Animal (2013), **7:5**, pp 705–713 © The Animal Consortium 2012
doi:10.1017/S1751731112002248

animal

Integrating genomic selection into dairy cattle breeding programmes: a review

A. Bouquet† and J. Juga

Department of Agricultural Sciences, University of Helsinki, PO Box 27, FI-00014, Helsinki, Finland

(Received 20 December 2011; Accepted 14 September 2012; First published online 3 December 2012)

Extensive genetic progress has been achieved in dairy cattle populations on many traits of economic importance because of efficient breeding programmes. Success of these programmes has relied on progeny testing of the best young males to accurately assess their genetic merit and hence their potential for breeding. Over the last few years, the integration of dense genomic information into statistical tools used to make selection decisions, commonly referred to as genomic selection, has enabled gains in predicting accuracy of breeding values for young animals without own performance. The possibility to select animals at an early stage allows defining new breeding strategies aimed at boosting genetic progress while reducing costs. The first objective of this article was to review methods used to model and optimize breeding schemes integrating genomic selection and to discuss their relative advantages and limitations. The second objective was to summarize the main results and perspectives on the use of genomic selection in practical breeding schemes, on the basis of the example of dairy cattle populations. Two main designs of breeding programmes integrating genomic selection were studied in dairy cattle. Genomic selection can be used either for pre-selecting males to be progeny tested or for selecting males to be used as active sires in the population. The first option produces moderate genetic gains without changing the structure of breeding programmes. The second option leads to large genetic gains, up to double those of conventional schemes because of a major reduction in the mean generation interval, but it requires greater changes in breeding programme structure. The literature suggests that genomic selection becomes more attractive when it is coupled with embryo transfer technologies to further increase selection intensity on the dam-to-sire pathway. The use of genomic information also offers new opportunities to improve preservation of genetic variation. However, recent simulation studies have shown that putting constraints on genomic inbreeding rates for defining optimal contributions of breeding animals could significantly reduce achievable genetic gain. Finally, the article summarizes the potential of genomic selection to include new traits in the breeding goal to meet societal demands regarding animal health and environmental efficiency in animal production.

Keywords: genomic selection, breeding programme, dairy cattle, genetic gain, inbreeding rates

Implications

The practical use of genomic information to select animals, or genomic selection, is currently revolutionizing the organization of dairy cattle breeding schemes. The success of this new technology is because of increased genetic progress on both the bull and cow genetic pathways by reducing costs compared with conventional selection schemes, but also by the potential of using this rich source of information to manage genetic resources. Practical results of the implementation of genomic selection in dairy cattle breeding schemes are of great importance for other livestock species in which genomic selection is now under consideration.

† E-mail: alban.bouquet@gmail.com

Introduction

Use of molecular information to make selection decisions in breeding programmes was envisaged decades ago (Smith, 1967; Soller, 1978). Although conceptually simple, the implementation of genetic markers into breeding programmes has been rather limited for technological reasons. The recent availability in dense panels of single nucleotide polymorphism (SNP) markers has offered new opportunities to do so. The basic concept underlying such an approach is to use dense genetic markers as proxies to detect genome regions involved in a trait – the quantitative trait loci (QTL) – and exploit them for selection purposes. SNP genotyping technology has enabled profiling many animals for thousands of marker loci in a single analysis, thus with a low cost per marker (Williams, 2005). Capitalizing on this rich source of information permitted estimation of breeding values for young candidates with higher

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

accuracy than before. The principle of genomic evaluation models is to take advantage of both genotypic and phenotypic data available in a training (also called 'reference') population to build prediction equations of the genetic merit of individuals (Meuwissen *et al*., 2001). These equations can then be applied to selection candidates having genotypes but no phenotypes. Diverse approaches have been proposed to estimate genomically enhanced breeding values (GEBV), as reviewed by Hayes *et al*. (2009).

The use of genomic information to make selection decisions, or genomic selection, was shown to greatly increase the technical and economic efficiency of dairy cattle breeding programmes (Schaeffer, 2006; König *et al*., 2009). Indeed, genetic progress was achieved in conventional progeny testing (PT) schemes via the wide use of the very best progeny-tested bulls, which was enabled by means of artificial insemination (AI). Because selection in dairy cattle is undertaken on traits expressed by females, the PT step was necessary to generate a daughter group whose performance was used to predict the genetic merit of bulls with high accuracy. However, PT implies long generation intervals and huge costs related to bull maintenance and progeny-group constitution. Furthermore, only a limited number of young sires can be progeny tested each year owing to structural constraints. The choice of young males was therefore critical, whereas the accuracy of mid-parental Best Linear Unbiased Prediction (BLUP) breeding values was low at a young age. Because genomic selection alleviates some of these costs and technical constraints, the dairy cattle breeding industry has rapidly integrated genomic information into selection programmes. Many efforts have been devoted (so far) to improving the statistical models used for genomic evaluations (Hayes *et al*., 2009). Research has also been undertaken to answer questions arising about the practical use of genomic selection in breeding programmes. Such questions are also likely to emerge in other populations in which application of genomic selection is being considered.

To date, two main approaches have been used to predict the response to genomic selection. Hence, the first objective of this review was to describe these methods, which rely on either infinitesimal or finite locus models (FLMs), and to discuss limitations in their application. The second aim was to summarize the main conclusions reached on the subject by considering the example of dairy cattle populations. Critical factors influencing the efficiency of genomic selection were reviewed, as were future challenges faced in designing breeding schemes. The present review mainly focused on factors affecting the technical efficiency of programmes that was assessed through three criteria, namely (i) the annual genetic gain (ΔG), (ii) the variability of this response and (iii) the impact of selection procedures on maintenance of genetic diversity accessed via inbreeding rates (ΔF).

Modelling the efficiency of genomic breeding programmes

Modelling genomic selection with the selection index theory

The selection index theory was proposed to model the overall gain in accuracy expected from using genomic information at some selection stages (Lande and Thompson, 1990; Dekkers, 2007). To do so, SNP information of genotyped individuals is summarized into the direct genomic value (DGV). In practice, DGV is modelled as an infinitesimal indicator trait, which is highly heritable and genetically correlated with the selected trait. This indicator trait is modelled with a heritability of 1, meaning that genotyping errors are ignored and genotyped animals will receive no information from DGV of relatives. In this approach, the genetic correlation considered between the DGV and the selected trait reflects the desired level of accuracy for genomic predictions (Dekkers, 2007). Assuming the infinitesimal model and multivariate normal distribution of genotypic and phenotypic values, the selection index theory offers a convenient framework to optimally combine both information sources. Because no molecular data are simulated in this approach, the inbreeding rate generated by a scheme must be assessed with analytical formulas in deterministic models or with pedigrees in stochastic simulations. The main appeal of this approximate approach lies in its computational efficiency and straightforward implementation into existing software using either deterministic prediction models or stochastic simulations. It was extended to multi-trait selection by Dekkers (2007), and was also referred to as *pseudo-genomic selection* by Buch *et al*. (2012a).

Simulations based on FLMs

Another approach entails simulating individuals and their genotypes at a finite number of markers and QTL. This approach allows modelling the dynamics of genetic diversity at neutral markers and genetic variance at QTL by explicitly considering the discrete nature of genotypes and the finite size of genomes (Dekkers *et al*., 2004). In FLM simulations, breeding programmes are generally not started from a founder population with genotypes sampled at random. It is recommended to first simulate an ancestral population that has reached equilibrium with respect to mutation and genetic drift to ensure that a realistic structure of linkage disequilibrium (LD) exists between syntenic loci (Muir, 2007). Before simulating the breeding scheme, a base population is constituted by sampling individuals from the ancestral population. The true breeding values (TBV) of individuals are obtained as the sum of effects of QTL alleles that they carry on their chromosomes plus a polygenic term, if simulated QTLs are not assumed to explain all of the genetic variance. Following conventional principles of stochastic simulations, phenotypic observations are reconstructed from TBV by adding a residual term sampled from known statistical distributions to achieve the desired heritability (Dekkers *et al*., 2004). GEBVs are predicted by performing a genomic evaluation with simulated phenotypes and genotypes as inputs and by using the variance components assumed in the simulation. Hence, the accuracy of genomic predictions is not a prior in the model, but rather a result of simulations.

Finally, inbreeding rates can be estimated in FLM simulations by using either recorded pedigrees or marker information. Although the pedigree-based estimator of inbreeding

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

ignores the effects of linkage and selection around QTL regions, its use is expected to give accurate estimations of ΔF for long genomes, that is, >10 M (Villanueva *et al.*, 2005). However, the use of pedigree-based inbreeding estimates does not provide information about reductions in genetic diversity around QTL regions under selection (Sonesson *et al.*, 2010). An advantage of FLM simulations is that they provide information at simulated markers to monitor the evolution of autozygosity along the genome. Indeed, the proportion of simulated marker loci that is homozygous by descent with respect to the base population gives an expectation of the 'realized' inbreeding coefficient.

Limits of application

Pseudo-genomic selection and FLM simulations rely on different genetic models. Given a certain genetic architecture, the FLM simulation is expected to be the most accurate approach because it models the effects of biological processes and selection procedures on both genetic diversity and genetic variance. Although most breeding programme simulations have focused on additive effects, sophisticated architectures can be simulated with QTL exhibiting dominance and epistatic effects. However, the choice of parameters is expected to influence the level and maintenance of genetic variance and diversity (Hu and Li, 2006). Hence, it is safe to test the sensitivity of predictions to different architectures when using this prediction method.

Pseudo-genomic selection is an approximate approach that requires a set of assumptions with various consequences on predictions. First, it assumes the infinitesimal model and multivariate normal distribution of DGV and phenotypic values to combine them optimally into a genomic selection index. This supposes that the genetic architecture of selected traits complies with the hypotheses of the infinitesimal model, that is, traits are controlled by a large number of genes with individually small additive effects (Bulmer, 1980). Research on model species confirmed the relevance of using the infinitesimal model to describe the genetic mean and variance of a population for most quantitative traits because QTL alleles with large effects are rare and most of genetic variation is generally due to many loci with small effects (Mackay *et al.*, 2009). Thus, this model is valid over a short to medium time period, ignoring long-term changes in genetic variance due to fluctuations of allele frequencies. In addition, when dense marker maps are used, DGVs are expected to follow a multivariate normal distribution as a result of the central limit theorem (Lande and Thompson, 1990).

Second, in pseudo-genomic selection, the accuracy of GEBVs is a prior in the model and assumed to remain constant over time and to be the same for all individuals. In practice, GEBV accuracy is expected to increase depending on the size of the training population. Fixing GEBV accuracies to currently achieved values consequently depicts a rather pessimistic scenario, which is safe when genomic schemes are compared with conventional schemes. However, accuracy of GEBVs was shown to be sensitive to many intermingled parameters comprising the genetic architecture of traits (Daetwyler *et al.*, 2010), the LD existing between markers and QTL (Goddard, 2009; Goddard *et al.*, 2011), the statistical model of genomic evaluation (Daetwyler *et al.*, 2010; Bastiaansen *et al.*, 2012), the frequency of updating prediction equations (Muir, 2007) and the composition of the training population and its relationship with selection candidates (Lillehammer *et al.*, 2011; Pszczola *et al.*, 2012). As a consequence, different designs of genomic schemes can lead to different evolutions of GEBV accuracy (Lillehammer *et al.*, 2011). Because pseudo-genomic selection ignores the influence of all of these parameters on GEBV accuracy, this approach may lead to biased comparisons of genomic schemes on the basis of predictions of genetic gain.

Finally, computational requirements to achieve calculations also influence the choice of prediction method. Pseudo-genomic selection makes the optimization of breeding schemes more tractable than in FLM simulations, especially when selection is carried out on a breeding objective that includes several genetically correlated traits. However, Meuwissen (2009) showed that it was possible in FLM simulations to scale down by the same factor the effective size of the population, the genome size and the number of training records without modifying the achieved level of GEBV accuracy. This allows comparison of expected genetic gain from different genomic selection schemes at lower computational expense.

Impact of genomic selection on selection response and its variability

Description of investigated genomic selection schemes

Most studies have compared genomic selection schemes with the conventional PT scheme, which was the norm in dairy cattle populations until recently. Genomic selection was integrated into breeding schemes in two different ways, leading to genomic pre-selection (PS) and genomic juvenile (JS) schemes. The PS scheme entails using GEBVs to preselect young males for PT. All subsequent steps for the selection of males remain the same as in PT schemes. Compared with PT schemes, PS schemes allow increasing the selection accuracy of young male candidates by using their genotypic information. In JS schemes, the use of genomic information is more aggressive. AI sires are selected by using GEBVs among young genotyped males able to produce semen. Although GEBVs of young sires are less accurate than conventional breeding values estimated for progeny-tested bulls, the loss in selection accuracy is compensated by a huge reduction in generation intervals, as there is no need for progeny testing. Genetic gain is then rapidly cumulated over shortened generations. In both genomic schemes, females can be genotyped to increase both accuracy and intensity for the selection of females.

Large variations in predicted ΔG were found across studies, as described by Pryce and Daetwyler (2012). Compared with conventional schemes, gains in ΔG ranged from +9% for PS schemes to more than +100% for JS schemes.

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

However, these figures must be interpreted with caution, considering the specific set of parameters used in each study, for example, the heritability of the trait, the accuracy of GEBV and the number of genotypings allocated to males and females. For instance, for a trait with moderate heritability (0.3) – typical of a production trait – setting up a PS scheme increased ΔG by 9% (Lillehammer *et al.*, 2011), 13% (Buch *et al.*, 2012a) and 16% (Pryce *et al.*, 2010) compared with the former PT scheme. JS schemes were shown to generate higher selection response, with increases of 29% (Lillehammer *et al.*, 2011), 59% (Pryce *et al.*, 2010), 65% (Buch *et al.*, 2012a) and 86% (Colleau *et al.*, 2009) compared with PT schemes.

Increases in ΔG because of genomic selection were the largest when selection was carried out on traits of low heritability, because genomic data added relatively more information to predict breeding values for these traits. Compared with a traditional PT scheme, Lillehammer *et al.* (2011) showed that ΔG was increased by 29%, 40% and 70% in JS schemes for heritability values of 0.30, 0.05 and 0.01, respectively. However, levels of GEBV accuracy remained generally moderate to low for traits of low heritability. Phenotypic information of an indicator trait genetically correlated with the selected trait and recorded on a large scale can be integrated in the genomic evaluation model to improve the accuracy of predictions for traits of low heritability when the genetic correlation between both traits is large, that is, >0.5 (Calus and Veerkamp, 2011; Buch *et al.*, 2012a).

GEBV reliabilities larger than 0.5 to 0.6 are already achieved in homogeneous dairy cattle populations with large reference populations (e.g. Holstein populations) for traits with moderate to high heritability (Harris and Johnson, 2010; Su *et al.*, 2010). By contrast, reliabilities of at least 0.8 are obtained for breeding values estimated for progeny-tested bulls. In breeds with large effective sizes or small reference populations, reliability of GEBVs is generally lower (<0.5) even for production traits.

More emphasis placed on bull dam selection

In schemes integrating genomic selection, an important part of genetic gain is achieved through the selection of bull dams (Schaeffer, 2006). Indeed, the reliability of breeding values is strongly increased for genotyped females to a level comparable with that achieved for genotyped males. Furthermore, GEBVs of bull dams are expected to be less biased, because genomic information reduces the weight attributed to own performance, which may be subject to preferential treatment. With a training population of sufficient size, large gains are expected from a more accurate selection of breeding cows, even when only a small proportion is genotyped (Sørensen and Sørensen, 2009). Therefore, it is relevant to define strategies to optimally allocate genotyping capacities between young males and females in order to maximize genetic gain. Sørensen and Sørensen (2009) showed that allocating larger proportions of genotypings to females than males resulted in larger selection responses. Indeed, the proportion of males kept for reproduction is small in dairy cattle populations. If extra genotyping capacity is spent on males, the second best males according to parent average will be genotyped, although they have a lower chance of being selected. On the contrary, if extra genotyping capacity is spent on females, there are greater chances of detecting interesting females among the cow population to generate high genetic merit bull calves.

Influence of reference population constitution on genomic prediction accuracies

The need for large reference populations to increase GEBV accuracy has encouraged collaboration between breeding organizations and countries to exchange genotypic data (Lund *et al.*, 2011; Wiggans *et al.*, 2011). Collaboration has quickly appeared to be one of the cheapest ways to achieve higher GEBV reliabilities, although competitive interests often impeded the process (VanRaden *et al.*, 2009). In such collaborations, all genotypes of selected and unselected candidates have to be shared to avoid biases in the estimation of genomic breeding values, at both the national and international scales (Patry *et al.*, 2011).

Constitution of the reference population also influences the level of GEBV accuracy (Habier *et al.*, 2007) and its persistence over time (Lillehammer *et al.*, 2011). Hence, breeding programmes should be designed to minimize the average relationship within the reference population and maximize relationships between candidates and the reference population (Pszczola *et al.*, 2012). On the basis of FLM simulations of scenarios with reference populations of equivalent sizes, Lillehammer *et al.* (2011) showed that PS and JS schemes led to different levels of GEBV accuracy. Indeed, in PS schemes, all sires of male candidates have been progeny tested and are included in the reference population. This implies larger genetic ties between candidates and reference bulls than in JS schemes, in which only grandsires of candidates and older ancestors are included. Reducing the number of progeny-tested bulls in PS schemes led to a decrease in GEBV accuracy because fewer bulls with daughter information could be added each year to the reference population. However, Lillehammer *et al.* (2011) showed that the resulting loss in ΔG could be compensated by increasing the progeny group size. Thus, in PS schemes, maintaining maximal testing capacity is required to ensure accurate EBVs for bulls in the training population, and subsequently, accurate GEBVs for candidates, especially for traits of low heritability.

Including cows in the reference population has recently been suggested to further increase the reliability of genomic predictions because the number of AI sires is generally limited in dairy cattle populations (McHugh *et al.*, 2011; Buch *et al.*, 2012b). Such a strategy is gaining some interest given the increasing popularity of genomic selection in the industry and the decreasing costs of genotyping (Strauss, 2010). Actually, the recent release of cheaper SNP platforms is likely to extend the use of genomic selection to a wider female population. Use of imputation techniques makes it feasible to use sparser SNP chips to estimate GEBVs with only moderate losses in prediction accuracy (Weigel *et al.*, 2010;

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

Dassonneville *et al*., 2011). However, before practical inclusion of female information in reference populations, the issue of preferential treatment, which frequently occurs with elite cows, must be tackled because it can introduce bias in the estimation of SNP effects, potentially decreasing the accuracy of genomic predictions.

Impact on variability of selection response

In PT schemes, the risk of using a bull with poor genetic merit is minimized because of PT. In PS schemes, this risk is further reduced owing to the genomic pre-selection step, which allows distinguishing with higher accuracy the best young males within families before PT (Buch *et al*., 2012a). In JS schemes, the accuracy of breeding values used to select AI sires is lower than that obtained after PT. Considering a fixed number of AI sires selected each year, the variability of selection response obtained by implementing a JS scheme is increased relative to conventional schemes (Buch *et al*., 2012a). Thus, in practice, increasing the number of AI sires in JS schemes is recommended to mitigate the risk of using bulls with poor genetic merit. Another recommendation is to use young sires in 'teams' because the reliability of the mean GEBV of a group of bulls increases with group size. Schefers and Weigel (2012) reported that a team of five young bulls, each having a GEBV with a reliability of 0.70, will have an average GEBV with a reliability of 0.94.

New opportunities to monitor and preserve genetic diversity

Impact of genomic selection on inbreeding rates

Exploiting genomic information enables estimation of the Mendelian sampling term of young individuals without any phenotypic information. Therefore, genomic selection is expected to reduce the weight of family information in selection decisions by placing the emphasis on Mendelian sampling information of young candidates (Daetwyler *et al*., 2007). The largest reductions in inbreeding rates due to the use of genomic selection were observed for traits of low heritability (Lillehammer *et al*., 2011) and when a large part of variance was explained by markers (de Roos *et al*., 2011). By screening a large population of candidates, genomic selection also facilitates the identification of the least related animals having high genetic merit with a higher accuracy than before. In particular, McHugh *et al*. (2011) showed that genotyping a large number of females had a very beneficial impact on ΔF reduction.

The use of genomic selection to pre-select males for PT (i.e. PS scheme) resulted in a clear reduction of per generation inbreeding rates compared with PT schemes, for only slight modifications of the generation interval (Pryce *et al*., 2010; de Roos *et al*., 2011; Lillehammer *et al*., 2011; Buch *et al*., 2012a). Setting up JS schemes also generally led to reductions in per generation ΔF compared with PT schemes. However, because the generation interval was reduced as well, annual ΔF may not be decreased compared with conventional schemes (Lillehammer *et al*., 2011; Buch *et al*., 2012a). Increasing the number of young AI sires used in JS schemes is an option to curb inbreeding rates without markedly reducing genetic gain (Lillehammer *et al*., 2011).

Differentiating the number of matings per young bulls on the basis of GEBVs is not a sustainable option; it leads to a slight increase in genetic gain at the expense of a drastic increase in ΔF (Sørensen and Sørensen, 2009). Finally, schemes with mixed use of proven and young genomic bulls were studied by Colleau *et al*. (2009) to mimic the situation in which some breeders did not accept the risk of using young AI bulls having a GEBV of moderate reliability. The authors showed that allocating 50% of inseminations in a JS scheme to the 20 best bulls having milking daughters led to a small increase in genetic gain and a large increase in inbreeding rate. Thus, this practice was harmful for genetic diversity unless strict rules were defined for the management of genetic resources.

The conclusions about the impact of genomic selection schemes on genetic diversity critically depend on the time unit chosen to express inbreeding rates. In practice, the inbreeding rate should be kept under 1% per generation in conservation and breeding schemes to avoid undesirable effects of inbreeding on fitness (Food and Agriculture Organization (FAO), 1998). This recommendation indicates that an effective population size of at least 50 individuals is required so that factors such as natural selection, recombination and mutation – which intervene at every generation – can counterbalance effects of inbreeding depression (FAO, 1998). When the selected trait and fitness were not genetically correlated, Meuwissen and Woolliams (1994) estimated that a critical effective size of 30 to 250 individuals was required in a selection scheme so that natural selection could overcome the effects of inbreeding depression on fitness, which is consistent with FAO recommendations. However, when the selected trait and fitness were negatively correlated, even weakly (-0.2), Meuwissen and Woolliams (1994) showed that it was almost impossible to prevent a decline in fitness. In such situations, shortening the generation interval makes the accumulation of inbreeding depression effects bigger. Thus, in our opinion, it is safe to compare breeding schemes by using annual inbreeding rates. In addition, the inbreeding rate can be considered a measure of the risk of a scheme. Therefore, it seems relevant to express it on the same time scale as genetic gain, that is, on an annual basis.

Although the design of breeding schemes is important to limit the rise in inbreeding in a population, there is also a need to develop efficient recording systems that are able to detect emerging genetic defects (Agerholm *et al*., 2001). In particular, when the generation interval is shortened, it is essential to take rapid measures to control the spreading of the defect by restricting the use of bulls carrying it. The example of Holstein BLAD, CVM and Brachyspina genetic defects illustrated how fast a lethal mutation can spread in a population of small effective size.

Including genomic information in the optimal contribution selection (OCS) method

In the near future, JS schemes will likely be preferred to other designs because they generate larger genetic gain at

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

moderate inbreeding rates, if practical recommendations for the use of young AI sires are followed. However, it may become difficult for breeding managers to select and mate animals while monitoring their relationships, as many more AI bulls are used and for a shorter time period. It is therefore important to develop and promote the use of tools to control inbreeding rates. The standard approach used to determine contributions of breeding animals, maximizing genetic gain at a desired rate of inbreeding, is referred to as OCS (Meuwissen, 1997). In OCS, the contribution of a parent is the result of a trade-off between its genetic merit and its relationship to other individuals (Fernández *et al*., 2011). Recently, Sonesson *et al*. (2010) showed with FLM simulations that OCS was not able to maintain genetic diversity across the whole genome in selected populations when relationship coefficients were estimated from pedigree information. In fact, it led to a strong reduction in diversity around QTL regions by favouring alleles with the largest effects. To circumvent this flaw, these authors proposed using marker-based relationships, because they more accurately reflect the genome sharing between individuals than pedigree-based expectations. Genomic Optimal Contribution Selection (GOCS) enabled a better control of genomic inbreeding across the genome. However, it also strongly reduced genetic gain. The study design may have influenced the results; in particular, the number of simulated QTL was relatively limited, which could have implied a strong effect of selection on QTL regions. Nevertheless, this result questions whether large ΔG can be achieved when an acceptable constraint on ΔF is considered. It also emphasizes the need to determine an acceptable rate of inbreeding in practical breeding programmes by integrating recent knowledge gained from genomic data. In addition, the ability of GOCS to maintain expected and observed heterozygosity in a conservation programme was shown to be sensitive to the estimator used to access genomic relationships (de Cara *et al*., 2011). This stresses the importance of identifying accurate marker-based relationship estimators to maximize the efficiency of GOCS in practical applications.

Finally, after having inferred the relationship between chromosome segments of individuals, one can envisage strategies to maintain genetic diversity at the chromosome level, for instance, to prioritize preservation in genome regions exhibiting acute reduction in diversity (Engelsma *et al*., 2011) or to maintain breed-specific genetic variation in a hybrid population. It would be interesting to assess whether benefits of these approaches are large relative to GOCS, which uses genome-wide relationship estimates.

Impact of using genomic selection with reproductive biotechnologies

One of the main advantages of genomic selection is to capitalize on SNP information to discriminate within a family the individuals with the largest Mendelian sampling terms. Hence, the greatest benefits of genomic selection are expected when selection is carried out in large full-sib families. Thus, in dairy cattle, there is enormous interest in increasing the number of offspring obtained from elite bull dams by means of reproductive biotechnologies such as multiple ovulation and embryo transfer (MOET) or ovum pick-up followed by *in vitro* fertilization. The advantages of using genomic selection together with MOET are twofold; it enables the discrimination of the best heifers to be flushed and the early differentiation of the best prospective males obtained from flushings (Sørensen and Sørensen, 2009; Pedersen *et al*., 2012). However, Pryce *et al*. (2010) showed that an aggressive use of such techniques in a closed nucleus could lead to large increases in inbreeding rates. Exploiting a closed rather than an open MOET scheme was advocated to be more cost-effective and safer with respect to sanitary status, because no heifers have to be purchased from the outside (Pryce *et al*., 2010). However, keeping the MOET nucleus open is generally recommended to maintain inbreeding rates at a reasonable level, especially for small-sized nuclei (Korpiaho *et al*., 2003).

A practical limitation in the use of embryo transfer technologies is the lack of large numbers of recipient cows (Humblot *et al*., 2010). Having GEBV for embryos could alleviate this limitation by discriminating the embryos with desirable genotypes before implantation. Colleau *et al*. (2008) assessed interest in genotyping embryos for the French marker-assisted selection scheme under a fixed cost constraint. Their results indicated that benefits of embryo genotyping were partially offset by the poor quality of genotypic data from small biopsies of embryos. However, in their study, embryos were genotyped for microsatellite markers. SNP technologies have proven to be more robust to genotyping errors and less demanding in terms of DNA quality and quantity than microsatellite technologies. Promising results were reported by Humblot *et al*. (2010) in terms of genotyping calling rates and pregnancy rates obtained from biopsied embryos under experimental conditions. If such results can be extrapolated to field conditions, embryo genotyping might facilitate exploitation of increased prolificacy of flushed females and reduce the costs related to maintenance of recipient cows in breeding nuclei. This could give rise to the 'velogenetics schemes', which were conceptualized 20 years ago by Georges and Massey (1991).

Finally, the use of sexed semen was shown to be able to increase genetic gain in conventional dairy cattle breeding schemes (Abdel-Azim and Schnell, 2007). However, most benefits were achieved in commercial herds by increasing selection intensity and reducing generation intervals on the cow-to-cow pathway. By simulating a great range of dairy cattle breeding schemes integrating genomic selection, Pedersen *et al*. (2012) confirmed that benefits of using sorted semen were slight or null relative to those of MOET in terms of genetic gain and preservation of genetic diversity. The most worthwhile scenario was obtained when Y-sorted semen and MOET were used on elite bull dams. No increase in genetic gain was observed, but the negative effects on inbreeding rates usually associated with MOET were partially offset because of increased within-family selection among larger full-sib families (Pedersen *et al*., 2012).

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

Defining selection objectives and integrating new selection criteria

Compared with conventional schemes, genomic selection has the potential of improving the balance of achieved genetic gain between production, conformation and functional traits. Accordingly, achieved genetic gain should better reflect economic weights considered in the total merit index (Boichard *et al*., 2010). However, a large difference in genetic gain is still expected between highly and lowly heritable traits if the latter receive insufficient weight in the selection objective (König and Swalve, 2009). This phenomenon can be strengthened if antagonistic genetic correlations exist between functional and production traits (Ytournel *et al*., 2011).

Although more complex objectives (e.g. environmental efficiency and animal welfare) are promoted by society for inclusion in breeding programmes, genomic selection offers an interesting opportunity to add new selection criteria to breeding objectives (Boichard and Brochard, 2012). Indeed, selecting new phenotypes seems to be more feasible, especially for traits that cannot be recorded on a large scale because they are difficult or expensive to measure. These new phenotypes are mainly related to health (lameness, disease resistance, metabolic and reproductive disorders), feed efficiency, milk composition or adaptation to different environments (Boichard and Brochard, 2012). Experimental herds, MOET herds or test stations could be exploited to record such phenotypes. The quantity and quality of phenotypic data are critical parameters for accurate genomic predictions, and hence for achieving genetic gain in the whole population. Veerkamp *et al*. (2011) showed the feasibility of combining data sets from several experimental herds to increase the accuracy of genomic predictions for a new trait. Even if herds displayed patent differences in recording and management, genomic predictions of reasonable accuracy were obtained. For new traits that can be recorded on a larger scale, Calus *et al*. (2011) estimated that selection accuracy could be greatly increased by including directly genotyped females in the training population. These studies suggested that breeding companies should contract out herds to expand the recording of new traits, but the issue of recording costs must also be considered.

World-wide use of genomic selection in dairy cattle breeding schemes

To date, genomic selection has only been implemented in a few countries and mainly in connection with breeding programmes of Holstein cattle (Pryce and Daetwyler, 2012). This breed was a good case study for developing genomic selection because it has been intensively selected for decades, hence strengthening statistical associations between markers and QTL. Furthermore, many large breeding programmes have been set up world wide, facilitating the constitution of large reference populations. Thus, international collaborations to exchange genotypes allowed further improving the accuracy of genomic predictions (Lund *et al*., 2011; Wiggans *et al*., 2011).

The first genomic evaluations were officially released for a few Holstein populations in 2009. Different strategies have since been adopted to integrate genomic selection into existing breeding programmes. Focusing on the Holstein breed, Pryce and Daetwyler (2012) summarized the strategies adopted in eight countries by April 2011. Advancements have been rapid and a few countries – such as Italy and the United Kingdom – have now released GEBV for large populations of Holstein bulls.

Utilization of genomic selection has accelerated among Holstein breeders (Strauss, 2010). For example, in France, genomically tested bulls without milking daughters represented 39% of market shares in 2011 (P. Le Mezec, personal communication). In most countries, PT continued to be carried out during this transition period (Pryce and Daetwyler, 2012). France was an exception, with PT officially ceasing in 2009 (Fritz *et al*., 2010; Pryce and Daetwyler, 2012). However, the number of AI sires used each year was increased, their diffusion was constrained to a few thousand straws per sire and farmers were encouraged to use teams of at least five young bulls (Fritz *et al*., 2010).

In other dairy cattle breeds, several difficulties have impeded the integration of genomic selection into breeding programmes. First, breeding programmes are generally of smaller size than Holstein programmes, which makes it more difficult to gather large reference populations. Several initiatives have, however, been created to exchange genotypes across countries, for example, the InterGenomics consortium for the Brown Swiss breed (Jorjani *et al*., 2011) and the collaboration between breeding schemes of the Nordic Red Dairy cattle in Denmark, Sweden, Finland and Norway. Second, some of these breeds may have larger effective population sizes, resulting in weaker associations between markers and QTL. Both of these factors tend to reduce the accuracy of genomic predictions, which are generally lower than in Holstein populations. The recent use of high-density SNP panels (e.g. Illumina Bovine HD (777K)) was a promising option to increase GEBV accuracy in such populations. Having large densities in SNPs reduces the physical distance between markers and QTL, and hence should strengthen the statistical association between them. At such marker density, associations between markers and QTL may be also maintained across breeds, making it possible to build across-breed prediction equations and to capitalize on reference populations of several breeds. However, preliminary analyses of high-density chips with the genomic BLUP evaluation model only resulted in marginal gains within breeds (Su *et al*., 2011) and across breeds (Harris *et al*., 2011).

Despite these difficulties, genomic selection has been integrated into the breeding programmes of a few large dairy cattle populations other than Holstein. In France, the first official GEBVs were released in 2009 for the Montbéliarde and Normande breeds (Fritz *et al*., 2010). Lower GEBV reliabilities were achieved in these two breeds than in Holstein because of the smaller size of their reference populations, comprising 1500 bulls for the Montbéliarde breed and 1250 bulls for the Normande breed in June 2010 (Fritz *et al*., 2010). The PT of bulls was discontinued in 2009,

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

and rules were defined for the commercialization of genomic sires (Fritz *et al*., 2010). The uptake of genomic selection was lower in these breeds than in the Holstein breed, with 21% of market shares achieved for young Normande genomic sires in 2011 (P. Le Mezec, personal communication).

Genomic selection has also recently been adopted into the breeding programme of Red dairy cattle breeds in Denmark, Sweden and Finland. Officially released in 2011, genomic breeding values displayed moderate reliability, given the size of the reference population, which included 4400 bulls in 2011 (Su *et al*., 2011). Indeed, effects of recurrent genetic admixture have weakened associations existing between SNPs and QTL in this composite population, hampering the accuracy of genomic predictions (Brøndum *et al*., 2011). Reliabilities of genomic breeding values are nevertheless comparable with those achieved for the French Normande and Montbéliarde breeds (Fritz *et al*., 2010). Given the moderate GEBV reliabilities, prudent use of genomic selection was chosen. Hence, the PT of bulls did not cease during this transition period, and the semen of the best 20 to 25 young sires was commercialized at the age of 2 years.

Conclusions

Genomic selection has been rapidly adopted by the dairy cattle breeding industry, particularly in the Holstein breed, and its use in other cattle breeds and livestock species is under investigation (Ibanez-Escriche and Gonzalez-Recio, 2011). In this transition period, strategies to implement genomic selection have been rather prudent, although the largest gains could be obtained by extensive use of genomic information. Success of this new selection method critically depends on the availability of large databases of phenotypes and genotypes. Hence, breeding organizations will have to work more closely with recording organizations to facilitate the large-scale recording of currently selected traits and to extend the recording of new traits. The link between breeding companies and breeders might also be strengthened to secure access to elite breeding animals.

Genomic selection also appears to be a promising tool to curb inbreeding rates because it allows exploitation of within-family genetic variance from the earliest selection stages. However, to be effective, breeders must consider using a diversified panel of genomically evaluated bulls, both to mitigate the risk of using bulls with poor merit and to limit the erosion of genetic diversity. To this end, the further development and application of tools, such as genomic optimal contribution selection, are critical for sustainable use of genetic resources.

Acknowledgement

This work was funded by FABA Oy (Vantaa, Finland) through TEKES grant no. 433/31/09.

References

Abdel-Azim G and Schnell S 2007. Genetic impacts of using female-sorted semen in commercial herds. Journal of Dairy Science 90, 1554–1563.

Agerholm JS, Bendixen C, Andersen O and Arnbjerg J 2001. Complex vertebral malformation in Holstein calves. Journal of Veterinary Diagnostic Investigation 13, 283–289.

Bastiaansen JWM, Costar A, Calus MPL, van Arendonk JAM and Bovenhuis H 2012. Long-term response to genomic selection: effects of estimation method and reference population structure for different genetic architectures. Genetics Selection Evolution 44, 3.

Boichard D and Brochard M 2012. New phenotypes for new breeding goals in dairy cattle. Animal 6, 544–550.

Boichard D, Ducrocq V, Fritz S and Colleau JJ 2010. Where is dairy cattle breeding going? A vision of the future. Interbull Bulletin 41, 63–68.

Brøndum RF, Rius-Vilarrasa E, Strandén I, Su G, Gulbrandtsen B, Fikse WF and Lund MS 2011. Reliabilities of genomic prediction using combined reference data of the Nordic Red dairy cattle populations. Journal of Dairy Science 94, 4700–4707.

Buch LH, Sørensen MK, Berg P, Pedersen LD and Sørensen AC 2012a. Genomic selection strategies in dairy cattle: strong positive interaction between use of genotypic information and intensive use of bulls on genetic gain. Journal of Animal Breeding and Genetics 129, 138–151.

Buch LH, Kargo M, Berg P, Lassen J and Sørensen AC 2012b. The value of cows in reference populations for genomic selection of new functional traits. Animal 6, 880–886.

Bulmer MG 1980. The mathematical theory of quantitative genetics. Oxford University Press, Oxford, UK. 254pp.

Calus MPL and Veerkamp RF 2011. Accuracy of multi-trait genomic selection using different methods. Genetics Selection Evolution 43, 26.

Calus MPL, de Haas Y, Pszczola M and Veerkamp RF 2011. Predicted response of genomic selection for new traits using combined cow and bull reference populations. Interbull Bulletin 44, 231–234.

Colleau JJ, Fritz S, Guillaume F, Baur A, Dupassieux D, Boscher MY, Journaux L, Eggen A and Boichard D 2009. Simulating the potential of genomic selection in dairy cattle breeding. Proceedings of the 16th Rencontres Recherches Ruminants, Institut de l'Elevage, Paris, France, 419pp.

Colleau JJ, Fritz S, Ponsart C, Le Bourhis D, Lacaze S, Tissier M, Mervant G, Amigues Y, Druet T, Malafosse A, Boichard D and Humblot P 2008. The value of embryo typing in a marker assisted selection program for dairy cattle. Proceedings of the 15th Rencontres Recherches Ruminants, Institut de l'Elevage, Paris, France, pp. 427–430.

Daetwyler HD, Villanueva B, Bijma P and Woolliams JA 2007. Inbreeding in genome-wide selection. Journal of Animal Breeding and Genetics 124, 369–376.

Daetwyler HD, Pong-Wong R, Villanueva B and Woolliams JA 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185, 1021–1031.

Dassonneville R, Brøndum RF, Druet T, Fritz S, Guillaume F, Guldbrandtsen B, Lund MS, Ducrocq V and Su G 2011. Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. Journal of Dairy Science 94, 3679–3686.

de Cara MAR, Fernandez J, Toro MA and Villanueva B 2011. Using genome-wide information to minimize the loss of diversity in conservation programmes. Journal of animal Breeding and Genetics 128, 456–464.

Dekkers JCM 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. Journal of Animal Breeding and Genetics 124, 331–341.

Dekkers JCM, Gibson JP, Bijma P and Van Arendonk JAM 2004. Design and optimization of animal breeding programmes. Lecture notes for the 'Animal Breeding Strategies' Course, Iowa State University. Retrieved December 19, 2011, from http://www.anslab.iastate.edu/Class/AnS652X/Chapter1,2.pdf

de Roos APW, Schrooten C, Veerkamp RF and Van Arendonk JAM 2011. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. Journal of Dairy Science 94, 1559–1567.

Engelsma KA, Veerkamp RF, Calus MPL and Windig JJ 2011. Consequences when prioritizing animals for conservation with pedigree or genomic information. Journal of Animal Breeding and Genetics 128, 473–481.

Fernández J, Meuwissen THE, Toro MA and Mäki-Tanila A 2011. Management of genetic diversity in small farm animal populations. Animal 5, 1684–1698.

Food and Agriculture Organization (FAO) 1998. Secondary guidelines for development of national farm animal genetic resources management plans: management of small populations at risk, FAO, Roma, Italy, 219 pp.

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press

Fritz S, Guillaume F, Croiseau P, Baur A, Hozé C, Dassonneville R, Boscher MY, Journaux L, Boichard D and Ducrocq V 2010. Implementing genomic selection in the three main French dairy cattle breeds. Proceedings of the 17th Rencontres Recherches Ruminants, Institut de l'Elevage, Paris, France, pp. 455–458.

Georges M and Massey JM 1991. Velogenetics, or the synergistic use of marker-assisted selection and germ-line manipulation. Theriogenology 35, 151–159.

Goddard ME 2009. Genomic selection: prediction of accuracy and maximization of long term response. Genetica 136, 245–257.

Goddard ME, Hayes BJ and Meuwissen THE 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. Journal of Animal Breeding and Genetics 128, 409–421.

Habier D, Fernando RL and Dekkers JCM 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177, 2389–2397.

Harris BL and Johnson DL 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. Journal of Dairy Science 93, 1243–1252.

Harris BL, Creagh FE, Winkelman AM and Johnson DL 2011. Experiences with the Illumina high-density Bovine BeadChip. Interbull Bulletin 44, 3–7.

Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 92, 433–443.

Hu XS and Li B 2006. Additive genetic variation and the distribution of QTN effects among sites. Journal of Theoretical Biology 243, 76–85.

Humblot P, Le Bourhis D, Fritz S, Colleau JJ, Gonzalez C, Guyader-Joly C, Malafosse A, Heyman Y, Amigues Y, Tissier M and Ponsart C 2010. Reproductive technologies and genomic selection in cattle. Veterinary Medicine International 2010, e192787.

Ibanez-Escriche N and Gonzalez-Recio O 2011. Promises, pitfalls and challenges in genomic selection in breeding programs. Spanish Journal of Agricultural research 9, 404–413.

Jorjani H, Zumbach B, Dürr J and Santus E 2011. Joint genomic evaluation in Brown Swiss populations. Interbull Bulletin 41, 8–14.

König S and Swalve HH 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. Journal of Dairy Science 92, 5292–5303.

König S, Simianer H and Willam A 2009. Economic evaluation of genomic breeding programmes. Journal of Dairy Science 92, 382–391.

Korpiaho PI, Strandén I and Mäntysaari EA 2003. Bull selection across age classes and variable female reproductive rates in an open nucleus breeding scheme of dairy cattle. Acta Agriculturae Scandinavia A: Animal Science 53, 74–82.

Lande R and Thompson R 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124, 743–756.

Lillehammer M, Meuwissen THE and Sonesson AK 2011. A comparison of dairy cattle breeding designs that use genomic selection. Journal of Dairy Science 94, 493–500.

Lund MS, de Roos APW, de Vries AG, Druet T, Ducrocq V, Fritz S, Guillaume G, Guldbrandtsen B, Liu Z, Reents R, Schrooten C, Seefried M and Su G 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. Genetics Selection Evolution 43, 43.

Mackay TFC, Stone EA and Ayroles JF 2009. The genetics of quantitative traits: challenges and prospects. Nature Reviews Genetics 10, 565–577.

McHugh N, Meuwissen THE, Cromie AR and Sonesson AK 2011. Use of female information in dairy cattle genomic breeding programs. Journal of Dairy Science 94, 4109–4118.

Meuwissen THE 1997. Maximizing the response to selection with a predefined rate of inbreeding. Journal of Animal Science 75, 934–940.

Meuwissen THE 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genetics Selection Evolution 41, 35.

Meuwissen THE and Woolliams JA 1994. Effective sizes of livestock populations to prevent a decline in fitness. Theoretical Applied Genetics 89, 1019–1026.

Meuwissen THE, Hayes BJ and Goddard ME 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

Muir WM 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. Journal of Animal Breeding and Genetics 124, 342–355.

Patry C, Jorjani H and Ducrocq V 2011. Implementation of genomic selection at national level: impact of pre-selection and biased national BLUP evaluations on international genetic evaluations. Interbull Bulletin 44, 67–73.

Pedersen LD, Kargo M, Berg P, Voergaard J, Buch LH and Sørensen AC 2012. Genomic selection strategies in dairy cattle breeding programmes: sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. Journal of animal Breeding and Genetics 129, 152–163.

Pryce JE and Daetwyler HD 2012. Designing dairy cattle breeding schemes under genomic selection: a review of internation research. Animal Production Science 52, 107–114.

Pryce JE, Goddard ME, Raadsma HW and Hayes BJ 2010. Deterministic models of breeding scheme designs that incorporate genomic selection. Journal of Dairy Science 93, 5455–5466.

Pszczola M, Strabel T, Mulder HA and Calus MPL 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. Journal of Dairy Science 95, 389–400.

Schaeffer LR 2006. Strategy for applying genome-wide selection in dairy cattle. Journal of Animal Breeding and Genetics 123, 218–223.

Schefers JM and Weigel KA 2012. Genomic selection in dairy cattle: integration of DNA testing into breeding programs. Animal Frontiers 2, 4–9.

Smith C 1967. Improvement of metric traits through specific genetic loci. Animal Production 9, 349–358.

Soller M 1978. The use of loci associated with quantitative effects in dairy cattle improvement. Animal Production 27, 133–139.

Sonesson AK, Woolliams JA and Meuwissen THE 2010. Maximising genetic gain whilst controlling rates of genomic inbreeding using genomic optimum contribution selection. Proceedings of the 9th WCGALP, Leipzig, Germany, 150 p.

Sørensen AC and Sørensen MK 2009. Genotyping both males and females is favourable in genomic dairy cattle breeding schemes. Interbull Bulletin 40, 94–97.

Strauss S 2010. Biotech breeding goes bovine. Nature Biotechnology 6, 540–543.

Su G, Guldbrandtsen B, Gregersen BR and Lund MS 2010. Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. Journal of Dairy Science 93, 1175–1183.

Su G, Brondum RF, Ma P, Guldbrandtsen B, Aamand GP and Lund MS 2011. Genomic prediction using high-density SNP markers in Nordic Holstein and Red. Interbull Bulletin 44, 157–161.

VanRaden PM, Wiggans GR, van Tassel CP, Sonstegaard TS and Schenkel F 2009. Benefits from cooperation in genomics. Interbull Bulletin 39, 67–72.

Veerkamp RF, Berry DP, Wall E, de Haas Y, McParland S, Coffey M and Calus MPL 2011. Use of phenotypes from research herds to develop genomic selection for scarcely recorded traits like feed efficiency. Interbull Bulletin 44, 249–254.

Villanueva B, Pong-Wong R, Fernandez J and Toro MA 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. Journal of Animal Science 83, 1747–1752.

Weigel KA, de Los Campos G, Vazquez AI, Rosa GJ, Gianola D and Van Tassel CP 2010. Accuracy of direct genomic values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. Journal of Dairy Science 93, 5423–5435.

Wiggans GR, VanRaden PM and Cooper TA 2011. The genomic evaluation system in the United States: past, present, future. Journal of Dairy Science 94, 3202–3211.

Williams JL 2005. The use of marker-assisted selection in animal breeding and biotechnology. Scientific and Technical Review 24, 379–391.

Ytournel F, Willam A and Simianer H 2011. Impact of genomic selection on functional traits in a dual-purpose cattle breeding program. Proceedings of the 62nd EAAP meeting, Stavanger, Norway, Wageningen Academic Publishers, The Netherlands, p. 213.

https://doi.org/10.1017/S1751731112002248 Published online by Cambridge University Press