Controlling Coancestry and Thereby Future Inbreeding by Optimum-Contribution Selection Using Alternative Genomic-Relationship Matrices

G. T. Gebregiwergis[1]*, Anders C. Sørensen[2], Mark Henryon[3,4] and Theo Meuwissen[1]

[1] Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, Ås, Norway, [2] Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark, [3] Seges, Copenhagen, Denmark, [4] School of Agriculture and Environment, University of Western Australia, Crawley, WA, Australia

We tested the consequences of using alternative genomic relationship matrices to predict genomic breeding values (GEBVs) and control of coancestry in optimum contribution selection, where the relationship matrix used to calculate GEBVs was not necessarily the same as that used to control coancestry. A stochastic simulation study was carried out to investigate genetic gain and true genomic inbreeding in breeding schemes that applied genomic optimum contribution selection (GOCS) with different genomic relationship matrices. Three genomic-relationship matrices were used to predict the GEBVs based on three information sources: markers ($\mathbf{G}_M$), QTL ($\mathbf{G}_Q$), and markers and QTL ($\mathbf{G}_A$). Strictly, $\mathbf{G}_Q$ is not possible to implement in practice since we do not know the quantitative trait loci (QTL) positions, but more and more information is becoming available especially about the largest QTL. Two genomic-relationship matrices were used to control coancestry: $\mathbf{G}_M$ and $\mathbf{G}_A$. Three genetic architectures were simulated: with 7702, 1000, and 500 QTLs together with 54,218 markers. Selection was for a single trait with heritability 0.2. All selection candidates were phenotyped and genotyped before selection. With 7702 QTL, there were no significant differences in rates of genetic gain at the same rate of true inbreeding using different genomic relationship matrices in GOCS. However, as the number of QTLs was reduced to 1000, prediction of GEBVs using a genomic relationship matrix constructed based on $\mathbf{G}_Q$ and control of coancestry using $\mathbf{G}_M$ realized 29.7% higher genetic gain than using $\mathbf{G}_M$ for both prediction and control of coancestry. Forty-three percent of this increased rate of genetic gain was due to increased accuracies of GEBVs. These findings indicate that with large numbers of QTL, it is not critical what information, i.e., markers or QTL, is used to construct genomic-relationship matrices. However, it becomes critical with small numbers of QTL. This highlights the importance of using genomic-relationship matrices that focus on QTL regions for GEBV estimation when the number of QTL is small in GOCS. Relationships used to control coancestry are preferably based on marker data.

Keywords: true inbreeding, genetic gain, genomic optimum contribution selection, genomic relationship matrices, prediction

BACKGROUND

Optimum contribution selection (OCS) is a selection method that maximizes genetic gain while controlling inbreeding (Meuwissen, 1997). It does this by optimizing the genetic contribution of selection candidates to the next generation using estimated breeding values and genetic relationships between candidates. A pedigree-based relationship matrix (**A**) was initially used to control inbreeding (Meuwissen, 1997). However, pedigree relationships have limitations. The **A**-matrix measures relationships and inbreeding at neutral, unlinked, and independent loci. But, genomic regions flanking quantitative trait loci (QTL) under selection lose more variation than neutral regions of the genome (Roughsedge et al., 2008). It also does not consider variation due to Mendelian sampling during gamete formation, assuming the same relationship between all full-sibs (Nejati-Javaremi et al., 1997; Avendaño et al., 2005). Dense panels of single nucleotide polymorphism (SNP) markers may be used to trace Mendelian segregation at marker loci (Hayes et al., 2009). Therefore, genomic markers might help to overcome some of the limitations imposed by pedigree.

There are several methods available to calculate genomic-relationships matrices (Nejati-Javaremi et al., 1997; Eding and Meuwissen, 2001; VanRaden, 2008; Yang et al., 2010; VanRaden et al., 2011). They have been used in different settings to realize high accuracies of genomic-prediction of breeding values and increase genetic gain (Jannink, 2010; Gómez-Romano et al., 2016). Moreover, genomic relationships that incorporate QTL information realize higher accuracies of genomic breeding values (GEBVs) than genomic relationships constructed based on markers only (Nejati-Javaremi et al., 1997; Zhang et al., 2010). In case of few QTL (100) and many records, accuracies of GEBVs close to one have been achieved (Fragomeni et al., 2017). Although, adding QTL information in the construction of genomic relationship matrices improved accuracies of prediction, there is no full understanding on the interaction of the use of alternative genomic relationship matrices in genomic optimum contribution selection (GOCS) schemes.

Optimum contribution selection can be extended into GOCS by using genomic information for both prediction of breeding values and estimation of relationship among selection candidates to manage group coancestry and thereby future inbreeding (Sonesson et al., 2012; Woolliams et al., 2015). Although, genomic relationship matrices that are based on dense SNP markers can reflect true relationship between individuals with a high degree of precision (Goddard, 2009), the covariances between additive genetic values of individuals for a specific trait are more accurately estimated using the relationships based on causal loci than SNP markers (Zhang et al., 2010; Habier et al., 2013; Yang et al., 2015). Moreover, there are limitations on our understanding of the way SNP markers are used to control group coancestry and thereby future inbreeding in GOCS (De Beukelaer et al., 2017; Henryon et al., 2019). Hence, our hypothesis is that the use of genomic relationship matrices based on QTL for the prediction of GEBV and marker based genomic relationships for OCS to control coancestry could show synergies regarding the rates of genetic improvement.

In this study, we investigated the use of alternative genomic relationship matrices for the prediction of GEBV and for the management of coancestry, where these relationship matrices are not necessarily identical. We investigated these combinations of relationship matrices by simulating three genetic architectures with 7702, 1000, and 500 QTLs. Alternative genomic relationship matrices were calculated using different genomic information sources such as SNP markers, QTLs, and both SNP markers and QTLs.

MATERIALS AND METHODS

We used stochastic simulations of breeding schemes to estimate rates of genetic gain realized by OCS at the same rate of true inbreeding with three matrices for prediction and two matrices to control of coancestry using three genetic architectures. The three prediction matrices were constructed using genetic markers ($\mathbf{G}_M$), QTL ($\mathbf{G}_Q$), and markers and QTL ($\mathbf{G}_A$). The two matrices to control coancestry were $\mathbf{G}_M$ and $\mathbf{G}_A$. $\mathbf{G}_Q$ was not used to control coancestry because allele frequency changes at the QTL were assumed desirable (increasing positive allele frequencies) in order to realize genetic gains. The six alternative genomic relationship matrix combinations for prediction and control of coancestry were A_A (both prediction and control of coancestry use $\mathbf{G}_A$), M_A (prediction of GEBV using $\mathbf{G}_M$ and control of coancestry using $\mathbf{G}_A$), Q_A (prediction of GEBV using $\mathbf{G}_Q$ and control of coancestry using $\mathbf{G}_A$), A_M (prediction of GEBV using $\mathbf{G}_A$ and control of coancestry using $\mathbf{G}_M$), M_M (both prediction and control of coancestry using $\mathbf{G}_M$), and Q_M (prediction of GEBV using $\mathbf{G}_Q$ and control of coancestry using $\mathbf{G}_M$). We investigated three genetic architectures with very many QTL (7702), a large number of QTL (1000), and a moderate number of QTL (500). For comparison, Wood et al. (2014) found that ∼2000, ∼3700, and ∼9500 SNPs explained ∼21, ∼24, and ∼29% of phenotypic variance of human height. QTL were randomly positioned on the genome of 18 chromosomes of equal length (167 cM), which resembles the pig genome.

Simulation of Genome and Population
Founder Populations

A schematic representation of the simulated breeding scheme is presented in **Figure 1**. Using 25 males and 25 females, a founder population was initiated and simulated for 1000 discrete generations. And, the effective-population size (Ne) of 50 was kept constant in each generation. The founder population genomes consisted of 3006 cM contained 30,000,000 equidistant monomorphic loci (both markers and QTLs; 1×10^4 loci per cM). The ratio of QTL loci to marker loci was 1:7. As a result 1/8 of the monomorphic loci (3.75×10^6) were QTL loci and the remaining loci were SNP markers. The mutation rate was assumed to be 4×10^{-6} per locus in order to generate bi-allelic polymorphism at mutated loci. The number of recombination sites for the ith animal was sampled from a Poisson distribution as n_{Qi} = Pois (λ), where λ = 30.06 is the length of the genome in Morgans. The n_{Qi} recombination sites were placed randomly on the genome.

FIGURE 1 | Schematic representation of the simulated breeding scheme.
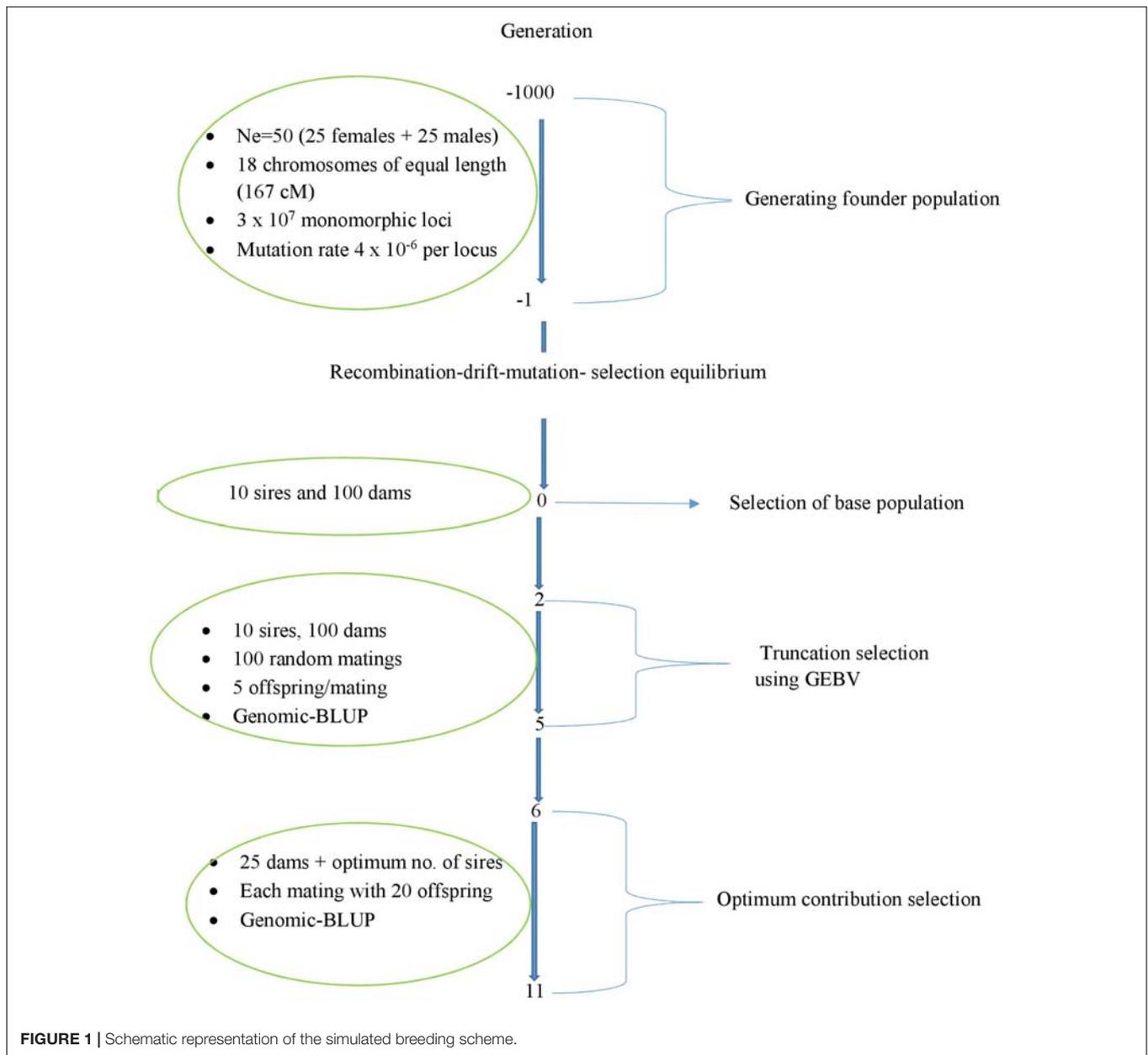
Recombination-drift-mutation-selection equilibrium of the founder population was reached after 1000 generations. Moreover, linkage disequilibrium between the QTL and markers alleles was established during the simulation of the founder population with a Fisher–Wright inheritance model (Fisher 1930, Wright 1931).

In each generation, male and female parents of the next generation were randomly sampled with replacement from the 25 males and 25 females of the current generation. The additive-genetic effect of the original allele at each QTL locus was set to zero. The additive-genetic effect of the mutant allele at each QTL was sampled from an exponential distribution and it was assumed to be positive with a probability of 0.1. Selection was introduced by culling and resampling approximately 5%

of animals with the lowest true breeding value (TBV) in each generation.

The TBV of an individual in the founder population was calculated as:

$TBV_i = \sum_{j=1}^{n} (x_{ij}\, g_j)$, where n is the number of QTL across the genome in the i-th founder animal, x_{ij} is the number of copies of the mutant allele that animal i inherited at the j-th QTL ($x_{ij} = 0, 1, 2$), and g_j is the additive effect of the mutant allele at the j-th QTL.

The average decay of linkage disequilibrium with distance between the segregating marker loci in generation 1000 of the founder population was similar to the average decay seen in the three commercial breeds of Danish pigs (Wang et al., 2013).

The founder population had 61,920 (7702 QTL and 54,218 marker) segregating loci at generation 0. All 54,218 segregating

marker loci were used in our breeding schemes. The number of segregating QTL used in the breeding schemes were all segregating QTLs (7702 QTL), 1000 or 500 QTLs. The number of segregating QTLs across the genome were reduced to a desired number, i.e., 1000 QTLs or 500 QTLs by random sampling from the 7702 QTLs.

The additive-genetic effects of the segregating mutant QTL alleles (7702, 1000, or 500) were standardized in order to get a total additive–genetic variance of 1 for the trait under selection in the founder population at generation 0.

We sampled a population of 100 chromosomes of chromosome number 1, representing the chromosome 1 chromosomes of 50 animals (two chromosomes per animal). The same sampling was performed for chromosome 2, 3, . . ., 18. These 50 animals formed the founder generation 0.

Base Populations

In each replicate of the simulation, a unique base population with a size of 110 animals (10 males and 100 females) was sampled from the pools of chromosomes of the founder population to initiate the breeding schemes. The genotype of each base animal was sampled from the 18 pools of chromosomes in the founder population. For chromosome j ($j = 1 \ldots 18$), two chromosomes were randomly sampled without replacement from the jth pool of 100 chromosomes. The sampled chromosomes were replaced before the next base animal was sampled. Base animals were assumed unrelated and non-inbred based on pedigree and identical by descent (IBD) alleles. They were genotyped, but not phenotyped.

Identical by Descent Markers

A total of 12,024 IBD loci were used to measure the true rate of inbreeding (ΔF_{true}) and were placed evenly across the genomes in the base population at 4 IBD loci per cM. Base animals were assigned unique alleles at each IBD locus (i.e., $2n$ distinct alleles at each IBD locus among the n animals in the base population), such that identical alleles in any later generation indicates that the loci are IBD with a single unique base population allele. The IBD loci were not involved in any way in the selection.

Simulation of Phenotypic Values

Phenotypic value of animal i, P_i, were simulated as: $P_i = TBV_i + e_i$, where e_i is an error term for individual i sampled from $e_i \sim N(0, \sigma_e^2 = 4)$ resulting in a trait heritability of 0.2; and TBV_i is the TBV of animals which was obtained as described above.

Genomic Estimated Breeding Values

The G-BLUP model (Meuwissen et al., 2001) was used to predict GEBVs:

$$y = 1\mu + Zg + e, \tag{1}$$

where y is a vector of phenotypes, μ is the overall mean, 1 is a vector of ones, Z is a design matrix allocating records to breeding values, g is a vector of breeding values for all animals with $Var(g) = G\sigma_g^2$, G is the genomic relationship matrix, and σ_g^2 is the additive genetic variance. The term e is a vector of normal independent and identical distributed residuals with variance σ_g^2.

Genomic-Relationship Matrices

Genomic relationship matrices (G) were computed using VanRaden method 2:

$$G = \frac{WD^{-1}W'}{L},$$

where W is a matrix of centered marker genotypes by subtracting the mean of the marker or QTL genotypes; L is the number of loci; D is a diagonal matrix with entries $2p_i (1-p_i)$; and p_i is frequency of the minor allele at locus i in the base population. All animals in the base population were used to calculate p_i to center and scale genotypes at locus i. After scaling, each locus obtained equal weight. The prediction and control of coancestry matrices, $\mathbf{G}_M$, $\mathbf{G}_Q$, and $\mathbf{G}_A$, were constructed using marker, QTL, and marker and QTL genotypes, respectively.

Truncation Selection

The base population animals were randomly mated to produce 500 offspring with equal sex ratio in generation 1. A truncation selection breeding program was conducted in generation 2–5 in order to mimic a population that had undergone selection. Ten sires and 100 dams were truncation selected based on genomic-estimated breeding values. Each selected sire was randomly mated with 10 dams and each mating produced five offspring. As the result of these matings, 500 offspring with an equal sex ratio were obtained.

Optimum Contribution Selection

Evolutionary algorithms (EVA) was used to optimize individual genetic contributions by maximizing the function U_t with respect to **c**:

$$U_t(c) = c'\hat{g} - \omega c'Gc, \tag{2}$$

where **c** is a n vector of genetic contributions of the current generation to the next which is proportional to the number of offspring each animal obtains; $\hat{g}$ is a n vector of genomic estimated breeding values, ω is a penalty applied on the average relationship of the selected parents for the next generation, and **G** is a $n \times n$ genomic relationship matrix among all animals in the population calculated as $\mathbf{G}_M$ or $\mathbf{G}_A$. In the above function, $c'\hat{g}$ and $c'Gc$ represent the average genetic value and average relationship of the new generation. For a detailed description of the EVA method see Henryon et al. (2015).

Optimum contribution selection was carried out in generations 6–11. A total of 25 matings were allocated to 500 selection candidates (approximately 250 males and 250 females) by OCS in each generation. Each male was allocated 0, 1, 2, . . ., or 25 matings in correspondence to their optimum contributions, **c**. Each of 25 selected female was allocated a single mating. The 25 sire and dam were mated randomly. Each mating produced 20 offspring, resulting in 25 full-sib families and 500 offspring. Offspring were assigned as males / females with a probability of 0.5.

Data Analyses

We plotted the rate of genetic gain against the rate of true inbreeding at different penalties (ω) for the schemes with 7702

and 500 QTL. The ω-values used were −50, −25, −10, and −5. For the scheme with 1000 QTL, we presented rates of genetic gain at 1 and 0.5% rates of true inbreeding. The 1 and 0.5% rates of true inbreeding were realized by calibrating the penalty, ω, in Eq. 2. We also compared the accuracies of males and females selection candidates for 1000 QTL at 1% rate of true inbreeding. Rates of genetic gain were calculated as the slope of the linear regression of G_t on t where G_t is the average true genetic value of animals born in generation t ($t = 6 \ldots 11$). Rates of true inbreeding (using the IBD markers) and rates of inbreeding based on pedigree were calculated as $1\text{-exp}(\beta)$, where β is a linear regression of $\ln(1-F_t)$ on t and F_t is the average coefficient of true inbreeding or pedigree inbreeding for animals born at generation t ($t = 6 \ldots 11$) (Sonesson et al., 2005). F_t for true inbreeding was calculated using the $d = 12,024$ IBD markers as $F_t = \frac{1}{n_t d} \sum_{i=1}^{n_t} \sum_{j=1}^{d} \delta_{ij}$, where n_t is the number of animals born in generation t and δ_{ij} is the IBD status at IBD-marker locus j ($j = 1 \ldots d$) for animal i ($i = 1 \ldots n_t$). δ_{ij} was equal to 1 if the alleles at IBD locus j for animal i were identical for a unique base-allele, and 0 otherwise.

Software

The simulations were run using the program ADAM (Pedersen et al., 2009). BLUP-breeding values were estimated using DMU6 (Madsen et al., 2006). OCS was carried out by EVA (Berg et al., 2006).
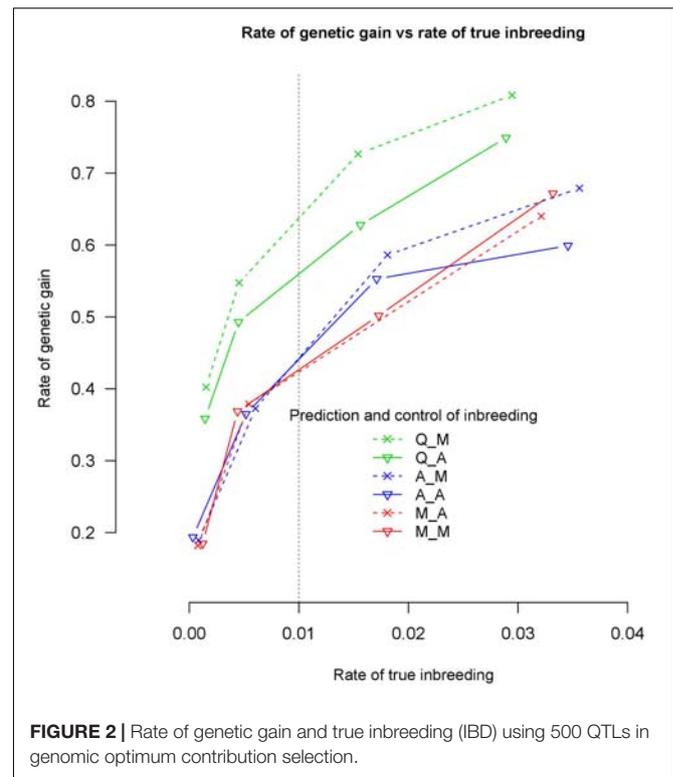
RESULTS

500 QTL

Quantitative trait loci based genomic relationship matrices for prediction of breeding values with marker based genomic relationship matrices to control coancestry, Q_M, realized more rate of genetic gain at a true inbreeding rate of 0.01 than the 5 alternative methods tested in this study (see **Figure 2**).

1000 QTL

Quantitative trait loci based genomic relationship matrices for prediction of breeding values with marker or all loci based genomic relationship matrices to control coancestry, Q_M and Q_A, realized more rate of genetic gain at a 1% rate of true inbreeding than 4 alternative methods of G-matrices. Q_M and Q_A realized between 21.5 and 29.7% more genetic gain than A_M, A_A, M_M, and M_A at 1% rate of true inbreeding (**Table 1**). At 0.5 % rate of true inbreeding, it realized between 29.9 and 53% more genetic gain than A_M, A_A, M_M, and M_A. Q_M realized almost the same rate of genetic gain as Q_A with both 1 and 0.5% rate of true inbreeding. Use of genomic relationship matrices computed based on QTLs (G_Q) to predict GEBVs gave higher accuracy of prediction than G_M or G_A at 1% rate of true inbreeding (**Table 2**). The accuracy of prediction of male selections using Q_M was 12.7% higher than M_M at 1% rate of true inbreeding (**Table 2**). This implies that the improvement in accuracy (12.7%) of the GEBVs explained only part of the increased in genetic gain (29.9%), i.e., 43% (12.7/29.9) of genetic gain was due to improvement in accuracy of GEBVs.

FIGURE 2 | Rate of genetic gain and true inbreeding (IBD) using 500 QTLs in genomic optimum contribution selection.

7702 QTL

With 7702 QTL, the six genomic relationships matrix combinations for prediction and control of coancestry realized almost the same rate of genetic gain at the same rate of true inbreeding (**Figure 3**). However, with an increase of the rate of true inbreeding >1%, the differences between the rate of genetic gain obtained by the different genomic relationship matrices became visible and the rate of genetic gain of Q_M and Q_A became slightly higher than the other four genomic relationships matrix combinations. However, all these differences in rate of genetic gain at the same rate of true inbreeding were not significant at 95% confidence intervals.

DISCUSSION

Our findings partly supported our hypothesis that prediction with QTL and control of coancestry with markers, Q_M, realizes more genetic gain at the same rates of true inbreeding than prediction and control of coancestry with markers, M_M. We found that prediction with QTL and control of coancestry with markers realized more genetic gain when 500 and 1000 QTL controlled the trait under selection. However, when the trait was controlled by 7702 QTL, prediction and control of coancestry with markers realized just as much genetic gain as prediction with QTL and control of coancestry with markers. These findings are important because they highlight that when traits under selection are controlled by small numbers of QTL, we need to select directly for the QTL to maximize genetic gain at pre-defined rates of true inbreeding. This implies that we need to know the exact number

TABLE 1 | Rate of genetic gain and rate of pedigree based inbreeding in genomic optimum contribution selection using different genomic information to predict GEBV and control of coancestry.

G-matrices	ΔIBD = 0.01		ΔIBD = 0.005	
	ΔG(SE)	ΔF (SE)	ΔG(SE)	ΔF (SE)
Q_M	0.677 (0.005)	0.012 (0.0002)	0.586 (0.007)	0.0086 (0.0003)
Q_A	0.672 (0.005)	0.012 (0.0002)	0.586 (0.006)	0.0085 (0.0003)
A_M	0.537 (0.005)	0.012 (0.0002)	0.433 (0.006)	0.0092 (0.0002)
A_A	0.544 (0.004)	0.012 (0.0002)	0.401 (0.005)	0.0093 (0.0003)
M_M	0.522 (0.005)	0.012 (0.0003)	0.383 (0.006)	0.0085 (0.0003)
M_A	0.557 (0.005)	0.012 (0.0002)	0.451 (0.007)	0.0086 (0.0003)

Rates of true inbreeding were 1 and 0.5% and there were 1000 QTL. ΔG, rate of genetic gain; ΔIBD, rate of true inbreeding; ΔF, rate of inbreeding based on pedigree; and SE, standard errors based on 100 replicates.

TABLE 2 | Accuracy of GEBVs for alternative genomic relationship matrices used to predict GEBVs and control of coancestry in genomic optimum contribution selection in generation 11 and at a rate of true inbreeding of 0.01.

G-matrices	Male accuracy (SE)	Female accuracy (SE)
Q_M	0.728 (0.005)	0.738 (0.005)
Q_A	0.721 (0.005)	0.720 (0.005)
A_M	0.661 (0.005)	0.661 (0.005)
A_A	0.656 (0.005)	0.660 (0.005)
M_M	0.646 (0.006)	0.646 (0.006)
M_A	0.659 (0.006)	0.659 (0.005)

The number of QTLs = 1000.

FIGURE 3 | Rate of genetic gain and true inbreeding (IBD) using 7702 QTLs.

of QTLs controlling the trait. On the other hand, we do not need to select directly for the QTL to maximize genetic gain when traits are controlled by large numbers of QTL. In this scenario, prediction using markers that are in linkage disequilibrium with the QTL (and control of coancestry using markers in LD with IBD alleles) is sufficient. Therefore, the method used in prediction and control of coancestry when using OCS depends on the number of QTL controlling the trait under selection.

Prediction with QTL and control of coancestry with markers only realized more genetic gain when 500 and 1000 QTL controlled the trait under selection for two reasons. First, prediction with QTL was more accurate than prediction with markers when small numbers of QTL controlled the trait. Prediction with QTL generated accurate breeding values because it had perfect knowledge of the true genetic (co)variance among individuals for the trait under selection. Prediction with markers was not as accurate because there was insufficient LD between the markers and QTL. Many of the markers where not located near QTL. With high numbers of QTL, prediction with markers had similar accuracy of prediction as prediction with QTL. There was more LD between the markers and QTL when many QTL were distributed across the genome. The second possible reason was that control of coancestry restricted changes in QTL-allele frequencies less when there were small numbers of QTL controlling the trait under selection. Control of coancestry in GOCS traced and penalized changes in marker-allele frequencies brought about by realized genetic drift and selection (Woolliams
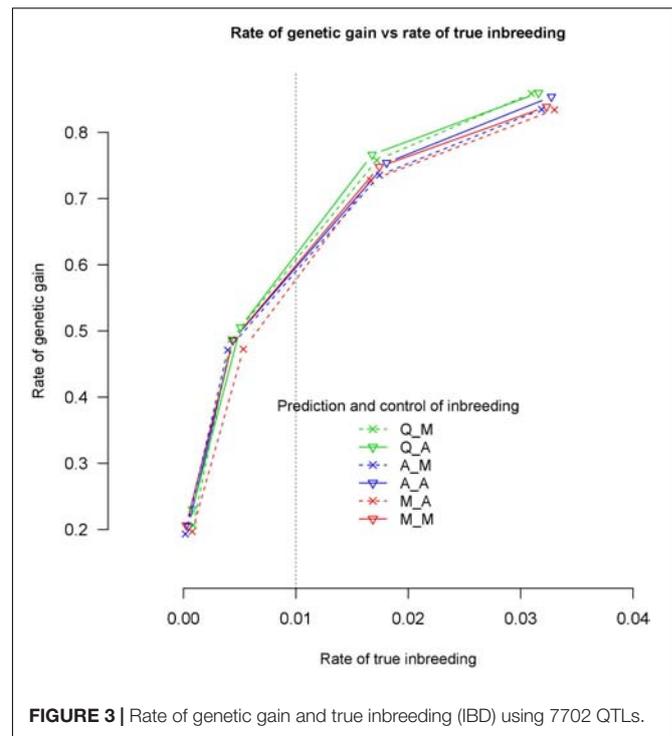
et al., 2015). It penalized changes in allele frequencies at all marker loci. Because these marker alleles were in linkage disequilibrium with QTL alleles, it restricted changes in QTL-allele frequencies. Inbreeding control is spread over the whole genome. With few QTL, much of the inbreeding control could be at regions of the genome that do not harbor QTL. With many QTL, inbreeding control is at all regions of the genome that harbor QTL – it penalized changes in allele frequencies at all loci, when we need to allow allele-frequency changes at some QTL loci. Therefore, genomic relationship matrices used to predict GEBV may differ from that used to control coancestry and thereby future inbreeding. In addition, genomic relationship matrices that consider the true genetic architecture of a trait under selection and allow differentiating the inbreeding rates at the QTL from the general rates at the genomic level could realize higher rates of genetic gain at the same rate of true inbreeding.

Prediction with QTL and control of coancestry with markers cannot be implemented directly in practical breeding schemes. However, this scheme does teach us some principles that apply to practical breeding schemes. Genomic relationship matrices based on QTL are not currently available in practice and it is unlikely to be available soon since exact number of QTLs controlling a trait are not known. Although, it is not possible to get a genomic-relationship matrix based on QTL, a trait specific genomic relationship matrix could be available from genome-wide association studies. Previous studies have shown that trait-specific genomic-relationship matrices realize higher prediction accuracy than marker-based genomic-relationship matrices but it realizes lower accuracy of prediction than the QTL-based genomic-relationship matrix (Nejati-Javaremi et al., 1997; Zhang et al., 2010;

Fragomeni et al., 2017). Until QTL-based genomic-relationship matrices become available, trait-specific genomic-relationship matrices could be used for prediction. Moreover, the inbreeding can be controlled using a selected panel of markers that has no association with the trait under selection to relax inbreeding control around the QTL regions. Therefore, when a trait under selection is controlled by a small number of QTLs, OCS that incorporates information about the trait for breeding value prediction realizes more genetic gain. Hence, there is an incentive for research work aiming to obtain more biological information about the QTL that underlie the breeding goal traits.

Our findings with 500 and 1000 QTL controlling the trait under selection are supported by several studies that assessed the role of alternative genomic relationship matrices on accuracy of genomic estimated breeding values (Nejati-Javaremi et al., 1997; Zhang et al., 2010; Hickey et al., 2013; Fragomeni et al., 2017). These studies addressed the effects of alternative genomic-relationship matrices on predictions without considering inbreeding control. However, our results are generally supported by their findings as we extended the study toward OCS and assessed both the effects on prediction and on control of coancestry. Fragomeni et al. (2017) showed that, adding causative QTN in an unweighted genomic relationship matrix improved the accuracy of prediction by 0.04. However, using a weighted genomic relationship matrix with weights obtained from genome-wide association studies, they reported an increase of accuracy by 0.1. These findings in general agree with our findings on (500 and 1000) QTL where Q_M realized the highest accuracy of prediction and genetic gain. Moreover, prediction with all loci (both markers and QTL) and inbreeding control with markers, A_M, realized higher accuracy of prediction than M_M. Our results also showed that the difference in rate of genetic gain obtained between Q_M and M_M became smaller as the number of QTL became larger, i.e., 500 and 1000 QTL. Moreover, this difference became insignificant as the number of QTLs further increased to 7702 QTL. Similarly, Nejati-Javaremi et al. (1997) reported higher accuracies and response to selection using genomic relationship matrices constructed based on QTL genotypes when a trait is controlled by a small number of loci (5 vs 25 and 100). Fragomeni et al. (2017) also reported higher accuracy using 100 QTL than 1000 QTL. This implies that the accuracy of prediction of the QTL based genomic relationship matrices increases as the number of QTLs controlling the traits decreases. Therefore, our findings that differences in rate of genetic gain between Q_M and

M_M increase as the number of QTL decrease from 1000 to 500 QTLs may hold also for very few QTLs controlling the traits as previous studies reported (Nejati-Javaremi et al., 1997; Fragomeni et al., 2017).

CONCLUSION

This study showed that as the number loci involved in the control of the trait of interest are large, genomic relationship matrices based on markers for both prediction and control of coancestry, M_M, perform as good as genomic relationship matrix constructed based on QTLs (Q_M) in GOCS. Whereas, when the trait is controlled by a small number of genes, genomic relationship matrix constructed based on QTLs (Q_M) realize higher rates of genetic gain than genomic relationships constructed based on markers (M_M) at the same rate of true inbreeding in GOCS. This improved rate of genetic gain of Q_M was partly explained by the increased prediction accuracy using G_Q, and partly by using a different relationship matrix for prediction as for coancestry control G_M, which creates opportunities for the genomic optimum contribution algorithm to enhance genetic gain.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this manuscript are not publicly available. Requests to access the datasets should be directed to gebreyohans.tesfaye.gebregiwergis@nmbu.no.

AUTHOR CONTRIBUTIONS

GG performed the study and drafted the manuscript. GG, TM, AS, and MH designed the study. MH modified the ADAM program. TM and AS planned and coordinated the whole study. TM and MH contributed to writing the manuscript. All authors read and approved the final manuscript.

FUNDING

REFERENCES

Avendaño, S., Woolliams, J., and Villanueva, B. (2005). Prediction of accuracy of estimated Mendelian sampling terms. *J. Anim. Breed. Genet.* 122, 302–308. doi: 10.1111/j.1439-0388.2005. 00532.x

Berg, P., Nielsen, J., and Sørensen, M. K. (2006). "EVA: realized and predicted optimal genetic contributions," in *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, 13-18 August, 2006*, (Belo Horizonte: Instituto Prociência), 27–29.

De Beukelaer, H., Badke, Y., Fack, V., and De Meyer, G. J. G. (2017). Moving beyond managing realized genomic relationship in long-term genomic selection. *Genetics* 206, 1127–1138. doi: 10.1534/genetics.116. 194449

Eding, H., and Meuwissen, T. (2001). Marker−based estimates of between and within population kinships for the conservation of genetic diversity. *J. Anim. Breed. Genet.* 118, 141–159. doi: 10.1046/j.1439-0388.2001.00290.x

Fragomeni, B. O., Lourenco, D. A., Masuda, Y., Legarra, A., and Misztal, I. (2017). Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genet. Sel. Evol.* 49:59.

Goddard, M. (2009). Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* 136, 245–257. doi: 10.1007/s10709-008-9308-0

Gómez-Romano, F., Villanueva, B., Fernández, J., Woolliams, J. A., and Pong-Wong, R. (2016). The use of genomic coancestry matrices in the optimisation of contributions to maintain genetic diversity at specific regions of the genome. *Genet. Sel. Evol.* 48:2. doi: 10.1186/s12711-015-0172-y

Habier, D., Fernando, R. L., and Garrick, D. J. (2013). Genomic BLUP decoded: a look into the black box of genomic prediction. *Genetics* 194, 597–607. doi: 10.1534/genetics.113.152207

Hayes, B. J., Visscher, P. M., and Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91, 47–60. doi: 10.1017/S0016672308009981

Henryon, M., Liu, H., Berg, P., Su, G., Nielsen, H. M., Gebregiwergis, G. T., et al. (2019). Pedigree relationships to control inbreeding in optimum-contribution selection realise more genetic gain than genomic relationships. *Genet. Sel. Evol.* 51:39.

Henryon, M., Ostersen, T., Ask, B., Sørensen, A. C., and Berg, P. (2015). Most of the long-term genetic gain from optimum-contribution selection can be realised with restrictions imposed during optimisation. *Genet. Sel. Evol.* 47:21. doi: 10.1186/s12711-015-0107-7

Hickey, J., Kinghorn, B., Tier, B., Clark, S., van der Werf, J., Gorjanc, G., et al. (2013). Genomic evaluations using similarity between haplotypes. *J. Anim. Breed. Genet.* 130, 259–269. doi: 10.1111/jbg.12020

Jannink, J.-L. (2010). Dynamics of long-term genomic selection. *Genet. Sel. Evol.* 42:35. doi: 10.1186/1297-9686-42-35

Madsen, P., Sørensen, P., Su, G., Damgaard, L. H., Thomsen, H., and Labouriau, R. (2006). "DMU-a package for analyzing multivariate mixed models," in *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, Belo Horizonte.

Meuwissen, T. (1997). Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.* 75, 934–940.

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Nejati-Javaremi, A., Smith, C., and Gibson, J. (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *J. Anim. Sci.* 75, 1738–1745.

Pedersen, L., Sørensen, A., Henryon, M., Ansari-Mahyari, S., and Berg, P. (2009). ADAM: a computer program to simulate selective breeding schemes for animals. *Livest. Sci.* 121, 343–344. doi: 10.1016/j.livsci.2008.06.028

Roughsedge, T., Pong-Wong, R., Woolliams, J., and Villanueva, B. (2008). Restricting coancestry and inbreeding at a specific position on the genome by using optimized selection. *Genet. Res.* 90, 199–208. doi: 10.1017/S0016672307009214

Sonesson, A. K., Woolliams, J. A., and Meuwissen, T. H. (2005). "Kinship, relationship and inbreeding," in *Selection and Breeding*

Programs in Aquaculture, ed. T. Gjedrem (Dordrecht: Springer), 73–87. doi: 10.1007/1-4020-3342-7_6

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

VanRaden, P. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

VanRaden, P., Olson, K., Wiggans, G., Cole, J., and Tooker, M. (2011). Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 94, 5673–5682. doi: 10.3168/jds.2011-4500

Wang, L., Sørensen, P., Janss, L., Ostersen, T., and Edwards, D. (2013). Genome-wide and local pattern of linkage disequilibrium and persistence of phase for 3 Danish pig breeds. *BMC Genet.* 14:115. doi: 10.1186/1471-2156-14-115

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* 46, 1173–1186. doi: 10.1038/ng.3097

Woolliams, J., Berg, P., Dagnachew, B., and Meuwissen, T. (2015). Genetic contributions and their optimization. *J. Anim. Breed. Genet.* 132, 89–99. doi: 10.1111/jbg.12148

Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A. A., Lee, S. H., et al. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* 47, 1114–1120. doi: 10.1038/ng.3390

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. doi: 10.1038/ng.608

Zhang, Z., Liu, J., Ding, X., Bijma, P., de Koning, D.-J., and Zhang, Q. (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PLoS One* 5:e12648. doi: 10.1371/journal.pone.0012648 doi: 10.1371/journal.pone.0012648