Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding

D. W. Bjelland,*[1] K. A. Weigel,* N. Vukasinovic,† and J. D. Nkrumah†
*Department of Dairy Science, University of Wisconsin, Madison 53706
†Pfizer Animal Genetics, Kalamazoo, MI 49007

ABSTRACT

The effects of increased pedigree inbreeding in dairy cattle populations have been well documented and result in a negative impact on profitability. Recent advances in genotyping technology have allowed researchers to move beyond pedigree analysis and study inbreeding at a molecular level. In this study, 5,853 animals were genotyped for 54,001 single nucleotide polymorphisms (SNP); 2,913 cows had phenotypic records including a single lactation for milk yield (from either lactation 1, 2, 3, or 4), reproductive performance, and linear type conformation. After removing SNP with poor call rates, low minor allele frequencies, and departure from Hardy-Weinberg equilibrium, 33,025 SNP remained for analyses. Three measures of genomic inbreeding were evaluated: percent homozygosity (F_{PH}), inbreeding calculated from runs of homozygosity (F_{ROH}), and inbreeding derived from a genomic relationship matrix (F_{GRM}). Average F_{PH} was 60.5 ± 1.1%, average F_{ROH} was 3.8 ± 2.1%, and average F_{GRM} was 20.8 ± 2.3%, where animals with larger values for each of the genomic inbreeding indices were considered more inbred. Decreases in total milk yield to 205 d postpartum of 53, 20, and 47 kg per 1% increase in F_{PH}, F_{ROH}, and F_{GRM}, respectively, were observed. Increases in days open per 1% increase in F_{PH} (1.76 d), F_{ROH} (1.72 d), and F_{GRM} (1.06 d) were also noted, as well as increases in maternal calving difficulty (0.09, 0.03, and 0.04 on a 5-point scale for F_{PH}, F_{ROH}, and F_{GRM}, respectively). Several linear type traits, such as strength (−0.40, −0.11, and −0.19), rear legs rear view (−0.35, −0.16, and −0.14), front teat placement (0.35, 0.25, 0.18), and teat length (−0.24, −0.14, and −0.13) were also affected by increases in F_{PH}, F_{ROH}, and F_{GRM}, respectively. Overall, increases in each measure of genomic inbreeding in this study were associated with negative effects on production and reproductive ability in dairy cows.

Key words: genomic, inbreeding depression, runs of homozygosity

INTRODUCTION

Inbreeding in US dairy cattle has increased steadily over the past several decades. The average inbreeding coefficient of Holsteins rose from 0.4% in 1970 to 5.8% in 2012 (USDA-AIPL, 2012), with the majority of this increase attributed to intense selection pressure on bulls. Some Holstein bulls used in AI have had as many as 250,000 milking daughters or 5,000 progeny-tested sons (Weigel, 2001). In addition, of the roughly 5,000 young bulls that were progeny tested each year globally at that time, almost 50% were offspring of the 10 most popular sires. Even with extensive pedigree records, avoidance of inbreeding is increasingly difficult.

The negative effects of inbreeding have been well documented and tend to fall into 2 categories. The first is an increased prevalence of rare lethal or harmful recessive disorders, such as bovine leukocyte adhesion deficiency (BLAD; Kehrli et al., 1990) or deficiency of uridine monophosphate synthase (DUMPS; Shanks et al., 1984), when closely related individuals are mated. The second is an overall decrease in functionality, performance, and profitability of inbred animals. Many studies have concluded that increased pedigree inbreeding in dairy cattle is associated with decreases in production (Miglior et al., 1995; Smith et al., 1998; Thompson et al., 2000), reproductive ability (Smith et al., 1998; Mc Parland et al., 2007), and survivability (Thompson et al., 2000; Mc Parland et al., 2007). Smith et al. (1998) indicated that for each 1% increase in inbreeding, lifetime total milk yield decreased by 94.5 kg, lifetime total fat yield decreased by 3.3 kg, and lifetime total protein yield decreased by 2.9 kg. Effects of inbreeding on reproductive traits and survival were measured by Mc Parland et al. (2007), and results indicated a 0.7-d increase in calving interval and a 0.3% decrease in survival to second lactation per 1% increase in inbreeding. Furthermore, Adamec et al. (2006) noted an increase in probability of maternal dystocia (0.42 and 0.30% for male and female calves, respectively) and

stillbirths (0.25 and 0.20% for male and female calves, respectively) per 1% increase in inbreeding.

With the recent development of high-throughput genomic tools, such as the Illumina Bovine SNP50 BeadChip (Illumina Inc., San Diego, CA), many new questions have arisen regarding inbreeding. Results from a simulation study by de Roos et al. (2011) noted that, with the reduction in generation interval from the use of genomic selection, the rate of genetic gain per generation could double at a given rate of inbreeding per generation. At the same time, however, the rate of inbreeding per year will increase, due to the reduction in generation interval. Daetwyler et al. (2007) noted that genomic selection will be able to predict the Mendelian sampling between full sibs without progeny testing. This will reduce the incidence of co-selection of siblings, as well as the overall genetic impact of single influential animals on the population, which may lead to slower accumulation of inbreeding.

Previous studies have considered only pedigree-based estimates of inbreeding, but with the availability of whole-genome marker panels the next logical step is to quantify inbreeding genomically. The inbreeding coefficient is defined as the probability that a pair of alleles is identical by descent (**IBD**). Historically, geneticists have estimated this probability using pedigree information. Using genomic information should lead to a more accurate depiction of inbreeding. For example, consider an organism whose parents are first cousins (with no other previous common ancestors). The pedigree inbreeding coefficient (F_{ped}) would be 6.25%; on average, 6.25% of this organism's genome would be identical, having originated from either of the common great-grandparents. Carothers et al. (2006) noted that this value varies greatly due to Mendelian sampling, with a standard deviation of 2.43%. This deviation depends on the recombination events that occurred during gamete formation in the parents, as well as the chance meeting of the successful gametes. Although this deviation is present when estimating inbreeding from pedigrees, genomic estimates of inbreeding should determine the actual product of the recombination events and provide a more accurate estimation.

In the dairy industry, genomic inbreeding coefficients of genotyped animals are currently calculated from a SNP-derived genomic relationship matrix (F_{GRM}). VanRaden et al. (2011b) compared F_{GRM} values to F_{ped} values, and reported correlations of 0.59, 0.68, and 0.61 for Holsteins, Jerseys, and Brown Swiss, respectively. Slightly higher correlations of 0.69 were obtained by Hayes and Goddard (2008) in a population of Australian Angus bulls. To date, no studies have determined the effects of F_{GRM} on lactation yield, fertility, or survivability in dairy cattle.

Increased levels of inbreeding would appear genomically as an increase in the frequency of homozygous alleles. One simple method to determine inbreeding genomically would be to look at the percentage of alleles that are homozygous. A problem with this method is that alleles that are IBD and identical by state (**IBS**) cannot be distinguished and are both included in this measure of inbreeding. An alternative method involving genomic runs of homozygosity (**ROH**) attempts to distinguish these differences and has been used in human genomic studies for almost a decade, examining population history (Li et al., 2008; Kirin et al., 2010) and the effects of inbreeding on disease risk (Simon-Sanchez et al., 2012). Keller et al. (2011) indicated that inbreeding estimates using ROH (F_{ROH}) are preferable to F_{ped} and other measures of genomic inbreeding, because it correlates strongly with homozygous mutation load. More specifically, at an effective population size similar to that of the Holstein cattle population, F_{ROH} had a correlation of about 0.6 to the homozygous mutation load, whereas F_{Ped} only had a correlation of about 0.25. Very few studies involving ROH have been performed in cattle, but Ferencakovic et al. (2011) noted positive correlations (0.61 to 0.68) between varying measures of F_{ROH} and F_{Ped} in a population of 500 Simmental bulls. The current study aims to quantify various measures of genomic inbreeding and determine their associations with economically important traits in dairy cattle.

MATERIALS AND METHODS

Data

A total of 5,853 animals were genotyped for 54,001 SNP markers. After editing SNP for minor allele frequency (<0.05), call rate (percent missing >0.1), and Hardy-Weinberg equilibrium ($P < 0.0001$), 33,025 SNP remained for analysis. Animals with more than 10% missing SNP were also removed from the analysis. Phenotypes were available for 2,913 cows in 9 herds from various geographical regions of the United States. Data compiled for each cow consisted of one lactation record through 205 d from either first, second, third, or fourth parity, with data from 854 cows in first lactation, 1,088 cows in second lactation, 592 cows in third lactation, and 379 cows in fourth lactation. All cows were born between July 1999 and December 2005. Lactations included in this study were initiated between June 2006 and March 2007.

Daily milk yields were collected from all animals from parturition through 205 DIM. From this, total milk yield to 205 DIM, average daily milk yield, and peak milk yield values were derived. Fat percentage, protein percentage, SCS, and MUN were recorded at

60-d intervals throughout the lactation and averaged over the lactation before analysis.

Days open, conception rate, DIM at first breeding, calf birth weight, and calving ease were derived from on-farm reproductive data. Cows were required to be at least 250 d postpartum if they had not been confirmed pregnant, and all cows with greater than 250 d open were set to 250 d, following the process of VanRaden et al. (2004) for routine genetic evaluation of daughter pregnancy rate. Conception rate for each cow was defined as 1 divided by the number of times bred (if confirmed pregnant), and zero otherwise. For periods of estrus in which more than 1 breeding occurred (e.g., 2 breedings within 3 d), only 1 breeding was counted toward the number of times bred. Calving ease was recorded on an ordinal scale from 1 (no assistance) to 5 (extremely difficult birth).

At an average of 97.3 ± 44.7 DIM, all cows were scored for linear type traits, including stature, strength, body depth, dairy form, rump angle, rump width, rear legs side view, rear legs rear view, foot angle, fore udder attachment, udder height, udder width, udder cleft, udder depth, front teat placement, rear teat placement, teat length, and udder tilt. All linear type traits were scored on a 50-point scale by trained evaluators.

Genomic Inbreeding Coefficients

The first measure of genomic inbreeding considered was the percent homozygosity ($\mathbf{F_{PH}}$) of all SNP. F_{PH} was derived with the following formula:

$$F_{PH} = \frac{N_{AA} + N_{BB}}{N_{AA} + N_{AB} + N_{BB}},$$

where N_{AA}, N_{AB}, and N_{BB} refer to the number of SNP that are classified as AA, AB, and BB, respectively.

Although the F_{PH} of an animal can provide some indication of its level of inbreeding, it does not distinguish between markers that are IBS and those that are IBD. One possible method to alleviate this problem is to consider genomic ROH. An ROH is defined as a specific number of consecutive SNP that are all homozygous. Inbreeding increases overall homozygosity in an individual, but this increase does not simply present itself as single randomly dispersed homozygous SNP, but rather as long runs of homozygous SNP that were inherited together. Furthermore, the length of the ROH correlates to the distance within the pedigree until the common ancestor is observed. Longer ROH indicate more recent common ancestors, because recombination has had fewer generations to break up the segments, whereas shorter ROH are indicative of common ancestors further back in the pedigree. If the minimum length

of the ROH is increased, the results would be focused on more recent inbreeding. Fisher (1954) noted that the expected length of the DNA segment that is IBD follows an exponential distribution with mean equal to $\frac{1}{2g}$ morgans, where g equals the number of generations since the common ancestor. Common ancestors occurring 10 generations back would have an average ROH length of 0.05 morgans, or 5 cM. At an average of 1.25 cM/Mb (Arias et al., 2009), converting the minimum length of ROH discovered in this study (about 4 Mb) to centimorgans results in an average ROH length of 5 cM. In practice, discovery of ROH is slightly more complex and was determined using the PLINK whole-genome association analysis toolset (Purcell et al., 2007). Figure 1 describes the manner by which ROH were discovered in PLINK using a sliding window of SNP along the chromosome. First, paternal and maternal chromosomal segments are presented. A sliding window of 10 SNP then moves along the chromosome 1 SNP at a time. This determines whether every SNP inside this window is homozygous. The number of completely homozygous windows, as well as the total number of windows, is summed for each SNP. If, at minimum, 10 consecutive SNP are determined to have greater than 5% of these windows homozygous, an ROH is called. Recommendations for many of the input parameters for ROH discovery were derived from Howrigan et al. (2011). No heterozygous SNP and 1 missing SNP were allowed within the sliding window. The minimum ROH length of 30 SNP (compared with 10 SNP in Figure 1) was used to capture inbreeding occurring in about the previous 10 generations in the current study (Fisher, 1954). This essentially means, compared with pedigree inbreeding, that F_{ROH} is calculated with 10 generations of complete pedigrees, or the base population used for determining F_{ROH} is 10 generations back. Because high linkage disequilibrium within given sections of DNA can lead to detection of ROH that are not truly IBD, linkage disequilibrium pruning was also performed on the SNP set before the ROH determination to increase power, as suggested by Purcell et al. (2007). Linkage disequilibrium pruning was performed using PLINK, and SNP that had a coefficient of determination >0.5 with all other SNP in a 50-SNP window were removed. This resulted in a total of 7,997 SNP being used for the ROH analysis.

The results of the ROH discovery were used to create an inbreeding coefficient for each animal, denoted as F_{ROH}, which was calculated by the following formula:

$$F_{ROH} = \frac{\sum_k \text{length} \left(ROH_k \right)}{L},$$
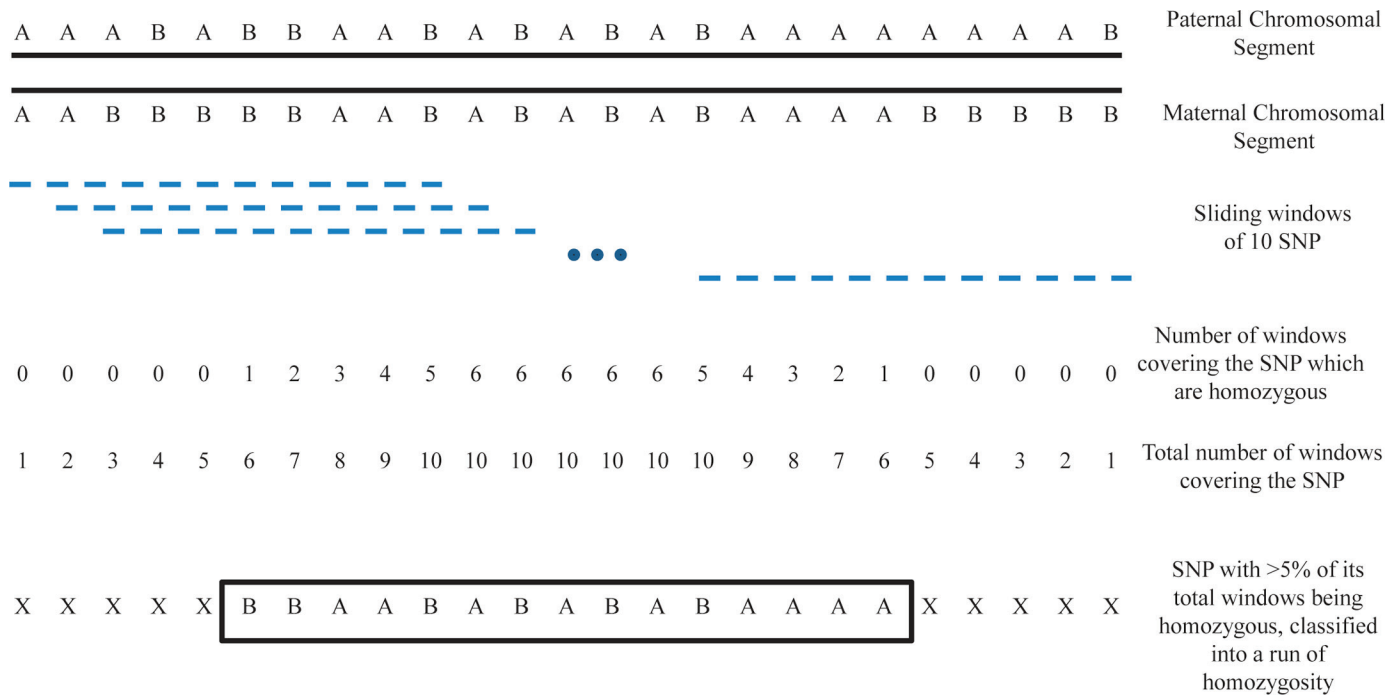
Figure 1. Description of the process for discovery of runs of homozygosity (ROH) using a sliding window of SNP markers along the chromosome, as implemented with PLINK software (Purcell et al., 2007). Color version available in the online PDF.

where k = number of ROH discovered for each animal, and L = total length of the genome. The length of ROH was measured in kilobases, with $L = 2,612,820$ kb (Zimin et al., 2009).

Measures of inbreeding from a genomic relationship matrix (**G**) were denoted as F_{GRM}, and were calculated using the method and programs described by Van-Raden et al. (2011b). This is the method used by the US Department of Agriculture Agricultural Research Service (**USDA-ARS**) Animal Improvement Programs Laboratory (Beltsville, MD) for routine genomic evaluation of US dairy cattle, as well as calculation of published genomic inbreeding values. The **G** matrix was calculated using the following formula:

$$G = \frac{ZZ'}{2\sum p(1-p)},$$

where **Z**, a matrix containing the subtraction of a base population allele frequency from the given marker values, contains the values $0 - 2p$ for homozygotes, $1 - 2p$ for heterozygotes, and $2 - 2p$ for opposite homozygotes, where p is the allele frequency. The F_{GRM} used in the depression analyses was calculated using $P = 0.5$, which is the current method used by the USDA-ARS Animal Improvement Programs Laboratory for their presentation of genomic inbreeding values. The matrix **Z** then contains values of 1 or -1 for homozygotes and 0 for heterozygotes. This essentially makes F_{GRM} a measure of homozygosity that has been transformed to follow a distribution similar to traditional F_{Ped}. The values on the diagonal of **G** denote the relationship of the animal to itself, or its genomic inbreeding coefficient. A second genomic inbreeding index (**F_{GRM-BP}**) was calculated by estimating the allele frequencies p in the base population with the algorithm of Gengler et al. (2007). This method uses the very limited pedigree information available as well as linear mixed-model equations to provide an estimate of the selection and drift of allele frequencies. This method had been previously used by the USDA-ARS Animal Improvement Programs Laboratory to report genomic inbreeding (VanRaden et al., 2011b), and will only be used to compare against F_{PH}, F_{ROH}, and F_{GRM} in the present study.

Statistical Analysis

All statistical analyses were performed using PROC MIXED of SAS (version 9.1; SAS Institute Inc., Cary, NC), and all phenotypic traits were analyzed using linear models that regressed the trait by the measures of inbreeding. All traits were adjusted for effects of herd-year-season and parity. Because some cows were missing up to 15% of daily milk records, the percentage of days with missing records was included in the model for

205-d total milk yield and average daily milk yield. The DIM at peak milk was included as a covariate when analyzing peak milk, and DIM at time of evaluation was included as a covariate when analyzing linear type traits. The type of birth (single/twins) and sex of the calf were included as covariates when analyzing BW of the calf and calving ease. Linear regression coefficients corresponding to the change in each trait per 1% increase in genomic inbreeding, as well as significance tests were derived from these models.

RESULTS

Figure 2 displays the distributions of F_{PH}, F_{ROH}, and F_{GRM}, respectively, with means of 60.5 ± 1.1%, 3.8 ± 2.1%, and 20.8 ± 2.3%. For each distribution, animals with smaller genomic inbreeding values are considered as the least-inbred animals in the population, whereas animals with larger genomic inbreeding values are considered as the most inbred. Correlations between the 3 measures of genomic inbreeding were large, with correlations between F_{PH} and F_{ROH} of 0.81, F_{PH} and F_{GRM} of 0.99, and F_{ROH} and F_{GRM} of 0.81. Furthermore, F_{GRM-BP} had more modest correlations to F_{PH} (0.77), F_{ROH} (0.55), and F_{GRM} (0.78).

Estimates of inbreeding depression for production traits are presented in Table 1. Increases in 1% of F_{PH}, F_{ROH}, or F_{GRM} resulted in decreases in 205-d milk yield of 53, 20, or 47 kg, respectively. Average daily milk yield also exhibited a decrease due to a 1% increase in F_{PH} (0.28 kg/d), F_{ROH} (0.11 kg/d), or F_{GRM} (0.25 kg/d). Furthermore, a slight decrease in MUN was observed when F_{PH} (0.06 mg/dL) or F_{GRM} (0.03 mg/dL) increased by 1%. Peak milk, average fat percentage,

average protein percentage, and average SCS were not affected by changes in F_{PH}, F_{ROH}, or F_{GRM}.

Table 2 displays the estimates of inbreeding depression for reproductive traits. A 1% increase in F_{PH}, F_{ROH}, or F_{GRM} resulted in an increase in days open of 1.76, 1.72, or 1.06 d, respectively. Conception rate also decreased with a 1% increase in F_{ROH} (−0.82%) or F_{GRM} (−0.53%), whereas increases in F_{PH} had no effect. Days in milk at first breeding was not affected by changes in any of the genomic inbreeding measures. Increases in F_{PH} or F_{GRM} resulted in decreases in the BW of the calves born from these cows of 0.4 or 0.2 kg/1% increase, respectively. Furthermore, calving ease scores (measured on a 5-point scale) increased per 1% increase in F_{PH} (0.09), F_{ROH} (0.03), or F_{GRM} (0.04).

Estimates of inbreeding depression for linear type traits are presented in Table 3. Stature, dairy form, rump width, rear legs side view, foot angle, fore udder attachment, udder height, udder width, and udder cleft did not change with an increase in any measure of genomic inbreeding. With each 1% increase in F_{PH}, F_{ROH}, or F_{GRM}, strength decreased (−0.40, −0.11, or −0.19), rear legs rear view tended toward closer hocks (−0.35, −0.16, or −0.14), front teats were closer together (0.35, 0.25, or 0.18), and teat length was shorter (−0.24, −0.14, or −0.13). Furthermore, an increase in F_{PH} resulted in more shallow body depth (−0.25) and a greater forward tilt to the udder (0.24). An increase in F_{ROH} resulted in higher udders (0.14), more forward placement of rear teats (0.25), and greater forward tilt to the udder (0.15). Increases in F_{GRM} resulted in a more shallow body depth (−0.14) and higher pins (−0.14). All values denoted are the estimated change in the linear type trait (measured on a 50-point scale) per
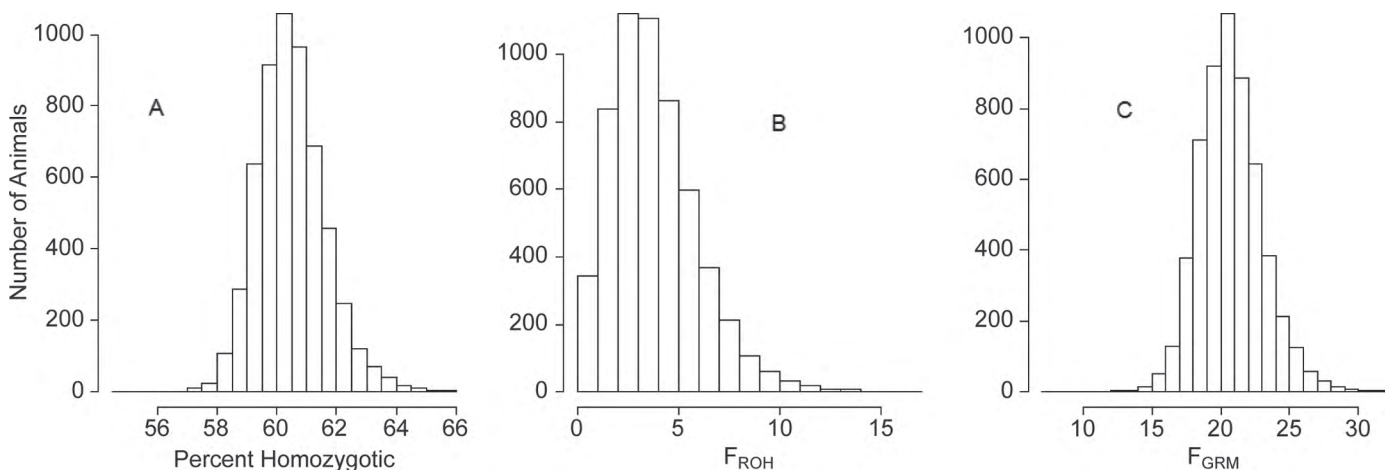
Figure 2. Frequency distribution of animals in the present study, according to (A) percent homozygosity (F_{PH}), (B) runs of homozygosity (F_{ROH}) with a minimum length of 30 SNP, and (C) inbreeding coefficient derived from a genomic relationship matrix (F_{GRM}) using the method of VanRaden et al. (2011b).

Table 1. Estimates of inbreeding depression for production traits, expressed as change in phenotype per 1% increase in percent homozygosity (F_{PH}), inbreeding coefficient derived from runs of homozygosity (F_{ROH}) with a minimum length of 30 SNP, and inbreeding coefficient calculated from a genomic relationship matrix (F_{GRM})

Item	Phenotypic mean	Phenotypic SD	F_{PH} Estimate	SE	F_{ROH} Estimate	SE	F_{GRM} Estimate	SE
205-d milk yield (kg)	8,473	1,586	−53***	19	−20**	10	−47**	22
Peak milk (kg)	64	13	−0.22	0.16	−0.14	0.09	−0.17	0.19
Average daily milk (kg)	44	8	−0.28***	0.10	−0.11**	0.05	−0.25**	0.11
Average fat (%)	3.63	0.59	−0.003	0.008	−0.001	0.005	−0.003	0.004
Average protein (%)	3.01	0.24	−0.002	0.004	0.001	0.002	−0.001	0.002
SCS ($\log_2$ cells/mL)	2.99	0.42	0.001	0.006	0.004	0.003	0.001	0.003
MUN (mg/dL)	13.5	2.4	−0.06**	0.03	−0.02	0.02	−0.03**	0.01

**$P < 0.05$; ***$P < 0.01$.

1% increase in the corresponding measure of genomic inbreeding.

To more accurately depict the expected differences in performance associated with changes in each of the 3 measures of genomic inbreeding, differences between cows with small and large genomic inbreeding coefficients were compared in Table 4. Predicted phenotypes are shown for cows with genomic inbreeding coefficients 2 standard deviations above or below the mean of the corresponding genomic inbreeding measurement (F_{PH}, F_{ROH}, or F_{GRM}). Because the F_{ROH} distribution is slightly skewed to the right, and the mean minus twice the standard deviation results in a negative number, 0 was used as the lower bound for F_{ROH}. Phenotypic values shown in Table 4 were calculated from the phenotypic mean and the estimated regression coefficient for the corresponding trait and genomic inbreeding measure. Cows with high (plus 2 standard deviations) inbreeding coefficients produced less total milk to 205 d (−242, −161, and −438 kg) and had a lower average daily milk yield (−1.28, −0.89, and −2.33 kg) for F_{PH}, F_{ROH}, and F_{GRM}, respectively, than cows with low (minus 2 standard deviations) inbreeding coefficients. Cows with high values for F_{PH} and F_{GRM} also had lower average MUN levels (−0.3 and −0.3 mg/dL, respectively). An increase in 8, 14, and 10 d open was noted between cows with high and low F_{PH}, F_{ROH}, and F_{GRM}, respec-

tively, whereas a decrease in conception rate of 6.6 and 4.9% was noted between cows with high and low F_{ROH} and F_{GRM} values, respectively. Cows with high F_{PH} and F_{GRM} tended to have calves that were lighter (−1.8 and −1.9 kg, respectively) than cows with low values. With a linear regression analysis, an increase of 0.41, 0.24, and 0.37 in average maternal calving ease scores was observed between cows with low and high F_{PH}, F_{ROH}, and F_{GRM}, respectively. Most linear type traits exhibited a difference of between 1 and 2 points when comparing cows with low and high genomic inbreeding coefficients. The largest difference (2.0) was observed for front and rear teat placement when the comparison was made with F_{ROH}, whereas the smallest difference of −0.9 was noted for the association between F_{ROH} and strength.

DISCUSSION

Because pedigree-based inbreeding coefficients were unavailable for animals in this study, comparisons must be based on previous studies. Comparison of genomic and pedigree inbreeding depression from the same animals would be preferred, but lack of reliable pedigree data is a common occurrence on most commercial dairy operations, as was the case with those used in the present study.

As expected, the correlation between F_{PH} and F_{GRM} was extremely high (0.99), as using $P = 0.5$ for an allele

Table 2. Estimates of inbreeding depression for reproductive traits, expressed as change in phenotype per 1% increase in percent homozygosity (F_{PH}), inbreeding coefficient derived from runs of homozygosity (F_{ROH}) with a minimum length of 30 SNP, and inbreeding coefficient calculated from a genomic relationship matrix (F_{GRM})

Item	Phenotypic mean	Phenotypic SD	F_{PH} Estimate	SE	F_{ROH} Estimate	SE	F_{GRM} Estimate	SE
Days open (d)	123	60	1.76*	1.00	1.72***	0.54	1.06**	0.52
Conception rate (%)	59.5	35.5	−0.82	0.60	−0.82**	0.33	−0.53*	0.31
DIM at first breeding (d)	72.6	19.6	0.27	0.24	0.20	0.13	0.14	0.12
Calf birth weight (kg)	40.1	4.8	−0.4**	0.2	−0.1	0.1	−0.2***	0.1
Calving ease (5-point scale)	1.7	0.9	0.09***	0.03	0.03**	0.02	0.04**	0.02

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

Table 3. Estimates of inbreeding depression for linear type traits on a 50-point scale, expressed as change in phenotype per 1% increase in percent homozygosity (F_{PH}), inbreeding coefficient derived from runs of homozygosity (F_{ROH}) with a minimum length of 30 SNP, and inbreeding coefficient calculated from a genomic relationship matrix (F_{GRM})

Item	Phenotypic mean	Phenotypic SD	F_{PH} Estimate	SE	F_{ROH} Estimate	SE	F_{GRM} Estimate	SE
Stature	31.7	8.7	−0.06	0.12	0.07	0.06	−0.06	0.06
Strength	29.0	8.8	−0.40***	0.11	−0.11*	0.06	−0.19***	0.06
Body depth	27.9	8.9	−0.25**	0.11	−0.01	0.06	−0.14**	0.06
Dairy form	29.3	8.1	−0.04	0.11	0.07	0.06	−0.03	0.05
Rump angle	22.6	9.5	−0.20	0.15	−0.05	0.08	−0.14*	0.08
Rump width	28.8	10.2	−0.06	0.13	0.07	0.07	−0.03	0.07
Rear legs side view	27.5	8.2	−0.06	0.13	−0.04	0.07	−0.06	0.07
Rear legs rear view	23.5	10.3	−0.35**	0.16	−0.16*	0.09	−0.14*	0.08
Foot angle	25.5	9.4	−0.16	0.14	−0.05	0.08	−0.04	0.08
Fore udder attachment	23.7	11.1	−0.13	0.17	0.06	0.09	−0.02	0.09
Udder height	26.0	10.9	−0.15	0.16	0.01	0.09	−0.05	0.08
Udder width	29.9	10.6	−0.21	0.15	−0.05	0.08	−0.05	0.08
Udder cleft	32.4	10.7	0.13	0.16	0.12	0.09	0.04	0.08
Udder depth	22.4	11.1	0.11	0.13	0.14**	0.07	0.06	0.07
Front teat placement	29.3	9.9	0.35**	0.15	0.25***	0.08	0.18**	0.08
Rear teat placement	29.6	11.9	0.27	0.18	0.25***	0.10	0.12	0.09
Teat length	27.0	9.1	−0.24*	0.13	−0.14*	0.07	−0.13*	0.07
Udder tilt	25.2	9.5	0.24*	0.14	0.15**	0.07	0.10	0.07

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

frequency in F_{GRM} is essentially a measure of homozygosity that has been adjusted to conform to a distribution similar to pedigree inbreeding. Using the results from Keller et al. (2011) in which F_{ROH} was determined as the optimal method of genomic inbreeding, the correlations between F_{ROH} and F_{GRM} (0.81) and the correlations between F_{ROH} and F_{GRM-BP} (0.55) would suggest that using a uniform base population allele frequency ($P = 0.5$) may be more beneficial than attempting to estimate a base population. These results are similar to those of VanRaden et al. (2011b), in which using allele

frequencies of 0.5 resulted in higher correlations with F_{Ped}. The results of the present study further demonstrate that the method of using allele frequencies of 0.5, which the USDA-ARS Animal Improvement Programs Laboratory currently uses, is preferable to estimating a base population.

Results from the present study are consistent with previous studies involving F_{Ped}. Decreases in total milk yield (Smith et al., 1998; Thompson et al., 2000; McParland et al., 2007), as well as decreases in overall reproductive ability (Smith et al., 1998; McParland

Table 4. Estimates of inbreeding depression for all significant traits, expressed as the difference (Diff) in predicted phenotype between ±2 SD from the mean for percent homozygosity (F_{PH}), inbreeding coefficient derived from runs of homozygosity (F_{ROH}) with a minimum length of 30 SNP, and inbreeding coefficient calculated from a genomic relationship matrix (F_{GRM})

Item	F_{PH} 58.20%	62.76%	Diff	F_{ROH} 0%	8.06%	Diff	F_{GRM} 16.10%	25.42%	Diff
205-d milk yield (kg)	8,594	8,352	−242	8,554	8,392	−161	8,692	8,254	−438
Average daily milk (kg)	45	43	−1.28	44	43	−0.89	45	43	−2.33
MUN (mg/dL)	13.7	13.4	−0.3	—	—	—	13.7	13.4	−0.3
Days open (d)	119	127	8	116	130	14	118	128	10
Conception rate (%)	—	—	—	62.8	56.2	−6.6	62.0	57.0	−4.9
Calf birth weight (kg)	41.0	39.2	−1.8	—	—	—	41.0	39.2	−1.9
Calving ease (5-point scale)	1.5	1.9	0.41	1.6	1.8	0.24	1.5	1.9	0.37
Strength	29.9	28.1	−1.8	29.5	28.6	−0.9	29.9	28.1	−1.8
Body depth	28.4	27.3	−1.1	—	—	—	28.5	27.2	−1.3
Rump angle	—	—	—	—	—	—	23.2	21.9	−1.3
Rear legs rear view	24.3	22.7	−1.6	24.2	22.9	−1.3	24.2	22.9	−1.3
Udder depth	—	—	—	21.8	23.0	1.1	—	—	—
Front teat placement	28.5	30.1	1.6	28.3	30.3	2.0	28.4	30.1	1.7
Rear teat placement	—	—	—	28.6	30.6	2.0	—	—	—
Teat length	27.5	26.4	−1.1	27.5	26.4	−1.1	27.6	26.3	−1.2
Udder tilt	24.6	25.7	1.1	24.6	25.8	1.2	—	—	—

et al., 2007) as inbreeding increases have been noted previously. Furthermore, Adamec et al. (2006) had also noted an increase in dystocia as inbreeding increased. Smith et al. (1998) provided estimates of expected differences in milk yield and calving interval between cows with 0% inbreeding and those with 12.5% inbreeding, which would correspond to a cow produced from a half-sib mating. At this inbreeding level, first-lactation mature equivalent milk yield would decrease by 464 kg. The genomic inbreeding measures of F_{PH} (-242 kg) and F_{ROH} (-161 kg) in Table 4 indicate less overall milk yield depression than that reported by Smith et al. (1998), but the estimate for F_{GRM} in Table 4 is similar (-438 kg) to that of Smith et al. (1998). Note that values presented in Table 4 span 4 standard deviations, whereas comparing F_{Ped} of 0 and 12.5% spans a range of 5 standard deviations. Also, the current study considered only milk yield to 205 d postpartum, whereas Smith et al. (1998) considered milk yield to 305 d. Much larger realized depression was observed in reproductive ability in the current study than in the study by Smith et al. (1998). When increasing F_{Ped} from 0 to 12.5%, an increase in calving interval of 3.3 d was predicted, but for the range of 4 standard deviations in F_{PH}, F_{ROH}, and F_{GRM} shown in Table 4, increases of 8, 14, and 10 d open were predicted.

Several linear type traits in the current study and in the study by Smith et al. (1998) shared significance (strength, body depth, udder depth, and front teat placement). For both studies, greater levels of inbreeding resulted in a narrower chest with less body depth, as well as a higher udder with closer front teats. Croquet et al. (2007) presented similar results, noting that more inbred cows tended to have narrower, smaller frame size with less body depth. Mc Parland et al. (2007) examined the same traits in Irish Holstein cows and reported opposite results with respect to strength and body depth, although they noted that this unexpected result may have been due to linebreeding or directional selection for these traits. The association with shallower udder depth in this study and others (Smith et al., 1998; Mc Parland et al., 2007) may reflect the fact that highly inbred animals tend to produce less milk, which may result in less volume and depth of the udder.

As shown in Figure 2, the 3 measures of genomic inbreeding considered in this study have very different probability density functions, means, and standard deviations. Furthermore, the number of SNP, size of the SNP chip, and selection criteria for the SNP used to determine the genomic inbreeding values can have a huge effect on these values. For example, the selection of SNP for the Illumina Bovine3K BeadChip focused on SNP that were more polymorphic than SNP on the Bovine SNP50 BeadChip (Illumina Inc., 2011). An index

for F_{PH} from the 3,000-marker (3K) chip would provide lower inbreeding values, simply due to preselection that has occurred among the SNP. Although each of the genomic inbreeding indexes in this study were associated with inbreeding depression, values presented to dairy producers should be consistent if genomic inbreeding is to be used effectively in selection decisions.

Using ROH may provide the most effective, consistent, and easily understood genomic inbreeding values. As shown in Figure 2B, the distribution of F_{ROH} values is very similar to what is normally seen with pedigree inbreeding. Changes in the number of SNP in a SNP panel can be accommodated easily by simply changing the minimum number of SNP in ROH determination. For example, when using 3K genomic data, a minimum length of 15 SNP may correlate closely to inbreeding that occurred in the previous 10 generations, as was the case for the minimum length of 30 SNP used in this study. Furthermore, the basic definition of F_{ROH}, the percentage of the genome that is IBD, is the definition of pedigree inbreeding as well. Determination of whether a homozygous SNP is IBD or IBS is important when examining genomic inbreeding, and using ROH is the most effective method presented herein to distinguish between IBD and IBS. In pedigree inbreeding, the determination of whether an allele is IBD is in reference to a base population. As has been mentioned, the estimation of a base population for use in a genomic relationship matrix is a difficult problem, and many of the methods may not provide a better estimation than simply using an allele frequency of 0.5 (VanRaden et al., 2011b). This same difficulty would occur when attempting to correct F_{PH} to a base population. This is alleviated with the use of ROH, though. Varying the minimum length of ROH discovered is analogous to changing the base population in pedigree inbreeding. A shorter minimum ROH would provide more ancient inbreeding (a base population occurring many generations previously), whereas a longer minimum ROH would only include more recent inbreeding (a base population of just several generation previously). Also, Keller et al. (2011) determined that F_{ROH} values correlate much higher to homozygous mutation load (0.6) than another measure of genomic inbreeding, which would be analogous to F_{PH} and F_{GRM} (0.45). Although all measures of genomic inbreeding presented had a higher correlation with homozygous mutation load than did F_{Ped} (0.25). Combining all of these aspects suggests that F_{ROH} would be the most effective and easily understood method of genomic inbreeding presented in the current study.

Several challenges may occur when using genomic inbreeding to predict inbreeding depression. One is that the traits analyzed to determine inbreeding depression are also traits that are under directional selection in

the population, meaning that increased homozygosity at some loci, or presence of some ROH, may actually be beneficial. Results of this study indicate, however, that overall increases in homozygosity (genomic inbreeding) are associated with decreased functionality and productivity. Much as a tradeoff between pedigree inbreeding and selection intensity existed in traditional selection, a tradeoff between genomic inbreeding and selection for homozygosity of favorable alleles may exist in genomic selection. Sonesson et al. (2012) noted that, when calculating genomic breeding values, a correction based on genomic inbreeding (instead of traditional pedigree inbreeding) is required. Cole and VanRaden (2010) previously demonstrated the selection of a "supercow," which would have the 30 best possible chromosomes. This cow would have a PTA for lifetime net merit of +$3,148, which is about 3.5 times greater than the highest living animal at that time ($911). Most likely, this "supercow" would be homozygous at large proportions of its genome, and although its breeding value would be superior, results of this study suggest that the actual production and reproductive ability would most likely be reduced due to inbreeding depression.

Another possible limitation is due to errors in genotyping, which could be exacerbated by the lack of pedigree information for cows in this study. Wiggans et al. (2011) noted that with pedigrees and genotypes from parents, misclassified SNP could often be corrected in the progeny. This type of correction, finding SNP labeled as homozygous that were actually heterozygous and vice versa, could provide more accurate estimates of genomic inbreeding, but correction of genotypes based on pedigrees may not be possible on many commercial dairies. Many genotyped animals on commercial farms would have genotyped sires, but few of the dams of these animals would have been genotyped. Errors in genotyping would most likely result in underestimation of F_{ROH}. Because no heterozygous SNP were allowed within an ROH, an error in which a homozygous SNP is mistakenly identified as heterozygous may result in a section of the genome that is not identified as IBD. This problem is less severe with F_{PH} and F_{GRM}, as those measures apply to single SNP.

VanRaden et al. (2011a) discovered the presence of 5 recessive defects (haplotypes affecting fertility) in the Holstein, Brown Swiss, and Jersey populations. These defects were discovered by noting that no homozygous recessive animals were present in the population, despite a large number of heterozygous animals, suggesting that individuals that were homozygous for the haplotype in question did not survive full term. This demonstrates both an effect of increased inbreeding and another possible way that genomic inbreeding measures presented in the current study may have been underestimated. If a sire and dam that were heterozygous for one of these haplotypes were mated, the resulting live progeny may be less genomically inbred than expected. As the present study focused on global genomic inbreeding, the results of VanRaden et al. (2011a) suggest that further examination into local inbreeding may also be beneficial when examining production, reproduction, and health traits.

CONCLUSIONS

The 3 methods to quantify genomic inbreeding discussed in this study all demonstrate inbreeding depression for economically important traits in dairy cattle. Lactation performance and reproductive ability were negatively affected when any measure of genomic inbreeding increased. Among the methods considered (F_{PH}, F_{ROH}, and F_{GRM}), only F_{ROH} could distinguish between markers that were IBS and markers that were IBD. This information is important when evaluating the impact of inbreeding and when attempting to control inbreeding using computerized mate selection algorithms. Furthermore, F_{ROH} has been previously shown to be the most correlated with homozygous mutation load and can also exploit the concept of a base population more effectively than the other measures of genomic inbreeding. Further work is needed to estimate the lifetime economic effects of increases in genomic inbreeding, as well as optimal strategies to balance genomic inbreeding and response to selection.

ACKNOWLEDGMENTS

REFERENCES

Adamec, V., B. G. Cassell, E. P. Smith, and R. E. Pearson. 2006. Effects of inbreeding in the dam on dystocia and stillbirths in US Holsteins. J. Dairy Sci. 89:307–314.

Arias, J. A., M. Keehan, P. Fisher, W. Coppieters, and R. Spelman. 2009. A high density linkage map of the bovine genome. BMC Genet. 10:18 http://dx.doi.org/http://dx.doi.org/10.1186/1471-2156/10/18.

Carothers, A. D., I. Rudan, I. Kolcic, O. Polasek, C. Hayward, A. F. Wright, H. Campbell, P. Teague, N. D. Hastie, and J. L. Weber. 2006. Estimating human inbreeding coefficients: Comparison of genealogical and marker heterozygosity approaches. Ann. Hum. Genet. 70:666–676.

Cole, J. B., and P. M. VanRaden. 2010. Visualization of results from genomic evaluations. J. Dairy Sci. 93:2727–2740.

Croquet, C., P. Mayeres, A. Gillon, H. Hammami, H. Soyeurt, S. Vanderick, and N. Gengler. 2007. Linear and curvilinear effects of inbreeding on production traits for Walloon Holstein cows. J. Dairy Sci. 90:465–471.

Daetwyler, H. D., B. Villanueva, P. Bijma, and J. A. Woolliams. 2007. Inbreeding in genome-wide selection. J. Anim. Breed. Genet. 124:369–376.

de Roos, A. P. W., C. Schrooten, R. F. Veerkamp, and J. A. M. van Arendonk. 2011. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. J. Dairy Sci. 94:1559–1567.

Ferencakovic, M., E. Hamzic, B. Gredler, I. Curik, and J. Sölkner. 2011. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. Agriculturae Conspectus Scientificus 76:325–328.

Fisher, R. A. 1954. A fuller theory of "junctions" in inbreeding. Heredity 8:187–197.

Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate the gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. Animal 1:21–28.

Hayes, B. J., and M. E. Goddard. 2008. Technical note: Prediction of breeding values using marker-derived relationship matrices. J. Anim. Sci. 86:2089–2092.

Howrigan, D. P., M. A. Simonson, and M. C. Keller. 2011. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics 12:460.

Illumina Inc. 2011. GoldenGate Bovine3K Genotyping BeadChip. Accessed Aug. 3, 2011. http://www.illumina.com/Documents/products/datasheets/datasheet_bovine3K.pdf.

Kehrli, M. E., Jr., F. C. Schmalstieg, D. C. Anderson, M. J. Van der Maaten, B. J. Hughes, M. R. Ackermann, C. L. Wilhelmsen, G. B. Brown, M. G. Stevens, and C. A. Whetstone. 1990. Molecular definition of the bovine granulocytopathy syndrome: Identification of deficiency of the Mac-1 (CD11b/CD18) glycoprotein. Am. J. Vet. Res. 51:1826–1836.

Keller, M. C., P. M. Visscher, and M. E. Goddard. 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189:237–249.

Kirin, M., R. McQuillan, C. S. Franklin, H. Campbell, P. M. McKeigue, and J. F. Wilson. 2010. Genomic runs of homozygosity record population history and consanguinity. PLoS ONE 5:e13996 http://dx.doi.org/10.1371/journal.pone.0013996.

Li, J. Z., D. M. Absher, H. Tang, A. M. Southwick, A. M. Casto, S. Ramachandran, H. M. Cann, G. S. Barsh, M. Feldman, L. L. Cavalli-Sforza, and R. M. Myers. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. Science 319:1100–1104.

Mc Parland, S., J. F. Kearney, M. Rath, and D. P. Berry. 2007. Inbreeding effects on milk production, calving performance, fertility, and conformation in Irish Holstein-Friesians. J. Dairy Sci. 90:4411–4419.

Miglior, F., E. B. Burnside, and B. W. Kennedy. 1995. Production traits of Holstein cattle: Estimation of nonadditive genetic variance components and inbreeding depression. J. Dairy Sci. 78:1174–1180.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, and M. A. R. Ferreira. 2007. PLINK: A toolset for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575.

Shanks, R. D., D. B. Dombrowski, G. W. Harpestad, and J. L. Robinson. 1984. Inheritance of UMP synthase in dairy cattle. J. Hered. 75:337–340.

Simón-Sánchez, J., L. L. Kilarski, M. A. Nalls, M. Martinez, C. Schulte, and P. Holmans., International Parkinson's Disease Genomics Consortium, Wellcome Trust Case Control ConsortiumGasser, T., J. Hardy, A. B. Singleton, N. W. Wood, A. Brice, P. Heutink, N. Williams, and H. R. Morris. 2012. Cooperative genome-wide analysis shows increased homozygosity in early onset Parkinson's disease. PLoS ONE 7:e28787 http://dx.doi.org/10.1371/journal.pone.0028787.

Smith, L. A., B. G. Cassell, and R. E. Pearson. 1998. The effects of inbreeding on the lifetime performance of dairy cattle. J. Dairy Sci. 81:2729–2737.

Sonesson, A. K., J. A. Woolliams, and T. H. E. Meuwissen. 2012. Genomic selection requires genomic control of inbreeding. Genet. Sel. Evol. 44:27.

Thompson, J. R., R. W. Everett, and N. L. Hammerschmidt. 2000. Effects of inbreeding on production and survival in Holsteins. J. Dairy Sci. 83:1856–1864.

USDA-AIPL (US Department of Agriculture-Animal Improvement Laboratories). 2012. Bovine inbreeding trends. Accessed Aug. 30, 2012. http://aipl.arsusda.gov/eval/summary/inbrd.cfm.

VanRaden, P. M., K. M. Olson, D. J. Null, and J. L. Hutchison. 2011a. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94:6153–6161.

VanRaden, P. M., K. M. Olson, G. R. Wiggans, J. B. Cole, and M. E. Tooker. 2011b. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 94:5673–5682.

VanRaden, P. M., A. H. Sanders, M. E. Tooker, R. H. Miller, H. D. Norman, M. T. Kuhn, and G. R. Wiggans. 2004. Development of a national genetic evaluation for cow fertility. J. Dairy Sci. 87:2285–2292.

Weigel, K. A. 2001. Controlling inbreeding in modern breeding programs. J. Dairy Sci. 84(E. Suppl.):E177–E184.

Wiggans, G. R., P. M. VanRaden, and T. A. Cooper. 2011. The genomic evaluation system in the United States: Past, present, and future. J. Dairy Sci. 94:3202–3211.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yorke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, *Bos taurus.* Genome Biol. 10:R42.