previous single-breed or multibreed evaluations have not produced large gains. Increasing the number of HD genotypes used for imputation above 1,074 did not improve the reliability of Holstein genomic evaluations.

Key words: genomic evaluation, imputation, marker density

INTRODUCTION

High-density (**HD**) genotypes provide markers closer to QTL, but missing alleles must then be imputed for animals genotyped at less than the highest density. Observed and imputed genotypes from chips of various marker densities are combined in 1 genomic evaluation to reduce costs and improve reliability. Many methods and programs are available to impute the missing genotypes, but programs developed for human genotypes often do not scale to the large data sets and general pedigrees in animal breeding (Chen et al., 2011; Johnston et al., 2011; Williams et al., 2012).

Simulations have forecast that increasing density much greater than 50,000 markers (**50K**) will give either no gains in evaluation reliability (Harris and

Received May 7, 2012.
Accepted September 7, 2012.

[1] Corresponding author: paul.vanraden@ars.usda.gov

Johnson, 2010), very small gains (VanRaden et al., 2011b), or large gains (Meuwissen and Goddard, 2010). Benefits from HD genotypes may be small if most genetic variation is from very small QTL effects (Clark et al., 2011). Imputation losses can also affect evaluation reliability if an insufficient number of animals have HD genotypes. Few studies have investigated the ability to impute from very low to very high density. Imputation of sequence data is now common in human genetics but is not yet common with bovines because of the limited number of full sequences available. Before investing in data collection, realistic simulations are useful in optimizing designs and developing efficient methods of analysis.

Two HD genotyping chips for cattle are currently available (Rincon et al., 2011): the BovineHD Genotyping BeadChip from Illumina (San Diego, CA) and the Axiom Genome-Wide BOS 1 Array Plate from Affymetrix (Santa Clara, CA). Early actual results with BovineHD genotypes in other populations have indicated small or no advantages in evaluation reliability as compared with 50K. Across-breed evaluations of Holsteins and Jerseys found no benefit from HD compared with 50K in New Zealand (Harris et al., 2011) or small benefits for Bayesian predictions that were not significant because of only 86 Jersey validation bulls in Australia (Erbe et al., 2012); the conclusion was that the 777,962 markers could be reduced to 329,329 by eliminating redundant markers (Harris et al., 2011) or to 58,532 by using markers only within transcribed DNA (Erbe et al., 2012). Reliability in the joint genomic evaluation of Denmark, Finland, and Sweden improved by an average of 0.5 percentage point when using 557 Holstein HD genotypes and by 1.0 percentage point when using 706 Red Dairy Cattle HD genotypes in separate within-breed analyses (Su et al., 2012). The use of HD genotypes for 384 Norwegian Red bulls increased correlations of predicted with observed data for milk, protein, and a mastitis trait by 7 to 9 percentage points but showed little or no increase for 4 other traits (Solberg et al., 2011). Those studies all tested the ability to predict merit of a recent generation of genotyped bulls by using genotypes and phenotypes of earlier generations.

Lower density chips have become widely used recently, particularly for genotyping of females. Three different densities <50K are now included in the North American genotype database. The Illumina Bovine3K BeadChip with 2,900 markers was introduced in August 2010 (Wiggans et al., 2012a) and was then replaced in November 2011 by the Illumina BovineLD BeadChip with 6,909 markers (Boichard et al., 2012). The GeneSeek Genomic Profiler (**GGP**), a customized version of the BovineLD chip with additional markers totaling 8,655, became available in February 2012 (Wiggans et al., 2012b). More animals are now genotyped at those lower densities than at 50K or HD. Accuracies for all densities should be optimized in genomic evaluations.

The objectives of this study were to examine 1) data quality, marker selection, and mapping issues using 1,741 BovineHD genotypes; 2) the accuracy of imputing HD from 50K and lower density genotypes; 3) the accuracy of imputing between HD and sequence data or between 2 different HD chips; 4) the effect of the number of Holstein HD genotypes used for imputation on within-breed evaluation; and 5) the reliability of HD genomic predictions in a very large Holstein population.

MATERIALS AND METHODS

Markers and Chips

Five densities of actual genotypes from Illumina were used in this study. The BovineHD chip contains 777,962 markers, including 1,509 on the Y chromosome and mitochondrial DNA, and is referred to here as HD. The BovineSNP50 v1 chip with 56,947 markers and v2 chip with 54,609 markers are both referred to here as 50K. The commercial v1 chip reported 54,001 of the 56,947 markers used here in the research v1. For simplicity, the GGP, BovineLD, and Bovine3K chips are referred to here as 8K, 6K, and 3K, respectively. The chips were designed as mostly nested subsets: the HD chip includes 91% of the 50K v2 markers, the 50K has 99% of the 8K markers, the 8K has 100% of the 6K markers, and the 6K has 75% of the 3K markers.

Edits for marker quality were automated because visual inspection of genotype clusters for each marker is too time consuming with HD chips. With the 50K, 8K, 6K, or 3K genotypes, cluster positions were hand adjusted to improve call rates and reduce parent-progeny conflicts. Markers on the HD chip were subjected to the same initial edits as on each other chip, which required a minor allele frequency (**MAF**) of >0.01 for at least 1 breed, a pooled test of within-breed Hardy-Weinberg equilibrium (**HWE**), <10% missing genotypes, and <2% parent-progeny conflicts; the 2 latter limits are reduced proportionally to MAF (Wiggans et al., 2012a). Observed heterozygotes were required to be between 0.3 and 1.3 times the expected number. Observed minor homozygotes were required to be between 0.1 and 10 times the expected number.

Redundant markers were eliminated by using a total of 1,741 HD genotypes from animals of 5 breeds, whereas only 1,510 HD genotypes of Holsteins were used to test the genomic evaluation because of having too few HD genotypes for imputation in most of the other breeds. The numbers of individuals with HD genotypes used for marker selection were 434 Ayrshire, 71 Brown Swiss, 61 Guernsey, 63 Jersey, and 1,112 Holstein. Markers were ordered by chromosome number and location on the chromosome from the UMD3.1 genome assembly (Center for Bioinformatics and Computational Biology, 2010).

Each marker was compared with the subsequent 349 markers on the chromosome assembly by using counts of the combinations of genotypes observed for each marker pair, and correlations were computed by using genotypes coded as BB = 0, AB = 1, and AA = 2. Pairs of loci were designated as redundant if the absolute value of the correlation exceeded 0.95 + 0.1 (0.5 − MAF) or if 95% + 10% (0.5 − MAF) of the animals had consistent calls. Thus, the threshold was raised linearly from 95 to 99 as the MAF decreased toward 0. If the correlation was positive, consistent calls were those for which both markers had AA, both had AB, or both had BB. If negative, consistent calls were those for which both markers had AB or opposite homozygotes. These edits were applied only to marker pairs for which at least 1 had MAF >0.05. In many cases, a group of nearby markers was mutually redundant. One was selected from each group, giving preference to the markers used for international parentage verification first and SNP on the BovineSNP50 chip second. The marker of the preferred type with the highest call rate was retained.

Map locations must be correct for imputation to work well. Several 50K markers were previously relocated in cooperation with researchers from the University of Missouri (R. D. Schnabel), the University of Maryland (J. R. O'Connell), and the University of Guelph (M. Sargolzaei and J. Johnston), but locations of HD markers had not yet been tested. Potential problems were identified from the largest counts of different haplotypes within 50-marker windows of usable HD SNP. Genotype correlations were then inspected visually by using heat maps within those segments. Instead of attempting to find correct locations on the UMD3.1 map, misplaced HD markers were simply deleted.

After edits, the numbers of markers used from each chip were 311,725 from the HD, 45,187 from the 50K, 8,031 from the 8K, 6,836 from the 6K, and 2,683 from the 3K. Nongenotyped dams (0K) were also included if >90% of haplotypes were imputable because the cow had ≥4 genotyped progeny. For HD evaluations, all densities were imputed to HD. For 50K evaluations, all densities were imputed to 50K and only the 50K subset of HD was included.

Animals Genotyped and Phenotyped

A total of 161,341 Holsteins had genotypes used in the genomic evaluation, including 65% females and 35% males. The numbers of animals genotyped with each density were 1,510 with HD, 82,358 with 50K, 1,797 with 8K, 20,177 with 6K, 52,270 with 3K, and 3,229 imputed dams with 0K. The 1,510 HD genotypes were from 305 US, 93 Canadian, 284 British, 460 Italian, and 196 other sires, from 138 cows from a US Department of Agriculture (Beltsville, MD) research herd, and from 34 other females. The markers were selected using genotypes for a subset of 109,205 animals available before the 6K and 8K chips were marketed.

Four evaluation studies were conducted with Holstein data. Three preliminary tests used earlier actual data sets and fewer animals with HD genotypes to impute the missing markers in low-density genotypes. The first study (December 2010 data) included only 342 HD animals but with 636,967 of the markers used (VanRaden et al., 2011a). The second and third studies (August 2011 data) both included 1,074 HD animals after adding genotypes from Great Britain. The second study computed evaluations using 636,967 markers, whereas the third used 311,725 markers to verify that results would be the same after eliminating redundant markers. The fourth (final) study was conducted with December 2011 data after adding genotypes from Italy and used 1,510 HD animals for imputation of the other 159,831 animals that had genotypes of lower density.

Genomic evaluations were computed for 28 traits using August 2008 data to predict December 2011 deregressed evaluations of 3,404 US bulls proven after 2008. Almost all those bulls had 50K genotypes. Observed reliability would be lower for animals genotyped with less density, or slightly higher if genotyped with HD to avoid imputation loss. The truncated reference population included 10,718 bulls with daughters and 5,124 cows with records. The reference cows had a mixture of genotype densities, but all had US yield trait phenotypes before August 2008. The reference bulls were from the combined genomic data of the United States, Canada, Italy, and Great Britain plus additional proven bulls from 15 other countries.

Imputation to HD from 50K, 6K, and 3K genotypes was tested by using a simulated chromosome with 1% of the genotypes randomly missing and 0.02% incorrect initially from each chip. To introduce incorrect genotypes, true homozygotes were set to observed heterozygotes and true heterozygotes were set to an even mixture of the 2 homozygotes. The simulation program (genosim.f90) and the methods of VanRaden et al. (2011b) generated linkage disequilibrium (**LD**) directly in chromosomes of the founding population (oldest members of the pedigree). Advantages are speed and simple control of allele frequencies as compared with strategies that require many generations of random mating followed by a marker selection step. A disadvantage is that the LD pattern may not mimic real data as well.

The parameter controlling correlation structure and LD in the founding population was set to 0.998 as in VanRaden et al. (2011b) for the 500,000 marker density. Underlying haplotype blocks ended between markers whenever a uniform random number exceeded the parameter. Underlying alleles were converted to observed alleles by using founding allele frequencies that were uniform between 0 and 1. Descendant chromosomes were created using a random recombination of parental chromosomes with no suppression of nearby crossovers.

The simulated pedigree and genotyping pattern exactly matched the earlier subset of actual data from August 2011 (the third study above) before the HD genotypes for Italian bulls were added. In the simulated data, unique chromosomes were assigned to unknown parents, whereas in actual populations, their chromosomes may be inherited from popular ancestors whose pedigree connection was not recorded. Therefore, for animals with incomplete pedigrees, imputation accuracy may be reduced in the simulated data compared with the actual data. However, other factors, such as incorrect pedigrees or remaining map problems, can reduce the accuracy in actual data as compared with simulated data.

When the simulation was conducted, actual BovineLD and GGP genotypes were not yet available and exact numbers of markers were not yet known, so only 5,130 markers were included in the imputation test for 6K genotypes, and 8K was not tested. The simulation generated 6K genotypes for 1,000 animals chosen randomly from the 39,441 that actually had 3K genotypes. In this imputation test, genotypes were simulated for 116,380 animals: 1,112 HD, 72,532 50K, 1,000 6K, 38,441 3K, and 3,295 0K (imputed dams). Among all animals, 89.9% of genotypes were missing initially.

Future Scenarios

Imputation between 2 different HD chips was tested to determine how many animals needed to be genotyped with both chips to provide sufficient overlap and how many needed to be genotyped with just 1 chip to reduce costs. The design had 61,615 animals with 50K genotypes and also included 1,000 genotypes from each HD chip. The bulls were ranked by highest reliability and the first n were double genotyped and the next $2{,}000 - 2n$ bulls were assigned alternately to chip 1 or chip 2, where $n = 0$, 50, 200, or 500. Thus, total expense was constant at 2,000 genotypes but with more or less overlap. The simulated chips each had 625,000 markers with 50,000 in common, for a total of 1.2 million. The parameter controlling LD was increased to 0.999. The Illumina BovineHD and Affymetrix BOS 1 chips have 107,945 markers in common, including those on the Illumina BovineSNP50 chip plus approximately 60,000 others. When this research was conducted, we were unaware of the 60,000 others in common or that only 311,725 HD markers would be used because of redundancy.

Sequence data also were simulated and imputed from genotypes based on 600,000 (**600K**) markers to test feasibility. The parameter controlling LD was increased to 0.9998 for 30 million marker (sequence) density. Allele frequencies in the founding population remained uniform between 0 and 1, but actual sequences would contain more low-MAF SNP than those selected to be placed on chips. The 500 bulls with the most daughters and with US registration codes had simulated sequences, and 500 randomly chosen US bulls born in 2009 had simulated 600K genotypes. Several generations sometimes separated the young bulls with 600K genotypes and the sequenced old bulls because the average birth year of the old bulls was 1987.

The sequences were simulated to contain 1 million polymorphic loci per chromosome; monomorphic loci were not generated and not needed. Only 1 chromosome was simulated instead of 30. Phasing and imputation of actual sequences both use genotypes as input data even though the sequences are originally read as haplotypes, because the read lengths are too short to reconstruct the long haplotypes directly. Individual animals are often sequenced at low coverage to reduce cost, but this research assumed high coverage per bull and that sequence genotypes were as accurate as chip genotypes. The goal was to examine computational feasibility rather than to forecast gains in reliability from sequence data.

Imputation and Evaluation Software

The programs findhap (VanRaden et al., 2011b) and FImpute (Sargolzaei et al., 2011) were tested for imputation of HD data. Both programs use deterministic methods to combine family and population information. Version 2 of findhap (VanRaden, 2011), which has improved imputation rates compared with version 1, uses both long segments to improve haplotype matches for close relatives and short segments to help detect matches from more remote ancestors. Several combinations of segment lengths were tested in findhap. The FImpute program also was modified recently to allow processing of HD genotypes. Population imputation is based on overlapping sliding windows and assumes that individuals are related to some degree. Overlapping of windows allows for consistency of haplotype phases across windows. The FImpute program also starts with a long window and gradually shrinks the window size in each sweep to a very short window.

Initial research showed improvements in imputation accuracy by combining results from 2 software packages (Johnston et al., 2011) or imputing genotypes in steps (first imputing 3K to 50K and then imputing 50K to HD) instead of doing all imputation in 1 step. For a simulated chromosome with 1,112 HD animals, final comparisons were carried out using findhap in 1 step, FImpute in 1 step, or FImpute in 2 steps (imputing lower densities to 50K and then 50K to HD). For actual data, the 3 preliminary evaluations were conducted using FImpute from lower density to 50K and findhap from 50K to HD. The final evaluation used imputed genotypes from FImpute in 1 step.

Genomic evaluations were computed by iteration for marker effects using both the linear model and the approximate Bayes A algorithm of VanRaden (2008). The exponential distribution by VanRaden contained a typographical error and should have shown that the normal variable was divided by $1.25^{[\mathrm{abs}(s)\,-\,2]}$, where s is the standard deviation of marker deviations such that estimated effects are decreased if standard deviations are <2 and increased if standard deviations are >2 as compared with linear model estimates. An optimal parameter of 1.12 instead of 1.25 was later estimated for use with US official evaluations (Cole et al., 2009) and with simulated HD data (VanRaden et al., 2011b) and was not estimated here again. Direct genomic values were then combined by selection index with traditional EBV and a subset EBV that was obtained by applying pedigree relationships to only the genotyped animals as in VanRaden et al. (2009).

RESULTS AND DISCUSSION

Marker Quality

Actual markers selected from 4 different Illumina chips are compared in Table 1. For usable markers, missing and parent-progeny conflict rates both were lowest for HD and were highest for 3K. Quality was best for the HD chip with the most markers, indicating rapid advances in technology to read DNA. However, a larger proportion of the HD markers did not pass the HWE edits because of fewer genotyped animals and more sampling variation in the estimated frequencies. The number of usable HD markers was 636,967 before applying the HWE edit and 614,012 after. The HWE edits made 3 markers on the 3K chip and about 400 markers on each of the 50K chips unusable compared with 22,955 markers on the HD chip. The 6K and 8K markers have very high quality because they were selected from 50K and HD markers with >98% call rate and <0.01% parent-progeny conflict and because chip chemistry is the same as for the 50K and HD markers (Boichard et al., 2012). Stricter edits, such as elimination of any marker for any problem, might be justified with HD chips because of no or less visual inspection of cluster quality and because more markers are available from which to choose.

Map location problems caused removal of 140 HD markers from chromosomes 1, 4, 6, 7, 29, and X. The markers removed were mostly in small contiguous sets of 5 to 30 SNP. The numbers of haplotypes within the segments were greatly reduced after removing these small blocks that had poor correlations with surrounding markers, which indicated that SNP deletion solved the map problems. Figure 1 provides an example heat map showing a block of SNP removed from chromosome 1.

About half of the Illumina HD markers were removed by the redundancy edit, which agrees with Harris et al. (2011). Markers identified as redundant included 426,718 detected as members of 105,305 groups, with the largest group containing 187 markers. Because only 1 marker was selected per group, 321,413 HD markers (~50%) were deleted as redundant. For comparison, only 1,755 of the 45,187 usable SNP (~4%) from the BovineSNP50 were redundant by this same criterion. This indicates a large degree of short-range LD and that addition of even more high-LD markers will not be helpful. The high LD observed in this study, even with genotypes from 5 breeds included, may explain why benefits were small in previous HD evaluations.

Table 1. Numbers and properties of usable markers selected from 4 Illumina[1] chips

		Markers			Usable markers	
Chip	Animals genotyped (no.)	Available (no.)	Usable (no.)	Redundant (no.)	Missing (%)	Conflicts (%)
BovineHD	1,170	777,962	614,012	321,413	0.2	0.004
BovineSNP50 v1	57,916	56,947	43,598	1,755	0.4	0.014
BovineSNP50 v2	15,270	54,609	43,293	1,755	0.4	0.011
Bovine3K	34,849	2,900	2,683	0	0.5	0.078

[1] Illumina, San Diego, CA.

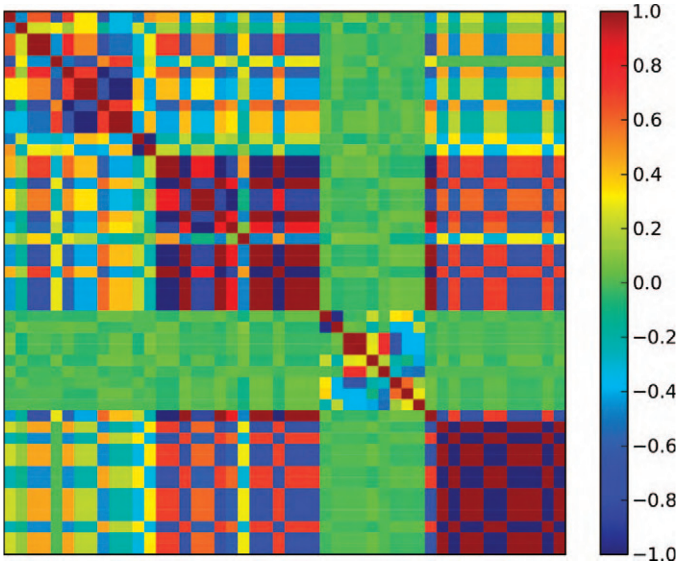

Figure 1. Heat map showing marker correlations in a segment of bovine chromosome 1 and the block of SNP removed. Color version available in the online PDF.

Imputation and Computation

Computer requirements were very reasonable for imputation, evaluation, and simulation. Imputation of 636,967 markers for 103,070 animals with findhap required 50 gigabytes of memory and 10 h using 6 processors. Imputation of 311,725 markers for 161,341 animals using version 2 of FImpute required 70 gigabytes of memory and 13 h using 5 processors. A more recent update of FImpute reduces the required memory to approximately 25 gigabytes for these data. Iteration for 311,725 marker effects for 29 traits (28 recorded traits plus net merit), using the densemap.f90 Fortran program of VanRaden (2008), required 30 gigabytes of memory and 2 d using 6 processors. Simulation of 1 million SNP on 1 chromosome (sequence data) for 1,000 genotyped animals plus 15,135 nongenotyped ancestors required 30 gigabytes of memory and 37 min with 1 processor. Imputation from 600K to sequence for 1 chromosome of the 1,000 animals required 4 gigabytes of memory and 15 min using 6 processors with findhap. Memory, time, and especially disk space all will become more limiting if many animals are imputed to sequence-level data.

Both imputation programs had high accuracy when imputing HD from 50K, but accuracy was less when imputing HD from lower densities, as expected (Table 2). After imputing missing markers with findhap, percentages of genotypes correct were 99.9% for HD, 99.0% for 50K, 94.6% for 6K, 90.5% for 3K, and 93.5% for 0K (imputed dams). With FImpute, 99.96% were correct for HD, 99.3% for 50K, 94.7% for 6K, 91.1% for 3K, and 95.1% for 0K. Accuracy further improved with imputation first to 50K and then to HD instead of all together for 6K (1.4 percentage points), 3K (2.6 percentage points), and 0K genotypes (1.6 percentage points). Smaller gains occurred when findhap was used to impute first to 50K and then to HD in an earlier data set (results not shown). The BovineLD chip with 6,836 usable markers should allow better imputation than the 5,130 markers tested.

A maximum length of 2,500 markers and a minimum of 100 markers yielded the best results for imputation to 330K when findhap was used to process all genotypes and all densities together. For the sequence data, a maximum length of 100,000 markers and a minimum of 2,000 markers yielded the best results. The programs findhap and FImpute both include options to adjust imputation algorithms for marker density.

Imputation with 2 different HD chips or from HD to sequence data both had high accuracy when using simulated genotypes and findhap. Imputation from 50K to the combined set of two 600K chips (1.15 million markers) was most accurate at 98.59% if 200 bulls were genotyped with both 600K chips (Table 3). However, advantages compared with 0 or 500 double-genotyped

Table 2. Percentage of simulated genotypes correctly called when imputing to high density from lower densities using findhap or FImpute

		Correctly imputed genotypes (%)		
Markers genotyped (no.), chip[1]	Animals genotyped (no.)	Findhap (all)	FImpute (all)	FImpute (2 steps[2])
330,000, HD	1,112	99.9	99.96	99.96
41,250, 50K	72,532	99.0	99.3	99.3
5,130, 6K	1,000	94.6	94.7	96.1
2,550, 3K	38,441	90.5	91.1	93.7
0, 0K	3,295[3]	93.5	95.1	96.7

[1] Illumina, San Diego, CA.

[2] Imputing lower densities to 41,250 markers and then imputing to 330,000 markers in a second step.

[3] Dams imputed from multiple progeny.

Table 3. Percentage of genotypes correctly imputed using 2 different simulated high-density chips and some bulls double genotyped

Chip used	Animals (no.)	Correctly imputed genotypes (%) for *n* bulls genotyped with both high-density chips			
		0	50	200	500
50K	61,615	98.53	98.56	98.59	98.52
600K, chip 1	1,000	98.96	99.12	99.38	99.49
600K, chip 2	1,000	98.02	99.38	99.39	99.43
Both 600K chips	*n*	—	99.76	99.83	99.88

bulls were both small. Obtaining overlap of the bulls genotyped may not be important because the chips already have sufficient overlap of markers to correctly match the haplotypes. Double genotyping could have other advantages, such as in comparing marker quality or map positions of markers on the 2 chips.

Imputation from 600K to 30 million SNP (sequence data) was 97.8% accurate on average. Accuracy may be lower in the sequence imputation than in the HD imputations because sequences were simulated for only 500 animals, whereas 1,000 simulated HD genotypes were used for imputation. In addition, a larger percentage of markers was missing going from 600K to sequence than from 50K to 600K. The LD may not be as high in the simulated sequences as would occur in real sequences, and other imputation strategies could be more accurate for sequence data than those used in findhap.

Allele frequencies in real and simulated data are compared in Table 4. The simulated frequencies closely match those of the 50K and HD markers. The actual markers on lower density chips were selected for high MAF, whereas the simulated markers were simply evenly spaced, and this affected imputation accuracy. If sequences were simulated with much lower MAF, as would occur in real data, imputation "accuracy" would be higher simply by guessing that the common allele is homozygous, but the correlations of estimations with true genotypes would be lower. Finally, actual sequences obtained from lower coverage will be less accurate than those simulated.

Genomic Evaluation

The average observed reliability over all traits for young bulls was 0.4 percentage point greater when using 311,725 markers as compared with 50K (Table 5). With 311,725 markers, the linear model gave an average reliability of 60.3%, only 0.8 percentage point less than the 61.1% for the nonlinear model in Table 5 and less than the difference of 1.6 percentage points estimated in simulation (VanRaden et al., 2011a). The realistic simulations of Harris and Johnson (2010) and VanRaden et al. (2011a) forecast small gains from more markers and nonlinear models, whereas the unrealistic simulation of Meuwissen and Goddard (2010) forecast large gains by assuming that all genetic variance was from 3 or 30 QTL on 1 chromosome. Unrealistic simulations hurt rather than help genomic selection by making breeders less confident in forecasts, whereas realistic simulations can be very valuable.

The largest marker effects were for HD markers at new locations for some traits, but for many other traits, the largest effects were still for the same 50K markers. This indicated that imputation loss may have prevented the new markers from contributing fully to overall reliability. Multibreed evaluation could produce larger gains than the single-breed evaluation that was investigated here, but it will also require more investment in HD genotypes for each breed.

The preliminary studies indicated that imputation losses were too large with only 342 HD genotypes (VanRaden et al., 2011b) and that 1,074 HD genotypes were sufficient, with no further advantage from 1,510 HD genotypes. The first study, with only 342 HD genotypes for imputation, gave an average decrease of 0.5 percentage point reliability as compared with the 50K reliability. The second study, with 1,074 HD genotypes and 636,967 markers, gave an average increase of 0.5 percentage point above the 50K reliability. The third

Table 4. Minor allele frequencies in Holstein for actual markers used from each chip and for simulated markers and sequences

Minor allele frequency range (%)	Actual markers used (% in each range)					Simulated markers
	2,683	6,836	8,031	45,187	311,725	
0–10	6	3	4	17	19	20
11–20	11	8	9	19	19	20
21–30	19	18	19	21	20	20
31–40	29	30	29	21	21	20
41–50	35	40	38	22	21	20

Table 5. Reliability of 45,187 (50K) and 311,725 (300K) marker genomic predictions from the nonlinear and linear models

Trait	Reliability (%)			Gain in reliability (percentage points)	
	Parent average	Nonlinear 50K	Nonlinear 300K	Nonlinear 300K − 50K	Nonlinear − linear 300K
Milk yield	38.6	65.5	65.2	−0.3	1.2
Fat yield	38.6	68.5	68.7	0.2	1.8
Protein yield	38.6	60.6	60.1	−0.5	0.3
Fat percentage	38.6	86.2	88.1	1.9	6.5
Protein percentage	38.6	81.4	83.5	2.1	3.6
Productive life	31.3	77.0	78.6	1.6	1.3
SCS	33.8	65.2	65.7	0.5	0.4
Daughter pregnancy rate	30.8	66.6	67.3	0.7	0.4
Sire calving ease	17.8	33.8	31.5	−2.3	1.9
Daughter calving ease	18.3	39.9	36.3	−3.6	1.0
Sire stillbirth	16.5	15.0	17.7	2.7	0.0
Daughter stillbirth	16.7	36.5	40.8	4.3	−2.6
Final score	29.4	55.2	55.0	−0.2	0.1
Stature	30.0	64.4	66.2	1.8	0.7
Strength	29.6	64.3	65.2	0.9	1.4
Dairy form	29.5	66.1	66.7	0.6	0.4
Foot angle	28.6	49.8	49.8	0.0	0.3
Rear legs (side view)	29.2	59.8	59.5	−0.3	0.1
Body depth	29.5	67.4	68.1	0.7	1.4
Rump angle	30.0	66.2	66.8	0.6	0.1
Rump width	29.3	62.4	62.1	−0.3	0.7
Fore udder attachment	29.3	71.6	72.0	0.4	0.4
Rear udder height	29.2	55.4	54.8	−0.6	0.8
Udder depth	29.9	76.3	76.8	0.5	0.3
Udder cleft	29.0	59.4	58.7	−0.7	−0.2
Front teat placement	29.5	68.6	67.0	−1.6	0.2
Teat length	29.7	67.0	68.5	1.5	0.7
Rear legs (rear view)	28.4	50.9	50.7	−0.2	0.1
Average	29.6	60.7	61.1	0.4	0.8

study, using 311,725 instead of 636,967 markers, gave reliabilities that were almost identical for all traits but averaged slightly (0.1 percentage point) greater, perhaps because imputation from lower densities to 636,967 markers was more difficult than to 311,725 markers or because the same genetic variance was explained by using fewer marker effects to estimate.

The final study, with 1,510 HD animals for imputation and 311,725 markers, gave reliability gains that were not greater and were actually slightly (0.2 percentage point) less than with 1,074 HD animals. In a previous simulation, reliability increased by 1.6 percentage points if all animals had HD but by only 0.9 percentage point when 1,406 animals had HD and 32,008 others were imputed from 50K (VanRaden et al., 2011a). The contrasting results are probably because improved imputation methods were used with the current actual data than with the previous simulated data. Consequently, fewer HD animals are now needed to reach high-level imputation accuracy from 50K, and further addition of HD animals to improve the imputation accuracy probably will not be profitable. However, a larger HD reference population may benefit imputation from low-density genotypes to HD.

The modest gains from HD as compared with 50K indicate that reliability for animals with lower density genotypes will actually decline instead of improve with HD evaluation because imputation is less accurate from lower density to HD than to 50K. This problem might be overcome by conducting 2 different routine evaluations and publishing 50K evaluations for animals genotyped at low density and HD evaluations only for animals genotyped at 50K. This strategy would require more computation. An alternative may be to select and genotype only the HD markers with the largest effects on future chips.

The HD markers do provide the benefits of locating a few new QTL and refining the positions of some QTL located less precisely with the 50K genomic evaluation. An example is a QTL on BTA18 with large effects on several traits (Cole et al., 2009). Figure 2 compares 50K and HD marker effects for productive life across all chromosomes. The largest marker effects are on BTA5, BTA6, BTA7, and BTA18 in both the 50K and

HD graphs, but the HD peaks are narrower on each of those chromosomes. The solid-colored areas at the bottom of each graph show that most HD marker effects are smaller than 50K because the HD prior distributes small genetic effects over more markers.

Figure 3 focuses on a 1-Mbase region of BTA18. The 50K marker reported by Cole et al. (2009) has a smaller effect in the HD evaluation, whereas 3 new markers from the HD chip have similar effects and locate the QTL further to the left. On BTA5, BTA6, and BTA7, effects of the 50K markers were also smaller, and new markers from the HD surpassed them. Across all chromosomes in Figure 2, the 6 markers with the largest effects were all from the HD chip. This may explain

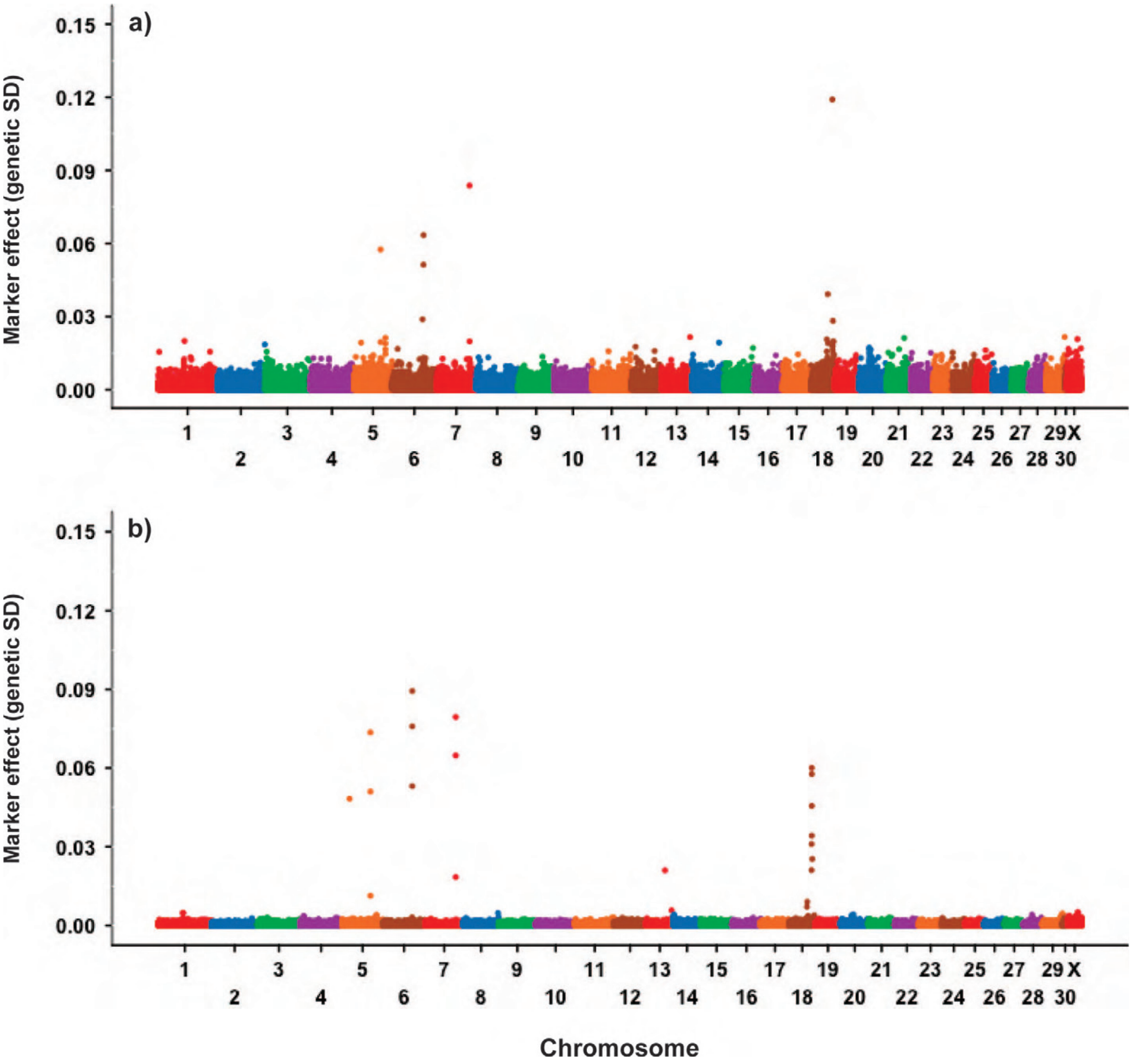

Figure 2. Effects (genetic SD) for productive life by chromosome from (a) 45,187 [BovineSNP50 Genotyping BeadChip (50,000 markers, 50K); Illumina Inc., San Diego, CA] and (b) 311,725 [BovineHD Genotyping BeadChip (high density, HD); Illumina] marker evaluations. Color version available in the online PDF.

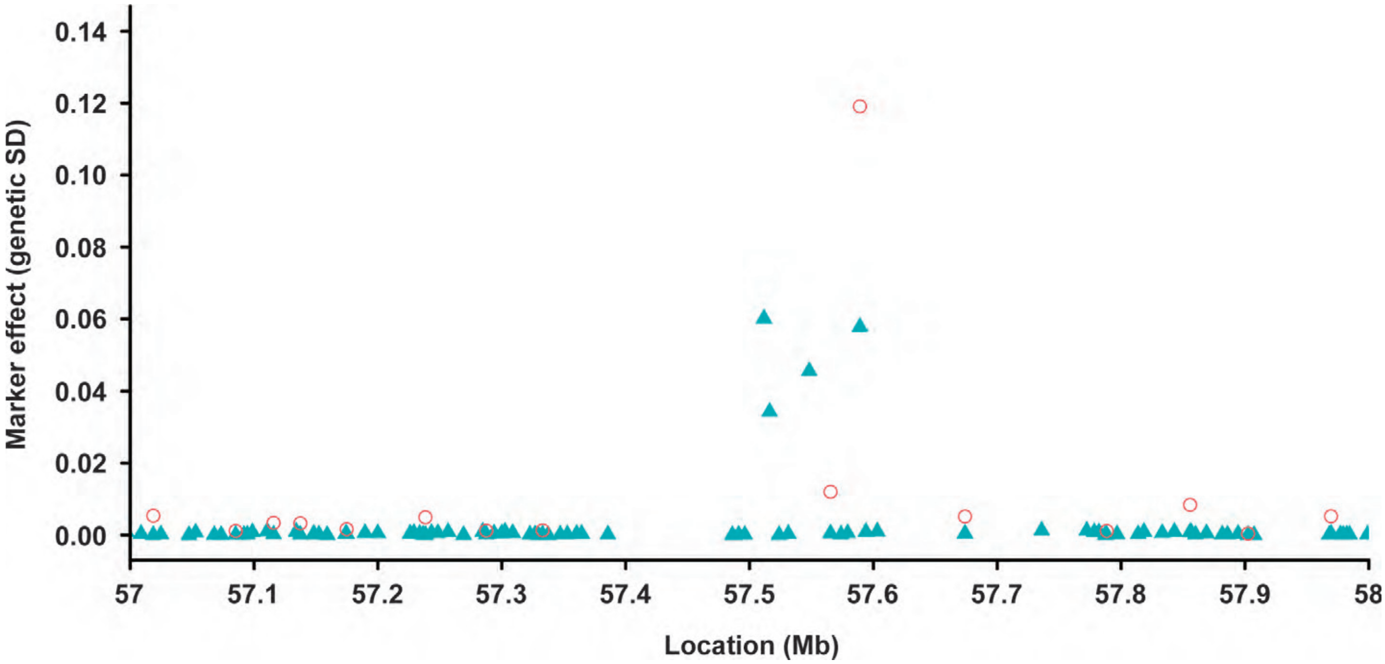

Figure 3. Effects on a 1 Mbase region of BTA18 from 45,187 (circles) and 311,725 (triangles) marker evaluations of productive life. Color version available in the online PDF.

why the gain of 1.6 percentage points in HD reliability over 50K for productive life in Table 5 was larger than for most other traits.

CONCLUSIONS

Genotypes from the Illumina BovineHD chip were of very high quality, but about half of the 777,962 markers were not used because of strong correlations with adjacent markers. Automated marker edits were developed to make visual inspection of cluster quality less necessary. Very large numbers of haplotypes within a few segments indicated that 140 markers were apparently mapped to incorrect positions, and those markers were removed. Accurate imputation is a key to ensuring that the benefits from more markers exceed the imputation loss because gains from HD are small. Imputation to HD gave 99.3% correct genotypes from 50K, 96.1% from 6K, and 93.7% from 3K. Imputation between 2 different HD chips or between HD and sequence data can be done accurately with reasonable computational effort. Increasing the number of markers gave only a 0.4 percentage point gain in average reliability of genomic predictions for HD compared with 50K, a little less than we expected from simulation, but in agreement with most other studies of actual data. The nonlinear model with heavy-tailed prior distribution for marker effects increased reliability by only 0.8 percentage point compared with a linear model. Reliability improved when the number of animals with HD genotypes used for imputation increased from 342 to 1,074 but did not improve with a further increase to 1,510 HD animals. Imputation and evaluation were both computationally affordable for 161,341 total animals currently genotyped, requiring about 10 h and 2 d, respectively, but benefits from higher density genotypes were small within the Holstein breed.

ACKNOWLEDGMENTS

The authors thank the Ministero delle Politiche Agricole Alimentari e Forestali (MIPAAF, Rome, Italy) for funding the HD genotypes contributed by the Innovagen project (DM 10750-7303-2011), Defra (London, UK) for funding the HD genotypes contributed by the United Kingdom as part of the Ruminant Genetic Improvement Network, Ezequiel Nicolazzi (Parco Tecnologico Padano, Lodi, Italy) for prioritizing Italian bulls to be HD genotyped, Kent Weigel (University of Wisconsin, Madison) for prioritizing bulls to be HD genotyped from the Cooperative Dairy DNA Repository, Alicia Beavers (Bovine Functional Genomics Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD) for processing the HD assays for the Beltsville dairy herd, and Kim Fletcher-Carroll and Suzanne Hubbard (Animal Improvement Programs

Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD) for technical editing and review of the manuscript.

REFERENCES

Boichard, D., H. Chung, R. Dassonneville, X. David, A. Eggen, S. Fritz, K. J. Gietzen, B. J. Hayes, C. T. Lawley, T. S. Sonstegard, C. P. Van Tassell, P. M. VanRaden, K. Viaud-Martinez, and G. R. Wiggans. 2012. Design of a bovine low-density SNP array optimized for imputation. PLoS ONE 7:e34130.

Center for Bioinformatics and Computational Biology. 2010. *Bos taurus* assembly. Accessed Jul. 9, 2012. http://www.cbcb.umd.edu/research/bos_taurus_assembly.shtml.

Chen, J., Z. Liu, F. Reinhardt, and R. Reents. 2011. Reliability of genomic prediction using imputed genotypes for German Holsteins: Illumina 3K to 54K bovine chip. Interbull Bull. 44:51–54.

Clark, S. A., J. M. Hickey, and J. H. J. van der Werf. 2011. Different models of genetic variation and their effect on genomic evaluation. Genet. Sel. Evol. 43:18.

Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J. Dairy Sci. 92:2931–2946.

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

Harris, B. L., F. E. Creagh, A. M. Winkelman, and D. L. Johnson. 2011. Experiences with the Illumina high density bovine beadchip. Interbull Bull. 44:3–7.

Harris, B. L., and D. L. Johnson. 2010. The impact of high density SNP chips on genomic evaluation in dairy cattle. Interbull Bull. 42:40–43.

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

Meuwissen, T., and M. Goddard. 2010. Accurate prediction of genetic values for complex traits by whole-genome resequencing. Genetics 185:623–631.

Rincon, G., K. L. Weber, A. L. Van Eenennaam, B. L. Golden, and J. F. Medrano. 2011. Hot topic: Performance of bovine high-density genotyping platforms in Holsteins and Jerseys. J. Dairy Sci. 94:6116–6121.

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., B. Heringstad, M. Svendsen, H. Grove, and T. H. E. Meuwissen. 2011. Genomic predictions for production- and functional traits in Norwegian Red from BLUP analyses of imputed 54K and 777K SNP data. Interbull Bull. 44:240–243.

Su, G., R. F. Brøndum, P. Ma, B. Guldbrandtsen, G. P. Aamand, and M. S. Lund. 2012. Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. J. Dairy Sci. 95:4657–4665.

VanRaden, P. 2011. findhap.f90. Accessed Jul. 9, 2012. http://aipl.arsusda.gov/software/findhap.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.

VanRaden, P. M., D. J. Null, G. R. Wiggans, T. S. Sonstegard, and E. E. Connor. 2011a. Genomic imputation and evaluation using 342 high-density Holstein genotypes. J. Dairy Sci. 94(E-Suppl. 1):533. (Abstr.)

VanRaden, P. M., J. R. O'Connell, G. R. Wiggans, and K. A. Weigel. 2011b. Genomic evaluations with many more genotypes. Genet. Sel. Evol. 43:10.

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

Wiggans, G. R., T. A. Cooper, P. M. VanRaden, K. M. Olson, and M. E. Tooker. 2012a. Use of the Illumina Bovine3K BeadChip in dairy genomic evaluation. J. Dairy Sci. 95:1552–1558.

Wiggans, G. R., P. M. VanRaden, T. A. Cooper, C. P. Van Tassell, T. Sonstegard, and B. Simpson. 2012b. Characteristics and use of the Illumina BovineLD BeadChip. J. Dairy Sci. 95(Suppl. 2):447. (Abstr.)

Williams, A. L., N. Patterson, J. Glessner, H. Hakonarson, and D. Reich. 2012. Phasing of many thousands of genotyped samples. Am. J. Hum. Genet. 91:238–251.