A review of genomic selection - Implications for the South African beef and dairy cattle industries

E. van Marle-Köster[1#], C. Visser[1] & D.P. Berry[2]

[1] Department of Animal & Wildlife Sciences, Faculty of Natural & Agricultural Sciences,
University of Pretoria, South Africa

[2] Animal and Grassland Research and Innovation Centre, Teagasc, Moorepark, Fermoy, Co. Cork, Ireland

Abstract

The major advancements in molecular technology over the past decades led to the discovery of DNA-markers, sequencing and genome mapping of farm animal species. New avenues were created for identifying major genes, genetic defects, quantitative trait loci (QTL) and ultimately applying genomic selection (GS) in livestock. The identification of specific regions of interest that affect quantitative traits aimed to incorporate markers linked to QTL into breeding programs by using marker assisted selection (MAS). Most QTL explained only a small proportion of the genetic variation for a trait with limited impact on genetic improvement. Single nucleotide polymorphism (SNP) markers created the possibility to genotype cattle in a single assay with hundreds of thousands of SNPs, providing sufficient genomic information to incorporate into breeding value estimation. Genomic selection is based on the principle of associating many genetic markers with phenotypic performance. A large database of genotyped animals with relevant phenotypes pertinent to a production system is therefore required. South Africa has a long history of animal recording for dairy and beef cattle. The challenge for implementation of GS would be the establishment of breed-specific training populations. Training populations should be genotyped using a high density SNP panel, and the most appropriate genomic prediction algorithm determined. The suitability of commercially available genotyping platforms to South African populations should be established. The aim of this review is to provide an overview of the developments that occurred over the past two decades to lay the foundation for genomic selection with special reference to application in the South African beef and dairy cattle industry.

Introduction

The past three decades were characterized by a number of major discoveries and technological developments in the field of molecular genetics. On a technological level the development of polymerase chain reaction (PCR) technology by Mullis (Fore *et al.*, 2006) was a major advancement for molecular research, followed by automated sequencing during the late ninety nineties. The discovery of the hypervariable region in the human genome (Tautz, 1989) paved the way for the discovery and mapping of different DNA markers (Dodgson *et al.*, 1997; Beuzen *et al.*, 2000), and these have been applied widely in several livestock species (Dekkers & Hospital, 2002; Dekkers, 2004; Pollak, 2005; Womack, 2005; Jeon *et al.*, 2006). The human genome was the first genome to be sequenced and contributed to improving sequencing and high throughput technologies (Adams, 2008; Eggen, 2012). Since the completion of the Human Genome project most farm animal species have been sequenced and mapped (Fan *et al.*, 2010) creating new opportunities for genetic improvement in livestock that was previously beyond researchers' reach.

Using DNA technology to perform genome mapping and sequencing, a number of useful single genes was mapped in small livestock (Montgomery & Kinghorn, 1997; Davis, 2004; Van der Werf, 2007) and in beef and dairy cattle (Anderson, 2001; Dekkers & Hospital, 2002). Diagnostic tests for genetic diseases such as bovine leukocyte adhesion deficiency (BLAD), deficiency of uridine monophospate synthase (DUMPS) and complex vertebral malformation (CVM) in dairy cattle have been applied to ensure that seed stock bulls can be identified as non-carriers or carriers of these autosomal recessive mutations (Robinson *et al.*, 1984; Schuster *et al.*, 1992; Thomsen *et al.*, 2006). In sheep, the Prp gene associated with scrapie has also been identified for individual testing (Belt *et al.*, 1995). Major genes affecting quantitative traits have been commercialized for application in the livestock industry, for example the *CAST* and *CAPN1* genes for meat tenderness and the *GDF8* gene for double muscling in beef cattle, as well as the *RYR* and *PRKAG3* genes affecting pork quality (Jeon *et al.*, 2006).

The identification of specific genomic regions of interest that affect economically important traits in farm animals held great interest for the livestock industry as it aims at incorporating genomic markers linked to quantitative trait loci (QTL) into breeding programs, by making use of marker assisted selection (MAS) (Anderson, 2001). In the search of QTL and major genes, different approaches have been investigated including genome wide scans using relatively high density panels of microsatellites or single nucleotide polymorphism (SNP) markers across the genome together with genome wide association studies or candidate gene approaches (Ron & Weller, 2007; Hayes & Goddard, 2010). However, it was found that most of the identified QTL explained only a small proportion of the genetic variation for a quantitative trait and was not expected to result in a significant increase in the rate of genetic improvement (Nicholas, 2006). It was postulated that DNA-markers and genomic information held the most potential for increasing genetic progress in quantitative traits with low heritability, traits that are difficult and/or costly to measure, and sex-limited traits (Dodds *et al.*, 2007; Hayes & Goddard, 2010).

Currently the possibility exists to genotype cattle in a single assay with up to 777 962 SNP markers with marker intervals of less than 3 kb, providing sufficient genomic information to be incorporated in breeding value estimation procedures as well as providing useful information for performing genome-wide association analyses. The aim of this review is to provide an overview of the developments that took place over the past two decades to lay the foundation for genomic selection with special reference to application in the South African beef and dairy cattle industry.

The bovine genome and the search for quantitative trait loci

To appreciate the molecular information currently available in cattle, it is first necessary to briefly summarise the first efforts to compile DNA marker maps of the bovine genome. The first genetic maps for cattle were constructed with 746 and 1250 microsatellite markers by Barendse *et al.* (1997) and Kappes *et al.* (1997), respectively. The map compiled by Kappes *et al.* (1997) had a genome coverage of 2990 cM with an average marker interval of 3.0 cM. This map was further improved by adding more than a thousand markers with improved coverage and decreased intervals between microsatellites (on average 1.4 cM) that resulted in a high density map with 3802 microsatellite markers (Ihara *et al.*, 2004). From this effort 880 000 genotypes were generated from the USDA MARC cattle reference families. The development of these high density linkage maps was essential for studying and fine mapping QTL and regions of interest.

In combination with molecular technology, software has been developed to incorporate large and complex pedigrees in the statistical analyses for identification of QTL (Seaton *et al.*, 2002; 2006). QTL identification studies have resulted in the identification of causative mutations such as DGAT for milk fat content in dairy cattle (Grisart *et al.*, 2002) and MSTN for double muscling in beef cattle (Charlier *et al.*, 1995). Genome scans were conducted for QTL associated with milk production, health and conformation traits in dairy cattle using granddaughter designs and outbred families (Zhang *et al.*, 1998; Heyen *et al.*, 1999; Schrooten *et al.*, 2000).

Potential advantages from incorporating QTL information in dairy breeding programs varied between studies. The French dairy industry genotyped more than 70 000 bulls for 14 chromosomal regions over a seven year period and the impact was sufficient to reduce the number of bulls for progeny testing by 15% (Boichard *et al.*, 2006). The putative QTL could, however, explain only a relatively small proportion of the genetic variance (0.1% to 13.5%) for most of the traits in the study. This limitation was even more obvious in a genome-wide association study (GWAS) performed in dairy cattle using approximately 43 000 SNP associated with feeding level and response in milk production to heat stress (Hayes & Goddard, 2010) which

only explained 1.5% and 2.0% of the genetic variance for the two traits. This phenomenon of unexplained genetic variation, termed the "missing heritability" (Maher, 2008), is not only confined to livestock, but is also present in human GWAS (Visscher, 2008).

Despite the efforts of performing GWAS, the effects of the individual putative QTL identified for the desired traits were small and MAS in livestock was much less effective than was initially expected (Hayes & Goddard, 2010). Furthermore, experiments with sufficient statistical power to detect QTL required several hundred or even thousands of individuals with both genotypic and phenotypic data (Van der Werf *et al*., 2007). It was thought that the uptake of this technology would increase if the costs could be shared by other applications (e.g. parentage testing, traceability). Nevertheless, it became clear that QTL identification would not be a feasible option for explaining a sizeable proportion of the genetic variance in complex traits. An alternative approach was therefore needed to exploit the information from a large number of genetic markers simultaneously.

The bovine genome sequence, completed in 2004, was the first genome with high SNP coverage of the Cetartiodactyla mammals (family Bovidae) to be sequenced. Single nucleotide polymorphisms are abundant, bi-allelic, single-locus markers located at approximately 3 kb intervals in the *Bos taurus* genome with an estimated total of nearly 40 million SNP which were identified during sequencing (Seidel, 2010). The frequency of more than 37 000 SNP markers were analyzed in 497 cattle by the Bovine HapMap Consortium and this set the scene for further expansion of SNP discoveries (Eck *et al*., 2009; Seidel, 2010). Van Tassel *et al*. (2008) added an additional 23 000 SNPs to the bovine collection, studying 66 cattle that included breeds such as the Holstein, Angus, Red Angus, Gelbvieh, Hereford, Limousin and Simmental. This resulted in the compilation of a commercial 54 001 SNP array that resulted in an additional tool for animal breeders that could be applied in genomic selection (Eck *et al*., 2009; Matukumali *et al*., 2009). The newly developed BovineHD beadchip from Illumina (http://www.illumina.com) features 777 962 relatively evenly spaced SNP over the entire bovine genome including the mitochondrial DNA.

Genomic selection

The potential use of genome-wide genetic marker information for use in animal breeding was originally proposed by Meuwissen *et al*. (2001). The principles of traditional MAS were to include a relatively small number of genetic markers in genetic evaluations (Fernando & Grossman, 1989). These markers originated from information generated in research studies, primarily meta-analyses of controlled experiments (Georges *et al*., 1995). Modern-day genomic selection, which is essentially a larger scale version of MAS, includes a considerably larger number of genetic markers; the "effects" of each marker are simultaneously estimated in the genomic selection process. The number of markers included in the genomic evaluations is dependent on the methodology used (discussed later) but the original set of genetic markers considered for inclusion is in the order of many thousands. Genomic selection generally assumes that all the genetic variation for a trait should be explained by markers. Polygenic effects are, however, sometimes included in the model to account for genetic variation that is unexplained by the genetic markers (Hayes *et al*., 2009). Genomic selection should ultimately lead to using genotypes defined by a set of polymorphisms to select for preferred phenotypes (Seidel, 2010).

Implementation of genomic selection in any population requires: 1) genotypes of a large population of animals that have, 2) pertinent phenotypes for the system of production where the genetic/genomic evaluations will be used, and 3) statistical methodologies for implementing efficient and accurate genomic predictions. This is based on the assumption that the breeding program in place is already optimal and includes a) an accurate genetic evaluation system based on access to relevant and heritable phenotypes, b) a pertinent breeding objective encompassing, as far as possible, all relevant traits optimally weighted within a breeding goal, and c) a breeding scheme to ensure long-term sustainable genetic gain.

Genotype information

Many genotyping platforms (commonly referred to as "SNP chips") are now commercially available, ranging in number of genetic markers from several SNPs to 777 962 SNPs (Illumina Inc., San Diego, CA), including copy number variants (CNVs) and indels. The editing of genotypes for inclusion in genomic evaluations (or genome-wide association studies) has been described in detail elsewhere (Wiggans *et al*., 2009; 2010; Berry & Kearney, 2011). A short summary is provided here: Firstly, samples with low genotype call rate (e.g. <90% [Wiggans *et al*., 2009] to <95% [Berry & Kearney, 2011]) are discarded as are parent-

offspring pairs with an unacceptable high (e.g. >0.5% [Berry & Kearney, 2011] to >0.7% [Wiggans *et al*., 2009]) level of autosomal SNP genotypes that do not adhere to Mendelian inheritance.

Monomorphic SNPs do not contribute any information to genomic predictions and may therefore be discarded to reduce future computational requirements, as can SNPs with low minor allele frequency (i.e. <2% [Wiggans *et al*., 2010; Berry & Kearney, 2011] to <2.5% [Hayes *et al*., 2009]). Restrictions on the extent of deviations of SNP frequency from Hardy-Weinberg equilibrium can also be imposed (Hayes *et al*., 2009; Wiggans *et al*., 2010; Berry & Kearney, 2011); care should, however, be taken if imposing these criteria in multi-breed or multi-strain populations. Even within breed, strict restrictions on the Hardy-Weinberg equilibrium statistic should be undertaken with caution since selection can cause departures from Hardy-Weinberg equilibrium, for example, loci harbouring lethal recessive genetic defects (Berry & Kearney, 2011).

The cost of generating the genotype of an animal has reduced considerably in recent years with advances in technologies. Improvement in accuracy of genomic predictions can be achieved with increased marker density (VanRaden *et al*., 2009) especially for across-breed evaluations (De Roos *et al*., 2008). The number of SNP necessary to achieve accurate genomic predictions depends on the extent of linkage disequilibrium in the species or breed, the length of the genome (L), and the effective population size (N_e) (Hayes & Goddard, 2010). It has been shown that approximately 50 000 SNPs are sufficient to obtain accurate genomic predictions (assuming a relatively large training population of genotyped and phenotyped animals) within the Holstein (Berry *et al*., 2009; VanRaden *et al*., 2009), but at least 300 000 SNP are required for reliable prediction in Jersey cattle, using a Holstein reference data of genotyped and phenotyped animals (De Roos *et al*., 2008). Improved accuracy of prediction with greater marker density is expected since the success of genomic selection is based on exploiting genetic markers in linkage disequilibrium with the causative mutations and therefore the greater the marker density the greater the likelihood of tighter linkage disequilibrium between the functional mutation and the genotyped markers. Nonetheless, improvements in the accuracy of genomic prediction from greater marker density are likely to be dependent on the algorithm used in the genomic predictions.

In addition to the ever-reducing cost of genotyping, the lower the number of genetic markers genotyped, the lower the likely cost. Imputation is a method of exploiting linkage analysis and/or linkage disequilibrium by deducing a higher density genotype from a lower density (and therefore lower cost) genotype. Current algorithms for imputation exploit linkage (findhap; VanRaden *et al*., 2011), population wide linkage disequilibrium (Beagle; Browning & Browning, 2007; 2009) or combined linkage analysis and linkage disequilibrium (FImpute; Sargolzaei *et al*., 2011). Berry & Kearney (2011) reported an average accuracy of imputation from 2 909 SNPs to 54 001 SNPs of 98% in Holstein-Friesian dairy cattle; similar imputation accuracies were reported elsewhere (Weigel *et al*., 2010). Therefore lower density genotyping, coupled with imputation, can be used to reduce the cost of genomic selection. Imputation approaches may also be used to impute the genotype of a un-genotyped ancestor with several genotyped progeny.

Several international initiatives in dairy cattle are underway to share genotypes thereby reducing the cost of genomic selection for each country (Cromie *et al*., 2010; Jorjani *et al*., 2011). The success of these initiatives in dairy cattle is due to the international nature of semen trade, and aided by international genetic evaluations undertaken by INTERBULL. Increased accuracy of prediction is achieved when the genomic selection reference population is closely related to the selection candidates (Habier *et al*., 2007); therefore providing genotyped back-pedigree (with phenotypes from INTERBULL) to the importing country can be beneficial to the exporting country.

Phenotypic information

Genomic selection is based on the principle of relating genetic markers to phenotypic performance. A large database is therefore required of genotyped animals with all relevant phenotypes pertinent to the system of production where the genomic predictions will be applied. This large database of phenotyped and genotyped animals is generally referred to as the reference population or training population and is used to estimate the genetic marker effects. Within breeds, the improvement in the accuracy of genomic predictions with increasing size of the reference population is non-linear and is dependent on how accurately the phenotypic measures reflect the true breeding value (i.e. heritability) of the animals (Daetwyler *et al*., 2008; Goddard, 2008) and the relatedness of the reference population of animals to the animals where the genomic prediction equations will be applied (Habier *et al*., 2007). The lower the accuracy of the phenotypes, as is, on

average, the case for lower heritable traits such as fertility (Veerkamp & Beerda, 2007) and health (Berry *et al.*, 2011), the lower the accuracy of genomic predictions for the same reference population size (Daetwyler *et al.*, 2008; Goddard, 2009). For the same reference population size, the greater the relatedness of the reference population to the population where the prediction equations will be applied, the greater will be the accuracy of the genomic predictions (Habier *et al.*, 2007).

Artificial insemination (AI) bulls constitute the majority of international dairy cattle training populations. This is because they generally have more accurate predictions of genetic merit and therefore fewer animals need to be genotyped to achieve the same accuracy of genomic predictions (Daetwyler *et al.*, 2008; Goddard, 2009). The disadvantage of using only AI sires is that the traits included in the genomic predictions are limited to those available nationally and used to estimate the sire breeding values. Facilitated by multiple trait-across country genetic evaluations (MACE; Schaeffer, 1994) undertaken by INTERBULL for dairy cattle, predictions of genetic merit of all international male AI animals on the scale of each member country are obtainable. Therefore, even though a male animal may have no progeny in a given country its MACE evaluation may be used as a phenotype for inclusion in genomic predictions (Berry & Kearney, 2011) weighted by a function of the MACE reliability. Nonetheless, the number of AI sires is limited and therefore natural mating bulls or cows must also be considered for inclusion in genomic selection reference populations. Cows are currently included in the genomic selection reference population in the United States (Wiggans *et al.*, 2011) while natural mating bulls are included in the dairy cattle genomic selection reference population in Ireland (Andrew Cromie, personal communication); the latter is particularly relevant for populations (e.g. some beef populations) where AI is used less frequently.

The phenotype included in all genomic evaluations of dairy cattle are either daughter yield deviations or deregressed estimated breeding values (Berry & Kearney, 2011) that remove pedigree contributions and reverse the effect of shrinkage during the BLUP evaluations. For some populations and traits (e.g. feed intake), however, a sufficiently large reference population may not be available for accurate genomic prediction. Intergenomics (Jorjani *et al.*, 2011) is an international collaboration to undertake international genomic evaluations in Brown Swiss cattle by pooling of international phenotypes (and genotypes). Veerkamp *et al.* (2012) combined data on Holstein-Friesian dairy cows from research herds in four countries to generate genomic predictions; a similar initiative was undertaken for genomic predictions for residual feed intake in growing dairy heifers in Austral-Asia (Pryce *et al.*, 2012).

Genomic prediction algorithms

The main challenge in genomic predictions is to estimate genetic marker effects because the number of genetic markers available exceeds the number of phenotypes available by several factors, although this phenomenon is changing in some countries (Wiggans *et al.*, 2011). Genetic markers may include SNPs, CNVs, indels but may also include haplotypes. Using simulations, Calus *et al.* (2007), however, documented that the advantage of using haplotypes over SNPs decreased as the linkage disequilibrium between adjacent SNPs increased, with an r^2 (i.e. measure of linkage disequilibrium) of 0.215 being where both gave similar accuracy. Within Holstein-Friesian cattle the average r^2 between adjacent SNPs with approximately 40,000 SNPs is expected to be >0.19 (Khatkar *et al.*, 2008) thus suggesting that with the Illumina Bovine50 beadchip, using SNPs rather than haplotypes is most sensible. SNPs are generally the marker of choice in national genomic predictions (Berry *et al.*, 2009; Hayes *et al.*, 2009; VanRaden *et al.*, 2009).

Irrespective of the genomic prediction algorithm, the general model used to estimate SNP effects is:

$$Y_i = \mu + \sum_{i=1}^{n} X_j g_{ij} + e_i$$

where Y_i is the phenotypic value of animal i for the trait under investigation, μ is the mean effect for the trait under investigation, X_j is the effect of locus j, g_{ij} is the genotype of animal i at locus j, and e_i is the residual term for animal i; a polygenic effect may also be included in the model (Calus, 2010).

Different approaches exist to estimate the SNP effects. Due to the number of parameters usually being considerably larger than the number of phenotypic records, simple multiple regression methods cannot be used, mainly because of a lack of degrees of freedom, but also because of other concerns such as high levels of multi-collinearity among SNPs. Several algorithms have been proposed for genomic selection and a summary of the assumptions for the different algorithms as described by Hayes & Goddard (2010) is shown in Table 1.

Table 1 Summary of assumptions for methods based on SNP markers for genomic EBV estimations (adapted from Hayes & Goddard, 2010)

Name	Assumed distribution of SNP effects	Implication
BLUP	Normal	A very large number of QTL of small effect each
BayesA	*t* distribution	A large number of QTL of small effect and a small proportion with moderate to large effect
BayesB	Mixture distribution of zero effects and *t* distribution of effects	A large number of genome regions with zero effect, a small proportion of QTL with moderate effects
Bayesian LASSO	Double exponential distribution of effects	Very large proportion of SNP with effect of close to zero, small proportion of moderate to large effect
BayesSSVS*	Mixture distribution of zero effects and *t* distribution of effects	A large number of genome regions with almost zero effect, a small proportion of QTL with moderate effects

* Stochastic search variable selection.

One approach for genomic predictions is to use least squares regression but on a considerably reduced number of SNPs that have passed certain criteria such as very strongly associated with the trait under investigation from a series of univariate analyses; machine learning techniques as well as many other approaches may also be used to select a subset of informative SNPs. The final set of SNPs included in the multiple regression model may be obtained using a stepwise algorithm. Dimension reduction techniques have also been proposed such as 1) principal component analyses of the SNPs and subsequent inclusion of the main principal components in a regression analysis, or 2) partial least squares analysis which is similar to principal component analysis but where the latent variables generated take cognizance of their ability to also capture variation in the dependent variable. Other dimension reduction techniques also exist.

Rather than include SNPs individually as fixed effects, SNPs may also be included as random effects. This approach was suggested by the original genomic selection paper (Meuwissen *et al.*, 2001) and is now commonly referred to as GBLUP. The GBLUP model assumes that each SNP contributes equally to the additive genetic variance of the phenotype under investigation and a normal distribution of the SNP effects is usually assumed. Treating each SNP simultaneously as a random effect is equivalent to constructing a genomic relationship matrix from the genotypes (VanRaden, 2008) and replacing the traditional numerator relationship matrix in the mixed model equations for genetic evaluations with the genomic relationship matrix. It is the latter description of GBLUP that is the commonly used GBLUP approach in national genomic evaluations (Berry & Kearney, 2011). However, the contribution of each SNP to the genomic relationship matrix can be weighted differently to account for different variances per SNP and differences in surrounding marker density (VanRaden, 2008).

Two Bayesian approaches were originally proposed by Meuwissen *et al.* (2001) which they termed BayesA and BayesB, although Gianola *et al.* (2006) questioned whether they were truly Bayesian approaches. In the BayesA approach of Meuwissen *et al.* (2001) the prior distribution of the variance of the SNP effects was chosen to represent a few SNPs with large effects and many SNPs with small effects (i.e. an inverted chi-square distribution). The SNP effects are sampled from a normal or t-distribution. BayesB, as proposed by Meuwissen *et al.* (2001), was largely similar to BayesA with the exception that, a priori, the proportion of SNPs not associated with the phenotype of interest was set and SNPs not entering the model within a Gibbs sampling chain were set to zero. Other Bayesian approaches have since been described (for example Habier *et al.*, 2011). BayesCπ (Habier *et al.*, 2011) as well as other modifications compared to BayesB (e.g. use of a multivariate t-distribution to describe the SNP effects) samples the proportion of SNPs that are associated with the phenotype, rather than this statistic having to be decided on prior to the analysis. Such mixture models, combining two or more distributions of SNP effects, can accommodate SNPs with a distribution of large effects and a distribution of small effects (Calus *et al.*, 2008). This approach is similar to

BayesB but avoids the requirement for the Metropolis-Hastings step thereby reducing computational requirements, but still facilitates the inclusion of many SNPs each with small effects (unlike BayesB where such effects were set to zero) thereby capturing any remaining unexplained genetic variance (Calus, 2010). Non-parametric kernel methods have also been proposed for genomic predictions (for review, see Calus, 2010).

Several studies have compared the efficiencies and accuracies of the different algorithms using either real data (Hayes *et al.*, 2009; VanRaden *et al.*, 2009) or simulated data (Calus *et al.*, 2008; Gredler *et al.*, 2009; Meuwissen & Goddard, 2010a; b; Meuwissen *et al.*, 2001; Pszczola *et al.*, 2011). Although not always consistent across studies, the Bayesian approaches described are generally more accurate for traits where large QTL exist (Hayes *et al.*, 2009; VanRaden *et al.*, 2009; Meuwissen & Goddard, 2010a). Genomic predictions of milk fat and protein composition generally perform better with a Bayesian approach assuming a t-distribution of SNP effects (Gredler *et al.*, 2009; Hayes *et al.*, 2009) and this is likely due to the large impact of the DGAT1 gene on milk fat and protein composition (Berry *et al.*, 2010).

Breeding scheme

The ability to accurately estimate the genetic merit of an individual from its DNA through genomic selection is causing a paradigm shift in dairy cattle breeding programs. Genomic selection is now implemented in many dairy cattle national genetic evaluations in the United States (VanRaden *et al.*, 2009), Europe (Berry *et al.*, 2009) and Australasia (Harris & Johnson, 2010). Devising and implementing breeding schemes to maximize the potential of this technology is required. Schaeffer (2006) compared a selection strategy using genomic selection to a traditional progeny testing scheme similar to that operated in Canada. He reported a two-fold increase in genetic gain using genomic selection with a 92% reduction in the cost of proving the bull. Other simulation studies suggest at least a 50% increase in annual genetic gain from implementation of genomic selection breeding programs compared to traditional breeding programs (Pryce *et al.*, 2010; Lillehammer *et al.*, 2011; McHugh *et al.*, 2011).

The fundamental change on how genomic selection will influence breeding programs is that genomic predictions can be obtained at a very young age, well before sexual maturity, thereby reducing the generation interval considerably. Therefore, research into the optimal use of reproductive technologies in breeding programs is likely to intensify, in particular the genotyping of embryos (Humblot *et al.*, 2010). Furthermore, two of the selection pathways in genetic gain originate with the dam (i.e. dams to produce sires and dams to produce dams). For low heritability traits in particular, accurate estimates of the genetic merit of the dams is difficult. Genomic selection will improve the accuracy of selection of the dams and increase genetic gain further. McHugh *et al.* (2011) documented a large increase in genetic gain achievable when females were also genotyped; not only were the candidate dams more accurately identified, but the contribution of the genotyped and phenotyped dams to the genomic selection training population increased the accuracy of selection even more.

Concerns

As with most new technologies, misuse of the technology, intentionally or not, can have unfavourable consequences. Genomic selection is no exception. Genetic gain is expected to increase by at least 50% with successful implementation of genomic selection (Pryce *et al.*, 2010; Lillehammer *et al.*, 2011; McHugh *et al.*, 2011). Rapid increases in genetic gain can reduce the ability to purge out unfavourable consequences of selection, including inbreeding depression (McParland *et al.*, 2009). One option to minimize (but not eliminate) this concern is to use bulls sparingly in their first year. In this way any congenital defects or calving difficulties may be identified earlier and the bull could be culled or used in appropriate matings (i.e. avoidance of carrier females if the bull is a carrier of a lethal recessive allele). Of course, if the bull is a carrier of a rare lethal recessive allele then this may not be observable in a small number of (unrelated) animals. Nonetheless, young bulls cannot produce sufficient semen to generate large progeny group sizes and therefore small first-crop progeny groups will be generated. Furthermore, an extensive phenotyping strategy must be implemented to rapidly identify any unfavourable consequences of selection. Sentinel herds could be used to compare national average genetic merit to elite genetic merit animals under contrasting systems of production (Wickham *et al.*, 2012).

Because the increase in accuracy of genomic predictions, and therefore genetic gain, is a function of the number of phenotyped animals and the accuracy of their phenotypes, genetic gain will tend to be greatest

in high heritability traits where ample phenotypic data are available (Amer, 2012). This will place greater selection pressure on traits such as milk production which are known to be antagonistically correlated with fertility (Berry *et al*., 2012) and health (Berry *et al.*, 2011) thereby possibly leading to a reduction in genetic gain for fertility and health traits or even reverting to a deterioration in genetic merit for the latter traits in populations where it is currently improving.

Genomic selection has intensified global competition in the international trade of dairy cattle germplasm and this can have serious consequences for inbreeding, and its associated deleterious effects (McParland *et al*., 2007). Several simulations (Daetwyler *et al*., 2007) have reported a lower accumulation of inbreeding per generation with genomic selection compared to breeding programs based on traditional BLUP genetic evaluations; however, because the generation interval is shortened in genomic selection breeding programs compared to traditional progeny test breeding programs, inbreeding per annum may be greater with genomic selection. Knowledge of the DNA of individual animals can, nevertheless, be used to reduce the accumulation of inbreeding by identifying least related animals at the genomic level. Stochastic simulations also show that the use of sexed semen in a genomic selection breeding program can help manage rates of inbreeding (Pedersen *et al.*, 2012). Moreover, genomic selection can be used to screen a much larger population of potential candidate animals thereby widening the genetic diversity and, if unrelated / lowly related candidates are genetically elite, then help manage the rates of inbreeding (Hayes & Goddard, 2010).

Implications of genomic selection for South African dairy and beef cattle
Genotypes
South African dairy and beef cattle breeders are already exploiting DNA technology through DNA-based parentage verification and diagnostic testing (Van Marle-Köster & Nel, 2003). Dairy bulls are routinely screened for known recessive disorders such as BLAD, DUMPS and CVM. Beef cattle can be tested for meat tenderness with various diagnostic kits as well as for a few genetic disorders, including dwarfism and certain translocations. This is, however, the only (limited) genotypic information available for cattle breeds and the quantity of information differs greatly among breeds. No official system for the collection and storage of biological samples of cattle currently exists in South Africa. A livestock identification system where breeders stored hair samples on a voluntary basis for use in the advent of livestock theft is the only collective repository for biological samples (ARC, 2007). Other biological material (including semen, blood and hair samples) is dispersed among AI companies, research institutions, laboratory facilities, etc.

The South African dairy industry has an advantage over the beef industry owing to stronger national and international genetic linkage with (internationally) superior AI sires. South Africa is a member of INTERBULL which is the body responsible for international genetic evaluations of dairy cattle (Mostert, 2007). In the beef industry there is, however, particular challenges regarding the implementation of genomic selection of which the main one is the limited recording of performance traits. Often the specific traits that could gain the most from genomic selection such as fertility, longevity and carcass quality have not been recorded in South Africa. There is also less genetic linkage between herds due to the limited use of AI by beef breeders and also lower availability of animals with accurate EBVs. These limitations are nonetheless not unique to South Africa and were also documented by Garrick (2011) for the US cattle industry. The structure of the beef industry in South Africa is also more complex with a relatively small seed stock sector where genetic selection takes place and which provides the genetic material (bulls) for commercial production. Also, there is a large number of beef breeds in South Africa compared to dairy breeds which has implications for the identification of suitable reference/training populations needed for the implementation of genomic selection in the entire beef sector.

The challenge for implementation of genomic selection therefore is to firstly collate all available biological samples (hair, semen or live animals for blood) and quantify whether they are of sufficient quality and quantity for the establishment of a breed-specific training population. DNA of animals that have made a large genetic contribution to the modern-day population should also be sourced. All animals should have accurate phenotypic information for use in genomic predictions but genotypes on prominent ancestors, even without phenotypic information, can also be useful to generate population haplotypes for more accurate imputation from lower to higher density genotypes. Once the training population has been established, all animals should be genotyped using a SNP panel, and the most appropriate genomic prediction algorithm

determined using the appropriate validation procedures. The suitability of commercially available genotyping platforms to South African populations should be established.

Phenotypes

Beef and dairy cattle in South Africa has a relatively long history of performance recording with official performance testing for beef cattle taking place since 1959 and milk recording since 1917 (Bergh, 2010). Participation in national performance recording schemes and private recording schemes differs, however, between breeds as performance testing has not been enforced by all breed societies. Participation in the recording scheme for beef cattle varies from as low as 32% to 100% (Scholtz, 2010), while participation in official milk recording schemes is estimated at only 24% (Scholtz & Grobler, 2009). The available recording schemes facilitate breeders to submit data on fitness (reproduction), production and quality traits. The average number of cows participating in milk recording in South Africa is summarized in Table 2 for the three major dairy breeds in South Africa.

Table 2 Number of dairy cows and bulls in animal recording, and traits recorded in South Africa (2012, R.R. van der Westhuizen, Pers. Comm., bobbie@studbook.co.za)

Breed	No of cows	No of bulls	Milk yield	Butter fat	Protein	SCC	Linear scores
Ayrshire	14604	126	x	x	x	x	
S A Jersey	56864	768	x	x	x	x	x
S A Holstein	71105	1552	x	x	x	x	x

x = trait recorded.
SCC: somatic cell count.

South Africa has approximately 30 different beef cattle breeds including British, European, composite and indigenous breeds. In Table 3 a summary is provided of the beef cattle breeds that have performance information on at least 1000 registered cows and bulls (older than 2 years).

Many of the South African dairy populations participate in INTERBULL and therefore genetic evaluation of all internationally available bulls exists on the South African scale for the majority of traits. These estimates of genetic merit (i.e. phenotypes) are freely available for use in national genomic evaluations. Much of the germplasm used in South African dairy herds originates from outside South Africa (Dürr & Jackobsen, 2009), either directly or indirectly. Therefore, these international estimates of genetic merit, which include performance information from South Africa itself, can be extremely beneficial for implementing genomic selection in a breed. Genotypes for many international AI bulls are currently available through international collaboration (Cromie *et al.*, 2010). Therefore, it is possible to at least undertake research on the potential to implement genomic selection in (some of) these populations.

Beef cattle breeds that may take advantage of genomic selection will need adequate numbers of phenotyped animals. Lower heritability traits such as fertility and health (Veerkamp & Beerda, 2007; Berry *et al.*, 2011) which are the type of traits that can generally benefit most from exploiting genomic information, require even larger populations of animals to achieve noticeable improvements in the accuracy of genomic predictions. The description of the Bonsmara (a composite breed developed in South Africa) detailing the available number of breeding animals with records for maternal, reproduction and growth efficiency traits is summarized in Table 4. The Bonsmara is the largest beef cattle breed in South Africa, with in excess of 81 000 cows. Participation in the beef cattle improvement scheme is compulsory for all Bonsmara breeders, and phenotypic recording is thus enforced. Assuming that sufficient biological material or DNA of appropriate quality is available on a large sample of these animals, the Bonsmara should have sufficient resources to implement genomic selection for at least a selection of performance traits. Some other beef breeds in South Africa (e.g. SA Angus and Beefmaster) should also have sufficiently large databases for the possible establishment of sufficiently sized training populations. For these breed, the resources (e.g. genotypes and phenotypes) available can also be improved through international collaboration.

Table 3 Major beef cattle breeds in South Africa and traits recorded (SA Studbook, May 2012)

	Registered Females	Registered Males	Birth weight dir	Birth weight mat	Weaning weight	Milk	Combined Mat	12month weight	18month weight	Post wean weight	Mature weight	ADG	Kleiber Ratio	FCR	Scrotum circumf.	AFC	ICP	Height	Length	Dressing%	Meat yield	Marbling	Growth Value	Cow Value	Production Value
Sanga:																									
Nguni	31400	5894	x	x	x	x		x	x		x	x	x												
Drakensberger	8167	1637	x	x	x	x				x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Afrikaner	4539	680	x	x	x	x		x	x		x	x	x	x	x			x	x						
Tuli	3747	824	x	x	x	x		x	x		x	x	x												
British breeds:																									
Angus SA.	10553	1723	x		x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Hereford	3789	624	x	x	x	x		x	x		x	x	x	x	x			x	x						
Sussex	4260	882	x	x	x	x		x	x		x	x	x	x	x			x	x						
European breeds:																									
Charolais	3291	570	x	x	x	x		x	x		x	x	x	x	x	x		x	x						
Braunvieh	1345	177	x	x	x	x		x	x		x	x	x					x	x						
Pinzgauer	1101	169	x	x	x	x		x	x		x	x	x					x	x						
Composite breeds:																									
Bonsmara	81881	14729	x	x	x	x				x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Beefmaster	24134	2992	x	x	x	x		x	x		x	x	x	x	x			x	x						
Santa Gertrudis	1825	1826	x	x	x	x		x	x		x	x	x	x	x			x	x						

	Registered Females	Registered Males	Birth weight	200d Weight	400d Weight	600d weight	Mature Cow Weight	Milk	Scrotal size	Days to calving	Gestation length	Calving ease	Carcass weight	EMA	Rib fat	Rump fat	Retail Beef Yield	IMF%	Feedlot index	Grassfed index	Terminal sire index	Weaner index
***Bos indicus*:**																						
Brahman	24542	11371	x	x	x	x	x	x	x				x	x	x	x	x	x				
European breeds:																						
Limousin	3949	1762	x	x	x	x	x	x	x				x	x	x	x	x	x				
Simmentaler	13755	4243	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
Composite breeds:																						
Brangus	4643	1704	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x				
Simbra	7388	2893	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x		x
Braford	1850	675	x	x	x	x	x	x	x	x			x	x	x	x	x	x				

The first challenge for genomic selection in South Africa is to evaluate which breeds have sufficient biological samples and phenotypic data to be able to establish a sufficiently sized training population to generate accurate genomic predictions. The suitability of the commercially available genotyping platforms for genomic evaluations needs also to be confirmed in these populations. This breed-specific training population can then be used to obtain a prediction equation with an ideally high correlation between the genotypic information and the "true" breeding values for the different traits. A number of proven bulls with a

large number of recorded progeny (with breeding value predictions with a very high reliability) should be genotyped to validate the accuracy of prediction. Once high predictive ability is achieved, information obtained from low density genotyping of animals can be imputed to calculate direct genomic values. The steps to be taken to achieve this are shown in Figure 1. The use of genomic selection should result in shorter generation intervals and an increased rate of genetic improvement for the South African cattle industry. Across-breed genomic evaluations may increase the accuracy of genomic predictions further, especially for breeds with smaller population sizes, but the potential to achieve high accuracy of genomic predictions using a multi-breed cattle population and the current genotyping platforms is currently unknown.

Table 4 Traits recorded and number of records per trait for the Bonsmara breed (April 2012)

Trait	Accuracy	Number in pedigree file (Maximum available)		Number alive (Minimum available)	
		Males	Females	Males	Females
Weight traits					
WW	>60	540853	577644	28992	82745
	>80	15585	83394	1485	15190
Milk	>60	441657	481033	22497	71031
	>80	11021	4800	975	523
Mature weight	>60	7108	45712	211	8475
	>80	913	20	18	0
Fertility traits					
AFC	>60	10657	89275	589	17424
	>80	1686	8	49	0
ICP	>60	883	64	16	4
	>80	50	0	2	0
Post-wean growth tests					
Extensive (ADG)	>60	148242	30213	10073	5588
	>80	3554	26	315	4
Intensive (FCR)	>60	20363	5320	1203	1091
	>80	806	19	61	4
Carcass traits					
EMA	>60	2362	17	526	13
	>80	61	0	24	0

WW: weaning weight; Milk: maternal WW; AFC: age at 1st calving; ICP: inter-calving period; ADG: average daily gain; FCR: feed conversion ratio; EMA: eye muscle area.

Conclusion

Application of genomic selection globally in both the dairy and beef industries is underway and smaller populations with fewer resources will have to collaborate and carefully plan breeding programs to remain internationally competitive. It is also important to note that the same principles of genomic selection discussed here for cattle also apply to sheep and goats. South Africa, with a significantly large small stock population, should therefore also investigate the potential of this technology for especially sheep where an internationally comparable animal recording system exists. It is envisaged that genomic selection will be important for South African dairy and beef cattle breeders and in the long term also for small stock.
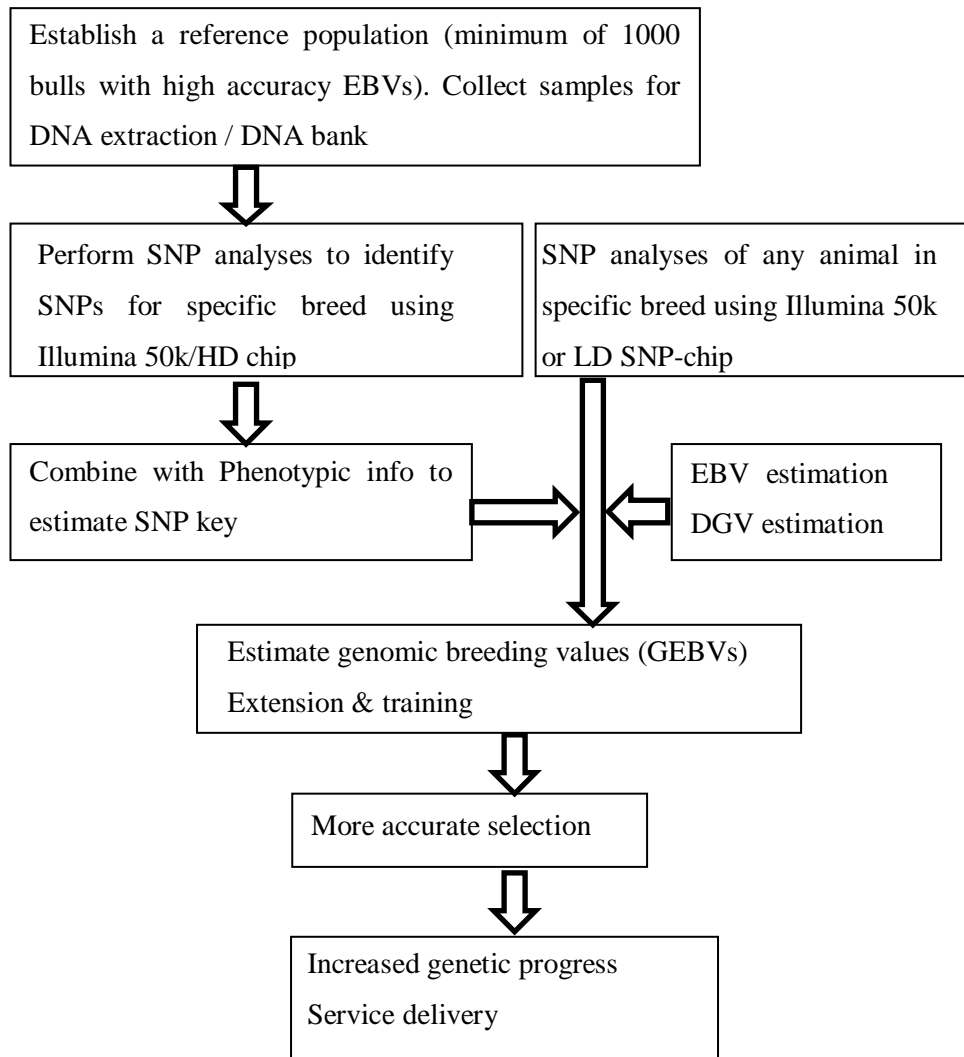
Establish a reference population (minimum of 1000 bulls with high accuracy EBVs). Collect samples for DNA extraction / DNA bank

⇩

Perform SNP analyses to identify SNPs for specific breed using Illumina 50k/HD chip

SNP analyses of any animal in specific breed using Illumina 50k or LD SNP-chip

⇩

Combine with Phenotypic info to estimate SNP key

⇨   ⇦

EBV estimation
DGV estimation

⇩

Estimate genomic breeding values (GEBVs)
Extension & training

⇩

More accurate selection

⇩

Increased genetic progress
Service delivery

Figure 1 Steps in the process to implement genomic selection.

References

Adams, J., 2008. Sequencing the human genome: The contributions of Francis Collins and Craig Venter. Nature Education 1 (1).

Amer, P.R., 2012. Turning science on robust cattle into improved genetic selection decisions. Animal 6, 551-556.

Anderson, L., 2001. Genetic dissection of phenotypic diversity in farm animals. Nature Rev. Genet. 2, 130-138.

ARC, Annual Report, 2006 – 2007. PO Box 8783, Pretoria, 0001.

Barendse, W., Vaiman, D., Kemp, S.J., Sugimoto, Y., Armitage, S.M., Williams, J.L., Sun, H.S., Eggen, A., Agaba, M., Aleyasin, S.A., Band, M., Bishop, M.D., Buitkamp, J., Byrne, K., Collins, F., Cooper, L., Coppettiers, W., Denys, B., Drinkwater, R.D., Easterday, K., Elduque, C., Ennis, S., Erhardt, G., Ferretti, L., Flavin, N., Gao, Q., Georges, M., Gurung, R., Harlizius, B., Hawkins, G., Hetzel, J., Hirano, T., Hulme, D., Jorgensen, C., Kessler, M., Kirkpatrick, B.W., Konfortov, B., Kostia, S., Kuhn, C., Lenstra, J.A., Leveziel, H., Lewin, H.A., Leyhe, B., Lil, L., Martin Burriel, I.., McGraw, R.A., Miller, J.R., Moody, D.E., Moore, S.S., Nakane, S., Nijman, I.J., Olsaker, I., Pomp, D., Rando, A., Ron, M., Shalom, A., Teale, A.J., Thieven, U., Urquhart, B.G.D., Vage, D.I., Van de Weghe, A.,

Varvio, S., Velmala, R., Vilkki, J., Weikard, R., Woodside, C., Womack, J.E., Zanotti, M. & Zaragoza, P., 1997. A medium-density genetic linkage map of the bovine genome. Mamm. Genome 8, 21-28.

Belt, P.B.G.M., Muileman, I.H., Schreuder, B.E.C., Bos-de Ruijter, J., Gielkens, A.L.J. & Smits, M.A., 1995. Identification of five allelic variations of sheep PrP gene and their association with natural scrapie. J. Gen. Virol. 76, 509-517.

Bergh, L., 2010. The national beef recording and improvement scheme. In: Beef Breeding in South Africa. 2nd ed. Ed. Scholtz, M.M., Publ.: Agricultural Research Council, Pretoria, South Africa. pp. 55-70.

Berry, D.P., 2008. Genetics – A tool to improve productivity and profitability. Int. J. Dairy Technol. 61, 30-35.

Berry, D.P. & Kearney, J.F., 2011. Imputation of genotypes from low-to high-density genotyping platforms and implications for genomic selection. Animal 5, 1162-1169.

Berry, D.P., Kearney, J.F. & Harris, B.L., 2009. Genomic selection in Ireland. Interbull Bulletin. Uppsala Sweden. 26-29 January 2009. pp. 29-34.

Berry, D.P., Howard, D., O'Boyle, P., Waters, S., Kearney, J.F. & McCabe, M., 2010. Associations between the K232A polymorphism in the diacylglycerol-O-transferace 1 (DGAT1) gene and performance in Irish Holstein-Friesian dairy cattle. Irish J. Agric. Food Res. 49, 1-9.

Berry, D.P., Bermingham, M., Good, M. & More, S.J., 2011. Genetics of animal health and disease in cattle. Ir. Vet. J. 64, 5.

Berry, D.P., Kearney, J.F., Twomey, K. & Evans, R.D., 2012. Genetics of reproductive performance in seasonal calving dairy cattle production systems. Irish J. Agric. Food Res. (in Press).

Beuzen, N.D., Stear, M.J. & Chang, K.C., 2000. Molecular markers and their use in animal breeding. Vet. J. 160, 42-52.

Boichard, D., Fritz, S., Rossignol, M.N., Boscher, A. Malafosse & Colleau, J.J., 2006. Implementation of marker-assisted selection in French dairy cattle. Proc. 8th World Congr. Genet. Appl. Livest. Prod., Commun. 2-11. Instituto Prociencia, Belo Horizonte, Brazil.

Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. Am. J. Hum. Genet. 81, 1084-1097.

Browning, B.L. & Browning, S.R., 2009. A unified approach to genotype imputation and haplotype phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84, 210-223.

Calus, M.P.L., 2010. Genomic breeding value prediction: Methods and procedures. Animal 4, 157-164.

Calus, M.P.L. & Veerkamp, R.F., 2007. Accuracy of breeding values when using and ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP for cM. J. Anim. Genet. 124, 362-368.

Calus, M.P.L., Meuwissen, T.H.E., De Roos, A.P.W. & Veerkamp, R.F., 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178, 553-561.

Charlier C., Coppieters, W., Farnir, F., Grobet, L., Leroy, P.L., Michaux, C., Mni, M., Schwers, A., Vanmanshoven, P. & Hanset, R., 1995. The mh gene causing double-muscling in cattle maps to bovine Chromosome 2. Mamm. Genome 6, 788-792.

Cromie, A.R., Berry, D.P., Wickham, B., Kearney, J.F., Pena, J., van Kaam, J.B.C.H., Gengler, N., Szyda, J., Schnyder, U., Coffey, M., Moster, B., Hagiya, K., Weller, J.I., Abernethy, D. & Spelman, R., 2010. International Genomic Co-Operation; Who, What, When, Where, Why and How? In: InterBull Bulletin No. 42. Riga, Latvia.

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

Daetwyler, H.D., Villanueva, B. & Woolliams, J.A., 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3, e3395.

Davis, G.H., 2004. Fecundity genes in sheep. Anim. Reprod. Sci. 82-83, 247-253.

Dekkers, J.C.M., 2004. Commercial application of marker-and gene-assisted selection in livestock: Strategies and lessons. American Society of Animal Science 82 (e.Suppl) E313-E328.

Dekkers, J.C.M. & Hospital, F., 2002. The use of molecular genetics in the improvement of agricultural populations. Nature Rev. Genet. 3, 22-32.

De Roos, A.P.W, Hayes, B.J., Spelman, R.J. & Goddard, M.E., 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics 179, 1503-1512.

Dodds, K.G., McEwan, J.C. & Davis, G.H., 2007. Integration of molecular and quantitative information in sheep and goat industry breeding programmes. Small Rumin. Res. 70, 32-41.

Dodgson., J.B., Cheng, H.H. & Okimoto, R., 1997. DNA marker technology: A revolution in animal genetics. Poult. Sci. 76, 1108-1114.

Dürr, J.W. & Jackobsen, J.H. 2009. Country Profiles Regarding the Use of Imported Dairy Bulls. InterBull Bulletin No. 40. Proceedings of the 2009 Interbull meeting Barcelona, Spain, August 21-24, 2009

Eck, S.H., Benet-Pages, A., Flisikowski, K., Meitinger, T., Fries, R. & Strom, T.M., 2009. Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery. Genome Biol. 10, 8 R82. pp. 1-8.

Eggen, A., 2012. The development and application of genomic selection as a new breeding paradigm. Animal Frontiers 2 (1), 10-15.

Fan, B., Du, Z-Q., Gorbach, D.M. & Rothschild, M.F., 2010. Development and application of high-density SNP arrays in genomic studies of domestic animals. Asian-Aust. J. Anim. Sci. 7, 833-847.

Fernando, R.L. & Grossman, M., 1989. Marker-assisted selection using best linear unbiased prediction. Genet. Sel. Evol. 21, 467-477.

Fore Jr., J., Wiechers, R. & Cook-Deegan, R., 2006. The effects of business practices, licensing and intellectual property on development and dissemination of the polymerase chain reaction: Case study. J. Biomed. Disc. Coll. 1-7.

Garrick, D.J., 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. Gen. Sel. Evol. 43, 17.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., Sargeant, L.S., Sørensen, A., Steele, M., Zhao, X., Womack, J.E. & Hoeschele, I., 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139, 907-920.

Gianola, D., Fernando, R.L. & Stella, A., 2006. Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173, 1761-1767.

Goddard, M.E., 2009. Genomic selection: prediction of accuracy and maximization of long term response. Genetics 136, 245-257.

Gredler, B., Nirea1, K.G., Solberg, T.R., Egger-Danner, C., Meuwissen, T.H.E. & Sölkner, J., 2009. A comparison of methods for genomic selection in Austrian dual-purpose Simmental cattle Proc. Assoc. Advmt. Anim. Breed. Genet. 18, 568-571.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M. & Snell, R., 2002. Positional candidate cloning of a QTL in dairy cattle identification of a missense mutation in the bovine DGAT1 gene with a major effect on milk yield and composition. Genome Res. 12, 222-231.

Habier, D., Fernando, R.L. & Dekkers, J.C.M., 2007. The impact of genetic relationship information on genome assisted breeding values. Genetics 177, 2389-2397.

Habier, D., Fernando, R.L., Kizilkaya, K. & Garrick, D.J., 2011. Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics 12, 186.

Harris, B.L. & Johnson, D.L., 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. J. Dairy Sci. 93, 1243-1252.

Hayes, B. & Goddard, M.E., 2010. Genome-wide association and genomic selection in animal breeding. Genome 53, 876-883.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. & Goddard, M.E., 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92, 433-443.

Heyen, D.W., Weller, J.I., Ron, M., Band, M., Beever, J.E., Feldmesser, E., Da, Y., Wiggans, G.R., Van Raden, P.M. & Lewin, H.A., 1999. A genome scan for QTL influencing milk production and health traits in dairy cattle. Am. Phys Soc. 165-175.

Humblot, P., Le Bourhis, D., Fritz, S., Colleau, J.J., Gonzalez, C., Joly, C.G., Malafosse, A., Heyman, Y., Amigues, Y., Tissier, M. & Ponsart, C., 2010. Reproductive technologies and genomic selection in cattle. Vet. Med. Int. Article ID 192787.

Ihara, N., Takasuga, A., Mizoshita, K., Takeda, H., Sugimoto, M., Mizoguchi, Y., Hirano, T., Itoh, T., Watanabe, T., Reed, K.M., Snelling, W.M., Kappes, S.M., Beattie, C.W., Bennett, G.L. & Sugimoto, Y., 2004. A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res. 14, 1987-1998.

InterBull Conference, Riga, Latvia, 31 May 2010, No. 42, 1 – 8 22315 A7.

Jeon, J.T., Lee, J.H., Kim, K.S., Park, C.K. & Oh, S.J., 2006. Application of DNA markers in animal industries. Aust. J. Exp. Agric. 46, 173-182.

Jorjani, H., Zumbach, B., Dürr, J. & Santus, E., 2011. Joint genomic evaluation of BSW populations. Proceedings of the InterBull International Workshop. Paris, France, March 4-5.

Kappes, S.M., Keele, J.W., Stone, R.T., McGraw, R.A., Sonstegard, T.S., Smith, T.P.L., Lopez-Corrales, N.L. & Beattie, C.W., 1997. A second generation linkage map of the bovine genome. Genome Res. 7, 235-249.

Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A.L., Barris, W., Schnabel, R.D., Taylor, J.F. & Raadsma, H.W., 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9, 187.

Lillehammer, M., Meuwissen, T.H.E. & Sonesson, A.K., 2011. A comparison of dairy cattle breeding designs that use genomic selection. J. Dairy Sci. 94, 493-500.

Maher, B., 2008. The case of the missing heritability. Nature 456, 18-21.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P.L., Sonstegard, T.S. & Van Tassell, C.P., 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS One, 2009. 4 (4), e5350.

McHugh, N., Meuwissen, T.H.E., Cromie, A.R. & Sonesson, A.K., 2011. Use of female information in dairy cattle genomic breeding programs. J. Dairy. Sci. 94, 4109-4018.

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

McParland, S., Kearney, F. & Berry, D.P., 2009. Purging of inbreeding depression within the Irish Holstein-Friesian population. Genet. Sel. Evol. 41, 16.

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

Meuwissen, T. & Goddard, M., 2010b. The use of family relationships and linkage disequilibrium to impute phase and missing genotypes in up to whole-genome sequence density genotypic data. Genetics 185, 1441-1449.

Meuwissen, T., Hayes, B. & Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819-1829.

Montgomery, G.W. & Kinghorn, B.P., 1997. Recent developments in gene mapping and progress towards marker-assisted selection in sheep. Aust. J. Agric. Res. 48, 729-741.

Mostert, B.E., 2007. The suitability of test-day models for genetic evaluation of dairy cattle in South Africa. PhD thesis, University of Pretoria, South Africa.

Nicholas, F.W., 2006. Discovery, validation and delivery of DNA markers. Aust. J. Exp. Agric. 46, 155-158.

Pedersen, L.D., Kargo, M., Berg, P., Voergaard, J., Buch, L.H. & Sørensen, A.C., 2012. Genomic selection strategies in dairy cattle breeding programmes: Sexed semen cannot replace multiple ovulation and embryo transfer as superior reproductive technology. J. Anim. Breed. Genet. 129, 152-163.

Pollak, E.J., 2005. Application and impact of new genetic technologies on beef cattle breeding: a "real world" perspective. Aust. J. Exp. Agr. 45, 739-748.

Pryce, J.E., Goddard, M.E., Raadsma, H.W. & Hayes, B.J., 2010. Deterministic models of breeding scheme designs that incorporate genomic selection. J. Dairy Sci. 93, 5455-5466.

Pryce, J.E., Arias, J., Bowman, P.J., Davis, S.R., MacDonald, K.A., Waghorn, G.C., Wales, W.J., Williams, Y.J., Spelman, R.J. & Hayes, B.J., 2012. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. J. Dairy Sci. 95, 2108-2119.

Pszczola, M., Strabel, T., Wolc, A., Mucha, S. & Szydlowski, M., 2011. Comparison of analyses of the QTLMAS XIV common dataset. I: genomic selection. BMC Proceedings 2011, 5 (Suppl 3), S1.

Rendel, J.M. & Robertson, A., 1950. Selection within a closed herd of dairy cattle. J. Genet. 50, 1-8.

Robinson, J.L., Dombrowski, D.B., Harpestad, G.W. & Shanks, R.D., 1984. Detection and prevalence of UMP synthase deficiency among dairy cattle. J. Hered. 75, 277-280.

Ron, M. & Weller, J.I., 2007. From QTL to QTN identification in livestock--winning by points rather than knock-out: a review. Anim. Genet. 38, 429-439.

Sargolzaei, M., Chesnais, J.P. & Schenkel, F.S., 2011. FImpute. An efficient imputation algorithm for dairy cattle populations. J. Anim. Sci. 89, 421.

Schaeffer, L.R., 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77, 2671-2678.

Schaeffer, L.R., 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123, 218-223.

Scholtz, M.M., 2010. Beef Breeding in South Africa. 2nd ed. Publ.: Agricultural Research Council, Pretoria, South Africa.

Scholtz, M.M. & Grobler, S.M., 2009. A systems approach to the South African dairy industry. S. Afr. J. Anim. Sci. 39, 116-120.

Schrooten, C., Bovenhuis, H., Coppieters, W. & Van Arendonk, J.A.M., 2000. Whole genome scan to detect quantitative loci for conformation and functional traits in dairy cattle. J. Dairy Sci. 83, 795-806.

Schuster, D.E., Kehrli Jr., M.E., Ackermann, M.R. & Gilbert, R.O., 1992. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. Proc. Natl. Acad. Sci. 89, 9225-9229.

Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M. & Visscher, P.M., 2002. QTL express: mapping quantitative loci in simple and complex pedigrees. Bioinformatics 18, 339-340.

Seaton, G., Hernandez, J., Grunchec, J.A., White, I., Allen, J., De Koning, D.J., Wei, W., Berry, D., Haley, C. & Knott, S., 2006. GridQTL: A grid portal for QTL mapping of compute intensive datasets. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, August, 13-18, 2006, Belo Horizonte, Brazil.

Seidel, G.E., 2010. Brief introduction to whole-genome selection in cattle using single nucleotide polymorphisms. Reprod. Fert. Develop. 22, 138-144.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 17, 16.

Thomsen, B., Horn, P., Panitz, F., Bendixen, E., Petersen, A.H., Holm, L., Nielsen, V.H., Agerholm, J.S., Arnbjerg, J. & Bendixen, C., 2006. A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. Genome Res. 16, 97-105.

Van der Werf, J.H.J., 2007. Marker assisted selection in sheep and goats. In: Marker-Assisted Selection: Current status and future perspectives in crops, livestock, forestry and fish. Eds. Guimarães, E., Ruane, J., Scherf, B., Sonnino, A. & Dargie, J. FAO, Rome, Italy. pp. 230-247.

Van Marle-Köster, E. & Nel, L.H., 2003. Genetic markers and their application in livestock breeding in South Africa: A review. S. Afr. J. Anim. Sci. 33, 1-10.

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

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

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

Van Tassell, C.P., Smith, T.P.L., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C. & Sonstegard, T.S., 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nature Methods 5, 247-252.

Veerkamp, R.F. & Beerda, B., 2007. Genetics and genomics to improve fertility in high producing dairy cows. Therio. 68S, S266-S273.

Veerkamp, R.F., Coffey, M.P., Berry, D.P., De Haas, Y., Strandberg, E., Bovenhuis, H., Calus, M.P.L. & Wall, E., 2012. Genome-wide associations for feed utilisation complex in primiparous Holstein-Friesian dairy cows from experimental research herds in four European countries. Animal 6, 1738-1749

Visscher, P.M., 2008. Sizing up human height variation. Nat. Genet. 40, 489-490.

Weigel, K.A., Van Tassell, C.P., O'Connell, J.R., VanRaden, P.M. & Wiggans, G.R., 2010. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using reference panels and population-based imputation algorithms. J. Dairy Sci. 93, 2229-2238.

Wickham, B.W., Amer, P.R., Berry, D.P., Burke, M., Coughlan, S., Cromie, A., Kearney, J.F., McHugh, N., McParland, S. & O'Connell, K., 2012. Industrial perspective: capturing the benefits of genomics to Irish cattle breeding. Anim. Prod. Sci. 52, 172-179.

Wiggans, G.R., Sonstegard, T.S., VanRaden, P.M., Matukumalli, L.K., Schnabel, R.D., Taylor, J.F., Schenkel, F.S. & Van Tassell, C.P., 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. J. Dairy Sci. 92, 3431-3436.

Wiggans, G.R., VanRaden, P.M., Bacheller, L.R., Tooker, M.E., Hutchison, J.L., Cooper, T.A. & Sonstegard, T.S., 2010. Selection and management of DNA markers for use in genomic evaluations. J. Dairy Sci. 93, 2287-2292.

Wiggans, G.R., Cooper, T.A., VanRaden, P.M. & Cole, J.B., 2011. Adjustment of traditional cow evaluations to improve the accuracy of genomic predictions. J. Dairy Sci. 94, 6188-6193.

Womack, J.E., 2005. Advances in livestock genomics: Opening the barn door. Genome Res. 15, 1699-1705.

Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., Pfister-Genskow, M., White, L.A., Grignola, F., Uimari, P., Thaller, G. & Bishop, M.D., 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics 149, 1959-1973.