J. Dairy Sci. 96:8014–8023
http://dx.doi.org/10.3168/jds.2013-6969

Mating programs including genomic relationships and dominance effects[1]

C. Sun,*[2] P. M. VanRaden,† J. R. O'Connell,‡ K. A. Weigel,§ and D. Gianola§
*National Association of Animal Breeders, Columbia, MO 65205
†Animal Improvement Programs Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD 20705-2350
‡School of Medicine, University of Maryland, Baltimore 21201
§Department of Dairy Science, University of Wisconsin–Madison, Madison 53706

ABSTRACT

Computerized mating programs using genomic information are needed by breed associations, artificial-insemination organizations, and on-farm software providers, but such software is already challenged by the size of the relationship matrix. As of October 2012, over 230,000 Holsteins obtained genomic predictions in North America. Efficient methods of storing, computing, and transferring genomic relationships from a central database to customers via a web query were developed for approximately 165,000 genotyped cows and the subset of 1,518 bulls whose semen was available for purchase at that time. This study, utilizing 3 breeds, investigated differences in sire selection, methods of assigning mates, the use of genomic or pedigree relationships, and the effect of including dominance effects in a mating program. For both Jerseys and Holsteins, selection and mating programs were tested using the top 50 marketed bulls for genomic and traditional lifetime net merit as well as 50 randomly selected bulls. The 500 youngest genotyped cows in the largest herd in each breed were assigned mates of the same breed with limits of 10 cows per bull and 1 bull per cow (only 79 cows and 8 bulls for Brown Swiss). A dominance variance of 4.1 and 3.7% was estimated for Holsteins and Jerseys using 45,187 markers and management group deviation for milk yield. Sire selection was identified as the most important component of improving expected progeny value, followed by managing inbreeding and then inclusion of dominance. The respective percentage gains for milk yield in this study were 64, 27, and 9, for Holsteins and 73, 20, and 7 for Jerseys. The linear programming method of assigning a mate outperformed sequential selection by reducing genomic or pedigree inbreeding by 0.86 to 1.06 and 0.93 to 1.41, respectively. Use of genomic over pedigree relationship information provided a larger decrease in expected progeny inbreeding and thus greater expected progeny value. Based on lifetime net merit, the economic value of using genomic relationships was >$3 million per year for Holsteins when applied to all genotyped females, assuming that each will provide 1 replacement. Previous mating programs required transferring only a pedigree file to customers, but better service is possible by incorporating genomic relationships, more precise mate allocation, and dominance effects. Economic benefits will continue to grow as more females are genotyped.

Key words: mating program, genomic relationship, dominance, genotype

INTRODUCTION

Phenotypic performance, animal viability, and dairy farm profitability can be affected negatively by decreased heterozygosity and increased frequency of harmful recessives that result from inbreeding. Computerized mating programs have helped breeders reduce pedigree inbreeding by identifying matings between animals with fewer than average ancestors in common (Weigel, 2001). In the genomic era, dense SNP markers across the whole genome have been widely used for genomic selection. Use of genomic relationships is the best way to reduce progeny homozygosity, even for other SNP that are not genotyped directly (Pryce et al., 2012; Sonesson et al., 2012). Breeders should use genomic relationships to control genomic inbreeding when selection is based on genomic EBV, just as pedigree-based relationships were used to control inbreeding when selection was based on traditional EBV computed from pedigree and phenotypic performance (Sonesson et al., 2012). Use of genomic rather than pedigree relationships in mating plans resulted in almost twice the reduction in progeny homozygosity compared with random mating; this additional reduction in genomic inbreeding of 1 to 2% was worth $5 to $10 for Australian Profit Ranking (Pryce et al., 2012). New programs to minimize genomic inbreeding by comparing genotypes of potential mates are needed by the dairy industry.

Received April 29, 2013.
Accepted August 16, 2013.

[1] The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the US Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

[2] Corresponding author: Chuanyu.Sun@ars.usda.gov

Genomic relationships can indicate the realized proportion of the genome that is identical by descent or by state (VanRaden, 2008; Hayes et al., 2009). Genomic mating programs can avoid inbreeding, even in commercial herds that have incomplete or no pedigrees (Bjelland et al., 2012). Genetic evaluations typically predict the additive value of the alleles of an animal but ignore the inbreeding depression that occurs for most traits when related parents are mated. Including a regression on inbreeding in the model can remove the effects of past inbreeding. The true genetic merit of an animal should include its additive value as well as an adjustment for mean relationship to the population of potential mates (VanRaden and Smith, 1999). If mean relationships can be calculated, the regression on inbreeding can be used both to remove past effects of inbreeding and to predict future effects. Mean relationships are also useful for identifying outcross animals.

Since 2008, the US dairy industry has received genomic inbreeding coefficients for animals as well as estimates of genomic future inbreeding, which is half the mean relationship of an animal to the current population (VanRaden et al., 2011b). Individual genomic relationships with potential mates also could be provided. Traditional US evaluations have been adjusted since 2005 for expected future inbreeding computed from pedigree relationships (VanRaden, 2005) but could be adjusted for genomic future inbreeding instead. Mating programs should remove any adjustment for expected or genomic future inbreeding from the PTA of each mate before including the inbreeding loss for the potential mating (VanRaden and Smith, 1999).

Mates can be assigned using simple methods such as sequential selection of least-related mates (Pryce et al., 2012) or linear programming (Jansen and Wilton, 1985; Weigel and Lin, 2000). Optimal contribution theory can minimize genomic inbreeding in the selection step (Sonesson et al., 2012) with an assumption of random mating, and then an additional mating step is required. In theory, combining the selection and mating steps could be better than separate steps. Linear programming uses simultaneous rather than sequential solving to avoid more-related pairs and find less-related pairs. Weigel and Lin (2000) reported that linear programming had a $17 advantage per mating in Holstein estimated lifetime profit over using a simple limit to avoid inbreeding and a $37 advantage per mating over random mating.

Dominance effects could also be included in mating programs to estimate inbreeding losses more precisely (Toro and Varona, 2010). Misztal et al. (1997) reported dominance variances for Holstein stature that were 11 to 16% of phenotypic variance. However, dominance effects have been rarely included in genetic evaluations because of computational complexity and lack of statistical reliability for estimates of variance components. Initially, most countries only genotyped bulls and a few females, but the increasing availability of cows with phenotypes and genotypes in the US lactation and genotype databases (Figure 1) now provides an opportunity to estimate dominance effects and include those in mating programs.

The objectives of this study were to investigate performance of mating programs that include genomic relationships and dominance effects, to develop a method of rapid delivery of genomic relationships to the industry, and to compare 3 mating strategies (linear programming, simple method, and random mating) for maximizing expected progeny value (**EPV**) for milk yield or lifetime net merit (**LNM**).

MATERIALS AND METHODS

Data

Genotypes were available from the US Department of Agriculture (Beltsville, MD) database as of October 2012 for 7,623 Brown Swiss, 28,618 Jerseys, and 233,482 Holsteins (Table 1). Genotyped females included 1,343 Brown Swiss, 21,767 Jerseys, and 165,540 Holsteins. Only 80 Brown Swiss, 287 Jersey, and 1,518 Holstein bulls were being marketed at that time. To estimate dominance variances and effects, 8,323 Jersey and 30,583 Holstein cows with both a genotype and a phenotype were employed. However, only a few hundred genotyped Brown Swiss cows also had phenotypes.

Genotypes were from 6 different SNP arrays: Illumina Bovine3K, Illumina BovineLD, Illumina BovineSNP50, and Illumina BovineHD (Illumina Inc., San Diego, CA) and GeneSeek Genomic Profiler and GeneSeek Genomic Profiler HD (GeneSeek, Lincoln, NE). All genotypes were restricted and imputed to a BovineSNP50 basis using findhap.f90 software (VanRaden et al., 2011a) before estimating genomic breeding values and dominance effects.

Genomic Relationships

Mating programs are often applied to individual herds, whereas genotypes are stored mainly in central databases. The genomic relationship matrix (**G**) for all Holsteins in this study was a 230,000 × 230,000 table. To reduce computing time and disk space required, relationships of young animals with each other are not currently computed or stored by the US Department of Agriculture in the genotype database. Operationally, genomic relationships of each female to currently marketed bulls could be computed within a central

Table 1. Numbers of animals used for calculating the genomic relationship matrix and dominance effect and used in mating programs by breed

Animals	Brown Swiss	Jersey	Holstein
Genotyped population	7,623	28,618	233,482
Animals in pedigrees of genotyped animals	35,193	138,247	656,079
Marketed males	80	287	1,518
Genotyped cows	1,343	21,767	165,540
Genotyped cows with phenotypes for dominance estimation	—	8,323	30,583
Mating programs			
Males	8	50	50
Cows	79	500	500

database and provided to cooperator databases for use in existing mating programs. Potential options for providing the genomic relationship matrix required for a genomic mating program include (1) computation of relationships only between requested females and bulls via a web query, (2) computation of relationships of all genotyped females with each marketed genotyped bull (e.g., >160,000 females and >1,500 bulls for Holsteins), or (3) computation of relationships between all genotyped animals.

For feasibility and usability, option 2 seems simplest, and option 3 is more difficult because of quadratic growth of the relationship table. Although the 230,000 animal × 45,187 SNP genotype incidence matrix (**Z**) for Holsteins fit in computer memory with 1 byte per genotype, **ZZ′** exceeded the 256 GB of available memory because multiplication needed 8 bytes (double precision). Therefore, **ZZ′** had to be computed in sections using the parallel subroutine DGEMM (Intel Corporation, 2013), and memory was reused again within each section.

For option 2, the genomic relationships of all genotyped females with each of the marketed genotyped bulls were computed and stored for later extraction of information from **G** for a subset of animals based on specified animal keys to simulate breeder requests for specific animals. For ease of use and simple interpretation, **G** is adjusted by regression to make mean diagonals and mean off-diagonals of **G** equal to mean diagonals and mean off-diagonals of the pedigree relationship matrix (**A**). From coefficients of **A** and **G**, a regression coefficient and an intercept were calculated using 2 equations and 2 unknowns and output during monthly calculation of **A** and **G**. The same estimated regression coefficient and intercept were used for all options to ensure consistent calculation, even if rela-

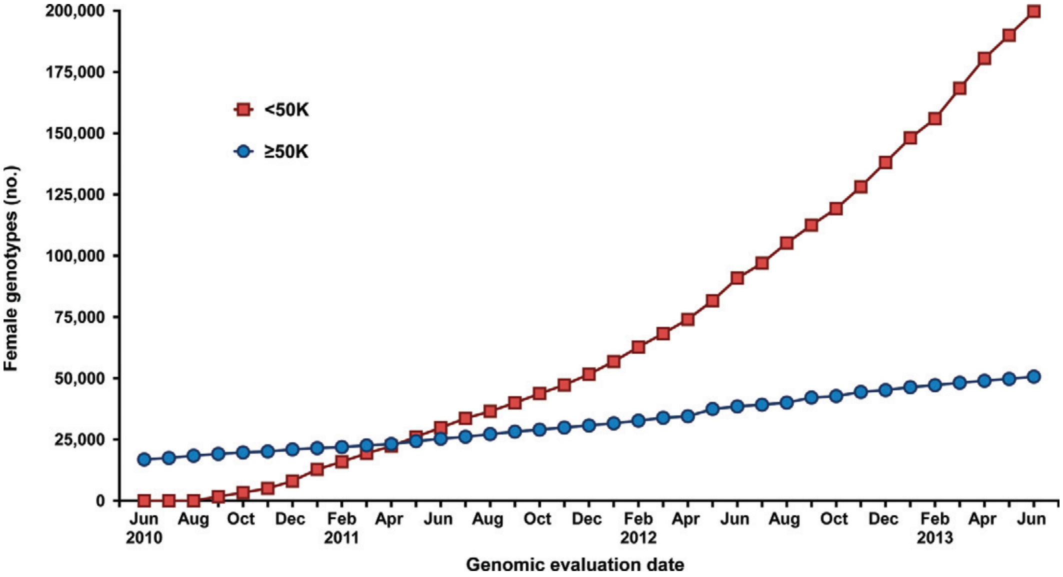

Figure 1. The number of female genotypes in the US database by SNP density of genotyping chip (<50K or ≥50K) and date of genomic evaluation. The <50K genotypes were from Illumina Bovine3K, Illumina BovineLD (Illumina Inc., San Diego, CA), and GeneSeek Genomic Profiler BeadChips (GeneSeek, Lincoln, NE); the ≥50K genotypes were from Illumina BovineSNP50, GeneSeek Genomic Profiler HD, and Illumina BovineHD BeadChips. Color version available in the online PDF.

tionships for only the subset of animals were computed for access through a web query. Genotype and pedigree files were prepared for requested animals, and a subset **G** was generated.

Models With Dominance

Linear mixed models were used to estimate additive and dominance variance components:

$$\mathbf{y} = \mathbf{1u} + \mathbf{W}_a\mathbf{a} + \mathbf{W}_d\mathbf{d} + \mathbf{e},$$

and to predict genetic effects:

$$\mathbf{y} = \mathbf{1u} + \mathbf{Z}_a\mathbf{a}^* + \mathbf{Z}_d\mathbf{d}^* + \mathbf{e},$$

where **y** is a vector of management group deviations for milk yield; **u** is the intercept; **a** and **d** are vectors of additive and dominance effects, respectively, for animals; **a*** and **d*** are vectors of additive and dominance effects, respectively, for SNP; $\mathbf{W}_a$, $\mathbf{W}_d$, $\mathbf{Z}_a$, and $\mathbf{Z}_d$ are incidence matrices; and **e** is the vector of random residuals. In addition, $\mathbf{Z}_a$ is a centered genotype matrix with each $\mathbf{Z}_a$ equal to a genotype code (0, 1, or 2) minus $2p_i$, where p_i is the frequency of the second of 2 alleles at locus i. For homozygous alleles, Z_d equals $0 - 2p_iq_i$, where q_i is the frequency of the first alleles at locus i; for heterozygous alleles, Z_d equals $1 - 2p_iq_i$ (Su et al., 2012a). Then, $\mathbf{a} \sim N\left(0, \mathbf{G}\sigma_a^2\right)$, $\mathbf{d} \sim N\left(0, \mathbf{D}\sigma_d^2\right)$, $\mathbf{a}^* \sim N\left(0, \mathbf{I}\sigma_a^2\right)$, $\mathbf{d}^* \sim N\left(0, \mathbf{I}\sigma_d^2\right)$, and $\mathbf{e} \sim N\left(0, \mathbf{I}\sigma_e^2\right)$, where **G** and **D** are additive and dominance genomic relationship matrices, **I** is an identity matrix, and σ_a^2, σ_d^2, and σ_e^2 are additive, dominance, and residual variances, respectively. The genomic relationship matrices were constructed based on information from genome-wide markers (VanRaden, 2008; Su et al., 2012a). Variance components were estimated using average information-REML (Gilmour et al., 1995) as implemented in the software MMAP (mixed models for pedigrees and populations; O'Connell, 2008, 2013); MMAP incorporates Intel MKL (Intel Corporation, 2013) for optimized parallel likelihood calculation.

Linear Programming

Linear programming is a technique for the optimization of a linear objective function that is subject to constraints of linear equality and inequality:

$$\begin{aligned} f_{\max}(z_{ij}) = {} & z_{11}x_{11} + z_{12}x_{12} + \ldots + z_{1c}x_{1c} + \\ & z_{21}x_{21} + z_{22}x_{22} + \ldots + z_{2c}x_{2c} + \\ & \vdots \\ & z_{b1}x_{b1} + z_{b2}x_{b2} + \ldots + z_{bc}x_{bc}, \end{aligned}$$

where z_{ij} is milk yield or LNM of progeny; b and c are the numbers of bulls and females used in the mating programs, respectively; constraints are

$$x_{i1} + x_{i2} + \ldots + x_{ic} \leq 10\ (i = 1, 2, \ldots, b);$$
$$x_{1j} + x_{2j} + \ldots + x_{bj} = 1\ (j = 1, 2, \ldots, c);$$

and variable bounds are $x_{ij} = 0$ or 1 ($i = 1, 2, \ldots, b$; $j = 1, 2, \ldots, c$).

Mating Programs

To compare mating programs using genomic and pedigree relationship matrices, the 500 youngest (i.e., born after October 1, 2010) genotyped females were selected for Jerseys and Holsteins from the largest herd for each breed. The top 50 marketed bulls of each breed for genomic LNM and traditional LNM as well as 50 randomly selected marketed bulls were used as potential mates. Only 79 females were in the largest Brown Swiss herd, and 8 bulls were selected as potential mates. Matings were limited to 10 females per bull and 1 bull per female.

Mates were assigned using linear programming, the sequential selection of least-related mates (Pryce et al., 2012), or random mating. Let B_{LNM} be defined as the loss of LNM per 1% inbreeding, **EFI** is expected future inbreeding, $G_{\mathrm{sire,dam}}$ is the genomic relationship between sire and dam, and n is the number of matings, which is equal to the number of genotyped females. For each method of choosing mates, mean EPV was

$$\frac{\sum_{j=1}^{n}\left\{\begin{matrix}(\mathrm{GLNM}_{\mathrm{sire}} + \mathrm{GLNM}_{\mathrm{dam}})_j + B_{\mathrm{LNM}}[(\mathrm{EFI}_{\mathrm{sire}} + \mathrm{EFI}_{\mathrm{dam}})_j] \\ -\ 100(0.5)B_{\mathrm{LNM}}[(G_{\mathrm{sire,dam}})_j]\end{matrix}\right\}}{n},$$

where GLNM is genomic LNM from official genomic evaluation, which includes a penalty for EFI (VanRaden, 2005). To avoid double counting EFI, the penalty for EFI should be removed from each mate (as above) before including the inbreeding loss for this specific mating (VanRaden and Smith, 1999). This replaces the mates' average relationship to the population used for selection with their relationship to each other used for mating. For all mating methods, a B_{LNM} of \$23.11 was assumed (Weigel and Lin, 2000). All combinations of genomic and traditional LNM and genomic and traditional inbreeding were examined.

Mating programs including dominance effects were investigated only for milk yield. For each mating, 3 probabilities (2 homozygotes and 1 heterozygote) of progeny genotypes were generated based on parent genotypes. The dominance effect for each progeny was obtained by summing over all loci and the 3 genotype probabilities, giving

$$D_{\text{progeny}} = \sum_{i=1}^{s} d_i^{*} \left(\sum_{j=1}^{3} Z_d s(\text{progeny genotype}|\text{parent genotypes}) \right),$$

where d_i^{*} is dominance deviation for SNP i, and s is the number of SNP. Mating strategies were the same as for LNM, but mean EPV for milk yield was

$$\frac{\sum_{j=1}^{n} \left\{ \begin{aligned} &(\text{GPTA milk}_{\text{sire}} + \text{GPTA milk}_{\text{dam}})_j \\ &+ B_{\text{milk}}[(\text{EFI}_{\text{sire}} + \text{EFI}_{\text{dam}})_j] \\ &- 100(0.5)B_{\text{milk}}[(G_{\text{sire,dam}})_j] + [(D_{\text{progeny}})_j] \end{aligned} \right\}}{n},$$

where GPTA is genomic PTA from official genomic evaluation, and B_{milk} is the loss of milk yield per 1% inbreeding, which was assumed to be 30.4 kg for all mating methods.

For the method of sequential selection of least-related mates (Pryce et al., 2012), a matrix **B** containing EPV of all potential matings was created first, with columns of the matrix corresponding to dams and rows to sires. Then, another matrix of selected mates **M** was constructed, where $M_{ij} = 1$ if the corresponding element B_{ij} was the highest value in the column $\mathbf{B}_j$; all other elements in that column were set to 0 and that dam was excluded from further matings. The maximum number of matings allowable for each sire was set to 10. If B_{ij} corresponded to a sire that had already been assigned to 10 other matings, then the sire-dam combination with the next highest value in $\mathbf{B}_j$ was selected. If that bull also already had 10 matings allocated, then again, the next highest element in $\mathbf{B}_j$ was selected, and so on.

RESULTS

Genomic Relationships

For genotyped Holstein females and marketed bulls, calculation and storage of **G** (Table 2) required >16 h and >400 GB, which indicates that providing **G** for the entire genotyped Holstein population to breeders is impractical. Nearly 2 h was required to extract a subset **G** for 1,817 Holsteins. The programs were written using Fortran 90; R and SAS languages would be slower for a file of >400 GB. For small data sets (such as Brown Swiss and Jerseys), computation of relationships of all genotyped females with each marketed genotyped bull (option 2) worked well.

Computing relationships only for specified animals was the best choice to save calculation time and disk storage. Even though regression coefficients from option 2 were used instead of direct regression of **A** on **G** while calculating **G** for the subset, the genomic relationships for the subset were verified to be identical for options 1 (computation of relationships only between requested females and bulls via a web query) and 2. The advantage of option 1 for large data sets (e.g., Holsteins and Jerseys) was obvious: 6 s to create **G** for 585 Jerseys and 31 s for 1,817 Holsteins.

Mating Program Without Dominance

Table 3 shows the EPV for matings to selected bulls using linear programming or sequential least-related selection relative to random selection and random mating. For each group of bulls, EPV was higher for linear programming than for the sequential method, and both of those methods were better than random mating. For all methods and groups of bulls, EPV was higher when the genomic rather than the pedigree relationship was used as the mate inbreeding source. The EPV was higher when mates were from the top 50 bulls for genomic LNM rather than the top 50 for traditional LNM or random bulls. For matings to the top 50 bulls for genomic LNM, genomic LNM as the mate EBV source, and a genomic source for mate inbreeding, the increase in EPV for the sequential method over random mating was only 64, 72, and 76% of the corresponding increase for linear programming for Brown Swiss, Holsteins, and Jerseys, respectively; if 50 randomly selected bulls

Table 2. Computation times and disk storage required for the genomic relationship matrix (**G**) for genotyped cows and marketed bulls (Table 1) and computation times for extraction or recalculation of **G** for a subset of animals

	G for genotyped cows and marketed bulls		**G** for subset of genotyped cows and marketed bulls		
				Computation time[2]	
Breed	Computation time (h:min:s)	Disk storage (Mbyte)	Animals (no.)	Extraction (h:min:s)	Recalculation (s)
Brown Swiss	00:00:13	31	338	00:00:01	4
Holstein	16:22:42	425,855	1,817	01:58:06	31
Jersey	00:17:11	7,422	585	00:01:46	6

were used instead, the corresponding percentages were 58, 51, and 59. By using genomic rather than pedigree inbreeding for Holstein matings based on linear programming, EPV were higher by \$32 for the top bulls for genomic LNM, \$30 for the top bulls for traditional LNM, and \$30 for randomly selected bulls.

Table 3 also shows the mean genomic inbreeding for progeny. When selection was among the top 50 bulls for genomic LNM and based on genomic relationships, mean genomic inbreeding of progeny using linear programming decreased 1.03, 0.86, and 1.06 percentage points compared with using the sequential method for Brown Swiss, Holsteins, and Jerseys, respectively; corresponding decreases when selection was based on pedigree relationships were 0.93, 1.41, and 1.40 percentage points. Regardless of how bulls were selected, the use of linear programming rather than the sequential method decreased progeny inbreeding, as did the use of genomic rather than pedigree relationships. The cost of computing time for linear programming can be ignored as computing time was only a few seconds when 50 bulls and 500 females were used.

Dominance

Genomic variance components and heritabilities for milk yield are given in Table 4. Additive and dominance variances were 14.9 and 4.1% of phenotypic variance, respectively, for Holsteins and 16.7 and 3.7% of phenotypic variance for Jerseys. Absolute values for Holstein SNP dominance deviations were larger than were those for Jerseys. The difference between breeds partly reflects scaling; the trait mean and variance were larger for Holsteins than for Jerseys.

Mating Program With Dominance

The results of mating programs that include a dominance effect are given in Table 5. Regardless of bull group, mating method, or inbreeding source, EPV for milk yield of Holsteins and Jerseys was higher when dominance effects were included in addition to additive effects and inbreeding. For the top 50 bulls for genomic PTA milk, EPV from linear programming increased 86 kg for Holsteins and 52 kg for Jerseys by including dominance effects compared with 96 kg and 50 kg, respectively, using the sequential method. Similar increases were found when bulls were selected randomly. For matings to the top 50 bulls for genomic PTA milk, a dominance effect included, and a genomic source for mate inbreeding, the increase in EPV for the sequential method over random mating was only 78 and 64% of the corresponding increase for linear programming for Holsteins and Jerseys, respectively; if 50 randomly selected bulls were used instead, the corresponding percentages were 69 and 62. The EPV from mating programs with dominance effects included did not improve much by using genomic rather than pedigree relationships.

When selection was among the top 50 bulls for genomic PTA milk and based on genomic inbreeding, mean genomic inbreeding of progeny (Table 5) using linear programming was 5.38% for Holsteins and 4.34% for Jerseys when a dominance effect was included and 4.62 and 3.63%, respectively, without the dominance effect. Corresponding percentages for 50 bulls selected randomly were 5.52 and 4.10% with dominance and 4.62 and 3.39% without dominance. Progeny inbreeding can be decreased by using linear programming instead of the sequential method and using genomic rather than pedigree relationships for the mating program with a dominance effect included.

DISCUSSION

Delivering genomic relationships from a central database to industry is a key step for implementing mating programs. Computation of relationships only between requested females and bulls via a web query (option 1) was the best solution, and the advantage was obvious for large data sets (i.e., Holsteins). Computing time to create a genomic relationship matrix for 1,817 animals was only 31 s. Option 1 not only saved computing time but also avoided storing the large $\mathbf{G}$. If all the requested animals have genotypes, generating $\mathbf{G}$ is easy with option 1.

What should breeders do if some animals to be mated have not been genotyped? With the single-step blending method (Legarra et al., 2009; Aguilar et al., 2010), a genetic relationship matrix can be constructed by combining SNP marker and pedigree information. The $\mathbf{G}$ for nongenotyped animals ($\mathbf{G}_0$) is $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{G}_{\alpha}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} + \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$, where $\mathbf{A}_{11}$ is a submatrix of $\mathbf{A}$ for genotyped animals, $\mathbf{A}_{22}$ is a submatrix of $\mathbf{A}$ for nongenotyped animals, $\mathbf{A}_{12}$ and $\mathbf{A}_{21}$ are submatrices of $\mathbf{A}$ describing the relationship between genotyped and nongenotyped animals, and $\mathbf{G}_{\alpha} = \alpha\mathbf{G} + (1 - \alpha)\mathbf{A}_{11}$ is the weighted genomic relationship for genotyped animals, where α is the relative weight placed on genomic relationships and $(1 - \alpha)$ is the weight placed on pedigree-based relationships; $\mathbf{G}_{\alpha}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ represents the genomic relationship between genotyped and nongenotyped animals. Su et al. (2012b) found that α had a small effect on reliabilities of genomic EBV but a large effect on the variation of genomic EBV. The effect of α on genomic relationships for nongenotyped animals and mating programs needs to be investigated.

Table 3. Expected progeny value (EPV) and mean genomic inbreeding of progeny from mating of marketed bulls selected for lifetime net merit (LNM) or randomly with the youngest genotyped cows[1] of the same breed in the same herd by mating method, mate EBV and inbreeding sources, and breed

Selected bulls	Mating method	Mate EBV source	Mate inbreeding source	EPV[2] ($) Brown Swiss	EPV[2] ($) Holstein	EPV[2] ($) Jersey	Progeny inbreeding (%) Brown Swiss	Progeny inbreeding (%) Holstein	Progeny inbreeding (%) Jersey
Top 50 for genomic LNM	Linear programming	Genomic LNM	Genomic	205	494	358	6.94	5.17	3.72
			Pedigree	184	462	326	7.87	6.58	5.12
	Sequential least-related[3]	Genomic LNM	Genomic	181	474	333	7.97	6.03	4.78
			Pedigree	175	450	312	8.27	7.09	5.70
	Random	—		138	422	255	9.83	8.31	8.17
Top 50 for traditional LNM	Linear programming	Traditional LNM	Genomic	158	393	307	6.11	4.87	3.41
			Pedigree	136	363	274	7.07	6.15	4.82
	Sequential least-related[3]	Traditional LNM	Genomic	127	372	278	7.45	5.79	4.66
			Pedigree	124	350	263	7.60	6.72	5.32
	Random	—	—	107	314	214	8.36	8.30	7.43
Random 50	Linear programming	Genomic LNM	Genomic	64	70	78	6.64	4.46	3.65
			Pedigree	43	40	42	7.56	5.77	5.22
		Traditional LNM	Genomic	64	70	78	6.64	4.46	3.65
			Pedigree	45	40	41	7.49	5.78	5.26
	Sequential least-related[3]	Genomic LNM	Genomic	37	36	46	7.83	5.97	5.04
			Pedigree	27	21	29	8.26	6.58	5.76
		Traditional LNM	Genomic	32	39	46	8.05	5.84	5.05
			Pedigree	22	24	27	8.47	6.48	5.86
	Random	—	—	0	0	0	9.30	7.51	7.04

[1] Born after October 1, 2010.

[2] Relative to randomly selected bulls that were randomly mated.

[3] Pryce et al. (2012).

Table 4. Variance components, heritabilities, and SNP dominance deviations for milk yield of Holsteins and Jerseys

Parameter	Holstein	Jersey
Variance (kg^2)		
Additive	1,981,596	1,122,446
Dominance	545,960	247,035
Residual	10,731,541	5,370,733
Heritability (%)		
Additive	14.9	16.7
Dominance	4.1	3.7
SNP dominance deviation (kg)		
Range	−18.17 to 15.25	−7.5 to 8.6
Mean (SD)	0.44 (3.38)	0.20 (1.67)

Traditionally, mating plans have constrained the inbreeding of predicted progeny through **A** (e.g., Kinghorn, 1998, 2011). Elements of **A** are the expected proportion of the genome that is identical by descent given pedigree relationships between individuals and are double the inbreeding coefficient for predicted progeny. Genomic information offers new possibilities to control progeny inbreeding, and mating programs using **G** can increase EPV and decrease progeny inbreeding compared with using **A**. The decrease in Holstein genomic inbreeding was 1.41% when mate genomic rather than pedigree inbreeding was used with linear programming and mate genomic LNM. The decrease was slightly larger than that estimated by Pryce et al. (2012) and had a total annual value of (\$494 − \$462)(120,989) = \$3,871,648 when applied to 120,989 females genotyped in the last 12 mo (ending June 2013), assuming that each will provide 1 replacement in its lifetime. The dollar advantage per mating was similar to the finding of Pryce et al. (2012) because the Australian Profit Ranking measures per lactation rather than lifetime net profit and the exchange rate for US and Australian dollars is close to 1.

Using optimal contribution selection (Wray and Goddard, 1994; Meuwissen, 1997), the effect of genomic selection on inbreeding rate was investigated by Sonesson et al. (2010). They found that both genomic and pedigree relationships were successful strategies to control inbreeding rate under genomic selection. However, the genomic inbreeding rate was around 3 times higher when using pedigree rather than genomic optimal contribution. Optimal contribution selection

Table 5. Expected progeny value (EPV) and mean genomic inbreeding of progeny from mating of marketed bulls selected for genomic PTA milk or randomly with the youngest genotyped cows[1] of the same breed in the same herd by mating method, dominance effect inclusion, and mate inbreeding source for Holsteins and Jerseys

Selected bulls	Mating method	Dominance effect included	Mate inbreeding source	EPV[2] (kg)		Progeny inbreeding (%)	
				Holstein	Jersey	Holstein	Jersey
Top 50 for genomic PTA milk	Linear programming	Yes	Genomic	964	732	5.38	4.34
			Pedigree	957	719	5.72	4.96
		No	Genomic	878	680	4.62	3.63
			Pedigree	763	604	6.11	5.11
	Sequential least-related[3]	Yes	Genomic	889	662	5.85	4.98
			Pedigree	881	649	6.11	5.48
		No	Genomic	793	612	5.60	4.83
			Pedigree	714	578	6.66	5.62
	Random	—		618	537	7.92	6.46
Random 50	Linear programming	Yes	Genomic	319	252	5.52	4.10
			Pedigree	313	237	5.83	4.84
		No	Genomic	214	198	4.62	3.39
			Pedigree	134	122	5.92	4.92
	Sequential least-related[3]	Yes	Genomic	220	155	6.08	5.08
			Pedigree	208	142	6.34	5.44
		No	Genomic	112	120	6.10	5.06
			Pedigree	65	92	6.74	5.61
	Random	—	—	0	0	7.57	7.51

[1] Born after October 1, 2010.

[2] Relative to randomly selected bulls that were randomly mated.

[3] Pryce et al. (2012).

based on pedigree relationships restricted inbreeding less at sites with QTL that had large effects, whereas optimal contribution selection based on genomic relationships resulted in a more evenly distributed increase in identity by descent across the genome. Therefore, using genomic instead of pedigree relationships appears to be better at constraining genomic inbreeding under genomic selection, causing fewer stretches of homozygosity and leaving more diversity across the genome and fewer footprints of selection. Changes in homozygosity for causative QTL may be smaller than those genotyped SNP due to incomplete linkage.

Estimation of nonadditive genetic effects in animal breeding is important. The inclusion of dominance effects could increase the accuracy of genomic selection, and predicted dominance effects could also be used to find mating pairs with good combining abilities by recovering inbreeding depression and utilizing possible overdominance (Wellmann and Bennewitz, 2012). However, the estimation of dominance effects requires the availability of direct phenotypes (i.e., genotypes and phenotypes for the same individuals) or single-step evaluation models and calculation of dominance probabilities from sire and maternal grandsire, as in Su et al (2012a). For dairy cattle, the number of bulls genotyped for dense genome-wide marker panels has been steadily increasing; however, availability of genotypes from cows has been limited in most countries. The US databases included many cows with genotypes and phenotypes (Figure 1) to estimate dominance effects, but the estimates of dominance variances were low and only 4.1 and 3.7% of total variance for Holstein and Jersey milk yield, respectively.

Of all the strategies focused on profiting from including dominance effects, mate allocation could be the easiest option. Optimal mate allocation relies on the idea that although selection should be carried out on estimated additive breeding values, animals used for commercial production should be the product of planned mating, which maximizes the overall (additive plus dominance effects) genetic merit of progeny. Mate allocation profits from dominance when the commercial population is created, but only additive effects are transmitted to the next generation. Although application of mate allocation has usually required 2 separate lines as in classical crossbreeding programs or reciprocal recurrent selection, it can be carried out within a single population.

This study indicated that mating programs that include dominance effects can increase EPV (86 kg for Holsteins and 52 kg for Jerseys when selection and mating use genomic PTA for milk, linear programming, and genomic relationships) compared with mating programs that only include additive genetic effects. Toro and Varona (2010) quantified the efficiency of mating allocation under a whole-genome evaluation scenario in terms of genetic response to selection in first and subsequent generations. They found that the advantage of genomic selection with dominance considered compared with dominance excluded ranged from 9 to 14% of expected response; in addition, using mate allocation provided an additional response that ranged from 6 to 22%. Mate allocation can improve the expected genetic response over random mating in the first generation, but gains do not accumulate across generations (Toro and Varona, 2010). Furthermore, the benefits of genomic selection are reduced each generation unless new phenotypic data are collected and genomic predictions updated. However, progeny inbreeding did not decrease by including a dominance effect. A possible reason may be that selection for dominance effects diluted the attempt to minimize genomic inbreeding. When dominance is included, EPV equals genomic PTA plus dominance effect minus inbreeding loss; general inbreeding loss may receive less attention in mate assignment as individual dominance effects receive more attention.

Linear programming was always better than the sequential selection of least-related mates (Pryce et al., 2012) and random mating for improving EPV and decreasing progeny inbreeding in this study. For the sequential method, each female was mated to the bull with the highest trait genetic merit if that bull had not exceeded a specified limit for number of matings. Different ordering of females would result in different mating pairs. Linear programming maximized the linear objective function, gave a globally optimized solution, and was not affected by mating order. In this study, the objective function was maximum expected lifetime profit (i.e., LNM minus inbreeding depression), not maximum LNM or minimum inbreeding only. Weigel and Lin (2000) compared the results of maximum expected lifetime profit and maximized net merit with several possible thresholds for controlling inbreeding; they found that when the objective function was maximum expected lifetime profit, mean Holstein inbreeding level was reduced by 1.8 percentage points relative to methods that ignored inbreeding. Holstein inbreeding also was 0.9, 1.2, and 1.4 percentage points less than when net merit was maximized with inbreeding thresholds of 5, 6, and 7%, respectively. Estimated lifetime profit increased an economic benefit per mating of \$16.66 relative to use of a 5% inbreeding threshold and \$37.37 relative to maximization of LNM, regardless of inbreeding.

In any mate assignment system, some attention to nonlinear economic values of traits (corrective mating) or positive assortative mating to increase variance of net merit could be better than simply minimizing inbreeding (Allaire, 1993). The economic value of net merit is nonlinear, with much higher prices for elite animals, but the current research did not include nonlinear economics or positive assortative mating to increase probability of obtaining elite progeny. Commercial mating programs typically need 1 or 2 alternative mate choices in addition to the optimal choice. A convenient method for obtaining alternative mates is to set the optimal solutions to much less favorable values and rerun the linear programming.

CONCLUSIONS

An effective method of transferring elements of **G** from a central database to customers was developed to allow implementation of genomic mating programs. Mating programs that include genomic relationships were more effective than those using pedigree relationships, and the expected decrease in inbreeding was worth >$3 million annually for US Holsteins. Extra benefit was gained when dominance effects were included in the mating program. Combining linear programming and genomic relationships was always better (i.e., largest EPV and lowest progeny inbreeding) than other methods regardless of the mates selected or the inclusion of a dominance effect. The best combination was mating to animals selected for predicted genomic merit and using linear programming, predicted genomic merit of mates adjusted for expected future inbreeding, genomic relationships, and inclusion of a dominance effect if available.

ACKNOWLEDGMENT

The MMAP software was supported by funding from NIH grant U01 HL084756 and National Research Initiative Competitive Grant no. 2007-35205-17883 from the USDA National Institute of Food and Agriculture (Washington, DC).

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752.

Allaire, F. R. 1993. Intrabreed mating pair selection methods for improvement of nonlinear merit from additive genetic inheritance: A review. J. Dairy Sci. 76:2308–2319.

Bjelland, D. W., K. A. Weigel, D. J. Nkrumah, and N. Vukasinovic. 2012. Effects of genomic inbreeding on production, reproduction, and udder health in Holstein dairy cows. J. Dairy Sci. 95(Suppl. 2):400. (Abstr.)

Gilmour, A. R., R. Thompson, and B. R. Cullis. 1995. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51:1440–1450.

Hayes, B. J., P. M. Visscher, and M. E. Goddard. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet. Res. (Camb.) 91:47–60.

Intel Corporation. 2013. Intel Math Kernel Library Reference Manual. Doc. No. 630813-061US, MKL 11.0, update 5. Accessed Apr. 6, 2013. http://download-software.intel.com/sites/products/documentation/doclib/mkl_sa/11/mklman/mklman.pdf.

Jansen, G. B., and J. W. Wilton. 1985. Selecting mating pairs with linear programming techniques. J. Dairy Sci. 68:1302–1305.

Kinghorn, B. P. 1998. Managing genetic change under operational and cost constraints. Pages 9–16 in Proc. 36th Natl. Congr. S. Afr. Soc. Anim. Sci., University of Stellenbosch, South Africa.

Kinghorn, B. P. 2011. An algorithm for efficient constrained mate selection. Genet. Sel. Evol. 43:4.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663.

Meuwissen, T. H. E. 1997. Maximizing the response of selection with a predetermined rate of inbreeding. J. Anim. Sci. 75:934–940.

Misztal, I., T. J. Lawlor, and R. L. Fernando. 1997. Dominance models with method R for stature of Holsteins. J. Dairy Sci. 80:975–978.

O'Connell, J. R. 2008. Optimizing measured genotype genome-wide association analysis for quantitative traits in pedigrees. Abstr. 2409 in 58th Annu. Mtg. Am. Soc. Hum. Genet., Philadelphia, PA. Accessed Sep. 23, 2013. http://www.ashg.org/2008meeting/abstracts/fulltext/f22593.htm.

O'Connell, J. R. 2013. MMAP User Guide. Accessed Apr. 29, 2013. http://edn.som.umaryland.edu/mmap/index.php.

Pryce, J. E., B. J. Hayes, and M. E. Goddard. 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. J. Dairy Sci. 95:377–388.

Sonesson, A. K., J. A. Woolliams, and T. H. E. Meuwissen. 2010. Maximising genetic gain whilst controlling rates of genomic inbreeding using genomic optimum contribution selection. Commun. no. 892 in Proc. 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. Gesellschaft für Tierzuchtwissenschaften e. V., Gießen, Germany.

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

Su, G., O. F. Christensen, T. Ostersen, M. Henryon, and M. S. Lund. 2012a. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. PLoS ONE 7:e45293.

Su, G., P. Madsen, U. S. Nielsen, E. A. Mäntysaari, G. P. Aamand, O. F. Christensen, and M. S. Lund. 2012b. Genomic prediction for Nordic Red cattle using one-step and selection index blending. J. Dairy Sci. 95:909–917.

Toro, M. A., and L. Varona. 2010. A note on mate allocation for dominance handling in genomic selection. Genet. Sel. Evol. 42:33.

VanRaden, P. M. 2005. Inbreeding adjustments and effect on genetic trend estimates. Interbull Bull. 33:81–84.

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

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

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., and L. A. Smith. 1999. Selection and mating considering expected inbreeding of future progeny. J. Dairy Sci. 82:2771–2778.

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

Weigel, K. A., and S. W. Lin. 2000. Use of computerized mate selection programs to control inbreeding of Holstein and Jersey cattle in the next generation. J. Dairy Sci. 83:822–828.

Wellmann, R., and J. Bennewitz. 2012. Bayesian models with dominance effects for genomic evaluation of quantitative traits. Genet. Res. (Camb.) 94:21–37.

Wray, N. R., and M. E. Goddard. 1994. Increasing long-term response to selection. Genet. Sel. Evol. 26:431–451.