Chapter 6
Current Status and Prospects of Genomic Selection in Legumes

Ankit Jain, Manish Roorkiwal, Manish K. Pandey, and Rajeev K. Varshney

6.1 Introduction

Availability of proper nutrition is of extreme importance as malnutrition at an early age may lead to reduced physical and mental development and limits the capacity to learn. UN World Food Program has reported that more than 900 million people in the world do not get nutritious food to eat. Global population has been growing at a fast pace, and feeding the ever increasing population with nutritious food is becoming more difficult day by day. This will continue until there is significant genetic gain by increasing crop productivity with enhanced nutrition. Although significant efforts have been focussing on enhancing the crop production to feed the world, still there are famines occurring in several parts of the world (http://www.latimes.com/world/africa/la-fg-southsudan-famine-20170220-story.html). Considering this alarming situation, the United Nations and other affiliated organizations have a challenge to eradicate hunger and malnutrition to ensure food and nutrition security by responding to nutritional needs, addressing emerging threats and meeting the zero hunger challenge. To overcome this devastating situation of malnutrition, legumes are expected to play significant role, and there is a dire need to enhance the productivity of these legumes.

Legumes have been cultivated since early civilizations and have been the major source of nutrition for humans and animals (Power 1987; Graham and Vance 2003; Varshney et al. 2013a; Rubiales and Mikic 2015; Pandey et al. 2016). Legumes have been recognized as most valuable food to meet the dietary requirements of undernourished or underserved global populations (Rebello et al. 2014). Research has shown that replacement of energy dense foods with legumes offers various

A. Jain • M. Roorkiwal (✉) • M.K. Pandey • R.K. Varshney
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana, India
e-mail: m.roorkiwal@cgiar.org

R.K. Varshney et al. (eds.), *Genomic Selection for Crop Improvement*,
DOI 10.1007/978-3-319-63170-7_6

health benefits (Tarawali and Ogunbile 1995). In addition, legumes have the ability to fix atmospheric nitrogen, which is vital for improving the soil nutritional profile, thereby reducing the requirement for nitrogen fertilizers enabling legumes more suited for crop rotation programs.

Legumes are among the important crop commodities and have high demand being a major supplement of protein, but the productivity is low compared with the increasing demand resulting from several biotic (Rubiales and Mikic 2015) and abiotic stresses (Araújo et al. 2015). The productivity trends for these legumes in the last five decades suggest very little improvement leading to low productivity in most of the legumes compared with cereal crops (FAOSTAT 2014). Nevertheless, several efforts made in these years identified the genetic variations for various traits of interest in these legumes to enhance the crop productivity. So far, limited success could be achieved with the application of conventional breeding approaches for enhancing the crop productivity by overcoming key constraints. It is time to adopt modern and new technologies for enhancing the rate of genetic gain, so that improved varieties can be developed faster and more precisely equipped with essential traits to face the climate and other stress factors.

A paradigm shift is required in approaches and breeding methodologies to develop superior varieties for the future. In this context, deployment of genomics tools and technologies has shown great potential in understanding the complex genetics and breeding problems. It has been realized that genomics-assisted breeding (GAB), with integration of conventional breeding is the key to overcome conventional breeding limitations (Varshney et al. 2013a). Further in the case of legumes, a journey from a status of orphan crops with a dearth of genomic resources a decade ago, to current well-enriched genomic resource crop status, opened the possibility of deployment of GAB for these crops. Additionally, recent advent of the next-generation sequencing (NGS) technologies had brought down the sequencing and genotyping cost significantly. As a result, draft genomes have become available for several legume crops including model legumes, i.e., *Medicago truncatula* (Young et al. 2011), *Lotus japonicus* (Sato et al. 2008) and crops such as *Glycine max* (Soybean) (Schmutz et al. 2010), *Cajanus cajan* (Pigeonpea) (Varshney et al. 2012), *Cicer arietinum* (Chickpea) (Varshney et al. 2013b; Jain et al. 2013); *Lupinus angustifolius* (Lupin) (Yang et al. 2013), *Vigna radiata* (Mung bean) (Kang et al. 2014) and *Arachis duranensis* and *A. ipaensis* (progenitors of cultivated groundnut) (Bertioli et al. 2016; Chen et al. 2016). Genome sequencing efforts followed by large scale re-sequencing efforts in each crop led to availability of millions of structural variations leading to availability of large numbers of genetic markers (see Varshney et al. 2013a; Bohra et al. 2014; Pandey et al. 2016).

Availability of large scale genome-wide genetic markers led to establishment of several high-throughput genotyping platforms, offering precise, rapid and cost-effective solutions to genotyping of large populations. For instance, informative single nucleotide polymorphisms (SNPs) with high genome density are being chosen and used to design assays/platforms for legumes such as in *Vigna unguiculata* (Egbadzor et al. 2014; Huynh et al. 2013; Lucas et al. 2013, Muñoz-Amatriaín et al. 2016), *Pisum sativum* (Deulvot et al. 2010; Bordat et al. 2011; Tayeh et al. 2015), *Lens culinaris* (Sharpe et al. 2013; Kaur et al. 2014a), *Vicia faba*

(Kaur et al. 2014b), soybean (Lee et al. 2015; Wang et al. 2016), chickpea (Gujaria et al. 2011; Hiremath et al. 2011; Roorkiwal et al. 2014), pigeonpea (Saxena et al. 2012) and groundnut (Pandey et al. 2017). Other alternative SNP detection systems like competitive allele-specific PCR (KASPar) (Cottage et al. 2012; Hiremath et al. 2012; Kumar et al. 2012; Saxena et al. 2012; Xu et al. 2012; Fedoruk 2013; Khera et al. 2013; Sharpe et al. 2013), custom-designed Illumina VeraCode assay (Deulvot et al. 2010; Roorkiwal et al. 2013, Duarte et al. 2014) have also been employed for various applications. The development and deployment of different genotyping platforms provide cost effective and precise genotyping solution to many legume crops leading to enhanced rate of progress in legume genomics. NGS-based genotyping by sequencing (GBS) allows simultaneous marker discovery as well as genotyping of the populations even in the absence of a reference genome (Davey et al. 2011). Among legumes, the GBS approach has been successfully used in lentil (Ates et al. 2016) and chickpea (Deokar et al. 2014; Jaganathan et al. 2015; Verma et al. 2015) for genome-wide SNP discovery and genetic mapping. Further, whole genome re-sequencing (WGRS) and restriction site-associated DNA (RAD) sequencing approaches have also been used to capture the variations in the genome and to understand diversity prevailing in the germplasm (see Varshney et al. 2013b).

GAB aims at to accelerate crop improvement by establishing and exploiting the relationships between genotype and phenotype. Of the three GAB approaches, marker-assisted backcrossing (MABC), marker-assisted recurrent selection (MARS) and genomic selection (GS), MABC has been deployed in most of the crops and proved to be an effective approach for development of improved varieties and lines in many legume crop plants (see Pandey et al. 2016). MABC uses markers linked to agronomical important traits and mainly aims at introgression of a limited number of alleles from one genetic background (donor) to other (recipient) (Hospital 2005). Further, the improved varieties developed as a result of MABC contain one or a few alleles at major gene/QTLs from the donor genotype, keeping intact the rest of the genome from recurrent parent (see Varshney et al. 2013a). For instance, one "*QTL-hotspot*" region having QTLs for several drought tolerance-related root traits was introgressed into JG11, a desi chickpea cultivar from the drought tolerant line ICC4958 (Varshney et al. 2013c). Similarly introgression lines developed using MABC for fusarium wilt (FW) and ascochyta blight (AB) resistance in the background of C214 have shown enhanced resistance for FW and AB (Varshney et al. 2014). In the case of groundnut, MABC has been exploited to introgress major QTLs for leaf rust resistance from GPBD 4, a leaf rust resistant cultivar into ICGV 91114, JL 24 and TAG 24 cultivars (Varshney et al. 2014). MABC along with MAS was further deployed in enhancing the oil quality by increasing oleic acid in three different groundnut varieties, viz. ICGV 06110, ICGV 06142 and ICGV 06420 (Janila et al. 2016). In the case of pea, Aphanomyces root rot resistance QTLs (Lavaud et al. 2015) and frost tolerance QTLs (Hascoët et al. 2014) were introgressed using MABC into different agronomically important genetic backgrounds. Likewise in soybean, MABC was deployed successfully to improve resistance to a defoliating insect (Zhu et al. 2007), bacterial leaf pustule

resistance (Kim et al. 2008) and to reducing a kunitz trypsin inhibitor (Kumar et al. 2015).

In order to address the limitations of MABC approach for improving multiple complex traits, MARS has been proposed for combining major and minor QTLs in several crops. In the case of MARS, the de novo QTL identification is carried out in a breeding population derived from the crosses of superior varieties followed by crossing genotypes with superior alleles for pyramiding targeted QTLs into one or more genetic backgrounds (Bernardo and Charcosset 2006). However, the MARS approach was not effective for increasing yield in chickpea (Pandey et al. 2016). MARS was suggested a method for improvement of drought tolerance in groundnut, however more than 100 main and epistatic effect QTLs were reported because handling these small effect QTLs through MABC was not possible (Gautami et al. 2012).

GS utilizes phenotypic as well as genome-wide marker data to predict the genomic-estimated breeding values (GEBV) for selecting the superior lines. In brief, two populations, training population and testing population (sometimes, it is part of training population, hence known as validation set as well) are used. Training population is the one with comprehensive phenotypic data under different environmental conditions, that is, different locations/seasons/treatments. Genome-wide genotypic and phenotypic data for the training population are used to train different statistical GS models. The training population can be subdivided into five to ten groups, and then, cross validation is used to evaluate the GS models and prediction accuracy. Trained models, are used to calculate GEBV of a testing or selection candidate population that has been genotyped but not phenotyped. The predicted GEBVs are used to select superior lines from the population. One of the advantages associated with GS is that it reduces the selection cycle length by eliminating the phenotyping that is required for multiple rounds of selection hence reducing time and cost, leading to genetic gain.

Genomic prediction is a key to success in GS breeding, and it depends on high-throughput and high-density genotyping along with accurate, multilocation phenotyping data. Availability of ample genomic resources and affordable high-density and high-throughput genotyping in several legumes will facilitate deployment of GS in legumes. This chapter briefly describes the critical factors determining the success of genomic selection and summarises the ongoing efforts to deploy genomic selection in legumes and further the existing possibilities by integrating available genomic resources to harness the full potential of modern breeding approaches.

6.2 Critical Factors in Deployment of Genomic Selection

High-precision prediction accuracies are the most critical point that determines the success of any GS breeding program. Multiple simulation and empirical studies involving estimation of prediction accuracies rely on multiple factors *viz*. number

and type of markers (Chen and Sullivan 2003; Poland and Rife 2012), population structure (Nakaya and Isobe 2012; Spindel et al. 2015), training population size (Daetwyler et al. 2008), heritability and architecture of target traits (Zhong et al. 2009; Zhang et al. 2014, 2016) and the relationship between training population and selection candidates.

Numerous GS models have been proposed to address the diverse requirements for achieving satisfactory prediction accuracies. Some of the routinely used GS models include Random Regression Best Linear Unbiased Predictor (RR-BLUP; Meuwissen et al. 2001; Liu et al. 2008; Zhang et al. 2010), Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani 1996; de los Campos et al. 2009a), semiparametric strategies (Kinship GAUSS), Bayesian approach viz. Bayesian Ridge Regression, Bayesian LASSO (de los Campos et al. 2009b; Legarra et al. 2011), Bayes A (Meuwissen et al. 2001), Bayes B (Meuwissen et al. 2001) and Bayes Cπ (Habier et al. 2011) and machine learning Random Forest Regression (RFR) (Breiman, 2001), and Support Vector Regression (SVR) (Drucker et al. 1997). Various comparative accounts have been drawn to assess the performances of these GS models among different organisms (Moser et al. 2009, Heslot et al. 2012, Resende et al. 2012a, b). Selection of an appropriate GS model varies from case to case, and hence, multiple models should be considered in any GS study.

Size of training population is another important factor that has significant impact on prediction accuracies. Bernardo and Yu (2007) suggested that a minimum size of the training population to be 100–150 genotypes to obtain the optimum prediction accuracy. In the case of genetically diverse populations, larger training populations are required to attain better prediction accuracies (Mujibi et al. 2011). Genetic relatedness of the individuals in the training and selection populations is known to affect the accuracies of GS studies (Asoro et al. 2011). Among cattle, GEBVs estimated within breed were found to be more accurate than the ones estimated across breeds (Hayes et al. 2009). Price et al. (2010) and Guo et al. (2014) demonstrated significant reduction in prediction accuracies in structured populations.

Application of genome-wide markers results in better prediction accuracies (Meuwissen et al. 2001; Calus and Veerkamp 2007). Higher marker density has been demonstrated to produce higher genomic prediction accuracy (Zhong et al. 2009; Asoro et al. 2011; Heffner et al. 2011; Poland et al. 2012; Heslot et al. 2013). Low marker densities in some cases result in lower prediction accuracies, that could be explained as lower probability of LD between markers and QTLs, because of the smaller fraction of variation (Solberg et al. 2008). Hickey et al. (2014) reported that a small number of markers (200–500) and phenotypes (1000) are required in a closely related biparental population to achieve effective prediction accuracies, whereas for a population that is unrelated to the selection candidates, a much larger number of markers and phenotypes are required for the same prediction accuracy. A large mixed training population set with higher marker density is recommendable to achieve high prediction accuracies rather than using multiple training populations representing one germplasm group (Asoro et al. 2011). In another study, De Roos

et al. (2009) suggested that a high marker density is required if training and selection populations are highly divergent.

High-throughput genotyping platforms such as DArT, SNP array and GBS are being used based on different needs. GBS has been deployed in almost all the crops in the initial genetic analysis as it provides a low cost option to plant species where there is no reference genome (Poland et al. 2012). A comparison made by Poland et al. (2012) using GBS for de novo genotyping of testing populations in case of the wheat (*Triticum aestivum* L.) genome showed higher prediction accuracies of 0.3–0.5 in comparison to established marker platforms.

Enhancing the marker numbers while imputing the missing marker data has been reported to improve in prediction accuracies. For instance, Poland et al. (2012) showed an improvement of prediction accuracies with the genotyping data set consisting of 35,000 SNPs with up to 80% missing data points, over the prediction accuracies estimated from 2000 DArT markers with missing data points up to 2%. In various studies including maize, wheat, barley and forest trees, a positive relationship between the trait heritability and prediction accuracies has been observed (Lorenzana and Bernardo 2009; Albrecht et al. 2011; Heffner et al. 2009, 2011; Grattapaglia et al. 2011; Guo et al. 2012; Combs and Bernardo 2013). In another study, Zhang et al. (2014) established higher prediction accuracies for less complex traits. Most of the results discussed here form the basis of ongoing efforts in legume genomic selection and serve as the guidelines for strategizing the future efforts. GS efforts in different legumes have been described below in detail.

6.3 Soybean (*Glycine max*)

Deployment of GS among legumes first started with improving yield and agronomic traits in soybean. A set of 301 elite breeding lines was genotyped with GBS and phenotyped for grain yield at multiple locations (Table 6.1) (Jarquín et al. 2014). By keeping a randomly selected set of 50 accessions for a validation population, a positive relationship was observed between the size of training population and prediction accuracy, which began to plateau at a training population size of 100; however, it continued to increase until the maximum available size. The study included the evaluation of three different imputation methods to impute the missing data for soybean. However, not many differences were obtained using these imputation methods. Although, random forest imputation produced the highest accuracies, no significant differences were observed. A high prediction accuracy (0.64) reflected high potential of GS for yield in soybean (Table 6.1) (Jarquín et al. 2014).

Further, exploiting the GAB, genotyping data for 31,045 SNPs on 309 soybean germplasm accessions were used to estimate the prediction accuracy for seed weight (SW) (Zhang et al. 2016). Five-fold cross validation (CV) was applied by randomly assigning 20% of the association panel as validation set and remaining

Table 6.1 Summary of key genomic selection studies in some legume crops

Legume crop	Population Size	Marker type	Traits	GS models	Reference
Soybean	301	GBS	1. Grain yield	A standard Genomic best linear unbiased prediction (G-BLUP) model including only additive effects, and an extended version of the G-BLUP model including additive-by-additive effects.	Jarquín et al. (2014)
Alfalfa	190	GBS (10,000 SNPs)	1. Single harvest biomass 2. Total biomass	Random Regression Best Linear Unbiased Predictor (RR-BLUP)	Li et al. (2015)
	278 (adapted to two different environment)	GBS	1. Dry matter yield	Support vector regression using linear and Gaussian kernel, RR-BLUP, random Forest regression and Bayes A, Bayes B and Bayesian lasso,	Annicchiarico et al. (2015)
Pea	372	331 SNP	1. Date of flowering 2. Number of seeds per plant 3. Thousand seed weight	LASSO (least absolute shrinkage ans selection operator), PLS (partial least squares), SPLS (sparse partial least squares), Bayes A, Bayes B and G-BLUP	Burstin et al. (2015)
	339	9824 SNPs (GenoPea 13.2 K SNP Array)	1. Date of flowering 2. Number of seeds per plant 3. Thousand seed weight	Kernel partial least squares regression (kPLSR), LASSO, G-BLUP, Bayes A, and Bayes B	Tayeh et al. (2015)
Chickpea	320	3000 DArT markers	1. Seed yield 2. 100 seed weight 3. Days to 50% flowering 4. Days to maturity	RR-BLUP, kinship GAUSS, Bayes Cπ, Bayes B, Bayesian LASSO, Random Forest (RF)	Roorkiwal et al. (2016)
Groundnut	188	2356 DArT markers	1. Days to flowering 2. Seed weight 3. Pod yield	RR-BLUP, kinship GAUSS, Bayes Cπ, Bayes B, Baysian LASSO and RF	Pandey et al. (2014b, 2015)

80% as the training set. Based on the number of SNPs used and the size of training population, the prediction accuracies were found to vary between 0.75 and 0.87. Like other studies (Asoro et al. 2011; Jarquin et al. 2014), on size of the training population, smaller populations resulted in lower prediction accuracies. Another observation was the prediction accuracy using all 2000 SNPs was found to be same, even reducing it to 500 SNPs. Higher prediction accuracies were observed compared to Jarquín et al. (2014) with same number of markers, similar population size, and broad sense heritability of traits, pointing towards the impact of genetic architecture of traits in populations under investigation.

6.4 Alfalfa (*Medicago sativa*)

Alfalfa is a perennial legume with a long breeding cycle, which limits crop improvement efforts. Selection cycle duration can be reduced by deploying GS for complex traits such as yield by using GS for predicting the breeding values (Li et al. 2015). Prediction accuracies were obtained using phenotyping data for yield traits during two selection cycles from three locations and using genotyping data for ~10,000 SNPs (Li et al. 2015). Varying levels of missing values from the marker data set were used for GS modelling using random forest method for missing values imputation. Validation of genomic prediction models was performed by cross validation, in which randomly selected 90% genotypes were used as training population and 10% was used for testing/validation. Marker data sets with more missing values resulted in a large number of markers and resulted in increased prediction accuracies. Prediction accuracies were validated for both the generation viz. cycle 0 and cycle 1. In individual generation analysis, prediction accuracies validated within locations were found to be much higher than prediction accuracies across the locations, possibility due to G × E interaction for biomass yield. Prediction accuracies of 0.43–0.66 for total biomass yield in a synthetic alfalfa breeding population showed the underlying potential of further application of GS in other complex traits (Li et al. 2015) (Table 6.1).

In total, 278 elite genotypes adapted to two different environments with a different genetic base were genotyped using GBS and phenotyped for dry matter yield of their densely planted half-sib progenies in separate environments (Annicchiarico et al. 2015). Prediction accuracies were higher using joint SNP calling in comparison to separate SNP calling for the two data sets. Random forest was used for missing marker imputation. A comparison of prediction accuracies within and across populations was performed with the same set of markers, and it was observed that within-population prediction accuracies were higher than across-population prediction accuracies, probably due to a high level of intra-population variation. Results indicated a greater than three-fold higher prediction for yield gain per unit time though GS in comparison to conventional selection (Annicchiarico et al. 2015) (Table 6.1).

6.5 Pea (*Pisum sativum*)

In the case of pea, SNP markers were used to predict the phenotypes using different statistical methods (Burstin et al. 2015). Phenotyping data for two seasons and genotyping data generated with 331 SNPs on >350 accessions representing various cultivars, diverse wild types, landraces, etc. were used to estimate the prediction accuracies (Table 6.1). To minimize the impact of population structure leading to spurious associations, authors used the approach recommended by Johnson et al. (2007). Thousand seed weight (TSW) was predicted better than the beginning of flowering (BegFlo) and number of seeds per plant (NSeed). During the same year, they reported deployment of a high-density genotyping platform for GS (Tayeh et al. 2015). Similarly, genotyping data from the GenoPea 13.2 K SNP Array on a collection of 339 accessions along with the phenotyping data for TSW, BegFlo and NSeed were used for estimating genomic prediction values using five different statistical methods (Tayeh et al. 2015). To estimate the impact of the training population size over the prediction accuracies, different sizes of training populations were selected randomly with multiple repetitions; however, the test set was fixed with 99 accessions. Similarly, to assess the effect of marker density on prediction accuracies, evenly distributed SNP subsets were selected for estimation. Of five models considered in the study, four showed equivalent performance, whereas performance of LASSO was less than others. Another highlight of the study was that no significant differences were observed whether or not the markers with low minor allele frequency (MAF) were included. The effect of a reduction in the size of the training population was reduction in accuracy of the prediction models (Q^2). In addition, reducing the marker density but retaining only a single marker per unique map position did not affect prediction accuracy. However, a further reduction in the number of markers led to reduced Q^2. Q^2 values obtained in Tayeh et al. (2015) were found to be higher than in Burstin et al. (2015).

6.6 Chickpea (*Cicer arietinum*)

In case of chickpea, there is only one report coming from ICRISAT about deploying GS breeding and conducting initial studies of standardizing different GS models (Roorkiwal et al. 2016). In this context, a training population containing 320 elite chickpea breeding lines consisting of desi and kabuli seed types, from the International Chickpea Screening Nursery (ICSN), was genotyped using the DArTseq platform. This platform generated 3000 polymorphic markers. Phenotyping data were generated for yield and yield-related traits *viz*. seed yield (SY), 100 seed weight (SDW), days to 50% flowering (DF) and days to maturity (DM), at two different locations during two different crop seasons for two different treatments, that is, rainfed and irrigated conditions. Six different statistical models were used to calculate prediction accuracies and perform five-fold cross validation to estimate

the prediction accuracies by randomly selecting 80% of the lines for the training population and the remaining 20% as the testing population (Roorkiwal et al. 2016). A large variation in prediction accuracies were observed among the traits undertaken in the study, but overall performance of the models were found to be similar for every trait. The effect of G × E interaction was observed in the prediction accuracies of individual traits. For instance, the best prediction accuracy was observed for SDW (trait least affected by G x E interaction and treatments, etc.); however, prediction accuracies were lower for SY trait, which is known to be affected by G × E. The impact of missing marker data and MAF on prediction accuracies was assessed for 100 seed weight, using nine different combinations of missing marker data and MAF (including markers in combination with 0%, ≤10% and ≤30% missing data, and 0%, ≥5% and ≥10% MAF). The results showed that the random forest model at 0% missing marker data and ≥5% MAF combination had the best prediction accuracy, whereas the Bayes B model with 0% missing marker data and ≥10% MAF produced lowest accuracies. This study also assessed the impact of population structure on GEBV prediction accuracy. Desi and kabuli seed types were undertaken as separate groups and also grouped together to calculate prediction accuracies. The results reflected a higher prediction accuracy using the complete set in comparison to different seed types considered separately, which might be attributed to a larger population size (Roorkiwal et al. 2016) (Table 6.1).

6.7 Groundnut (*Arachis hypogaea*)

In case of groundnut, ICRISAT has taken some initiatives towards deploying GS breeding and conducting initial studies of standardizing different GS models (Pandey et al. 2016).While undertaking deployment of GS in groundnut, the focus of the study was to assess the impact of associated markers on prediction accuracies for three important traits viz. days to flowering (DF), seed weight (SW) and pod yield (PY) with different heritabilities (Pandey et al. 2014a, b; Pandey et al. 2015). Six seasons of phenotyping data for these traits and genotyping of the reference set with 2356 DArT markers were used for GS analysis (Table 6.1). When comparing the prediction accuracy for total and associated markers, the impact of population size and two different approaches were used to estimate the prediction accuracies. In the first approach, the whole population set was considered as a training population, and a part of the training population was considered as validation set to calculate the prediction accuracies. However, in another approach, the whole population was fractioned into five random smaller sets, of which one set was used to train the GS model, hence acted as training population, and the rest four were used as validation sets. Associated markers were compared with using all markers and the associated marker set showed higher prediction accuracies. However in a second approach where randomly selected smaller sets were used to genotype the training population, prediction accuracies obtained with associated

markers were less predictive than all genome-wide markers. Overall, only marginal differences were observed between the prediction accuracies estimated using total genome-wide markers by both the approaches. As expected, the traits with higher heritability showed higher prediction accuracies in comparison to those with lower heritability. A positive relation between the heritability and prediction accuracies was observed, supporting similar observations in maize, wheat, barley, etc. (Lorenzana and Bernardo 2009; Albrecht et al. 2011; Heffner et al. 2011; Guo et al. 2012; Combs and Bernardo 2013). So far, the lack of a high-throughput genotyping platform to generate high-density genotyping data has been the major obstacle in deploying the GS breeding in groundnut. However, the availability of genome sequences of a diploid progenitor species and 58 K Axiom_*Arachis* SNP (Pandey et al. 2017) array during 2016 will further boost the deployment of GS breeding in groundnut.

6.8 Conclusions

The majority of legume crops lacked the attention of researchers for generating genomic resources for a longer time compared with cereal crops. Nevertheless, the speedy development in NGS technologies and assembly methodologies made generating genomic resources affordable and technically sound over the time. The legume crops have made much progress from poor resource to highly enriched genomic resourced crops. This has provided many opportunities to implement advanced genomic-assisted breeding. GS breeding has demonstrated its great value to the ongoing conventional breeding programs of cattle and in some plant species. This approach is gaining attention from other crop breeders including legumes as it promises greater genetic gain by improving complex traits in less time with more precision. Seeing the benefits achieved in the maize and wheat breeding programs, legume crops are now looking forward to deploying GS breeding to address its some of the most complex problems that are the key obstacles in achieving higher productivity. Selected studies conducted so far in legumes have suggested the possibility of achieving high prediction accuracies. These preliminary studies also indicated the potential role of GS in developing superior varieties with enhanced genetic gain and ability to overcome various stresses, hence ensuring food security with higher productivity. Currently, the majority of the legume crops are in the process of deploying GS in their breeding program; however, it will take a few years for GS to become routine similar to other major crop breeding programs.

References

Albrecht T, Wimmer V, Auinger H, Erbe M, Knaak C et al (2011) Genome-based prediction of testcross values in maize. Theor Appl Genet 123:339–350

Annicchiarico P, Nazzicari N, Li X, Wei Y, Pecetti L et al (2015) Accuracy of genomic selection for alfalfa biomass yield in different reference populations. BMC Genomics 16:1020

Araújo SS, Beebe S, Crespi M, Delbreil B, González EM et al (2015) Abiotic stress responses in Legume crops: strategies used to cope with environmental challenges. Crit Rev Plant Sci 34:237–280

Asoro FG, Newell MA, Beavis WD, Scott MP, Jannink J (2011) Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. The Plant Genome 4:132–144

Ates D, Sever T, Aldemir S, Yagmur B, Temel HY et al (2016) Identification QTLs Controlling Genes for Se Uptake in Lentil Seeds. PLOS ONE 11(4): e0154054

Bernardo R, Charcosset A (2006) Usefulness of gene information in marker-assisted recurrent selection: a simulation appraisal. Crop Sci 46:614–621

Bernardo R, Yu J (2007) Prospects for genome-wide selection for quantitative traits in maize. Crop Sci 47:1082–1090

Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD et al (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat Genet 48:438–446

Bohra A, Pandey MK, Jha UC, Singh B, Singh IP et al (2014) Genomics-assisted breeding in four major pulse crops of developing countries: present status and prospects. Theor Appl Genet 127:1263–1291

Bordat A, Savois V, Nicolas M, Salse J, Chauveau A et al (2011) Translational genomics in legumes allowed placing in silico 5460 unigenes on the pea functional map and identified candidate genes in Pisum sativum L. Genes Genome Genet 1:93–103

Breiman L (2001) Random forests. Mach Learn 45:5–32. doi:10.1023/A:1010933404324

Burstin J, Salloignon P, Chabert-Martinello M, Magnin-Robert JB, Siol M et al (2015) Genetic diversity and trait genomic prediction in a pea diversity panel. BMC Genomics 16:105

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. J Ani Breed Genet 124:362–368

Chen X, Sullivan PF (2003) Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenomics J 3:77–96

Chen X, Li H, Pandey MK, Yang Q, Wang X et al (2016) Draft genome of the peanut A-genome progenitor (Arachis duranensis) provides insights into geocarpy, oil biosynthesis, and allergens. Proc Nat Acad Sci 113:6785–6790

Combs E, Bernardo R (2013) Accuracy of genomewide selection for different traits with constant population size, heritability and number of markers. The Plant Genome 6:1

Cottage A, Gostkiewicz K, Thomas JE, Borrows R, Torres AM et al (2012) Heterozygosity and diversity analysis using mapped SNPs in a faba bean inbreeding programme. Mol Breed 30:1799–1809

Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS One 3:e3395

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM et al (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–510

de los Campos G, Gianola D, GJM R (2009a) Reproducing kernel Hilbert spaces regression: a general framework for genetic evaluation. J Anim Sci 87:1883–1887

de los Campos G, Naya H, Gianola D, Crossa J, Legarra A et al (2009b) Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182:375–385

de Roos APW, Hayes BJ, Goddard ME (2009) Reliability of genomic breeding values across multiple populations. Genetics 183:1545–1553. doi:10.1534/genetics.109.104935

Deokar AA, Ramsay L, Sharpe AG, Diapari M, Sindhu A et al (2014) Genome wide SNP identification in chickpea for use in development of a high density genetic map and improvement of chickpea reference genome assembly. BMC Genomics 15:708

Deulvot C, Charrel H, Marty A, Jacquin F, Donnadieu C et al (2010) Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. BMC Genomics 11:468

Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V (1997) Support vector regression machines. Adv Neural Info Process Syst 9:155–161

Duarte J, Rivière N, Baranger A et al (2014) Transcriptome sequencing for high throughput SNP development and genetic mapping in pea. BMC Genomics 15:126

Egbadzor KF, Ofori K, Yeboah M, Aboagye LM, Opoku-Agyeman MO et al (2014) Diversity in 113 cowpea [Vigna unguiculata (L) Walp] accessions assessed with 458 SNP markers. Springer Plus 3:541

Fedoruk M (2013) Linkage and association mapping of seed size and shape in lentil. Thesis (Masters of Science), University of Saskatchewan, Saskatoon

Gautami B, Pandey MK, Vadez V, Nigam SN, Ratnakumar P et al (2012) Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (Arachis hypogaea L.) Mol Breed 30:757–772

Graham PH, Vance CP (2003) Legume crops: importance and constraints to greater use. Plant Physiol 131:872–877

Grattapaglia D, Resende MDV, Resende MR, Sansaloni CP, Petroli CD et al (2011) Genomic selection for growth traits in eucalyptus: accuracy within and across breeding populations. BMC Proc 5:O16. doi:10.1186/1753-6561-5-S7-O16

Gujaria N, Kumar A, Dauthal P, Dubey A, Hiremath P et al (2011) Development and use of genic molecular markers (GMMs) for construction of a transcript map of chickpea (Cicer arietinum L.) Theor Appl Genet 122:1577–1589

Guo Z, Tucker D, Lu J, Kishore V, Gay G (2012) Evaluation of genome-wide selection efficiency in maize nested association mapping populations. Theor Appl Genet 124:261–275

Guo Z, Tucker DM, Basten CJ, Gandhi H, Ersoz E et al (2014) The impact of population structure on genomic prediction in stratified populations. Theor Appl Genet 127:749–762

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011) Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics 12:186

Hascoët E, Jaminon O, Devaux C, Blassiau C, Bahrman N, Bochard A-M et al (2014) Towards fine mapping of frost tolerance QTLs in pea, in 2nd PeaMUST Annual Meeting (Dijon)

Hayes B, Bowman P, Chamberlain A, Goddard M (2009) Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci 92:433–443

Heffner EL, Me S, Jannink JL (2009) Genomic selection for crop improvement. Crop Sci 49:1–12

Heffner EI, Jannink JL, Iwata H, Souza E, Sorrells ME (2011) Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51:2597–2606

Heslot N, Yang HP, Sorrells ME, Jannink JL (2012) Genomic selection in plant breeding: a comparison of models. Crop Sci 52:146–160

Heslot N, Rutkoski J, Poland J, Jannink JL, Sorrells ME (2013) Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. PLoS One 8:e74612

Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R et al (2014) Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Sci 54:1476–1488

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H et al (2011) Large-scale transcriptome analysis in chickpea (Cicer arietinum L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnol J 9:922–931. doi:10.1111/j.1467-7652.2011.00625.x

Hiremath PJ, Kumar A, Penmetsa RV, Farmer A, Schlueter JA et al (2012) Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. Plant Biotechnol J 10:716–732

Hospital F (2005) Selection in backcross programmes. Philos Trans Roy Soc Lond B Biol Sci 360:1503–1511

Huynh BL, Close TJ, Roberts PA, Hu Z, Wanamaker S et al (2013) Gene pools and the genetic architecture of domesticated cowpea. The Plant Genome 6:3

Jaganathan D, Thudi M, Kale S, Azam S, Roorkiwal M et al (2015) Genotyping-by-sequencing based intra-specific genetic map refines a "QTL-hotspot" region for drought tolerance in chickpea. Mol Gen Genomics 290:559–571

Jain M, Misra G, Patel RK, Priya P, Jhanwar S et al (2013) A draft genome sequence of the pulse crop chickpea (Cicer arietinum L.) Plant J 74:715–729. doi:10.1111/tpj.12173

Janila P, Pandey MK, Shasidhar Y, Variath MT, Sriswathi M et al (2016) Molecular breeding for introgression of fatty acid desaturase mutant alleles (*ahFAD2A* and *ahFAD2B*) enhances oil quality in high and low oil containing peanut genotypes. Plant Sci 242:203–213

Jarquín D, Kocak K, Posadas L, Hyma K, Jedlicka J et al (2014) Genotyping by sequencing for genomic prediction in a soybean breeding population. BMC Genomics 15:740

Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics 8:118–127. doi:10.1093/biostatistics/kxj037

Kang YJ, Kim SK, Kim MY, Lestari P, Kim KH et al (2014) Genome sequence of mungbean and insights into evolution within Vigna species. Nat Commun 5:5443

Kaur S, Cogan NO, Stephens A, Noy D, Butsch M et al (2014a) EST-SNP discovery and dense genetic mapping in lentil (Lens culinaris Medik.) enable candidate gene selection for boron tolerance. Theor Appl Genet 127:703–713

Kaur S, Kimber RBE, Cogan NOI, Materne M, Forster JW et al (2014b) SNP discovery and high-density genetic mapping in faba bean (Vicia faba L.) permits identification of QTLs for ascochyta blight resistance. Plant Sci 217–218:47–55

Khera P, Upadhyaya HD, Pandey MK, Roorkiwal M, Sriswathi M et al (2013) SNP-based genetic diversity in the reference set of peanut (Arachis spp.) by developing and applying cost-effective KASPar genotyping assays. Plant Genome 6:1–11

Kim KH, Kim MY, Van K, Moon JK, Kim DH et al (2008) Marker-assisted foreground and background selection of near isogenic lines for bacterial leaf pustule resistant gene in soybean. J Crop Sci Biotechnol 11:263–268

Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next-generation sequencing and its applications. Int J Plant Genomics 15

Kumar V, Rani A, Rawal R, Mourya V (2015) Marker assisted accelerated introgression of null allele of kunitz trypsin inhibitor in soybean. Breed Sci 65:447–452

Lavaud C, Lesne A, Piriou C, Le Roy G, Boutet G et al (2015) Validation of QTL for resistance to Aphanomyces euteiches in different pea genetic backgrounds using near-isogenic lines. Theor Appl Genet 128:2273–2288

Lee YG, Jeong N, Kim JH, Lee K, Kim KH et al (2015) Development, validation and genetic analysis of a large soybean SNP genotyping array. Plant J 81:625–636

Legarra A, Robert-Granie P, Croiseau G, Guillaume F, Fritz S (2011) Improved LASSO for genomic selection. Genet Res 93:77–87

Li X, Wei Y, Acharya A, Hansen JL, Crawford JL et al (2015) Genomic prediction of biomass yield in two selection cycles of a tetraploid alfalfa breeding population. Plant Genome 8

Liu XQ, Rong JY, Liu XY (2008) Best linear unbiased prediction for linear combinations in general mixed linear models. J Multivariate Analysis 99:1503–1517

Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120:151–161

Lucas MR, Ehlers JD, Huynh BL, Diop NN, Roberts PA et al (2013) Markers for breeding heat-tolerant cowpea. Mol Breed 31:529–536

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

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. Genet Sel Evol 41:56

Mujibi FD, Nkrumah JD, Durunna ON, Stothard P, Mah J et al (2011) Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. J Anim Sci 89:3353–3361

Muñoz-Amatriaín M, Mirebrahim H, Xu P, Wanamaker SI, Luo M et al (2016) Genome resources for climate-resilient cowpea, an essential crop for food security. Plant J. doi:https://doi.org/10.1101/059261

Nakaya A, Isobe SN (2012) Will genomic selection be a practical method for plant breeding? Ann Bot 110:1303–1316

Pandey MK, Rathore A, Das RR, Khera P, Upadhyaya HD et al (2014a) Selection of appropriate genomic selection model in an unstructured germplasm set of peanut (Arachis hypogaea L.). 6th international Food Legumes Research conference & 7th international conference on Legume Genetics and Genomics on 7–11 July 2014, Saskatoon

Pandey MK, Upadhyaya HD, Rathore A, Vadez V, Sheshshayee MS et al (2014b) Genome-wide association studies for 50 agronomic traits in peanut using the 'reference set' comprising 300 genotypes from 48 countries of semi-arid tropics of the world. PLoS One 9:e113326

Pandey MK, Agarwal G, Rathore A, Janila P, Upadhyaya HD, et al. (2015). Development of high density 60K "Axiom_Arachis" SNP Chip and optimization of genomic selection model for enhancing breeding efficiency in peanut. Proceedings of 8th international conference of the Peanut Research Community on "Advances in Arachis through Genomics and Biotechnology", Brisbane, 5–9 Nov 2015

Pandey MK, Roorkiwal M, Singh VK, Ramalingam A, Kudapa H et al (2016) Emerging genomic tools for legume breeding: current status and future prospects. Front Plant Sci 7

Pandey MK, Agarwal G, Kale SM, Clevenger J, Nayak SN et al (2017) Development and evaluation of a high density genotyping 'Axiom_Arachis' array with 58K SNPs for accelerating genetics and breeding in groundnut. Nat Sci Rep 7:40577. doi:10.1038/srep40577

Poland J, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5:92–102

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S et al (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103–113

Power JF (1987) Legume crops: their potential role in agricultural production. Am J Alt Agri 2:69–73

Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. Nat Rev Genet 11:459–463

Rebello CJ, Greenway FL, Finley JW (2014) A review of the nutritional value of legumes and their effects on obesity and its related co-morbidities. Obesity Rev 15:392–407

Resende MDV, Resende MFR, Sansaloni CP, Petroli CD, Missiaggia AA et al (2012a) Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytol 194:116–128

Resende MFR, Munoz P, Resende MDV, Garrick DJ, Fernando RL et al (2012b) Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda l.) Genetics 190:1503–1510

Roorkiwal M, Rathore A, Das RR, Singh MK, Jain A, Srinivasan S, Gaur PM, Chellapilla B, Tripathi S, Li Y, Hickey JM, Lorenz A, Sutton T, Crossa J, Jannink J-L, Varshney RK (2016) Genome-Enabled prediction models for yield related traits in chickpea. Front Plant Sci 7

Roorkiwal M, Sawargaonkar SL, Chitikineni A, Thudi M, Saxena RK et al (2013) Single nucleotide polymorphism genotyping for breeding and genetics applications in chickpea and pigeonpea using the BeadXpress platform. Plant Genome 6

Roorkiwal M, Von Wettberg EJ, Upadhyaya HD, Warschefsky E, Rathore A et al (2014) Exploring germplasm diversity to understand the domestication process in Cicer spp. using SNP and DArT markers. PLoS One 9(7):e102016

Rubiales D, Mikic A (2015) Introduction: legumes in sustainable agriculture. Crit Rev Plant Sci 34:2–3

Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T et al (2008) Genome structure of the legume, *Lotus japonicus*. DNA Res 15:227–239

Saxena RK, Penmetsa RV, Upadhyaya HD, Kumar A, Carrasquilla-Garcia N et al (2012) Large-scale development of cost-effective single-nucleotide polymorphism marker assays for genetic mapping in pigeonpea and comparative mapping in legumes. DNA Res 19:449–461

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Sharpe AG, Ramsay L, Sanderson LA, Fedoruk MJ, Clarke WE et al (2013) Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil. BMC Genomics 14:192

Solberg TR, Sonesson AK, Woolliams JA (2008) Genomic selection using different marker types and densities. J Anim Sci 86(10):2447–2454

Spindel J, Begum H, Akdemir D, Virk P, Collard B et al (2015) Genomic selection and association mapping in rice (Oryza sativa): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet 11:e1005350

Tarawali G, Ogunbile OA (1995) Legumes for sustainable food production in semi-arid savannahs. ILEIA Newslett 11(4):18–23

Tayeh N, Aluome C, Falque M, Jacquin F, Klein A et al (2015) Development of two major resources for pea genomics: the GenoPea 13.2 K SNP Array and a high-density, high-resolution consensus genetic map. Plant J 84:1257–1273

Tibshirani R (1996) Regression shrinkage and selection via the lasso. J Roy Stat Soc Series B 58:267–288

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK et al (2012) Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30:83–89

Varshney RK, Mohan SM, Gaur PM, Gangarao NVPR, Pandey MK et al (2013a) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31:1120–1134

Varshney RK, Song C, Saxena RK, Azam S, Yu S et al (2013b) Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat Biotechnol 31:240–246. doi:10.1038/nbt.2491

Varshney RK, Gaur PM, Chamarthi SK, Krishnamurthy L, Tripathi S et al (2013c) Fast-track introgression of "*QTL-Hotspot*" for root traits and other drought tolerance traits in JG 11, an elite and leading variety of chickpea. Plant Genome 6:3. doi:10.3835/plantgenome2013.07.0022

Varshney RK, Mohan SM, Gaur PM, Chamarthi SK, Singh VK et al (2014) Marker-assisted backcrossing to introgress resistance to Fusarium wilt (FW) race 1 and Ascochyta blight (AB) in C 214, an elite cultivar of chickpea. Plant Genome. doi:10.3835/plantgenome2013.10.0035

Verma S, Gupta S, Bandhiwal N, Kumar T, Bharadwaj C et al (2015) High-density linkage map construction and mapping of seed trait QTLs in chickpea (Cicer arietinum L.) using genotyping-by-sequencing (GBS). Sci Rep 5:17512

Wang J, Chu S, Zhang H, Zhu Y, Cheng H et al (2016) Development and application of a novel genome-wide SNP array reveals domestication history in soybean. Sci Rep 6

Xu P, Wu XH, Wang BG, Luo J, Liu YH et al (2012) Genome wide linkage disequilibrium in Chinese asparagus bean (Vigna unguiculata ssp. sesquipedialis) germplasm: implications for domestication history and genome wide association studies. Heredity 109:34–40

Yang H, Tao Y, Zheng Z, Zhang Q, Zhou G et al (2013) Draft genome sequence, and a sequence-defined genetic linkage map of the legume crop species Lupinus angustifolius L. PLoS One 8:e64799

Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524

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

Zhang Z, Ober U, Erbe M, Zhang H, Gao N et al (2014) Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. PLoS One 9: e93017

Zhang J, Song Q, Cregan PB, Jiang GL (2016) Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). Theor Appl Genet 129:117–130

Zhong S, Dekkers JC, Fernando RL, Jannink JL (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. Genetics 182:355–364

Zhu S, Walker DR, Warrington CV, Parrott WA, All JN et al (2007) Registration of four soybean germplasm lines containing defoliating insect resistance QTLs from PI 229358 introgressed into 'Benning'. J Plant Reg 1:162–163