Advances and Challenges in Genomic Selection for Disease Resistance

Jesse Poland[1] and Jessica Rutkoski[2,3]

[1]Wheat Genetics Resource Center, Department of Plant Pathology and Department of Agronomy, Kansas State University, Manhattan, Kansas, 66506; email: jpoland@ksu.edu

[2]Plant Breeding and Genetics Section, Cornell University, Ithaca, New York, 14853; email: jer263@cornell.edu

[3]International Maize and Wheat Research Center (CIMMYT), Texcoco, Estado de México, 56237 Mexico

Annu. Rev. Phytopathol. 2016. 54:79–98

The *Annual Review of Phytopathology* is online at phyto.annualreviews.org

This article's doi:
10.1146/annurev-phyto-080615-100056

Keywords

quantitative disease resistance, genomic selection, plant breeding

Abstract

Breeding for disease resistance is a central focus of plant breeding programs, as any successful variety must have the complete package of high yield, disease resistance, agronomic performance, and end-use quality. With the need to accelerate the development of improved varieties, genomics-assisted breeding is becoming an important tool in breeding programs. With marker-assisted selection, there has been success in breeding for disease resistance; however, much of this work and research has focused on identifying, mapping, and selecting for major resistance genes that tend to be highly effective but vulnerable to breakdown with rapid changes in pathogen races. In contrast, breeding for minor-gene quantitative resistance tends to produce more durable varieties but is a more challenging breeding objective. As the genetic architecture of resistance shifts from single major R genes to a diffused architecture of many minor genes, the best approach for molecular breeding will shift from marker-assisted selection to genomic selection. Genomics-assisted breeding for quantitative resistance will therefore necessitate whole-genome prediction models and selection methodology as implemented for classical complex traits such as yield. Here, we examine multiple case studies testing whole-genome prediction models and genomic selection for disease resistance. In general, whole-genome models for disease resistance can produce prediction accuracy suitable for application in breeding. These models also largely outperform multiple linear regression as would be applied in marker-assisted selection. With the implementation of genomic selection for yield and other agronomic traits, whole-genome marker profiles will be available for the entire set of breeding lines, enabling genomic selection for disease at no additional direct cost. In this context, the scope of implementing genomics selection for disease resistance, and specifically for quantitative resistance and quarantined pathogens, becomes a tractable and powerful approach in breeding programs.

INTRODUCTION

Types of Disease Resistance

Disease resistance is often classified into qualitative and quantitative resistance (84). The scope of qualitative disease resistance is generally resistance conditioned by a single resistance (R) gene recognizing avirulence factors in a classic gene-for-gene mechanism, and the inheritance is said to be qualitative or Mendelian. In contrast, quantitative resistance is usually conditioned by many genes of small effect, and the inheritance is said to be quantitative or polygenic (56). However, the dichotomy between these two classifications of disease resistance is not clear-cut, with many examples of pathosystems having a genetic architecture between qualitative genes with major effect and purely quantitative resistance (59).

From a breeding perspective, quantitative resistance is considered to be more durable when deployed in varieties for agricultural production (44). Unlike resistance based on R genes, quantitative resistance generally does not appear to be race specific (38). Both qualitative and quantitative resistance have been, and continue to be, targets of selection for breeding programs. In the scope of quantitative genetics and genomics-assisted breeding, differential reaction to different populations or races of a given pathogen (race-specific resistance) contributes to genotype-by-environment (G × E) interaction, which can be observed as differences in resistance levels across different environments and may consist of scale differences or changes in rank of the varieties tested. With race-specific resistance, the prevalent pathogen race(s) is a component of the environment, leading to greater observed G × E. By comparison, minor-gene resistance that has no or minimal gene-for-gene interaction leads to much less G × E. In breeding, yield stability is the ideal target, especially because good performance across years is desired. Stability in resistance, resistance with minimal G × E, is important for achieving yield stability, particularly in areas prone to epidemics. Breeding for race-nonspecific minor-gene resistance is one way to minimize G×E of resistance. In addition to minimizing G × E, there are also quantitative genetic and genomic prediction models that can help improve breeding efficiency when G × E is present, as long as there is some genetic correlation between environments (12, 41).

Genetic Architecture of Disease Resistance

The genetic architecture of a trait includes the number of gene(s) controlling the trait, the genomic location and allele substitution effect of these genes, and the overall heritability of the trait (49). Important agronomic traits such as grain yield or lodging tolerance are generally considered to have a complex genetic architecture, meaning that the trait is controlled by many genes, each with a small effect. There are notable examples of important, large-effect genes in breeding programs. However, these genes are rapidly selected to fixation and are therefore inconsequential for future gain from selection. The breeding methodologies employed for different traits largely depend on the genetic architecture, the heritability, and the economic importance of that trait (25, 31).

Breeding methods for disease resistance vary depending on whether the resistance is considered to be qualitative or quantitative. It follows that the genetic architecture of disease resistance is closely tied to whether the resistance is quantitative or qualitative, and hence both the phenotypic and molecular breeding approaches must be matched accordingly. As can be expected, qualitative resistance conditioned by single, major genes does not have a complex genetic architecture and hence is more suitable for identifying and mapping single resistance genes of large effect. In contrast, quantitative resistance can be approached similarly to other agronomic traits, using breeding methodology and whole-genome prediction models developed for quantitative traits.

BREEDING FOR DISEASE RESISTANCE

As a selection objective, breeding for disease resistance is often completed in early generations with selection of lines or single plants in dedicated disease nurseries. In a typical selected bulk or pedigree program, many hundreds of populations are developed from which individual plants or lines are selected for general agronomic performance (e.g., height, maturity) and disease resistance. Another strategy is to generate lines using doubled haploids or single-seed descent, and then phenotype lines for disease resistance and all other traits of interest prior to selection. This is suitable if heritability is low, making single plant selection too inaccurate to warrant its application.

For molecular breeding, marker-assisted selection (MAS) has been implemented in breeding for selection and stacking (gene pyramids) of major resistance genes. As outlined in detail below, the use of molecular markers and MAS can facilitate gene pyramids for more durable resistance, as multiple genes that are effective against the pathogen isolates can be combined into a single breeding line or a variety in a manner that would not be possible with only phenotypic selection.

Breeding for Qualitative Disease Resistance

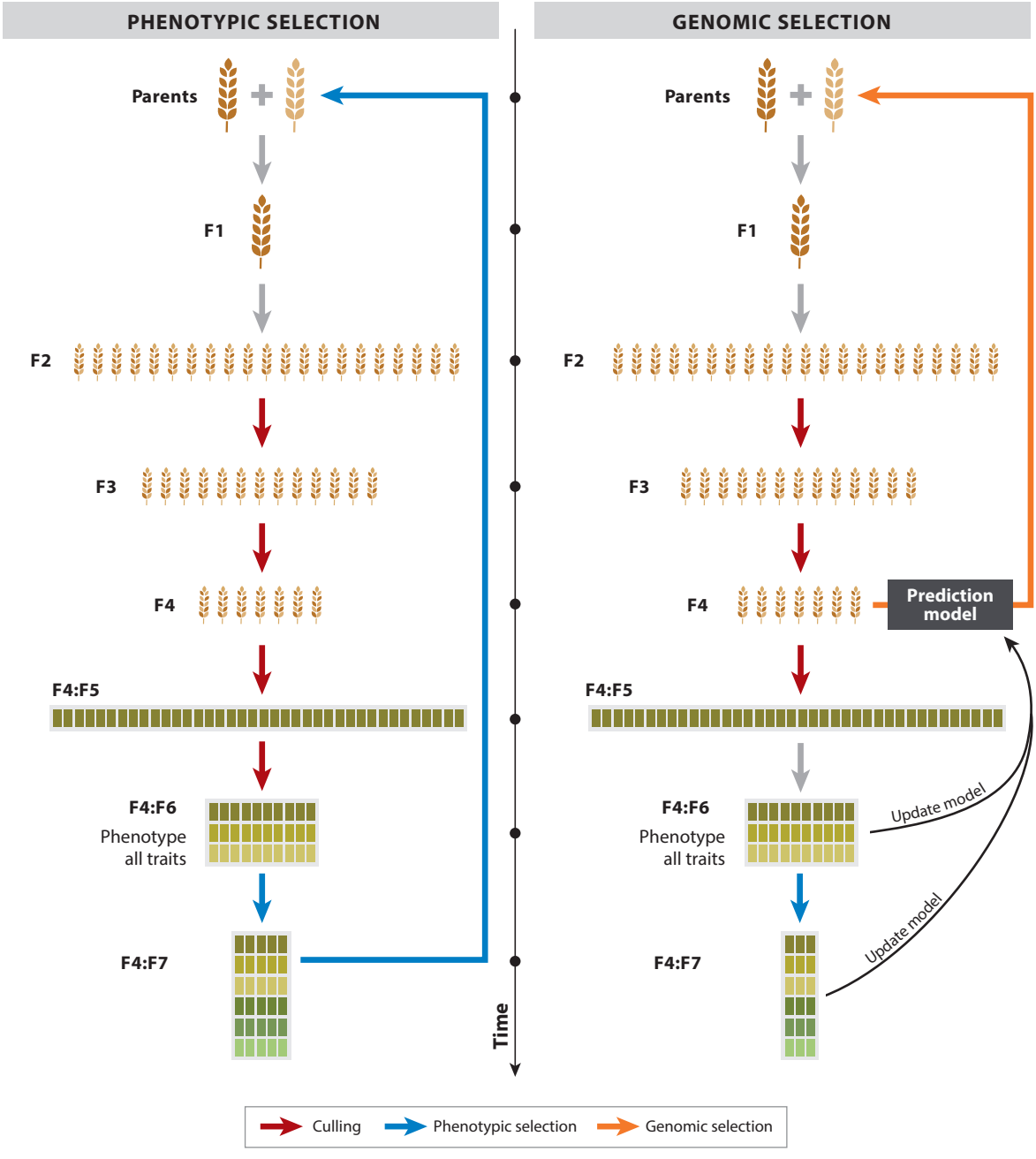
Breeding for qualitative resistance usually involves screening large numbers of progeny in early generations to identify and discard susceptible individuals that do not carry the resistance gene allele(s) of interest. This can be done phenotypically in a greenhouse, often at very early growth stages, or in the field on adult plants. Resistant plants or lines are then selected and advanced in the breeding program. The same process can be done using markers linked to the resistance genes, enabling screening at very early stages in the absence of the pathogen. By screening large numbers of individuals in early generations, the targeted resistance allele(s) can be fixed or enriched in the population (7) before individuals are advanced in the breeding program and subjected to selection for other traits. Large population sizes are important so that many selection candidates remain after screening for resistance, enabling selection for other key traits.

Backcrossing is another possible breeding strategy that can be implemented for disease resistance. With backcrossing, a resistant parent is crossed to an elite but susceptible parent. Resistant progeny are then crossed to the elite parent, and the process is continued until a line is developed that has the desired resistance allele in the elite parent background. The selection imposed during backcrossing can be based on either phenotype or a marker linked to the resistance gene. To expedite the process, markers distributed on all chromosomes can be used to select for the elite parent background (33). One of the key features of breeding for qualitative disease resistance is that resistance alleles can be fixed in relatively few cycles of breeding, and there is a clear endpoint at which selection for a resistance allele is no longer necessary (20).

Breeding for Quantitative Disease Resistance

Unlike qualitative resistance, breeding for quantitative resistance requires multiple cycles of breeding, leading to a gradual improvement in resistance over time. Breeding methodology for quantitative resistance may be the same as those for other quantitative traits, such as yield, and selection is generally imposed on all quantitative traits at once, with each trait weighted either implicitly or explicitly according to economic value or other criteria (30). In a typical breeding program that aims to develop lines (**Figure 1**), a breeding cycle begins with the identification and crossing of parents with the highest breeding values, taking into account all target traits. The progeny of the cross are then self-pollinated to create an F_2 population that will segregate for most or all traits. The F_2 population is grown, and selection can be imposed for highly heritable traits,

possibly including quantitative disease resistance. The selected F_2 plants are self-pollinated, and advanced to generate the F_3 population. Selection and advancement are continued in the F_3 and F_4 populations. In many programs, single plants from the F_4 are harvested individually, and all the seed is kept and stored separately. Seed from the F_4 plants is planted, giving rise to the F_4:F_5 generation. At this point, selection among F_4:F_5 lines for multiple high-heritability traits is usually imposed. All seed from the F_4:F_5 lines are harvested and then planted for advanced testing (F_4:F_6



generation). The next generation is planted the next year for another year of advanced testing. The self-pollination stages from F_2 to F_4 are required for the development of pure-breeding lines that may be needed for commercialization, depending on the crop, and for enabling accurate measurement of low-heritability traits at realistic planting densities. Various modifications to the scheme outlined above are common, and for crops that do not require line development, as in the case of clonally propagated crops, the selfing stages may be replaced with clone or seed-increase generations where selection may be imposed. Likewise, the use of double haploid (DH) technology can enable advancement directly from F_1 plants to fixed, inbred lines with selection on DH populations employing techniques equivalent to F_5 lines in the program (19). The development of hybrid cultivars follows an equivalent scheme to generate inbred lines with the notable exception that the agronomic and yield testing are done with test-cross hybrids generated from the inbred lines.

Selection for quantitative disease resistance can occur at any stage of the breeding process, including the early and late generation stages, as long as its evaluation is feasible. However, because yield is generally considered the most important trait, selection for new parents to initiate a new cycle of breeding generally occurs after yield testing, during late generations.

Gain from selection can be predicted using the theoretical formula $R = ir\sigma_a$(18) where i is the selection intensity, r is the accuracy of selection, and σ_a is the additive genetic standard deviation among the selection candidates. When selection occurs at several stages throughout the breeding cycle, the total gain from selection is the sum across all selection events, $R_{total} = \sum R$. Thus, selecting at multiple stages can be beneficial; however, it can be less effective than expected for several reasons. First, selecting in early generations is generally less accurate because it must be based on a single plant without replication. Second, σ_a in early generations prior to inbreeding is partitioned both within and among lines, limiting the effectiveness of selection as progeny segregate from the parental phenotype. Third, selection reduces σ_a(11), and there is less potential gain in subsequent stages.

Gain from selection per unit time is formalized as

$$\Delta G = \frac{R_{total}}{L},$$

where L is the length of the breeding cycle. The gain per unit time is of ultimate importance because it indicates the rate of genetic improvement and can be used to evaluate breeding schemes

Figure 1

Examples of phenotypic and genomic selection schemes enabling population improvement and line development. In both the phenotypic and genomic selection schemes, inbred parents are intermated, giving rise to the F_1 generation. F_1 plants are self-pollinated, giving rise to F_2 seed that becomes the F_2 population. Undesirable F_2 plants are culled based on highly heritable traits, possibly including disease resistance. One or multiple seeds are harvested from the remaining F_2 plants and then planted to give rise to the F_3 generation. Culling and harvesting are carried out as before, giving rise to the F_4 generation. In phenotypic selection (*left*), selected F_4 plants are harvested and stored separately. The set of F_5 seeds, each derived from a single F_4 plant (F_4:F_5), is planted in a small plot, and culling selection among plots is imposed. Seed from each remaining plot is harvested and stored, giving rise to the F_4:F_6 generation. The F_4:F_6s are phenotyped for all traits, and the best lines are selected for advancement. The next generation, F_4:F_7, is phenotyped again, and new parents are identified for crossing based on all phenotypic data. In genomic selection (*right*), F_4 plants are genotyped, their breeding values for all traits of interest are predicted, and new parents for crossing are identified. Seed from F_4 plants is harvested and stored separately, giving rise to the F_4:F_5 generation. An optimal subset of the F_4:F_5 population is planted in small plots. Seed from each plot is harvested and stored, giving rise to the F_4:F_6 generation. The F_4:F_6s are phenotyped for model updating and also for variety release. F_4:F_6s are harvested and the next generation, F_4:F_7, is phenotyped for model updating and variety release. Each stage in the breeding pipeline represents a decrease in the overall number of lines and an increase in the level of replication and phenotypic testing after the F4 stage.

of different cycle lengths. The length of the breeding cycle can have a major impact on the rate of genetic gain, and thus efforts to reduce *L* without drastically reducing gain per cycle are key for increasing the rate of genetic gain. It is important to note that in plant breeding programs that target multiple traits, *L* is usually determined by the trait that takes the longest to evaluate, which is generally yield. Furthermore, when selecting for multiple traits the gain for each individual trait is usually less than that observed when breeding for only one trait (25). Thus, breeding for quantitative disease resistance is effectively slowed by the need to improve other important traits to develop commercially viable varieties.

Breeding for Exotic and Quarantined Diseases

There are a number of important exotic and quarantined diseases for which preemptive breeding efforts are warranted. For example, highly virulent stem rust races in wheat from the Ug99 lineage have received considerable breeding effort in regions where the races are not found (76, 77). This is facilitated through shuttle breeding by moving cycles of selection between locations where the disease/races are present and locations that represent the environment(s) targeted by the breeding program (77). Also, resource intensive screening can be conducted in quarantined pathogen facilities, such as the USDA-ARS BL3 (United States Department of Agriculture–Agricultural Research Service Biosafety Level 3) Cereal Disease Laboratory (53). Although these types of approaches are effective, they ultimately impose a limit on the size of breeding populations that can be screened. Owing to the limitations and challenges in evaluating breeding lines for resistance to exotic or quarantined diseases, genomics-assisted breeding becomes an increasingly attractive methodology.

Breeding for Multiple Disease Resistance

Likewise, breeding for multiple diseases presents a challenge for breeders (8). Implementation of a single disease nursery that gives effective screening for the suite of important pathogens can be difficult, and high pressure from multiple diseases can confound scoring and selection. Selection for any set of diseases that would require multiple locations or replications is not possible until later in the breeding cycle when inbred lines are developed. For crops and target locations where multiple disease resistance is important, molecular breeding approaches become a tractable approach for identifying superior lines.

Molecular Breeding for Disease Resistance

In the scope of using molecular breeding for improving disease resistance, there has been considerable effort focused on the identification of markers linked with major genes and mapping quantitative trait loci (QTLs) for disease resistance. With molecular markers for R genes, direct selection with these markers for disease resistance can be implemented in the breeding programs. Many hundreds of R genes have been mapped across important crops, including rice (e.g., 10, 82, 86), wheat (e.g., 42, 57, 73), maize (13, 14, 36, 43, 60, 64, 87), soybean (e.g., 4, 15), and potato (e.g., 3, 81) as well as numerous other crop pathosystems (e.g., 39, 67, 92). Work in *Arabidopsis thaliana* has identified many genes and provided a strong fundamental foundation in molecular plant-pathogen interactions.

However, given the economic importance of these diseases and the overall effort invested in genetic markers, there are relatively few examples of large-scale implementation of MAS for disease resistance in applied breeding programs. As noted by Miedaner & Korzun (51) in relation to

wheat and barley breeding, the lack of markers applied in commercial breeding for disease resistance could be due to having few diagnostic markers, the prevalence of QTL background effects, and overall economic constraints from a low return of investment for implementing markers. Furthermore, few monogenic resistances are durable, and only a few QTLs with high effects have been successfully transferred into elite breeding material, limiting the practical use of MAS for R genes (51).

Following mapping and implementation of markers for multiple R genes, pyramiding of selected genes is possible (e.g., 34, 61, 78). However, the practical implementation of MAS for stacking multiple disease resistance genes adds an additional layer of complexity to applied breeding programs, as the population size needed for maintaining and fixing multiple genes quickly outgrows the reasonable available resources for MAS. For example, in a simple F_2 population, an expected 25% of the lines would be fixed for the favorable allele at any given locus and only 1.5% of the plants would be fixed for all favorable alleles at three different loci. It is possible with a reasonable population size of a few hundred to identify plants through MAS with a three-gene pyramid. However, this must be put in the context of breeding programs in which breeders want to have many hundreds of lines for advancement to yield testing. It would take a population size of 10,000 with MAS to identify 150 lines carrying a three resistance gene pyramid that could be advanced for yield testing, which is still far too few for making progress for yield. Although the probability in the F_2 increases to 42% if just maintaining the favorable allele combination (in a heterozygous or homozygous condition), further rounds of marker testing are needed in subsequent generations.

Backcross introgression is another option for combining multiple R genes into an elite background. Although varieties have been developed using this approach (5, 75, 79), it often fails because by the time the backcrossing is completed, the recurrent parent is no longer one of the highest yielding lines (91). Thus, this strategy works best if the backcrossing can be done very quickly, possibly requiring foreground and background selection, or if the market demands a specific variety that is very difficult to improve upon, such as the case with malting barley.

GENOMIC SELECTION

Genomic selection (GS) was first proposed by Meuwissen et al. (50) as an approach to capture the total additive genetic variance using genome-wide molecular markers and to apply molecular breeding for difficult quantitative traits. This approach represented a convergence between pedigree predictions using traditional estimates of relationships between individuals and the prediction of genetic effects using molecular markers. GS extends from currently implemented methods for molecular breeding of complex traits in that the GS prediction models forgo significance testing and simultaneously estimate the effect of all markers. Contrasting to MAS, the use of whole-genome prediction models generally has greater power to capture small-effect loci that would be missed by MAS because of limited power for declaring significant marker effects (29).

In the scope of molecular breeding, the preferred and best performing models for prediction using markers can be thought of as following a continuum based on the genetic architecture of the trait. On one end of the spectrum a very simple genetic architecture in which a trait is conditioned by a single gene of large effect and the other end represented by highly polygenic traits that are controlled by hundreds or thousands of genes each with very small effects. GS is ideal for complex traits with lower heritability and a complex genetic architecture. With this genetic architecture, whole-genome prediction models as employed in GS have consistently shown superior prediction accuracy over models that use significance testing to first identify individual markers and then model the effect of these markers (46).

GS using only molecular information prior to phenotyping could be especially useful for increasing the rate of genetic gain by reducing the breeding cycle time, increasing the selection intensity, and possibly increasing selection accuracy. GS as proposed by Meuwissen (50) and reviewed in plants by Lorenz et al. (46) involves predicting the total breeding values of selection candidates based on genome-wide markers. To implement GS, a training population consisting of breeding lines phenotyped for target traits and genotyped with genome-wide markers is used to train a prediction model. That prediction model is then applied to a new set of selection candidates that have been genotyped with genome-wide markers. One of the key features of GS is that all markers, rather than a subset of markers deemed to be significant, are used for prediction modeling (50). By using all markers, GS attempts to capture all the additive genetic variance for the traits of interest. As a result, GS is more accurate than MAS approaches that capture only a portion of the total genetic variance (26, 27, 35). GS can be used to predict breeding values for candidates with or without phenotypic information of their own. In the former case, GS can increase the selection accuracy and in the latter case GS can decrease the breeding cycle length because phenotypic data on candidates are no longer needed prior to selection. It is also possible to predict breeding values for a very large number of candidates that could not be feasibly evaluated phenotypically, thus leading to increased selection intensity. GS can also be applied at multiple stages in the breeding process to increase R_{total}.

Models for Genomic Selection

There are a variety of statistical models used to estimate breeding values in GS. Some are considered suitable for highly quantitative traits, whereas others are better suited for traits that fall between qualitative and quantitative inheritance. Two of the most commonly used models for purely quantitative traits are genomic best linear unbiased prediction (G-BLUP) and ridge-regression BLUP (RR-BLUP). G-BLUP is a mixed linear model, with individuals as random effects, and the covariance among individuals is assumed to be proportional to the genomic relationship matrix estimated with genome-wide markers. This leads to the utilization of information from relatives for breeding value prediction, with closer relative information weighted more heavily. G-BLUP is a modification of the conventional BLUP model (31), which uses pedigree relationships rather than genomic relationships. When using G-BLUP for prediction, both the model-training and validation individuals are included in the relationship matrix, but only the model-training individuals have phenotypic data. RR-BLUP is also a mixed linear model, but markers are considered random effects (88). Covariance between markers is considered to be zero, and the marker variance is assumed to be the total genetic variance divided by the number of markers. This assumes that variance is equal for all markers, which enables many more marker effects to be estimated than there are phenotypic records. To avoid overfitting, creating a model that fits the test data perfectly but performs poorly when applied to another data set, RR-BLUP imposes shrinkage on the marker effect estimates, which pushes marker estimates closer to zero and reduces variance of the estimates. To use RR-BLUP for prediction, the model is fit using the training set to estimate marker effects and then the marker effect estimates are multiplied by the marker genotypes of the validation individuals to generate predicted breeding values. RR-BLUP is equivalent to G-BLUP (22), which indicates that the marker effects with RR-BLUP are actually capturing genomic relationships. Both RR-BLUP and G-BLUP assume that the trait is conditioned by an infinitesimally large number of additive loci. Although not true, this assumption is the foundation for many quantitative genetics principles and works quite well for prediction. However, in cases in which there is a large-effect locus or loci, assuming unique marker variances can lead to better prediction accuracies.

Genomic Selection with Major Genes

When there is a large-effect locus present, a simple way to allow it to have a unique variance is to consider it as a fixed effect in G-BLUP or RR-BLUP (6). This has been shown to increase prediction accuracy for quantitative stem rust resistance in wheat (72), and, based on simulation, modeling a locus as fixed is expected to improve accuracy when the locus explains more than 10% of the variation (6). Other suitable models for traits that fall between quantitative and qualitative inheritance are the Bayesian models BayesA, BayesB (50), BayesCπ (23), and Bayesian LASSO (55). With BayesA, each marker is assumed to have a unique variance. BayesB is an extension of BayesA and allows some markers to have no effect. Because Bayesian models use a different approach for parameter estimation compared with mixed linear models, they are able to estimate unique marker variances when there are fewer phenotypic records than markers. BayesCπ assumes all markers have an equal variance, but a certain proportion of markers are assumed to have no effect. In Bayesian LASSO, markers are assumed to have an equal variance, but shrinkage of marker effects is more severe on small-effect markers, whereas larger-effect loci shrink less. These models essentially lead to variable selection, as some loci are estimated to have near zero effect. Most studies comparing GS models for quantitative disease resistance in crops have found that RR-BLUP performs about as well as Bayesian LASSO (47, 68, 72) and BayesCπ (47, 52, 72, 74). In one study of GS for wheat rust resistance, Bayesian LASSO was found to perform slightly better than RR-BLUP (54), whereas a study of GS for Fusarium head blight (FHB) resistance in wheat found that Bayesian LASSO performed poorly compared with RR-BLUP (1).

Thus far, the models discussed have the underlying assumption that the genetic variance is additive. This is key for prediction of breeding value, which is by definition the sum of the additive loci effects. However, in some instances it may be useful to predict total genetic value, and in these cases models that can capture nonadditive effects are useful. Two such models are Reproducing Kernel Hilbert Space (RKHS) (21) and Random Forest (RF) (9). RKHS is similar to G-BLUP, but the relationship matrix is defined as the squared Euclidean distance divided by a scale parameter that influences how quickly the genetic covariance decays with distance. This enables more weight placed on more related individuals. Because closer relatives are known to share more nonadditive genetic variance, this leads to capturing nonadditive effects for prediction. RF is a machine-learning algorithm, unlike any other model yet discussed. With RF, a series of regression trees are grown using the model-training set. The trees are then applied to new data, and results are averaged to generate a final prediction. The tree structure of the prediction model essentially allows the effect of markers to vary depending on the genotypes of other markers, which implicitly captures nonadditive effects. Some studies of GS for quantitative disease resistance have found either RF or RKHS to lead to higher prediction accuracies compared with linear models (52, 68).

Selection Accuracy and Implementation of Genomic Selection

Factors that affect the accuracy of GS include the heritability of the trait, the rate of linkage equilibrium (LD) decay, the marker density, and the number of individuals in the training population (16). When LD decays more rapidly with physical distance, then a greater number of markers and individuals for model training are needed. When predicting across populations, as in the case where previous breeding candidates are used for model training to predict new selection candidates, the relationship between the model-training population and the selection candidates is important (62). To maintain a close relationship between the training and selection candidate individuals, the GS prediction model should be updated each time new phenotypic data are available.

GS accuracy is calculated using validation studies in which a set of individuals is used to train the model and a separate set of individuals is predicted. There are different types of validation methods, the most common are k-fold cross-validation, resampling-based cross-validation, forward validation, and progeny validation. In k-fold cross validation a data set is randomly split into k number of folds. Then prediction models are fit to k times, each time leaving out the kth fold, which serves as the validation set. In resampling-based cross-validation, a random sample of individuals is selected and used for model training, and the prediction model is then used to predict a specific validation set or the remaining individuals not sampled for model training. Forward validation attempts to mimic a breeding program scenario by training the model with data from a series of years and then predicting a new set of individuals observed in a new year. Progeny validation uses one generation for model training, and the progeny of that generation for validation. For each of the validation methods, predictive ability is estimated as the Pearson's correlation (r) between the breeding values estimated using phenotype and the breeding values predicted from GS. The prediction accuracy can then be calculated by dividing the predictive ability by the square root of the line mean heritability (r/h) in the validation set. For k-fold and resampling based cross-validation, average predictive ability or prediction accuracy is reported. GS prediction accuracy is a useful metric because it can be used directly in the gain from selection formula; however, many report predictive ability when estimates of line-mean heritability are thought to be inaccurate.

A GS breeding scheme amenable for line development (**Figure 1**) may be similar to a typical line development breeding scheme. However, GS could be applied for selection for quantitative disease resistance and all other target traits among plants or lines at the F_4 or F_4:F_5 stage (or DH), for example. If phenotypic information for a given individual is available for some traits at the time of GS, that information can be integrated in the prediction model for better prediction accuracy. For example, if quantitative disease resistance data can be measured on the F_4 plants or F_4:F_5 lines, then that information can be used in the prediction model. Thus, GS need not be purely based on marker-only predictions. Using both markers and available phenotypes can increase selection accuracy compared with GS based on markers only or purely phenotypic selection, especially if measured traits are correlated with or are themselves traits of interest. After selection of new parents, the F_4-derived lines can be phenotyped for all target traits to update the prediction model. For traits that are expensive to phenotype, an optimal subset of individuals can be selected for model updating. In the case of quantitative disease resistance, individuals with major race-specific loci should be excluded from the model-training set because these loci interfere with phenotyping quantitative resistance by masking effects of other resistance loci.

GENOMIC SELECTION FOR DISEASE RESISTANCE

Over the past several years, there have been a number of studies demonstrating the utility of whole-genome prediction models and GS approaches for disease resistance in crops (**Table 1**). Overall, these studies have demonstrated the effectiveness of the current GS models to capture and predict the genetic variation for disease resistance, particularly quantitative disease resistance.

Genomic Selection for Wheat Rusts

Perhaps one of the best-studied pathosystems for using GS models in disease resistance breeding is rusts in wheat (*Triticum aestivum* L.). Broadly grouped together, the major rust pathogens of wheat include *Puccinia graminis* (stem rust), *Puccinia striiformis* (yellow/stripe rust), and *Puccinia triticina* (leaf/brown rust) and are obligate biotrophic fungal pathogens that are endemic throughout the wheat-growing world and historically have been the most severe pathogens of wheat.

Table 1 Compiled publications demonstrating genomic selection for disease resistance

Reference	Pathosystem[a]	Pheno[b]	h^2	Validation type	*r/h* (*r*)[c]
Ornella et al. 2012 (54)	Stem rust/wheat	Severity	0.72–0.9	Cross-validation within full-sib family	(0.39–0.85)
				Cross-validation across family	(0.14–0.67)
	Yellow rust/wheat	Severity	0.45–0.66	Cross-validation within full-sib family	(0.14–0.63)
				Cross-validation across family	(0.14–0.63)
Rutkoski et al. 2014 (72)	Stem rust/wheat	Severity	0.82	k-fold cross-validation	(0.56–0.62)
Rutkoski et al. 2015 (70)	Stem rust/wheat	Severity	0.61	Forward validation	(0.2–0.4)
				Cross-validation	(0.55)
Lorenz et al. 2012 (47)	FHB/barley	Severity	0.44–0.56	Forward validation	0.41–0.68
		DON	0.55–0.76	Forward validation	0.47–0.76
Rutkoski et al. 2012 (68)	FHB/wheat	Severity	–	k-fold cross-validation	(0.59–0.64)
		DON	–	k-fold cross-validation	(0.19–0.41)
Sallam et al. 2015 (74)	FHB/barley	Severity	~0.35–0.7	Progeny validation	~0.2–1
		DON	~0.5–0.65	Progeny validation	~0.75–0.85
Jiang et al. 2014 (37)	FHB/wheat	FHB-index	0.91	k-fold cross-validation	0.68–0.74
Mirdita et al. 2015 (52)	FHB/wheat	Severity	0.4–0.8	k-fold cross-validation	(~0.46–0.64)
	Septoria tritici blotch/wheat	Severity	0.2–0.6	k-fold cross-validation	(~0.36–0.59)
Arruda et al. 2015 (1)	FHB/wheat	Severity	–	k-fold cross-validation	0.4–0.48
		DON	–	k-fold cross-validation	0.53–0.64
Technow et al. 2013 (83)	NCLB/maize	Severity	–	Cross-validation	0.64–0.71
			–	Cross-validation across heterotic group	0.11–0.29
Riedelsheimer et al. 2013 (65)	Gibberella ear rot/maize	Severity	0.7–0.83	Cross-validation within full-sib family	~0.65–0.7
				Cross-validation across family	~-0.25–0.6
		DON	0.64–0.8	Cross-validation within full-sib family	~0.65–0.7
				Cross-validation across family	~-0.05–0.7
Ly et al. 2013 (48)	Cassava mosaic disease	Severity	0.63	k-fold cross-validation	0.503
	Cassava anthracnose disease	Severity	0.17	k-fold cross-validation	0.17

[a]The crops disease pathosystem in the reported study.
[b]The disease phenotype measured and modeled.
[c]The observed prediction accuracy as reported by correlation divided by square root of heritability or in parenthesis as the direct correlation.
Abbreviations: DON, deoxynivalenol; FHB, Fusarium head blight; NCLB, Northern corn leaf blight.

Resistance to wheat rusts can be either qualitative or quantitative, with numerous race-specific R genes identified in these pathosystems (77). The R genes can be detected in the seedling stage, whereas quantitative resistance loci generally confer resistance only in the adult plants, and quantitative resistance is therefore also referred to as adult plant resistance (APR). APR is often based on minor effect loci that have not been shown to be race specific. APR has been shown to vary to some degree across environments, which could be due to race, temperature, or other unknown environmental factors. However, heritability of APR across environments is usually high. Although resistance breeding efforts using major R genes for rust have produced highly resistant varieties, these varieties are often short lived as they are subject to the boom-and-bust cycles of deploying race-specific R genes (77). Therefore, considerable effort has focused on breeding minor-gene (quantitative) resistance as a deployment strategy for durable resistance. In addition to being race nonspecific, the quantitative resistance to rust can sometimes be effective against more than one rust species or other biotrophic pathogens (32, 40, 42, 89).

Implementing GS for disease resistance was proposed in wheat rust by Rutkoski et al. (71) with the scope of reducing cycle time by up to twofold while facilitating pyramiding of major genes with APR genes. Ornella et al. (54) found that whole-genome prediction accuracy using biparental populations was correlated to the broad-sense heritability of the disease screening. They observed moderate to high prediction accuracies ranging from 0.3 to 0.8 and were able to predict both within and across environments with comparable accuracy, indicating that $G \times E$ for predicting rust resistance was minimal in these populations, likely because the resistance was mostly race nonspecific. However, the accuracy for predictions in new populations was considerably lower, confirming previous studies that relatedness of the training population and selection candidates has a strong influence on the accuracy (24, 47). This study also found that linear models outperformed nonlinear models, and Bayesian LASSO led to only slightly higher prediction accuracy compared to RR-BLUP. Rutkoski and colleagues further advanced the work on predicting stem rust resistance while evaluating realized gain from GS (69, 70, 72). They found that GS could be implemented as an effective breeding method for quantitative stem rust resistance even with a relatively small training population. They also found that G-BLUP, Bayesian LASSO, and Bayes Cπ led to very similar accuracies, but incorporating previously discovered large-effect loci as fixed effects in G-BLUP led to the highest prediction accuracy overall (6, 72). Based on this, it has been proposed that for disease resistance in which the genetic architecture is often a combination of a few large-effect QTLs combined with a polygenic background, these types of mixed models using gene information can be valuable.

Genomic Selection for Fusarium Head Blight in Wheat

There are also a number of examples of exploring GS models for predicting FHB in wheat. FHB is a serious disease throughout the world, particularly in regions with cropping systems in rotation with maize and high humidity and moisture through heading and maturity. It is primarily caused by the fungal pathogen, *Fusarium graminearum* Schwabe, which infects the spikes of wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.). The disease results in yield loss and reduction of grain quality (17). More seriously, the infections lead to shriveled and discolored grain and contamination from the mycotoxin deoxynivalenol (DON), resulting in the whole crop being unusable (58). From a number of classic genetic studies, it is known that resistance to FHB in wheat is quantitatively inherited (45, 85) and that genetic variation for FHB resistance is predominantly additive (2, 80), indicating that accumulation of resistance genes may be possible. As described above, this combination of quantitatively inherited

resistance and a challenging, complex phenotype makes resistance to FHB an excellent target for GS approaches.

Rutkoski et al.'s (68) first report evaluating GS models for FHB resistance in wheat found that various prediction models have accuracy high enough to be useful in breeding, GS models always significantly outperform MLR, and the nonlinear models RHKS and RF consistently lead to higher accuracies compared with linear models. This study also examined the benefit of genotyping FHB QTL-linked markers in addition to whole-genome profiling and found that QTL-targeted genotyping significantly improved prediction accuracy for DON concentration, whereas for other resistance traits there was no improvement. The authors also found that accuracy for DON concentration could be further improved by including correlated trait data on the selection candidates in the prediction model.

Arruda et al. (1) also evaluated RR-BLUP as well as a nonlinear and a variable selection model on a diverse set of breeding lines for FHB. They confirmed that the various traits associated with FHB resistance, including infection incidence and severity and DON levels, could be predicted with moderate to high accuracy. However, in contrast to the previous study by Rutkoski et al. (68), Arruda et al. (1) found that simple RR-BLUP with genome-wide markers led to high accuracies for DON and outperformed other models tested for all resistance traits. They also confirmed that prediction accuracies using genomic information were significantly higher than when using only pedigree-based information. Mirdita et al. (52) evaluated prediction accuracies for FHB resistance using RR-BLUP, Bayes Cπ, RKHS, and an additional nonadditive model. The authors reported that compared to the linear models, nonlinear models led to a 10% improvement in accuracy.

Jiang et al. (37) compared prediction accuracies using three different marker platforms at varying densities and three different prediction models, including RR-BLUP, BayesCπ, and RKHS. The authors found that marker density had only a small impact on accuracy, and there was no difference between accuracies from the prediction models tested. Overall, half of the studies of GS for FHB resistance have found that the simple, linear RR-BLUP model performs as well as other more complex models, whereas the other half have found nonlinear models to perform best, indicating that for FHB resistance the choice of prediction model should depend on the genetic background.

Genomic Selection for Leaf Blight in Maize

Northern leaf blight (NLB) of maize, caused by *Setosphaeria turcica*, is a pathosystem with genetic architecture and complexities, including both major R-gene and quantitative resistance similar to wheat rusts. NLB is a hemibiotrophic fungus with resistance primarily conferred by quantitative genes of small effect, but a number of large-effect, race-specific loci are known (90). In addition, a number of studies have implicated numerous QTLs with multiple alleles in the complex genetic architecture of NLB resistance (60). As such, resistance breeding for NLB is another good target for GS approaches.

Using a standard G-BLUP model, Technow et al. (83) observed prediction accuracies across different heterotic groups up to 0.706 (dent genotypes) and 0.690 (flint genotypes). Finding that the genetic background had little effect, they were able to combine training sets from the different heterotic groups to increase accuracy. It was observed that the cycle time for GS needed to be only 80% or less of the phenotypic cycle to have superior gains from implementing GS. With the ability to run off-season nurseries and generate marker profiles on seedlings, this is very obtainable in a commercial breeding program.

CURRENT AND FUTURE PROSPECTS

Potential of Genomics-Assisted Breeding for Disease Resistance

Given the rapid escalation of overall molecular breeding and GS approaches in breeding programs, there is great potential for applying GS for disease resistance across a range of crops and pathosystems. An important consideration for applying GS as a breeding methodology for disease resistance is recognizing that yield remains the primary objective and that GS for yield is the primary driver of generating whole-genome molecular profiles. It is significant to note that once the genotyping is completed, prediction models can be applied for no additional cost to any trait for which sufficient phenotypic data are available on the training sets. Therefore, the economic considerations weighing the direct costs of phenotypic selection versus GS of disease resistance are greatly diminished.

One of the major benefits to applying GS for quantitative disease resistance breeding is that more breeding candidates can be evaluated, leading to higher selection intensity. With phenotyping selection, there is a limit to the number of entries one can measure in an inoculated disease nursery. Typically, there are few locations where one can reliably create artificial epidemics for a given disease, and inoculation can be labor intensive. This limits the number of entries that a given breeding program can screen, unless a substantial investment in infrastructure and staff is made, and, even so, land in environments favorable for the disease may not be available. Given the increasing availability of whole-genome profiling, it is feasible to genotype many more individuals than can be phenotyped in a disease nursery. This is especially important for selection prior to advanced testing when there are tens of thousands of selection candidates and when disease testing is severely constrained, as in the case of exotic or quarantined diseases. The ability to select among such a large number of candidates prior to advanced testing leads to not only increased selection intensity but decreased breeding cycle time.

Decreasing breeding cycle time can have a major impact on increasing rates of genetic gain. GS can be applied to select among single plants in early generations or among lines prior to advanced testing for all traits of interest. In either case, the breeding cycle time is reduced if new parents are selected based on predictions prior to advanced testing. Although it is true that many disease resistances can be evaluated in early generations, it is important to be able to select for yield when selecting new parents, because yield is generally the most important trait. Thus, yield testing becomes the rate-limiting step in breeding that limits progress in all other traits of interest regardless of when these other traits can be measured. By applying GS, all traits can be selected upon at any stage, and yield testing is no longer the bottleneck.

Another advantage to using GS to select for quantitative disease resistance is that quantitative resistance can be selected even when R genes are present. Many R genes confer a very high level of resistance and are epistatic in that they can largely mask the effect of other resistance QTLs in the background. In fact, the use of R genes in breeding programs can lead to the vertifolia effect, which is an erosion of quantitative resistance due to the inability to phenotype for quantitative resistance in the presence of R genes (84). In many breeding populations, quantitative resistance loci and R genes are both present, and the ideal genotype may be one with at least one effective R gene along with a high level of quantitative resistance. This would lead to a very high level of resistance as long as the R gene is effective, and when the R gene is rendered ineffective, quantitative resistance in the background would still provide a good level of protection against yield loss and a pathogen population boom. With GS, quantitative resistance can be predicted in any individual or line regardless of whether R genes are present; however, it is important that the training population for quantitative resistance be devoid of major R genes.

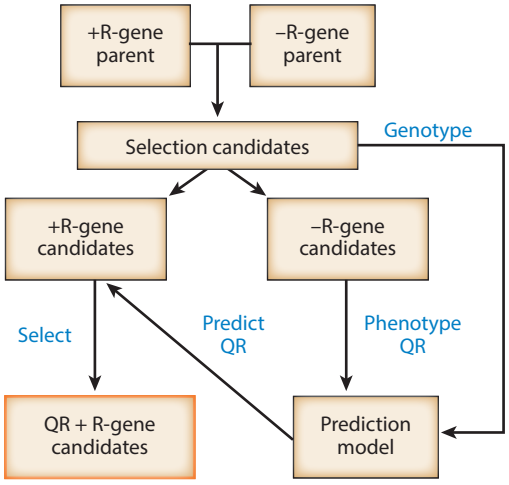


Figure 2

A genomic selection scheme for combining resistance (R) genes and quantitative resistance (QR). First, a parent positive (+) for a race-specific R gene is crossed with a parent absent (−) for the same R gene to generate a segregating population of selection candidates. The selection candidates are genotyped and split into a +R-gene group and a −R-gene group. The −R-gene candidates are phenotyped for QR, and their phenotypic and genotypic data are used for genomic prediction model training. The prediction model is then applied to the +R-gene candidates that were genotyped previously. Candidates with high levels of QR in combination with the R gene (QR + R-gene candidates) are then selected for use in crossing or further testing.

The implementation of genomics-assisted breeding for disease resistance opens opportunities for combined selection of both R-gene and quantitative resistance (**Figure 2**). The process begins with crossing two parents that segregate for R gene–based resistance to generate a population of selection candidates. The selection candidates are then genotyped with genome-wide markers. The selection candidates without effective R genes are then phenotyped for quantitative resistance. The phenotypes and genotypes are used to train the model, which is then used to predict the level of quantitative resistance in the individuals with the effective R genes. The individuals with R genes with the highest level of quantitative resistance are then selected for advancement and/or for crossing.

Selection for quarantined pathogens also presents a unique challenge for both phenotypic selection and GS because phenotyping can only occur in countries or states where the disease or pathogen strain has already been reported, or in expensive biocontainment facilities. For phenotypic selection, seed of the selection candidates would need to be imported to a specific region for evaluation; for GS, seed of the training population would need to be imported for evaluation. This may lead to an increased cost of phenotyping and a decreased phenotyping capacity. For GS, an optimal subset of training individuals could be selected so that prediction accuracy is maximized, given a fixed phenotyping capacity (66, 69). Then a large number of selection candidates could be predicted and selected using a training population that fits within the phenotyping constraints.

Challenges in Genomic Selection for Disease Resistance

High heritability makes GS more accurate but also less competitive with phenotypic selection. One of the major challenges with GS for disease resistance is that many disease resistances are

highly heritable, which makes phenotypic selection hard to beat in both per-cycle and per-unit genetic gain (28, 70). In order for GS to outperform phenotypic selection for quantitative disease resistance, it may be necessary to increase the selection intensity in addition to decreasing cycle time. We expect that this will be possible in most breeding programs because when selecting at earlier stages there are naturally more selection candidates that have not yet been culled. In some GS breeding schemes, genomic prediction models may include both phenotype of the candidates per se and the phenotypes of the candidates' relatives. For low-heritability traits, using genotype in addition to phenotype can substantially improve selection accuracy (63). However, for high-heritability traits, such as many disease resistances, there may very little benefit. The direct cost-to-benefit ratio of genome profiling (or markers) may not outweigh phenotypic selection for disease resistance; however, as noted, the genomic profile will be already generated for predicting yield and present no additional cost for the genotypic data needed to predict disease resistance.

CONCLUSIONS

With the implementation of routine whole-genome profiling in breeding programs, prediction models for any trait can be applied for selection and advancement. The primary driver for implementing GS in a breeding program is selection for yield, to offset and reduce the time and costs associated with yield testing. However, any successful variety must have a complete package of disease resistance, agronomic performance, and end-use quality.

Within the breeding program, disease screening nurseries can generally be run very efficiently with single plants or on single rows and with large numbers of entries, making the cost per line relatively low. In a direct comparison, the cost of disease phenotyping will likely be lower than whole-genome genotyping. However, in contrast to phenotypic measurements that are relevant only to the trait, the single set of genotypic information can be used for any and all traits. This quickly amortizes the cost of genotyping, leaving only a marginal cost to be offset by any single trait. Therefore, viewed in the scope of a holistic breeding program and objectives, there are only marginal costs associated with predicting disease resistance. As such, GS for disease resistance becomes a very tractable implementation to a breeding program.

There has been considerable effort and focus on genetically mapping important disease resistance genes and identifying markers for MAS. However, for many genes for many diseases, MAS quickly becomes very complex and intractable in an applied breeding program, leading to bottlenecks in population sizes and challenges in practical implementation. Designing an array for MAS in many traits likewise quickly turns into a whole-genome profile.

Overall, there is a strong foundation for further implementation of genomics-assisted breeding for disease resistance. A quickly growing number of R genes are being identified, cloned, and characterized. This refined biological information can be incorporated into whole-genome prediction models to simultaneously select for gene pyramids in a background with high levels of quantitative resistance. A number of studies have demonstrated that current whole-genome prediction models can be effectively applied to predict and select for quantitative disease resistance. Reduced disease phenotyping is unlikely to directly offset the cost associated with genotyping; however, the genotyping cost can be offset by the sum of all traits that are targeted in the breeding program. As breeders implement GS to decrease the breeding cycle time, increase selection intensity, and increase the overall rate of gain, they will continue to select for the whole suite of traits that compose a successful variety. Disease resistance will need to be one of the key traits that is predicted for selection.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Arruda MP, Brown PJ, Lipka AE, Krill AM, Thurber C, Kolb FL. 2015. Genomic selection for predicting head blight resistance in a wheat breeding program. *Plant Genome* doi:10.3835/plantgenome2015.01.0003
2. Bai G-H, Shaner G, Ohm H. 2000. Inheritance of resistance to *Fusarium graminearum* in wheat. *Theor. Appl. Genet.* 100:1–8
3. Ballvora A, Ercolano MR, Weiß J, Meksem K, Bormann CA, et al. 2002. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBL/LRR class of plant resistance genes. *Plant J.* 30:361–71
4. Bao Y, Vuong T, Meinhardt C, Tiffin P, Denny R, et al. 2014. Potential of association mapping and genomic selection to explore PI 88788 derived soybean cyst nematode resistance. *Plant Genome* doi: 10.3835/plantgenome2013.11.0039
5. Belcher AR, Zwonitzer JC, Santa Cruz J, Krakowsky MD, Chung C-L, et al. 2012. Analysis of quantitative disease resistance to southern leaf blight and of multiple disease resistance in maize, using near-isogenic lines. *Theor. Appl. Genet.* 124:433–45
6. Bernardo R. 2014. Genomewide selection when major genes are known. *Crop Sci.* 54:68–75
7. Bonnett D, Rebetzke G, Spielmeyer W. 2005. Strategies for efficient implementation of molecular markers in wheat breeding. *Mol. Breed.* 15:75–85
8. Boyd LA, Ridout C, O'Sullivan DM, Leach JE, Leung H. 2013. Plant-pathogen interactions: disease resistance in modern agriculture. *Trends Genet.* 29:233–40
9. Breiman L. 2001. Random forests. *Mach. Learn.* 45:5–32
10. Bryan GT, Wu K-S, Farrall L, Jia Y, Hershey HP, et al. 2000. A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *Plant Cell Online* 12:2033–45
11. Bulmer M. 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–11
12. Burgueno J, de los Campos G, Weigel K, Crossa J. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Sci.* 52:707–19
13. Carson ML, Stuber CW, Senior ML. 2004. Identification and mapping of quantitative trait loci conditioning resistance to Southern leaf blight of maize caused by *Cochliobolus heterostrophus* race O. *Phytopathology* 94:862–67
14. Collins NC, Webb CA, Seah S, Ellis JG, Hulbert SH, Pryor A. 1998. The isolation and mapping of disease resistance gene analogs in maize. *Mol. Plant-Microbe Interact.* 11:968–78
15. Cregan PB, Mudge J, Fickus EW, Danesh D, Denny R, Young ND. 1999. Two simple sequence repeat markers to select for soybean cyst nematode resistance coditioned by the rhg1 locus. *Theor. Appl. Genet.* 99:811–18
16. Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA. 2010. The impact of genetic architecture on genome-wide evaluation methods. *Genetics* 185:1021–31
17. Dexter J, Clear R, Preston K. 1996. Fusarium head blight: effect on the milling and baking of some Canadian wheats. *Cereal Chem.* 73:695–701
18. Falconer DS, Mackay TFC. 1996. *Introduction to Quantitative Genetics*. Upper Saddle River, NJ: Pearson Educ. 4th ed.
19. Forster BP, Thomas WT. 2005. Doubled haploids in genetics and plant breeding. *Plant Breed. Rev.* 25:57–88
20. Frisch M, Melchinger AE. 2005. Selection theory for marker-assisted backcrossing. *Genetics* 170:909–17
21. Gianola D, Van Kaam JB. 2008. Reproducing Kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178:2289–303
22. Habier D, Fernando RL, Dekkers JCM. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177:2389–97

23. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinform.* 12:186
24. Hayes B, Bowman P, Chamberlain A, Verbyla K, Goddard M. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:51
25. Hazel LN. 1943. The genetic basis for constructing selection indexes. *Genetics* 28:476–90
26. Heffner EL, Jannink J-L, Iwata H, Souza E, Sorrells ME. 2011. Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 51:2597–606
27. Heffner EL, Jannink J-L, Sorrells ME. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *Plant Genome* 4:65–75
28. Heffner EL, Lorenz AJ, Jannink J-L, Sorrells ME. 2010. Plant breeding with genomic selection: gain per unit time and cost. *Crop Sci.* 50:1681–90
29. Heffner EL, Sorrells ME, Jannink J-L. 2009. Genomic selection for crop improvement. *Crop Sci.* 49:1–12
30. Henderson CR. 1963. Selection index and expected genetic advance. *Stat. Genet. Plant Breed.* 982:141–63
31. Henderson CR. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–47
32. Herrera-Foessel SA, Lagudah ES, Huerta-Espino J, Hayden MJ, Bariana HS, et al. 2011. New slow-rusting leaf rust and stripe rust resistance genes *Lr67* and *Yr46* in wheat are pleiotropic or closely linked. *Theor. Appl. Genet.* 122:239–49
33. Hillel J, Schaap T, Haberfeld A, Jeffreys A, Plotzky Y, et al. 1990. DNA fingerprints applied to gene introgression in breeding programs. *Genetics* 124:783–89
34. Huang N, Angeles E, Domingo J, Magpantay G, Singh S, et al. 1997. Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and pcr. *Theor. Appl. Genet.* 95:313–20
35. Jannink J-L, Lorenz AJ, Iwata H. 2010. Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genom.* 9:166–77
36. Jenkins MT, Robert AL. 1952. Inheritance of resistance to the leaf blight of corn caused by *Helminthosporium turcicum*. *Agron. J.* 44:136–40
37. Jiang Y, Zhao Y, Rodemann B, Plieske J, Kollers S, et al. 2015. Potential and limits to unravel the genetic architecture and predict the variation of Fusarium head blight resistance in European winter wheat (*Triticum aestivum* L.). *Heredity* 114:318–26
38. Johnson R. 1992. Past, present and future opportunities in breeding for disease resistance, with examples from wheat. *Euphytica* 63:3–22
39. Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG. 1994. Isolation of the tomato *cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266:789–93
40. Juliana P, Rutkoski JE, Poland JA, Singh RP, Murugasamy S, et al. 2015. Genome-wide association mapping for leaf tip necrosis and pseudo-black chaff in relation to durable rust resistance in wheat. *Plant Genome* 8:1–12
41. Kelly AM, Smith AB, Eccleston JA, Cullis BR. 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Sci.* 47:1063–70
42. Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, et al. 2009. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323:1360–63
43. Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, et al. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* 43:163–68
44. Lindhout P. 2002. The perspectives of polygenic resistance in breeding for durable disease resistance. *Euphytica* 124:217–26
45. Liu S, Abate Z, McKendry A. 2005. Inheritance of Fusarium head blight resistance in the soft red winter wheat Ernie. *Theor. Appl. Genet.* 110:454–61
46. Lorenz AJ, Chao S, Asoro FG, Heffner EL, Hayashi T, et al. 2011. Genomic selection in plant breeding: knowledge and prospects. *Adv. Agron.* 110:77
47. Lorenz AJ, Smith KP, Jannink J-L. 2012. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. *Crop Sci.* 52:1609–21
48. Ly D, Hamblin M, Rabbi I, Melaku G, Bakare M, et al. 2013. Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection: a study in cassava. *Crop Sci.* 53:1312–25

49. Mackay TF. 2001. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* 35:303–39
50. Meuwissen THE, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–29
51. Miedaner T, Korzun V. 2012. Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* 102:560–66
52. Mirdita V, He S, Zhao Y, Korzun V, Bothe R, et al. 2015. Potential and limits of whole genome prediction of resistance to Fusarium head blight and Septoria tritici blotch in a vast central European elite winter wheat population. *Theor. Appl. Genet.* 128:2471–81
53. Olivera Firpo P, Newcomb M, Szabo L, Rouse MN, Johnson JL, et al. 2015. Phenotypic and genotypic characterization of race TKTTF of *Puccinia graminis f. sp. Tritici* that caused a wheat stem rust epidemic in Southern Ethiopia in 2013/14. *Phytopathology* 105:917–28
54. Ornella L, Singh S, Perez P, Burgueño J, Singh R, et al. 2012. Genomic prediction of genetic values for resistance to wheat rusts. *Plant Genome* 5:136–48
55. Park T, Casella G. 2008. The Bayesian lasso. *J. Am. Stat. Assoc.* 103:681–86
56. Parlevliet JE. 2002. Durability of resistance against fungal, bacterial and viral pathogens; present situation. *Euphytica* 124:147–56
57. Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, et al. 2013. The gene *Sr33*, an ortholog of barley *mla* genes, encodes resistance to wheat stem rust race *Ug99*. *Science* 341:786–88
58. Pestka JJ, Smolinski AT. 2005. Deoxynivalenol: toxicology and potential effects on humans. *J. Toxicol. Environ. Health Part B* 8:39–69
59. Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. 2009. Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* 14:21–29
60. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *PNAS* 108:6893–98
61. Pradhan SK, Nayak DK, Mohanty S, Behera L, Barik SR, et al. 2015. Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice* 8:19
62. Pszczola M, Strabel T, Mulder H, Calus M. 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389–400
63. Pszczola M, Veerkamp R, De Haas Y, Wall E, Strabel T, Calus M. 2013. Effect of predictor traits on accuracy of genomic breeding values for feed intake based on a limited cow reference population. *Animal* 7:1759–68
64. Ramakrishna W, Emberton J, Ogden M, SanMiguel P, Bennetzen JL. 2002. Structural analysis of the maize *Rp1* complex reveals numerous sites and unexpected mechanisms of local rearrangement. *Plant Cell* 14:3213–23
65. Riedelsheimer C, Endelman JB, Stange M, Sorrells ME, Jannink J-L, Melchinger AE. 2013. Genomic predictability of interconnected biparental maize populations. *Genetics* 194:493–503
66. Rincent R, Laloë D, Nicolas S, Altmann T, Brunel D, et al. 2012. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics* 192:715–28
67. Ruffel S, Dussault MH, Palloix A, Moury B, Bendahmane A, et al. 2002. A natural recessive resistance gene against *Potato virus Y* in pepper corresponds to the eukaryotic initiation factor 4e (eIF4e). *Plant J.* 32:1067–75
68. Rutkoski J, Benson J, Jia Y, Brown-Guedira G, Jannink J-L, Sorrells M. 2012. Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Genome* 5:51–61
69. Rutkoski J, Singh RP, Huerta-Espino J, Bhavani S, Poland J, et al. 2015. Efficient use of historical data for genomic selection: a case study of stem rust resistance in wheat. *Plant Genome* doi: 10.3835/plantgenome2014.09.0046
70. Rutkoski J, Singh RP, Huerta-Espino J, Bhavani S, Poland J, et al. 2015. Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. *Plant Genome* doi: 10.3835/plantgenome2014.10.0074
71. Rutkoski JE, Heffner EL, Sorrells ME. 2011. Genomic selection for durable stem rust resistance in wheat. *Euphytica* 179:161–73

72. Rutkoski JE, Poland JA, Singh RP, Huerta-Espino J, Bhavani S, et al. 2014. Genomic selection for quantitative adult plant stem rust resistance in wheat. *Plant Genome* doi: 10.3835/plantgenome2014.02.0006
73. Saintenac C, Zhang W, Salcedo A, Rouse MN, Trick HN, et al. 2013. Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. *Science* 341:783–86
74. Sallam AH, Endelman JB, Jannink J-L, Smith KP. 2015. Assessing genomic selection prediction accuracy in a dynamic barley breeding population. *Plant Genome* doi: 10.3835/plantgenome2014.05.0020
75. Singh AK, Singh VK, Singh A, Ellur RK, Pandian R, et al. 2015. Introgression of multiple disease resistance into a maintainer of basmati rice CMS line by marker assisted backcross breeding. *Euphytica* 203:97–107
76. Singh RP, Hodson DP, Jin Y, Huerta-Espino J, Kinyua MG, et al. 2006. Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* 1:13
77. Singh RP, Hodson DP, Jin Y, Lagudah ES, Ayliffe MA, et al. 2015. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology* 105:872–84
78. Singh S, Sidhu JS, Huang N, Vikal Y, Li Z, et al. 2001. Pyramiding three bacterial blight resistance genes (*xa5*, *xa13* and *Xa21*) using marker-assisted selection into indica rice cultivar PR106. *Theor. Appl. Genet.* 102:1011–15
79. Singh VK, Singh A, Singh S, Ellur RK, Choudhary V, et al. 2012. Incorporation of blast resistance into "PRR78", an elite basmati rice restorer line, through marker assisted backcross breeding. *Field Crops Res.* 128:8–16
80. Snijders C. 1990. Fusarium head blight and mycotoxin contamination of wheat: a review. *Neth. J. Plant Pathol.* 96:187–98
81. Song J, Bradeen JM, Naess SK, Raasch JA, Wielgus SM, et al. 2003. Gene *RB* cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. *PNAS* 100:9128–33
82. Song W-Y, Wang G-L, Chen L-L, Kim H-S, Pi L-Y, et al. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 270:1804–6
83. Technow F, Bürger A, Melchinger AE. 2013. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3* 3:197–203
84. Vanderplank JE. 2012. *Disease Resistance in Plants*. Philadelphia: Elsevier
85. Van Ginkel M, Van Der Schaar W, Zhuping Y, Rajaram S. 1996. Inheritance of resistance to scab in two wheat cultivars from Brazil and China. *Plant Dis.* 80:863–67
86. Wang GL, Mackill DJ, Bonman JM, McCouch SR, Champoux MC, Nelson RJ. 1994. RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. *Genetics* 136:1421–34
87. Webb CA, Richter TE, Collins NC, Nicolas M, Trick HN, et al. 2002. Genetic and molecular characterization of the maize *Rp3* rust resistance locus. *Genetics* 162:381–94
88. Whittaker JC, Thompson R, Denham MC. 2000. Marker-assisted selection using ridge regression. *Genet. Res.* 75:249–52
89. William M, Singh R, Huerta-Espino J, Islas SO, Hoisington D. 2003. Molecular marker mapping of leaf rust resistance gene *Lr46* and its association with stripe rust resistance gene *Yr29* in wheat. *Phytopathology* 93:153–59
90. Wisser RJ, Balint-Kurti PJ, Nelson RJ. 2006. The genetic architecture of disease resistance in maize: a synthesis of published studies. *Phytopathology* 96:120–29
91. Xu Y, Crouch JH. 2008. Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.* 48:2
92. Zhou J, Loh Y-T, Bressan RA, Martin GB. 1995. The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by *Pto* and is involved in the hypersensitive response. *Cell* 83:925–35