Annals of Botany **110**: 1303–1316, 2012
doi:10.1093/aob/mcs109, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

REVIEW: PART OF A HIGHLIGHT ON BREEDING STRATEGIES FOR FORAGE AND GRASS IMPROVEMENT

Will genomic selection be a practical method for plant breeding?

Akihiro Nakaya[1] and Sachiko N. Isobe[2,*]

[1]*Center for Transdisciplinary Research, Niigata University, 1-757 Asahimachi-dori, Chuo-ku, Niigata 951-8585, Japan and* [2]*Kazusa DNA Research Institute, 2-6-7 Kazusa Kamatari, Kisarazu, Chiba 292-0818, Japan*
**For correspondence. E-mail: sisobe@kazusa.or.jp*

Received: 12 January 2012 Returned for revision: 9 March 2012 Accepted: 11 April 2012 Published electronically: 29 May 2012

• *Background* Genomic selection or genome-wide selection (GS) has been highlighted as a new approach for marker-assisted selection (MAS) in recent years. GS is a form of MAS that selects favourable individuals based on genomic estimated breeding values. Previous studies have suggested the utility of GS, especially for capturing small-effect quantitative trait loci, but GS has not become a popular methodology in the field of plant breeding, possibly because there is insufficient information available on GS for practical use.
• *Scope* In this review, GS is discussed from a practical breeding viewpoint. Statistical approaches employed in GS are briefly described, before the recent progress in GS studies is surveyed. GS practices in plant breeding are then reviewed before future prospects are discussed.
• *Conclusions* Statistical concepts used in GS are discussed with genetic models and variance decomposition, heritability, breeding value and linear model. Recent progress in GS studies is reviewed with a focus on empirical studies. For the practice of GS in plant breeding, several specific points are discussed including linkage disequilibrium, feature of populations and genotyped markers and breeding scheme. Currently, GS is not perfect, but it is a potent, attractive and valuable approach for plant breeding. This method will be integrated into many practical breeding programmes in the near future with further advances and the maturing of its theory.

Key words: Genomic selection, plant breeding, marker assisted selection, genetic model, linkage disequilibrium.

INTRODUCTION

Genomic selection or genome-wide selection (GS) has been highlighted as a new approach for marker-assisted selection (MAS) in recent years. GS is a form of MAS that selects favourable individuals based on genomic estimated breeding values (GEBVs). Breeding values have not been a popular index in plant breeding, although they are frequently used in animal breeding. They are defined as 'the sum of the estimate of genetic deviation and the weighted sum of estimates of breed effects' (Van Vleck *et al.*, 1992), which are predicted using phenotypic data from family pedigrees based on the additive infinitesimal model (Fisher 1918). Several statistical approaches have been proposed for the prediction of estimated breeding values (EBVs), such as best linear unbiased prediction (BLUP) (Henderson, 1975) and a Bayesian framework (Gianola and Fernando 1986). Furthermore, an innovative method for predicting breeding values was proposed based on genome-wide dense DNA markers, known as the GEBV (Meuwissen *et al.*, 2001). When the idea of GEBV was proposed, it was regarded as an unrealistic approach because of the lack of large-scale genotyping technologies at the period. However, it has become a feasible approach with recent advances in high-throughput genotyping platforms. The term 'GS' was first introduced by Haley and Visscher at the 6th World Congress on Genetics Applied to Livestock Production at Armidale, Australia in 1998 according to Meuwissen (2007), although it was not used in the main text of Meuwissen *et al.* (2001). However, the overall MAS programme using GEBV was later referred to as GS.

The general processes of GS and traditional MAS used for quantitative traits (QTs) are shown in Fig. 1. The main frameworks of the two approaches are similar, where both GS and traditional MAS consist of training and breeding phases. In the training phase, phenotypes and genome-wide (GW) genotypes are investigated in a subset of a population, i.e. the training population in GS and the mapping population in traditional MAS. Within populations, significant relationships between phenotypes and genotypes are predicted using statistical approaches. In the breeding phase, genotype data are obtained in a breeding population, before favourable individuals are selected based on the genotype data obtained. Three obvious differences between the two approaches are apparent: (1) in the training phase, quantitative trait loci (QTLs) are identified in traditional MAS while formulae for GEBV prediction are generated in GS, known as GS models; (2) in the breeding phase, genotype data are only required for targeted regions in traditional MAS, whereas GW genotype data are considered to be necessary in GS; (3) in the breeding phase, favourable individuals are selected based on the genotypes of markers in MAS, whereas GEBVs are used for selection in GS. Thus, GS jointly analyses all the genetic variance of each individual by summing the marker effects of GEBV (Heffner *et al.*, 2009), and it is expected to address small effect genes that cannot be captured by traditional MAS (Hayes *et al.*, 2009).

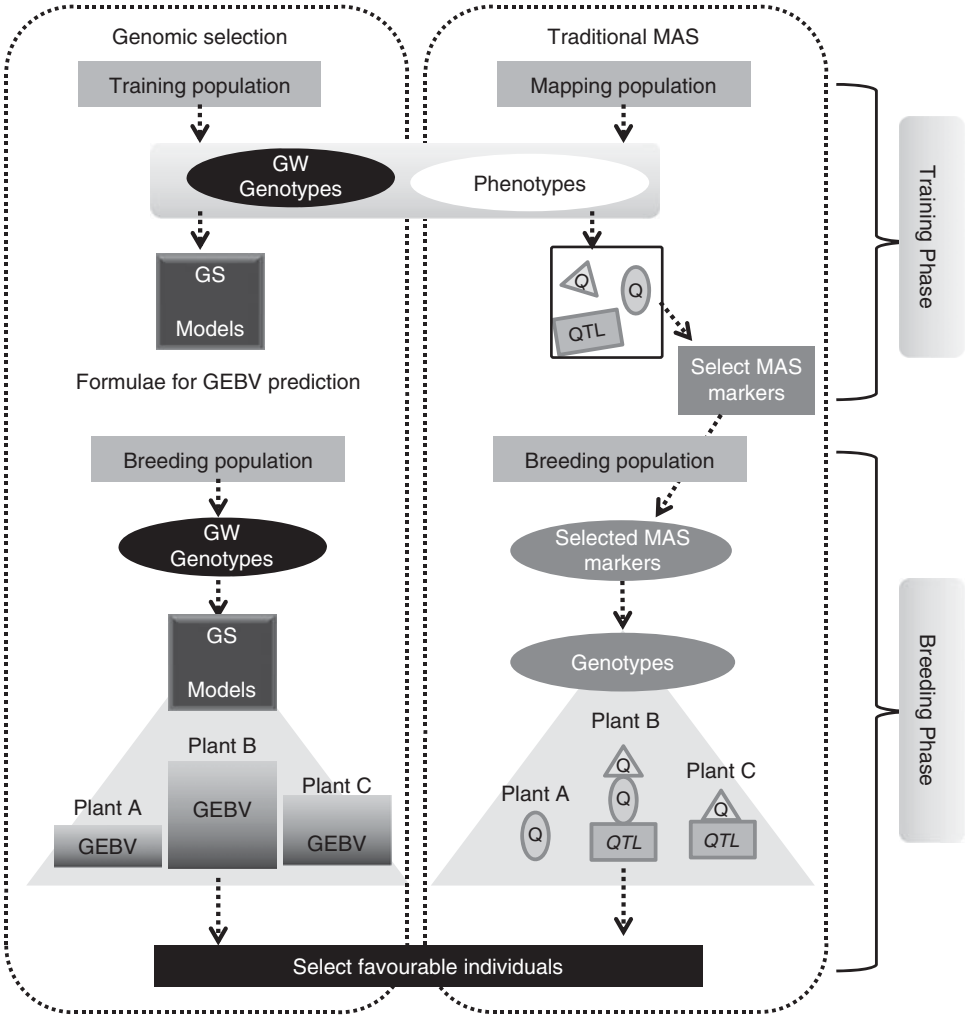


FIG. 1. Schemes of genomic selection (GS) (left) and traditional MAS for the selection of quantitative traits (right). Both GS and traditional MAS contained training and breeding phases. In the training phase, quantitative trait loci (QTLs) are identified in traditional MAS to produce formulae for genomic estimated breeding value (GEBV) prediction, i.e. GS models. In the breeding phase, favourable individuals are selected based on the genotypes of the selected markers in MAS, whereas GEBVs are used for selection in GS.

Since GS was first propounded by Meuwissen *et al.* (2001), many reports have indicated the usability of GS for breeding for QTs. However, GS has still not become a popular methodology in the field of plant breeding. We consider that a major obstacle is the availability of insufficient knowledge of GS for practical use. Indeed, most fields of GS studies have dealt with statistics and simulation that are discussed in terms of formulae, which are often too specific for breeders and molecular biologists to understand. To initiate further discussions on the applicability of GS in plant breeding, here our aim is to discuss GS from a practical breeding viewpoint. First, the statistical approaches used in GS are briefly explained to understand the essence of this approach. Second, we survey recent progress in GS studies from the areas of animal and plant science, mainly addressing those dealing with empirical data. Third, we describe several specific factors that require careful consideration before practicing GS in plant breeding. Finally, we discuss future prospects for the further advancement of GS and MAS programmes overall.

STATISTICAL CONCEPTS USED IN GS

All GS, traditional MAS and pedigree-based phenotypic selection (PS) methods are reliant on a common selection framework, i.e. finding a causal relationship between genetic factors and target traits based on putative genetic factors underlying the phenotypic distribution (in PS) or observed marker genotypes (in GS and traditional MAS) in a training population. Before describing the statistical approaches used for GEBV prediction, we briefly review the general statistical concepts that are commonly used in PS, traditional MAS and GS.

Genetic models and variance decomposition

A genetic model of QTs is generally constructed based on an assumption that only effects caused by genetic factors are inherited across the generations. A simple but frequently used genetic model is that the phenotypic value of an individual (P) is expressed as the summation of the genetic value (G)

and the residual environmental effect (E):

$$P = G + E$$

where the genetic value G includes additive genetic effect, dominance effect and epistasis. If we suppose that there is no correlation between G and E (i.e. no $G \times E$ effect), the covariance between G and E can be set at zero [$\mathrm{Cov}(G, E) = 0$]. The phenotypic variance, $V(P)$, is then expressed as the summation of the genetic variance, $V(G)$, and the environmental variance, $V(E)$:

$$V(P) = V(G) + V(E) + 2\mathrm{Cov}(G, E) = V(G) + V(E).$$

Heritability

Heritability is a measure for evaluating the degree to which the phenotypic characteristics of a population are inherited to the next generation, and it is represented as the ratio of genetic variance to phenotypic variance. Broad sense heritability (H^2) focuses on the total genetic effects, G, including the additive, dominance and epistatic effects, whereas narrow sense heritability (h^2) counts only additive genetic effects. Therefore, for h^2, the genetic model ($P = G + E$) can be rewritten using the additive genetic effect, A:

$$P = A + E'$$

Here, E' represents the residual effects that are not included in the additive genetic effect, A. Note that the dominant and epistatic effects are in E'. If we suppose that there is no correlation between A and E', then the phenotypic variance $V(P)$ can be broken down into the additive genetic variance, $V(A)$, and the residual effects variance, $V(E)$:

$$V(P) = V(A) + V(E').$$

Because h^2 is defined by the ratio of $V(A)$ to $V(P)$, it is represented as follows:

$$h^2 = V(A)/V(P).$$

In GS, $V(A)$ is broken down again into the variances explained by multiple DNA markers, $V(A_1)$, $V(A_2)$, …, $V(A_M)$ under the assumption that DNA markers are not correlated with each other (Meuwissen *et al.*, 2001).

Breeding value (BV)

The BV of an individual i in a population is defined as follows based on the narrow sense heritability, h^2:

$$BV_i = m_0 + h^2(y_i - m_0) = m_0 + (y_i - m_0)V(A)/V(P).$$

Here, y_i is the phenotypic value of individual i, while m_0 is the mean phenotypic value of the population. Because $V(A)$ cannot be directly observed, h^2 has been conventionally estimated by a comparison of the phenotypic values of parents and their offspring. The BVs that are predicted based on an estimated heritability are known as EBVs. By contrast, phenotypic value y_i and the $V(A)$ in GS are estimated based on the flux of the genotype effects of GW markers. Thus, the BV predicted in GS is known as the genomic EBV (GEBV). Note the residual effect variance $V(E')$ is ignored in BV prediction, because narrow sense heritability is employed.

Linear model for marker effects

In many implementations of GS, the causal relationship between the phenotype and genotype is represented as a linear model or its extension, which is then used to infer the GEBV of an individual in a breeding population. Thus, the linear model is a fundamental model employed in GS. Here, we assume there are N individuals and M bi-allelic markers in the training population, and we focus on one of the markers. Let (y_i, x_{1i}) denote the pair of the observed phenotype and genotype of the marker of the ith individual, i.e. (y_1, x_{11}), (y_2, x_{12}), …,(y_N, x_{1N}). In addition, let us suppose that the bi-allelic genotypes are encoded by 0 and 1, respectively, and that the phenotypes of the N individuals are distributed as shown in Fig. 2. Because an individual gains additional phenotypic value, β_1, depending on its marker genotype, the phenotype can be modelled as follows:

$$y_i = \beta_0 + x_{1i}\beta_1 + \varepsilon_i \quad (1)$$

where β_0 and β_1 are the parameters to be determined, and ε_i is an error term that is usually assumed to have a normal distribution with a mean of zero. This model is represented as a linear combination of the terms, known as a 'linear model', showing that the phenotypes of individuals with genotypes 0 and 1 are normally distributed around β_0 and $\beta_0 + \beta_1$, respectively. The parameters of a linear model may be determined by least-squares estimation, such that the summation of ε_i^2, i.e. an error function $E = \Sigma_i\,(y_i - \beta_0 - x_{1i}\,\beta_1)^2$, is minimized and the line is fitted to the phenotype. The linear model (1) represents the relationship between the genotype and phenotype for a single marker, but it can be extended to include all the M

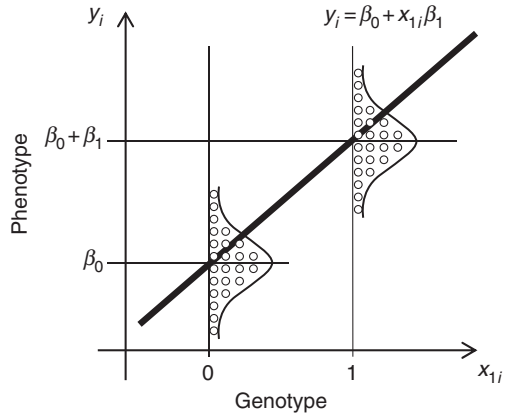


FIG. 2. Relationships between marker genotypes (x_{1i} : 0 and 1) and phenotypes (y_i) of the individuals (open circles) in a training population. If the marker genotype is correlated with the phenotype, segregation is modelled using the bold line ($y_i = \beta_0 + x_{1i}\,\beta_1$, where β_0 and β_1 are parameters to be determined.).

markers as follows:

$$y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 \ldots + x_{Mi}\beta_M + \varepsilon_i$$
$$= \Sigma_{j=0}^{M} x_{ji}\beta_j + \varepsilon_i \quad (2)$$

where x_{ji} is the genotype of the jth marker in the ith individual, coefficient β_j is its effect on the phenotype, and $x_{0i} = 1$ is a dummy variable. Similarly, the coefficients are determined by minimizing an error function, $E = \Sigma_{i=1}^{N} (y_i - \sum_{j=0}^{M} x_{ji}\beta_j)^2$.

Because GW genotype data are used in GS, a problem often arises, known as 'large p, small n problem', when the linear model (2) is employed for GEBV prediction (p and n are the numbers of markers and individuals, respectively). That is, a linear model that consists of p markers is too complicated for the prediction of BVs of n individuals. Thus, it can cause over-fitting and the linear model only works well in the training population. To avoid over-fitting, a penalty term is introduced in the error function, i.e. $E = \Sigma_{i=1}^{N} (y_i - \Sigma_{j=0}^{M} x_{ji}\beta_j)^2 + \lambda\Sigma_{j=0}^{M} |\beta_j|^q$, where λ is a parameter that controls the effects of the penalty term. Note that setting a high β_j inhibits the minimization of the error function. Setting $q = 1$ and $q = 2$ are known as LASSO (least absolute shrinkage and selection operator) and ridge regression (RR in Table 2), respectively. Ridge regression forces all the coefficients to shrink toward zero equally, while LASSO can set several coefficients that are unrelated to the phenotype to zero. Therefore, if the phenotype is controlled by many markers with small effects, ridge regression will capture those effects (Heffner *et al.*, 2009), whereas LASSO will capture large effects with a small number of markers. If the coefficients of the markers are set to zero or a low value in the training phase, they are excluded from the model and their genotype information is not required during the breeding phase.

For example, 'least-square estimation' and 'BLUP estimation' for effects of markers or chromosome segments in Meuwissen *et al.* (2001) adopt similar linear models. Here, BLUP stands for best linear unbiased prediction of a parameter. As Heffner *et al.* (2009) summarizes, the methods using ridge regression assume that effects of markers have an equal variance. On the other hand, Bayesian methods that are known as BayesA and BayesB of Meuwissen *et al.* (2001) can make relaxed assumptions to estimate the variances of the effects of markers separately. In a Bayesian framework, effect of a marker is represented by distribution of a random variable that is determined by its prior distribution according to some assumptions. Actually, BayesA and BayesB adopt different prior distributions for the variance of the effects of markers; that of the latter is defined to allow a part of markers to have no effects on a phenotypic value. Although simultaneous evaluation of markers and no need for marker selection are advantageous characteristics of GS, decreasing the number of markers required in the breeding phase might be preferable from the economic viewpoint.

RECENT PROGRESS IN GS STUDIES

The most important factor determining the success of GS is the accurate prediction of GEBVs. The accuracy of the predicted GEBVs is often estimated based on the correlation between the observed phenotypic value and GEBVs. To produce accurate GEBVs, several studies have applied comparative statistical approaches to GEBV prediction. In addition, simulations studies have been widely used to investigate the affect of the number of QTLs, markers, individuals and other variables. These studies were reviewed recently by Heffner *et al.* (2009) and Jannink *et al.* (2010), and so are not described further in this section. Instead, we focus on recent progresses in GS based on empirical data to understand better the practical use of GS.

Animal science

Studies of GS are more common in the field of animal science than plant science. The BV concept was used in animal breeding long before the emergence of GS, so the GS approach was more readily accepted by animal scientists. In addition, the lower diversity of the targeted species and fewer effects of environmental factors during the growing stage might have contributed to the rapid introduction of GS in animal science. The first empirical GS study in animal science was reported by Legara *et al.* (2008) using mice (Table 1). A total of 1884 individuals were generated from eight inbred lines and genotyped using 10 946 single nucleotide polymorphism (SNP) markers, before predicting the GEBVs for four traits related to body sizes. A comparison of the predictive ability and accuracy of GEBVs generated with or without SNP genotypes and polygenetic effects demonstrated that GW genetic evaluation and selection provided better accuracy and predictive ability than the classical polygenic model.

The most advanced progress in GS has been observed in dairy cattle. In Table 1, the results of three GS studies in dairy cattle are summarized (Hayes *et al.*, 2009; Luan *et al.*, 2009; Van Raden *et al.*, 2009). In addition to the three reports in Table 1, seven empirical GS studies of dairy cattle were also reported and reviewed by Hayes *et al.* (2009), Moser *et al.* (2009) and Calus (2010). Of the three cattle studies in Table 1, a total of 500–5335 individuals were used for GEBV prediction using 18 991–38 416 SNPs. GEBVs for various QTs related to milk production, cattle body size and fertility were predicted using several different methods, where the accuracy of GEBVs ranged from 0·14 to 0·69. Rolf *et al.* (2010) and Mujibi *et al.* (2011) reported GEBV prediction in beef cattle. Parentally identified steers and sires of 2405 Angus cattle were genotyped using 41 028 SNPs in a study by Rolf *et al.* (2010), while an admixture population consisting of Angus, Charolais and hybrid bulls was genotyped using 37 959 SNPs for 721 individuals in a study by Mujibi *et al.* (2011). GEBVs for traits related to daily gain and daily intake were investigated, and the estimated accuracies ranged from −0·07 to 0·48. In chickens, Wolc *et al.* (2011) tested 16 traits related to eggs and chicken body size with 23 356 SNP genotypes using 2708 individuals derived from a single blown egg-layer line. The accuracy of GEBVs estimated ranged from 0·2 to 0·7. In addition, Calenge *et al.* (2011) reported GS studies on *Salmonella* carrier-state resistance in chickens (not shown in Table 1).

TABLE 1. *Features of test populations, number of genotyped SNPs and ranges of GEBV accuracy in empirical animal GS studies*

Species	Population type	Size of test population	Training: validating[a]	No. of genotyped SNPs	Accuracy of GEBVs[b]	Models for GEBV prediction[c]	Traits	Reference
Mice	A heterogeneous population derived from eight inbred lines	1884	942:942	10 946	0·16–0·25[f], 0·27–0·67[g]	Linear mixed model with SNP genotypes, not polygenetic effects	Weight, growth slope, body length, body mass index	Legarra *et al.* (2008)
Dairy cattle; Australian Holstein-Friesian	Bull progeny tested by Genetics Australia	730	Bulls born in 1998–2002 : bulls born in 2003	38 259	0·14–0·55	BayesA	Breeding value, profit ranking, selection value, protein yield and protein percentage	Hayes *et al.* (2009)
Dairy cattle; Norwegian Red	34 sires and 466 sons	500	100: 400[d], 34–100 : 400–466[e]	18 991	0·20–0·61[d], 0·15–0·62[e]	G-BLUP	Milk yield, fat yield, protein yield, clinical mastitis, calving ease	Luan *et al.* (2009)
Dairy cattle; North American Holstein	American Holstein bulls born between 1952 and 2002	5335	3576 : 1759	38 416	0·33–0·69	Linear mixed model	27 traits for milk production, body size, shape and fertility	Van Raden *et al.* (2009)
Beef cattle; Angus	Parental identified steers and sires	2405	85–2405 : 84–2405	41 028	0·23–0·44	Genomic relationship matrices	Average daily feed intake, residual feed intake, average daily gain	Rolf *et al.* (2010)
Beef cattle; Angus, Charolais, University of Alberta hybrid bulls	Admixture population	721	198–203 : 721	37 959	−0·07–0·48	RR-BLUP (with top 200 SNPs)	Average daily gain, dry matter intake, residual feed intake	Mujibi *et al.* (2011)
Chicken; blown-egg layer line	Five consecutive generations in a single line	2708	768–2167 : 274–289	23 356	0·2–0·7[h]	G-BLUP and Bayes-C-π	13 traits for eggs and 3 traits for chicken bodies	Wolc *et al.* (2011)

Only studies that investigated the accuracy of GEBVs based on the correlation between observed phenotypic values and GEBVs are listed.

[a] Number of individuals used for GEBV prediction (training population) versus that used for validation (validating population).

[b] Correlation between observed phenotypic values and GEBVs.

[c] Models with the highest or higher accuracy of GEBVs when multiple methods were used for GEBV prediction. G-BLUP, Best linear unbiased prediction; RR-BLUP, random regression best linear unbiased prediction.

[d] Random masking.

[e] Cohort masking.

[f] Across families.

[g] Within families.

[h] Data cited from Figs 1 and 2.

The populations used in the empirical studies mentioned above were usually divided into two, i.e. training and validating populations. Training populations were used to develop GS models based on genotypic and phenotypic data, whereas the validating populations were used for investigating the GEBV accuracy by estimating the correlation between the GEBVs predicted by the GS models and the observed phenotypic values. Validation is not theoretically essential for a GS scheme (Fig. 1), although it is practically important to confirm the adequacy of a GS model before moving onto the breeding phase. Of the seven studies listed in Table 1, five considered pedigree relationships when the populations were divided into training and validating populations (Hayes *et al.*, 2009; Luan *et al.*, 2009; Van Raden *et al.*, 2010; Rolf *et al.*, 2010; Wolc *et al.*, 2011). Thus, these studies reflected the entire GS process better compared with the others, because the breeding phases in GS were demonstrated virtually by the verification of GS models using the progeny of the training populations.

The reported studies used different materials and statistical methods for GEBV prediction, but many of these studies showed that the accuracy of GEBV was higher than that of traditional EBV and it was increased with a larger population size, larger numbers of genotyped SNPs, and higher heritability of the targeted traits. The details are not described here, but some of the studies compared different statistical methods for GEBV prediction. Note that the best approaches with the highest accuracy of GEBVs were different in each case (Table 1). The accuracy of GEBVs estimated in empirical studies fell below 0·7 (Table 1), which was lower than that suggested by many simulation studies such as 0·85 in Meuwissen *et al.* (2001). Calus (2010) indicated that the distribution of QTL effects in real data is generally lower than that assumed in simulation studies. If this is true, the lower accuracy estimated by real data might be affected by a lower number of QTL effects as well as other factors, such as the non-additive effects of QTLs and environmental factors.

Plant science

Plant breeding targets a diversity of species with different reproduction systems, generation times, genome structures and utilized organs. Thus, various methods are used in conventional breeding, i.e. PS and traditional MAS, to adapt to the demands of different targeted species and breeding objectives. Like conventional breeding, GS should be adapted to the fit different types of plant species and breeding objectives.

Reports on plant species that specified 'genomic selection' or 'genomewide selection' have been published since 2007. Piyasatin *et al.* (2007) simulated the efficiency of GS in a cross of inbred lines, which is common in plant breeding but not in animal breeding. However, no specific plant species was considered as the targeted species in this paper. Simulation studies of specific species were firstly published for maize (Bernardo and Yu 2007), where a comparison between GS and marker-assisted recurrent selection (MARS) was demonstrated for three cycles of selection of doubled haploid lines (DHLs). The response of GS was 18–43 % greater than that of MARS with different numbers of QTLs (20, 40 and 100). Moreover, simulation studies using maize were performed to determine the advantages of using DHLs compared with F_2 populations in GS and MARS (Mayor and Bernardo, 2009), and to develop a methodology for the rapid introgression of exotic germplasms in an adapted line of maize via GS (Bernardo, 2009). In addition to maize, two GS simulations were performed with the oil palm, which is an outcrossing species that requires 19 years for one cycle of (PS) (Wong and Bernardo, 2008), and with a self-pollinated crop, barley (Bernardo, 2010).

While these studies simulated biparental cross populations, three studies also reported GS simulation using multiple inbred lines in barley based on real genotype data obtained mainly from SNPs and diversity array technology (DArT) (Zhong *et al.*, 2009; Jannink, 2010; Iwata and Jannink, 2011). Zhong *et al.* (2009) compared the accuracy of four GS prediction methods that were affected by marker density, level of linkage disequilibrium (LD), QTL number, and sample size, where the level of replication in populations was generated using 42 multiple inbred lines of two-row spring barley with the genotypes of 1933 loci obtained from SNP, DArT and classical markers. They concluded that the GS prediction method with the highest accuracy changed with different levels of LD between the marker and QTLs, QTL effects, and generations of individuals. Moreover, Iwata and Jannink (2011) simulated the accuracy of GS using more large-scale data, consisting of 1325 SNPs in 863 breeding lines of barley derived from nine breeding programmes in the USA. Seven methods were used for GEBV prediction and the mean of the predictions in all methods was more accurate than predictions based on any single method under medium and high heritability. Jannink (2010) simulated the dynamics of long-term GS using 192 breeding lines from an elite six-row spring barley programme with genotypes identified by 983 polymorphic markers. The results suggested that losing favourable alleles with weak LD with markers during selection cycles was inevitable, while placing additional weight on low-frequency favourable alleles was important for long-term GS.

Investigations of the accuracy of GEBV predictions using empirical data have been reported for maize, barley, wheat and *Arabidopsis thaliana* (Table 2). It was first demonstrated by Lorenzana and Bernardo (2009) for maize, *A. thaliana* and barley. All the test populations were generated from biparental crosses where the number of test progeny and markers ranged from 119 to 415 and 69 to 1339, respectively. *Arabidopsis thaliana* had the highest accuracy of GEBVs, although the number of polymorphic markers used for genotyping was the lowest. This study was followed by demonstrations of GS using empirical data in maize by Piepho (2009), Crossa *et al.* (2010) and Guo *et al.* (2011), as shown in Table 2. Piepho (2009) compared the performance of nine models using a series of experiments with DHLs derived from a single cross conducted in five environments, and suggested the need to appropriately model genotype–environment interactions and to employ an independent estimate of error. Crossa *et al.* (2010) demonstrated GS using a genetically diverse population [300 lines bred in CIMMYT (The International Maize and Wheat Improvement Center)] and 1148 SNPs, with a predicted accuracy of GEBVs ranging from 0·42 to 0·79 by ridge regression BLUP. The largest-scale analysis of maize was performed by Guo *et al.* (2011), which

used 4699 progeny derived from 25 nested association mapping populations with genotypes for 1106 SNPs. While a common line, 'B73', was used as the maternal line across the 25 mapping populations, the paternal lines were all different. Interestingly, the accuracy of the predicted GEBVs was different in the 25 crosses, although the study used almost the same SNPs, targeted traits and population sizes.

GS studies using empirical data from wheat were first reported by Crossa *et al.* (2010) using 1279 DArT genotypes and 599 wheat lines bred in CIMMYT. The targeted trait was grain yield and GEBVs predicted by reproducing kernel Hilbert spaced regression ranged from 0·48 to 0·61. In addition, Heffner *et al.* (2011) reported empirical results for wheat using 209 (CC population) and 174 (FKQ population) progeny of DHLs of biparental crosses with 399 and 574 polymorphic genotypes, respectively. The accuracy of GEBVs in the CC and FKQ populations ranged from 0·32 to 0·84 and 0·41 to 0·73, respectively (RR-BLUP, sample size was 96).

GS of perennial crops is considered to be more effective than annual crops because of their long generation times. GEBV predictions based on empirical data were presented for Loblolly pine and eucalyptus at the IUFRO Tree Biotechnology Conference 2011 (Table 2; Grattapaglia *et al.*, 2011; Isik *et al.*, 2011; Resende *et al.*, 2011). All cases used full-sib families as test populations and the number of individuals ranged from 149 to 920. In the two studies of Loblolly pine, 3406–3938 SNP markers were used for genotyping, while 3120–3564 DArT markers were used in the study of eucalyptus. The GEBV accuracy of all studies ranged from 0·3 to 0·77.

Interestingly, the ranges of accuracies in empirical studies were higher in plant studies than animal studies, although most plant studies employed lower numbers of genotyping markers. This might be due to the lower genetic diversity caused by a small number of parental lines and a greater bottleneck in the breeding materials. Note that the numbers of markers used for woody species was higher than that used for annual plant species. Empirical plant GS studies show that GS is a potential method for plant breeding and that it can be performed with realistic sizes of populations and markers when the populations used are carefully chosen.

THE PRACTICE OF GS IN PLANT BREEDING

Linkage disequilibrium (LD)

LD has a major affect on the operability of GS, so it has to be well understood before performing GS. LD is defined as the non-random association of alleles at different loci (Williams and Cummings, 1997). The intensity of LD between two loci is measured based on the frequency of alleles, using indexes such as D, D' and r^2, and it ranges from completely random ($|D| = |D'| = r^2 = 0$) to complete LD ($|D| = 0{\cdot}25$, $|D'| = r^2 = 1$) (Gaut and Long, 2003). The LD intensity decays with greater distance between two markers. Although it is difficult to delineate, a significant LD intensity is commonly considered to be $r^2 > 0{\cdot}1$ (Remington *et al.*, 2001; Garris *et al.*, 2003; Palaisa *et al.*, 2003). In general, the distance between two markers with significant LD intensity ($r^2 > 0{\cdot}1$) is found to be greater in outcrossing species than selfing species, although it varies with different species, population structure and genome regions (Gupta *et al.*, 2005). For example, observed marker intervals with significant LD intensity in outcrossing species are reported to be 100–150 bp in Loblolly pine, >500 bp in grape and 0·4–7·0 kbp in maize, whereas those in selfing species are >50 kbp in soybean, 100 kbp in rice and 250 kbp in *A. thaliana* (reviewed by Gupta *et al.*, 2005).

The number of markers required for GS modelling is determined based on the marker interval with a significant LD intensity in targeted populations. In a case of Loblolly pine, the genome size exceeds 20 Gbp (Wakamiya *et al.*, 1993) and the marker interval with a significant LD intensity was between 100 and 150 bp (González-Martínez *et al.*, 2004) in 435 unrelated individuals. If the 435 individuals were used for GS modelling, the number of markers required would be at least 200 M (20 Gbp per 100 bp). However, significant GEBVs with 0·3–0·83 accuracy were obtained using 3406–3938 SNPs in full-sib families with Loblolly pine (Table 2; Isik *et al.*, 2011; Resende *et al.*, 2011). This large disparity in the number of required markers is caused by the different length in the marker interval with a significant LD intensity in an unrelated mapping population and full-sib families. In other words, employing a population that originated from a few parental lines is effective in reducing the number of markers required, especially for species whose LD intensities decay rapidly among unrelated individuals (see Fig. 3).

Relationship between training and breeding populations

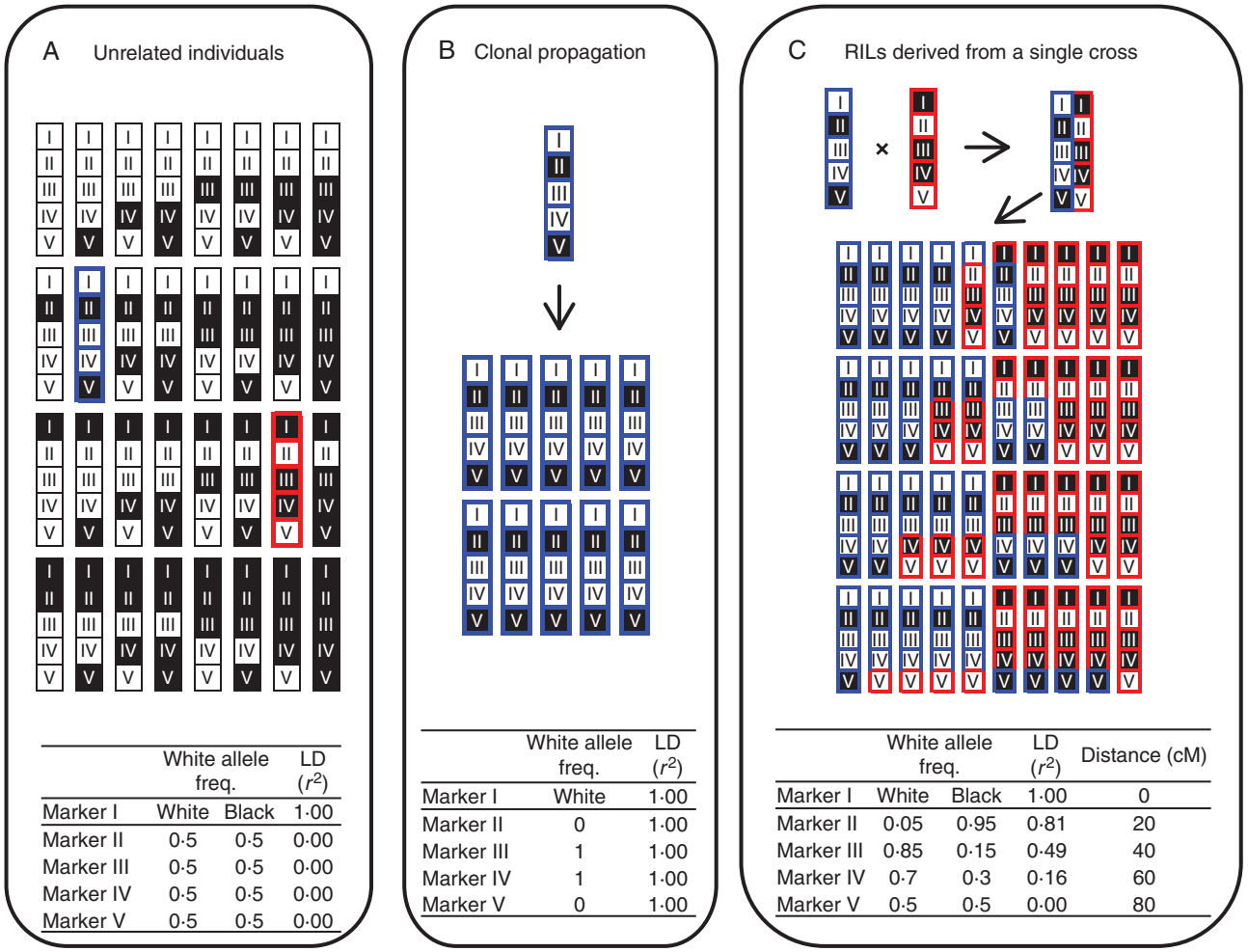
In traditional MAS, a marker that is confirmed to have tight linkage with a target QTL or gene can be used as a selection marker in most breeding populations of that species. Therefore, breeders have not had to seriously consider the relationship between mapping populations and breeding populations. However, in GS, the relationship between training and breeding populations must be carefully considered with the single exception of a marker set where adjacent markers have significant LD intensities across unrelated individuals in a pool of breeding materials genotyped for the training populations.

Suppose that two pairs of lines used for biparental crosses are selected from a pool of breeding materials (Fig. 4). The genotypes of the flanking markers (II and IV) of a targeted gene/allele (yellow-coloured G) are 'white' in cross 1, while those in cross 2 are 'black'. This indicates that allele types with significant LD with the targeted genes are not kept across different crosses. When this happens in traditional MAS, we usually have to explore the markers nearest to the targeted genes to avoid false positive selection. However, because GW markers are used in GS, it is almost impossible or meaningless to explore the nearest markers to each GW marker. Thus, establishing a GS model based on a training population does not work in a breeding population if the genetic structures of both populations are different, except for the case described in the preceding paragraph. Indeed, in most reported GS studies, the training populations were assumed to consist of ancestors or randomly selected individuals in a breeding population. Harris *et al.* (2008) reported that SNP estimates calculated from a Holstein–Friesian

TABLE 2. *Features of test populations, number of genotyped loci, and ranges of GEBV accuracy investigated in empirical plant GS studies*

Species	Population type	Size of population used	Training population ratio*	No. of genotyped markers†	Accuracy of GEBVs‡	Models for GEBV prediction¶	Traits	Reference
Maize	RILs derived from single cross	223	0·43, 0·65, 0·80	1339 SSRs and RFLPs	0·48–0·73	BLUP	8 morphological traits, 3 chemical components, grain moisture	Lorenzana and Bernardo (2009)
Maize	RILs derived from single cross	119	0·80	1339 SSRs and RFLPs	0·40–0·50	BLUP	5 morphological traits, grain moisture	
Maize	F_2 derived from single cross	349	0·08, 0·13, 0·20	160 SSRs	0·59–0·72	BLUP	3 morphological traits, grain moisture	
Maize	Testcrosses of DHLs	371	0·13, 0·26, 0·32	125 SNPs	0·31–0·55	BLUP	3 morphological traits, grain moisture	
Arabidopsis thaliana	RILs derived from single cross	415	0·12, 0·23, 0·29, 0·32	69 SSRs	0·90–0·93	BLUP	Flowering time, dry matter, free amino acids	
Barley	DHLs derived from single cross	150	0·36, 0·64, 0·80	223 RFLPs	0·64–0·83	BLUP	Plant height, grain yield, 3 chemical components	
Barely	DHLs derived from single cross	140	0·34, 0·69, 0·80	107 RFLPs and AFLPs	0·66–0·85	BLUP	Plant height, two chemical components	
Maize	DHLs derived from single cross	208	1·00	136 SNPs and SSRs	1·00§	RR, POW, EXP, GAU, SPH	Kernel dry weight	Piepho (2009)
Wheat	Lines bred in CIMMYT	599	0·10	1279 DArTs	0·48–0·61	PM-RKHS	Grain yield	Crossa *et al.* (2010)
Maize	Lines bred in CIMMYT	300	0·90	1148 SNPs	0·42–0·79	M-BL	Grain yield, female flowering, male flowering, anthesis-silking interval	
Wheat	DHLs derived from single cross	209	0·11,0·23,0·46	399 SSRs, DArTs, AFLPs, TRAPs, STS	0·32–0·84	RR-BLUP	8 grain quality	Heffner *et al.* (2011)
Wheat	DHLs derived from single cross	174	0·14, 0·28, 0·55	574 DArTs	0·41–0·73	RR-BLUP	8 grain quality	
Maize	25 nested association mapping populations	(126–196) × 25 populations	0·20, 0·40, 0·60, 0·80	1106 SNPs	0·26–0·57	RR-BLUP	Three flowering traits	Guo *et al.* (2011)
Loblolly pine	61 full-sib families derived from 32 parents	790 – 840	not shown	3938 SNPs	0·64–0·77	BLUP	Diameter at breast height, total height	Resende *et al.* (2011)
Loblolly pine	Full-sib offspring	149	Not shown	3406 SNPs	0·3–0·83	Pedigree model	Growth and quality traits	Isik *et al.* (2011)
Eucalyptus	43 full-sib family 11 interspecific hybrids	783	0·90	3120 DArTs	0·53–0·69	BLUP	Height, diameter at breast height, wood density, pulp yield, lignin content, *Puccinia* rust resistance	Grattapaglia *et al.* (2011)
Eucalyptus	75 full-sib family 55 elite parents (hybrids)	920	0·90	3564 DArTs	0·54–0·62	BLUP		

* Percentages of number of individuals in training populations to whole populations.
† SSRs, Simple sequence repeat markers; RFLPs, restriction fragment length polymorphism markers; SNP, single nucleotide polymorphic markers; DArTs, diversity array technology markers; AFLPs, amplified fragment length polymorphism markers; TRAPs, target region amplification polymorphism markers; STS, sequence tagged site marker.
‡ Correlation between observed phenotypic values and GEBVs.
§ Correlation between adjusted mean and GEBVs. Error variance is not fixed.
¶ Models with the highest or higher accuracy of GEBVs when multiple methods were used for GEBV prediction. BLUP, Best linear unbiased prediction. Spatial models using: POW, power; EXP, exponential; GAU, Gaussian; SPH, spherical models. PM-RKHS, Pedigree plus molecular marker model using reproducing kernel Hilbert space regression. M-BL, Regression model using the Bayesian LASSO; RR, ridge regression.



	White allele freq.		LD (r^2)
Marker I	White	Black	1·00
Marker II	0·5	0·5	0·00
Marker III	0·5	0·5	0·00
Marker IV	0·5	0·5	0·00
Marker V	0·5	0·5	0·00

	White allele freq.	LD (r^2)
Marker I	White	1·00
Marker II	0	1·00
Marker III	1	1·00
Marker IV	1	1·00
Marker V	0	1·00

	White allele freq.		LD (r^2)	Distance (cM)
Marker I	White	Black	1·00	0
Marker II	0·05	0·95	0·81	20
Marker III	0·85	0·15	0·49	40
Marker IV	0·7	0·3	0·16	60
Marker V	0·5	0·5	0·00	80

FIG. 3. Variation of LD intensity in different populations of a single species. (A) Allele frequency and LD indexes (r^2) between marker I and others in an unrelated population. Roman numerals represent markers mapped on a linkage group with 20-cM intervals. The two allele types, white and black, are represented in white and black. White allele freq. means the frequency of white alleles for markers II–V, in each case where the marker I allele is white or black. In this example, the white allele frequencies of markers II, III, IV and V are all 0·5, while the LD indices (r^2) between marker I and other markers are all zero (completely random). (B) A population of clonally propagated individuals. Assume that an individual is selected from an unrelated population (outlined in blue in population 'A') and clonally propagated. All individuals in population 'B' share the same genotype. Thus, the r^2 between marker I and the other markers are all 1·0 (complete LD). (C) Suppose two individuals are selected from population 'A' (outlined in blue and red) and RILs (recombinant inbred lines) are developed based on a cross between the two individuals. Recombination occurs during meiotic division in the F_1, so the white allele frequency varies depending on the distances between marker I and other markers. Then, LD decays are observed in the RILs.

training population did not produce accurate GEBVs in a Jersey population. Toosi *et al.* (2009) simulated the accuracy of GEBVs in admixed and cross-bred livestock populations, and found that the accuracy was greatly reduced when genes from the target pure breed were not included in the admixed and cross-bred population.

Population size

Several reports of simulation and empirical GS studies suggest that a larger training population size improves the accuracy of GEBV predictions. For example, Heffner *et al.* (2011) reported that the average ratio of GS accuracy to PS accuracy for grain quality traits in biparental wheat populations containing 174 or 209 individuals were 0·66, 0·54 and 0·42 for training population sizes of 96, 48 and 24, respectively. The ratio of the number of individuals in the training to the breeding populations varied in different studies. For example, it ranged from 0·08 to 1·00 in empirical studies of plants (Table 2). Although the appropriate ratio varied depending on the genetic diversity, population size, heritability of traits and the number of QTLs, it can be suggested that a higher training : breeding population ratio is required with greater genetic diversity, smaller-sized breeding populations, lower heritability of traits and larger numbers of existing QTLs to obtain GEBVs with high accuracy. In addition, the balance of the population size and the genotyped marker is also important. When the population size is small and the genotype data are large, this often causes an overestimation of the genotype effect, which exaggerates minor flux in the data, i.e. the 'large p, small n' issue (Jannink *et al.*, 2010).

The empirical studies indicated that the sizes of training populations in plant GS studies were often smaller than those of animal studies (Tables 1 and 2). Two factors are

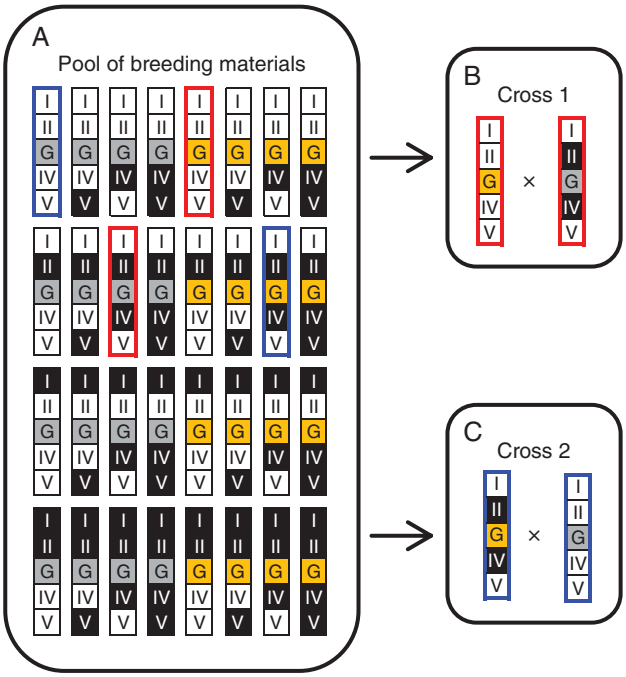


FIG. 4. Allele types of flanking markers for a targeted gene. Roman numerals represent the markers (I, II, IV and V) mapped on a linkage group. 'G' indicates a targeted gene. Distances between adjacent markers and the gene G are 20 cM. White and black represent the allele types of the markers, while grey and yellow indicate the allele types of a targeted gene. Suppose that the yellow allele is a favourable genotype on a targeted gene G. The LD between gene G, marker II and marker IV is completely random in a pool of breeding materials (unrelated population) while significant LD ($r^2 = 0{\cdot}8$) is observed in RILs developed from biparental crosses (1 and 2), as shown in Fig. 3. When the two individuals outlined in red are selected for a biparental cross (B: cross 1), the genotypes of the flanking markers (II and IV) linked to gene G/yellow are white. By contrast, when the two individuals outlined in blue are selected for a biparental cross (C: cross 2), the genotypes of the flanking markers (II and IV) linked to gene G/yellow are black. This example indicates that the allele types with significant LD with the targeted genes are different between the two crosses.

expected to affect the size differences of training populations. The first factor is the narrow genetic diversity of plant populations, which is mainly caused by self-crossing reproduction and/or the smaller number of parental lines used for generating tested populations (biparental crosses have often been used). Because populations having greater genetic diversity require larger population sizes to obtain GEBVs with high accuracy (Mujibi *et al.*, 2011), smaller sizes of training populations are used in plant GS studies, especially for self-crossing reproduction species and/or biparental cross-derived populations. The second factor is the existence of a large quantity of legacy data about the phenotypes of pedigrees, which have been used to estimate traditional BVs in animal breeding (Hayes *et al.*, 2009). The accumulated phenotype data should make performing GS studies possible with low cost. As with animal studies, pooling phenotypes of plant populations in which multiple regions have been investigated would be a promising approach for achieving success in plant GS studies, satisfying both high-accuracy GEBV and low experimental cost.

Number of markers

Generally, a greater number of markers is required for a population where the marker intervals with a significant LD intensity are shorter. In addition, empirical and simulation studies suggest that a larger number of markers improves the accuracy of GEBVs. For example, Solberg *et al.* (2008) found that the simulated accuracy of GEBVs was improved by increasing the marker density from 0·25 to 8 SNP markers per centimorgan in 100 unrelated animals. Furthermore, in an inbred population derived from a biparental cross, Bernardo and Yu (2007) demonstrated that the response of the GEBV improves when decreasing the adjacent marker intervals from 28·0 to 7·0 cM, whereas no differences were observed with marker intervals of 7·0, 3·5 and 2·3 cM when the total length of the linkage map was assumed to be 1794 cM. The heritability of targeted traits is also affected by the relationship between the density of markers and the accuracy of GEBV. Calus and Veerkamp (2007) demonstrated that an adjacent marker r^2 of 0·15 was sufficient for a trait with a heritability of 50 %, while the GEBV accuracy was improved by increasing the r^2 to 0·2 for a trait with a heritability of 10 %. However, we have to consider that too many markers often leads to a loss in GEBV accuracy, as described in the section on population size.

One of the obvious differences between GS and traditional MAS is the number of markers required for genotyping in a breeding population. In most GS studies, the whole set of markers used in the training population was also applied to the breeding or validating population. For example, suppose that the numbers of individuals in the training and breeding populations are 200 and 1000, respectively, and that the number of genotyped markers is 1000, then the genotype data points are 200×1000 in the training population and 1000×1000 in the breeding population. This is quite different from traditional MAS, which requires a few selected marker genotypes that are related to a targeted trait in the breeding phase, except when investigating the GW genetic backgrounds of a breeding population in MARS.

Several reports suggest that advances in genotyping technologies will resolve the cost issue of the large number of genotype data points required in a breeding population and this idea might be correct. However, it is still necessary to conduct large-scale genotyping when performing GS in many breeding programmes, especially for non-major crops. To overcome this obstacle, several studies have used decreasing numbers of genotyped markers. For example, Habier *et al.* (2009) proposed a panel of evenly spaced low-density SNPs for tracking the effects of high-density SNP alleles within families based on the utilization of cosegregation information. Iwata and Jannink (2010) determined the imputation scores of untyped markers in a low-density genotyped panel by referencing a high-density panel in barley. Both studies were based on a common idea of predicting the interval genotypes of a population using low-density allelic data. By contrast, Cleveland *et al.* (2010) discussed the performance of GEBV prediction by reducing the density of marker panels. It was found that low-density and evenly spaced SNPs performed poorly when predicting GEBV, whereas SNPs selected

based on their additive-effect size yielded accuracies similar to those at a high density.

Heffner *et al.* (2009) proposed a model for a genomic selection breeding programme, which consisted of a model training cycle and a line development cycle. It suggested that the most immediate impact of selecting an elite line by GEBVs would be a marked increase in the speed of the cycles. Shorter selection cycles of populations would lead to a rapid change of genetic diversity in the breeding populations and would affect GEBV accuracy during long-term selection. In addition, novel recombinants generated during selection cycles would cause LD decay between markers and QTLs. This would be a more serious issue when lower-density markers are used for GEBV predictions. Goddard (2009) and Jannink *et al.* (2010) surveyed the dynamics of long-term selection responses by performing simulation studies, and concluded that GS leads to a more rapid decline in the selection response than PS unless new markers are continually added to the prediction of breeding value. They also suggested that placing additional weight on low-frequency favourable alleles, especially at the beginning of GS, was important for maximizing the long-term response in GS.

Types of markers

Most GS studies use SNP, DArT and simple sequence repeat (SSR) markers for genotyping. Results based on other types of markers in high-throughput genotyping systems will be reported in the near future, such as restriction site-associated DNA (RAD) (Miller *et al.*, 2007) and genotyping by sequencing (GBS) (Elshire *et al.*, 2011) marker systems. The DArT, RAD and GBS markers identify polymorphisms by hybridization or sequencing digested DNAs using restriction enzymes, so they are dominant markers except in the case where high coverage genome data are obtained for each individual using RAD and GBS markers. Li *et al.* (2007) showed that the LD detection power of a dominant marker is less than that of a co-dominant marker, and it was improved with a three-locus LD analysis. The results suggest that dominant markers lead to a lower accuracy of GEBV prediction than co-dominant markers and employing haplotypes would improve the accuracy.

DNA markers are also categorized as bi-allelic markers and multi-allelic markers. The former includes SNP, DArT, GBS and RAD markers while the latter include SSR, RAPD (random amplified polymorphic DNA) and RFLP (restriction fragment length polymorphism) markers. Solberg *et al.* (2008) demonstrated the accuracy of GEBV prediction with SNP and SSR markers using 100 unrelated animals, and concluded that the SNP markers required two to three times greater density compared with using an SSR marker to achieve similar accuracy. With a bi-allelic marker, additional consideration must be given to the genetic sources used in marker development. Barendse *et al.* (2009) compared Australian and Bovine HapMap samples, and found differences in the presumptive selective signatures when different breeds or SNPs were used. Based on these results, they suggested that using the same SNP is necessary when comparing the selection signatures among studies.

RAD and GBS marker systems that can scan GW polymorphisms in *de novo* would bypass the need for prior marker development and rather allow direct genotyping of the training and breeding populations.

Traits

The main advantage of MAS is considered to be the lack of a requirement for phenotyping during selection cycles (Heffner *et al.*, 2009). More strictly, there is no need for the phenotyping of traits that were previously investigated in a training population. In conventional breeding, multiple expressed traits are investigated during a whole growing period and after harvesting. Thus, all traits of interest to breeders should be investigated during the training phase to exclude phenotyping during the breeding cycle if the gain of 'selection' is regarded as equivalent between MAS and conventional breeding. In traditional MAS, only a few selected markers are used during the breeding phase, whereas GW genotypes are used in GS. Therefore, MAS for multiple traits is performed more systematically in GS, because there is no need to change the marker set used during the breeding phase. To our knowledge, no reports have been published on the selection of traits where trade-off relationships are observed in breeding materials, such as stress tolerance and quality. However, we considered that the selection of trade-off traits in GS will be a major issue in the near future, because the use of GW genotypes might help break up the trade-off relationships of targeted traits, although current GS does not consider the weight of each marker effect in the result.

Breeding scheme with GS

According to published reports, GS is not assumed to be a perfect replacement for PS in plant breeding and instead it is proposed as a method for accelerating part of a whole breeding programme. For example, Bernardo and Yu (2007) proposed using GS during the off-season for the selection of random mating DHLs that are pre-selected for their test-crossing ability in the regular season by PS. By contrast, Heffner *et al.* (2009, 2010) and Jannink (2010) proposed using GS for parental selection to generate the breeding population in the next selection cycle. For example, 288 inbred (F_5) lines of winter wheat were assumed to be created and genotyped by single-seed descent (Heffner *et al.*, 2010). F_5-derived lines were grown in the field to increase the seeds, which were then selected for advanced testing based on their phenotypes and GEBV. Small numbers of F_5-derived lines were selected based on GEBV to start recombining for the next cycle. In addition, phenotypic data from F_5-derived lines were used for GS modelling of the next cycle. The proposed scheme suggested that GS fits well with recurrent selection approaches that are not usually employed in the conventional breeding of selfing crop species. Interestingly, Heffner *et al.* (2010) also proposed using traditional MAS for important QTLs in the F_2 and F_3 generations, before GS in the F_5 generation. This eliminates unnecessary marker scoring and greenhouse space for lines that do not carry essential QTL alleles. These propositions suggest the importance of flexible GS

introduction into breeding programmes and combining it with other approaches, i.e. traditional MAS and PS.

Computer package for GS modelling

An R-Package for GS is available on http://www.r-project.org/. No user-friendly software has yet been developed, such as QTL Cartographer (Wang *et al.*, 2011) and MapQTL (Van Ooijen, 2004) that are used in QTL analysis. The development of a user-friendly software package is required to enhance the general application of GS.

FUTURE PERSPECTIVES IN GS

Like conventional breeding and traditional MAS, GS cannot be used for the selection of low heritability (in the narrow sense) traits. Narrow-sense heritability is defined as the ratio of the genetic variance of additive genetic effects to the phenotypic variance. Thus, low heritability traits are caused by the high variance of non-additive genetic effects, such as environmental factors, G × E interactions, and dominant and epistatic genetic effects.

Previous studies of QTL identification suggest that the magnitudes of G × E are unequal, with some QTLs expressed in all tested environments and others expressed in a particular environment (Xu and Crouch, 2008). With GS, Goddard and Hayes (2007) indicated that different animals tend to be selected for the two environments when the genetic correlation between production in the two environments is $<0{\cdot}8$. The only solution for overcoming the G × E issue is considered to be the accumulation and comparison of phenotypes investigated in different environments. Recently, Thomas (2010) reviewed potent approaches for gene–environment–wide association studies, i.e. mining GW association data for G × E interactions. For example, in an approach known as two-phase case-control, the whole case-control dataset was divided into subgroups based on different categories, such as age and gender, before correlations between SNPs and traits in the subgroups were identified. The approaches are different from GS in plants and they cannot be applied directly, but the essence of the idea, i.e. dividing groups based on environmental conditions, might be applicable in the future development of GS methodology. Heffner *et al.* (2009) also suggested that GS is feasible for phenotypic data accumulation, because once phenotypes have been evaluated in particular environmental conditions (e.g. severe winter once per a decade), the phenotypic values can be included in the GS model and used for selection. This idea is better suited to crops where the genetic structures of breeding materials are relatively fixed, and that have been investigated for a long time. In addition, careful consideration should be given to phenotypic values investigated under extreme environmental conditions. If such phenotypic values are mixed with other values investigated under normal conditions and then used for GS modelling, the GEBV accuracy will be decreased. The comparison of phenotypic values evaluated in different environmental conditions has not yet matured and further study will be needed.

GEBVs are predicted based on additive genetic effects, so current GS does not consider both dominant genetic effects and epistasis. In traditional MAS, the dominant effects are considered because interval mapping and LD mapping can predict the dominant effect of QTLs. Thus, the consideration of dominant effects in GS modelling is expected to be achieved in the near future by improving modelling algorithms. By contrast, the consideration of epistasis in GS is more challenging. Previous studies indicate that the magnitude of epistatic effects depends on the species, population structure and targeted traits, although sometimes it is negligible (Xu and Jia, 2007) whereas in other cases it is more important than the additive effects (Malmberg *et al.*, 2005; Mei *et al.*, 2005; Dudley and Johnson, 2009). Isobe *et al.* (2007) developed a QTL mapping approach that searches for QTL interactions in genetic variation and they demonstrated that QTL interactions among small effect QTLs sometimes produce larger effects than single main-effect QTLs. Recently, Hu *et al.* (2011) used an empirical Bayesian method for GEBV prediction that had been used previously for the identification of epistatic QTLs by Xu and Jia (2007). The results showed that including epistatic effects greatly increased the accuracy of GEBV prediction compared with the non-consideration of epistatic effects. Epistasis demands vast amounts of calculation for its identification, but the consideration of epistatic effects in GS is an issue that needs to be addressed in the future.

In conclusion, GS may be regarded as a potent, attractive and valuable approach for plant breeding. The main contribution of GS to breeding MAS might be in providing a concept for the conversion of genotypic value to phenotypic value. With this idea, we are free from the pyramiding of QTLs and we can enjoy designing the ideal genotype based on the results of one or a few test trials. At the same time, GS is not a perfect method and several issues demand careful attention and improvement. Moreover, as van der Werf (2007) suggested, GS leaves an understanding of the underlying biology behind a black box. In our opinion, the main weakness of current GS might be lack of value of contexts in genome sequences. Current GS algorithms have not been connected with previous and current studies of genetics and genomics, such as QTLs and (candidate) gene identification. By integrating the essence of GS with other fields of genetics and genomic studies, it might be possible to escape the black box. Meuwissen (2007) noted that GS was considered a crazy idea when he and his colleagues proposed it. However, it has now become a realistic approach in plant breeding and we have validated its availability with empirical data. GS is not a final solution of MAS, but it is a turning point on the road that leads us to the next phase of MAS. GS will be integrated into many practical breeding programmes in the near future as it becomes more advanced and its theory matures.

LITERATURE CITED

Barendse W, Harrison BE, Bunch RJ, Thomas MB, Turner LB. 2009. Genome wide signatures of positive selection: the comparison of independent samples and the identification of regions associated to traits. *BMC Genomics* **10**: 178. http://dx. doi.org/10·1186/1471-2164-10-178.

Bernardo R. 2009. Genomewide selection for rapid introgression of exotic germplasm in maize. *Crop Science* **49**: 419–425.

Bernardo R. 2010. Genomewide selection with minimal crossing in self-pollinated crops. *Crop Science* **50**: 624–627.

Bernardo R, Yu J. 2007. Prospects for genomewide selection for quantitative trait in maize. *Crop Science* **47**: 1082–1090.

Calenge F, Legarra A, Beaumont C. 2011. Genomic selection for carrier-state resistance in chicken commercial lines. *BMC Proceedings* **5** (Suppl. l4): S24. http://dx.doi.org/10.1186/1753-6561-5-S4-S24.

Calus MP. 2010. Genomic breeding value prediction: methods and procedures *Animal* **4**: 157–164.

Calus MPL, Veerkamp RF. 2007. Accuracy of breeding values when using and ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP per cM. *Journal of Animal Breeding and Genetics* **124**: 362–368.

Cleveland MA, Forni S, Deeb N, Maltecca C. 2010. Genomic breeding value prediction using three Bayesian methods and application to reduced density marker panels. *BMC Proceedings* **4** (Suppl 1): S6. http://dx.doi.org/10.1186/1753-6561-4-S1-S6.

Crossa J, Campos G, de L, Pérez P, *et al.* 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **186**: 713–724.

Dudley JW, Johnson GR. 2009. Epistatic models improve prediction of performance in corn. *Crop Science* **49**: 763–770.

Elshire RJ, Glaubitz JC, Sun Q, *et al.* 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**: pe19379. http://dx.doi.org/10.1371/journal.pone.0019379.

Fisher RA. 1918. The correlation between relatives on the supposition on Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**: 399–433.

Garris AJ, McCouch SR, Kresovich S. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa5* locus of rice (*Oryza sativa* L.) *Genetics* **165**: 759–769.

Gaut BS, Long AD. 2003. The lowdown on linkage disequilibrium. *The Plant Cell* **15**: 1502–1506.

Gianola D, Fernando RL. 1986. Bayesian methods in animal breeding theory. *Journal of Animal Science* **63**: 217–244.

Goddard M. 2009. Genomic selection: prediction of accuracy and maximisation of long term selection. *Genetica* **136**: 245–257.

Goddard ME, Hayes BJ. 2007. Genomic selection. *Journal of Animal Breeding and Genetics* **124**: 323–330.

González-Martínez SC, Brown GR, Ersoz E, *et al.* 2004. Nucleotide diversity, linkage disequilibrium and adaptive variation in natural populations of loblolly pine. *Plant & Animal Genomes XII Conference*, 10–14 January 2004, San Diego, CA, W3.

Grattapaglia D, Resende MDV, Resende MR, *et al.* 2011. Genomic selection for growth traits in *Eucalyptus*: accuracy within and across breeding populations. *BMC Proceedings* **5** (Suppl. 7): O16.

Guo Z, Tucker DM, Lu J, Kishore V, Gay G. 2011. Evaluation of genome-wide selection efficiency in maize nested association mapping populations. *Theoretical and Applied Genetics* **124**: 261–275.

Gupta PK, Pawan S, Kulwal PL. 2005. Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Molecular Biology* **57**: 461–485.

Habier D, Fernando RL, Dekkers CM. 2009. Genomic selection using low-density marker panels. *Genetics* **182**: 343–353.

Harris BL, Johnson DL, Spelman RL. 2008. Genomic selection in New Zealand and the implications for national genetic evaluation. *Proceedings of the Interbull Meeting*, 16–19 June 2008, Niagara Falls, Canada.

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

Heffner EL, Sorrells ME, Jannink J-L. 2009. Genomic selection for crop improvement. *Crop Science* **49**: 1–12.

Heffner EL, Lorenz AJ, Jannink J-L, Sorrells ME. 2010. Plant breeding with genomic selection: gain per unit time and cost. *Crop Science* **50**: 1681–1690.

Heffner EL, Jannink J-L, Iwata H, Souza E, Sorrells ME. 2011. Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Science* **51**: 2597–2606.

Henderson CR. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31**: 423.

Hu Z, Li Y, Song X, *et al.* 2011. Genomic value prediction for quantitative traits under the epistatic model. *BMC Genetics* **12**: 15. http://dx.doi.org/10·1186/1471-2156-12-15.

Isik F, Whetten R, Zapata-Valenzuela J, Ogut F, McKeand S. 2011. Genomic selection in loblolly pine – from lab to field. *BMC Proceedings* **5** (Suppl. 7): I8.

Isobe S, Nakaya A, Tabata S. 2007. Genotype Matrix Mapping (GMM): searching for QTL interactions in genetic variation in complex traits. *DNA Research* **14**: 217–225.

Iwata H, Jannink J-L. 2010. Marker genotype imputation in a low-marker-density panel with a high-marker density reference panel: accuracy evaluation in barley breeding lines. *Crop Science* **50**: 1269–1278.

Iwata H, Jannink J-L. 2011. Accuracy of genomic selection prediction in barley breeding programs: a simulation study based on the real single nucleotide polymorphism data of barley breeding lines. *Crop Science* **51**: 1915–1927.

Jannink J-L. 2010. Dynamics of long-term genomic selection. *Genetics Selection Evolution* **42**: 35.

Jannink J-L, Lorenz AJ, Iwata H. 2010. Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics and Proteomics* **9**: 166–177.

Legarra A, Robert-Granié C, Manfredi E, Elsen JM. 2008. Performance of genomic selection in mice. *Genetics* **180**: 611–618.

Li Y, Li Y, Wu S, *et al.* 2007. Estimation of multilocus linkage disequilibria in diploid populations with dominant markers. *Genetics* **176**: 1811–1821.

Lorenzana RE, Bernardo R. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics* **120**: 151–161.

Luan T, Woolliams JA, Lien S, Kent M, Svendsen M, Meuwissen TH. 2009. The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. *Genetics* **183**: 1119–1126.

Malmberg RL, Held S, Waits A, Mauricio R. 2005. Epistasis for fitness-related quantitative traits in *Arabidopsis thaliana* grown in the field and in the greenhouse. *Genetics* **171**: 2013–2027.

Mayor PJ, Bernardo R. 2009. Genomewide selection and marker-assisted recurrent selection in doubled haploid versus F_2 populations. *Crop Science* **49**: 1719–1725.

Mei HW, Li KZ, Shu QY, *et al.* 2005. Gene actions of QTLs affecting several agronomic traits resolved in a recombinant inbred rice population and two backcross populations. *Theoretical and Applied Genetics* **110**: 649–659.

Meuwissen THE. 2007. Genomic selection: marker assisted selection on a genome wide scale. *Journal of Animal Breeding and Genetics* **124**: 321–322.

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

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17**: 240–248.

Moser G, Tier B, Crump RE, Khatkar MS, Raadsma HW. 2009. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. *Genetics Selection Evolution* **41**: 56.

Mujibi FDN, Nkumah JD, Durunna ON, *et al.* 2011. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *Journal of Animal Science* **89**: 3353–3361.

Palaisa KA, Morgante M, Williams M, Rafalski A. 2003. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *The Plant Cell* **15**: 1795–1806.

Piepho HP. 2009. Ridge regression and extensions for genomewide selection in maize. *Crop Science* **49**: 1165–1176.

Piyasatin N, Fernando RL, Dekkers JCM. 2007. Genomic selection for marker-assisted improvement in line crosses. *Theoretical and Applied Genetics* **115**: 665–674.

Remington DL, Thornsberry JM, Matsuoka Y, *et al.* 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the USA* **98**: 11479–11484.

Resende MFR Jr, Valle PRM, Acosta JJ, Resende MDV, Grattapaglia D, Kirst M. 2011. Stability of genomic selection prediction models across ages and environments. *BMC Proceedings* **5** (Suppl. 7): O14.

Rolf MM, Taylor JF, Schnabel RD, *et al.* 2010. Impact of reduced marker set estimation of genomic relationship matrices on genomic selection for feed efficiency in Angus cattle. *BMC Genetics* **11**: 24. http://dx. doi.org/10·1186/1471-2156-11-24.

Solberg TR, Sonesson AK, Woolliams JA, Meuwissen THE. 2008. Genomic selection using different marker types and densities. *Journal of Animal Science* **86**: 2447–2454.

Thomas D. 2010. Gene-environment-wide association studies: emerging approaches. *Nature Reviews Genetics* **11**: 259–272.

Toosi A, Fernando RL, Dekkers JCM. 2009. Genomic selection in admixed and crossbred populations. *Journal of Animal Science* **88**: 32–46.

Van Ooijen. 2004. MapQTL® 5, software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, The Netherlands.

Van Raden PM, Van Tassell CP, Wiggans GR, *et al.* 2009. Invited review: reliability of genomic predictions for North American Holsein bulls. *Journal of Dairy Science* **92**: 16–24.

Van Vleck LD, Hakim AF, Cundiff LV, Koch RM, Crouse JD, Boldman KG. 1992. Estimated breeding values for meat characteristics of crossbred cattle with an animal model. *Journal of Animal Science* **70**: 363–371.

Wakamiya I, Newton R, Johnston JS, Price HJ. 1993. Genome size and environmental factors in the genus *Pinus*. *American Journal of Botany* **80**: 1235–1241.

Wang S, Basten CJ, Zeng Z-B. 2011. *Windows QTL Cartographer 2·5*. Department of Statistics, North Carolina State University, Raleigh, NC USA.

van der Werf J. 2007. Animal breeding and the black box of biology. *Animal Breeding and Genetics* **124**: 101.

Williams KS, Cummings MR. 1997. *Concepts of genetics*, 5th edn. Englewood Cliffs, NJ: Prentice-Hall.

Wolc A, Stricker C, Arango J, *et al.* 2011. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genetics Selection Evolution* **43**: 5.

Wong CK, Bernardo R. 2008. Genomewide selection in oil palm: increasing selection gain per unit time and cost with small populations. *Theoretical and Applied Genetics* **116**: 815–824.

Xu S, Jia Z. 2007. Genomewide analysis of epistatic effects for quantitative traits in barley. *Genetics* **175**: 1955–1963.

Xu Y, Crouch JH. 2008. Marker-assisted selection in plant breeding: from publication to practice. *Crop Science* **48**: 391–407.

Zhong S, Dekkers JCM, Fernando RL, Jannink J-L. 2009. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. *Genetics* **182**: 355–364.