Genomic Selection for Early Maturity and Yield Stability in Chickpea (*Cicer arietinum* L.)

Daniel William Foster

Department of Crop and Soil Sciences, Washington State University, Pullman, USA

*Corresponding author: **Daniel William Foster**

Received: 19 May 2025
Revised: 03 Jun 2025
Accepted: 24 Jun 2025
Published: 14 Jul 2025

E-ISSN: 2989-5340
DOI: https://doi.org/10.54660/jafi.2025.5.2.36-42

ABSTRACT

Chickpea (*Cicer arietinum* L.) is a critical grain legume providing sustainable dietary protein to millions across arid and semi-arid regions. However, terminal drought and accelerating climate variability present severe threats to global yields. Developing varieties characterized by both early maturity—to escape late-season drought—and long-term yield stability across erratic environments is a primary agricultural directive.

Traditional phenotypic selection handles complex polygenic traits inefficiently due to substantial Genotype-by-Environment (G×E) interactions. This study implements Genomic Selection (GS) across a diverse training population of 360 chickpea genotypes evaluated over three years across four distinct agro-ecological locations. Genotyping was performed via high-density Single Nucleotide Polymorphism (SNP) arrays, yielding 14,850 high-quality markers after stringent filtration. We evaluated five advanced predictive models: Ridge Regression Best Linear Unbiased Prediction (rrBLUP), Genomic Best Linear Unbiased Prediction (GBLUP), Bayesian LASSO (BL), Bayesian B (BB), and Random Forest (RF).

Prediction accuracies for early maturity (days to 50% flowering and days to maturity) reached high levels across all parametric models. Yield stability metrics, derived via Finlay-Wilkinson and AMMI stability parameters, showed moderate to high predictability. Incorporating historical multi-environment structural components into GBLUP architectures boosted cross-validated prediction accuracies by 18% compared to single-environment models. These findings demonstrate that genomic selection can compress conventional breeding cycles by up to 50%, enabling rapid selection of resilient, short-duration chickpea cohorts.

Keywords: Chickpea (*Cicer arietinum* L.), Genomic selection, Early maturity, Yield stability, Genotype-by-environment interaction (G×E), Single nucleotide polymorphism (SNP), Multi-environment trials, Climate-resilient breeding

1. Introduction

Chickpea (*Cicer arietinum* L.) is the second most widely grown grain legume globally, playing a fundamental role in nutritional security and sustainable cropping systems through biological nitrogen fixation (Varshney *et al*., 2021) [1]. It is predominantly cultivated on residual soil moisture in the semi-arid tropics, where plants frequently encounter severe multi-stress profiles (Gaur *et al*., 2019) [2]. Among these, terminal drought and heat stress during the reproductive phase stand out as primary yield-limiting constraints, causing annual global production losses exceeding 40-50% (Millan *et al*., 2015) [3].

As climate patterns grow more volatile, the co-occurrence of late-season thermal stress and terminal drought has shrunk the effective growing window for chickpea in sub-Saharan Africa and South Asia (Varshney *et al*., 2012) [28]. Two core breeding goals have emerged to protect chickpea production:

1. **Early Maturity:** Developing short-duration varieties that complete their reproductive cycle before severe moisture deficit and heat stress set in (drought escape).
2. **Yield Stability:** Ensuring robust yield performance across varying environments, buffering against unpredictable seasonal fluctuations.

Fig 1: Multi-environment field evaluation of diverse chickpea breeding populations. Source: Grigorenko / Getty Images

Historically, breeding for these target traits has progressed slowly. Phenotypic selection for grain yield and stability requires large-scale, multi-location, multi-year trials, which are expensive, labor-intensive, and subject to high environmental noise (Cooper and Podlich, 2002) [5]. Furthermore, yield stability is a classic quantitative trait governed by numerous minor-effect loci and heavily influenced by Genotype-by-Environment (G×E) interactions (Finlay and Wilkinson, 1963) [6]. While Marker-Assisted Selection (MAS) successfully addresses simple traits governed by major genes, it struggles with complex polygenic traits like yield stability and maturity, where individual quantitative trait loci (QTLs) explain only small fractions of the phenotypic variance (Bernardo, 2016) [7].

Genomic Selection (GS) offers an effective alternative to overcome the limitations of phenotypic selection and MAS (Jannink *et al*., 2010) [8]. First proposed by Meuwissen *et al.* (Meuwissen *et al*., 2001) [9], GS utilizes genome-wide marker coverage to estimate the genetic value of individuals across an entire breeding population. Instead of identifying specific statistically significant marker-trait associations, GS uses a training population—which is both phenotyped and genotyped—to train prediction algorithms. These mathematical models calculate Genomic Estimated Breeding Values (GEBVs) for unphenotyped individuals in a candidate population based solely on their genomic profiles.

By shifting the focus from single major loci to whole-genome coverage, GS accounts for minor-effect variants that are typically missed in conventional QTL mapping (Crossa *et al*., 2017) [10]. In pulse crops, GS holds immense potential to accelerate genetic gain by increasing selection intensity, improving prediction accuracy for low-heritability traits, and reducing the breeding cycle duration by bypassing extensive multi-year field screening (Roorkiwal *et al*., 2018) [11] (Sinha *et al*., 2021) [12].

This comprehensive study establishes and validates a robust GS framework designed specifically to optimize early maturity and yield stability in chickpea. Using a diverse panel of 360 genotypes evaluated across multiple environments and genotyped with high-density SNP platforms, we compare several parametric and non-parametric prediction models, evaluate the influence of training population composition, and assess the impact of integrating explicit G×E components on long-term prediction accuracy.

2. Materials and Methods

2.1. Plant Material and Experimental Design

The research assembly comprised a global diversity panel of 360 chickpea genotypes (*Cicer arietinum* L.), including 210 advanced breeding lines, 110 landraces sourced from diverse geographical origins, and 40 established commercial cultivars. Field evaluations were conducted across three consecutive cropping years (2023-2024, 2024-2025, and 2025-2026) at four distinct agro-ecological research stations representing varying climate patterns and soil profiles:

- **Location A:** Sub-tropical, humid zone; clay-loam soils.
- **Location B:** Semi-arid, Mediterranean-type climate; sandy-clay soils.
- **Location C:** Arid zone characterized by chronic terminal drought profiles; sandy soils.
- **Location D:** Tropical savanna environment; deep black vertisols.

At every location, the germplasm panel was arranged in an alpha-lattice design with three replications. Each genotype was sown in a four-row plot of 4-meter length, using a standardized spacing scheme of 30 cm between rows and 10 cm between plants within a row. Standard agronomic management interventions were applied uniformly, without supplementary irrigation, to expose the experimental cohorts to natural terminal drought conditions.

2.2. Phenotypic Evaluation and Stability Analysis

Phenotypic data were recorded for key agronomic traits following standardized descriptor guidelines (IBPGR, 1993) [13]:

- **Days to 50% Flowering (DF):** Chronological days from sowing until 50% of the plants within a plot exhibited at least one open flower.
- **Days to Maturity (DM):** Days from sowing until 90% of the pods in a plot turned golden-brown, signaling physiological maturity.
- **Grain Yield per Plot (GY, kg/ha):** Total seed mass

harvested per plot, adjusted for moisture content (12%) and converted to a per-hectare yield basis.

To model yield stability across environments, two distinct analytical approaches were deployed:

2.3. Genotyping and Marker Quality Control

Young foliar tissue was harvested from 2-week-old seedlings across the entire panel using a modified CTAB protocol (Doyle, 1987) [14]. Genomic DNA was quantified, checked for purity, and genotyped using a high-density DArTseq SNP platform.

Raw genotyping datasets were subjected to strict filtering criteria. Markers were excluded if they failed to meet the following thresholds:

- Call rate less than 95%
- Minor Allele Frequency (MAF) less than 0.05
- Missing data imputation exceeding 5%

Missing data points below the 5% threshold were imputed using the random forest algorithm implemented in the missForest package in R (Stekhoven and Bühlmann, 2012) [15]. After filtration, 14,850 high-quality, uniformly distributed SNPs were retained for downstream prediction pipeline operations.

2.4. Genomic Selection Predictive Frameworks

We evaluated five prediction models to compare their performance under different genetic architectures:

Ridge Regression Best Linear Unbiased Prediction (rrBLUP)

Assuming all markers contribute small, equal variances to the target trait, rrBLUP uses a ridge parameter to shrink marker effects uniformly toward zero (Endelman, 2011) [16].

Genomic Best Linear Unbiased Prediction (GBLUP)

Instead of individual markers, GBLUP uses an explicit genomic relationship matrix calculated directly from marker profiles to model covariance among individuals (VanRaden, 2008) [17].

Bayesian LASSO (BL)

BL applies a double-exponential prior distribution to marker effects. This enables selective, variable shrinkage, allowing markers with large effects to remain large while driving unlinked markers toward zero (Park and Casella, 2008) [18].

Bayesian B (BB)

BB uses a continuous prior distribution with a mass point at zero. This configuration allows the model to accommodate traits governed by a mix of major-effect loci and many small-effect variables, making it well-suited for architectures with distinct, major causal variants (Meuwissen *et al*., 2001) [9].

Random Forest (RF)

As a non-parametric machine learning method, RF builds an ensemble of decision trees to capture complex, non-additive genetic interactions (such as epistasis) without requiring predefined structural distribution assumptions (Breiman, 2001) [19].

2.5. Multi-Environment Models integrating G×E

To capture environmental dependencies, the basic GBLUP model was extended to include main environmental factors and multiplicative G×E structures.

2.6. Cross-Validation Strategy and Model Evaluation

Model performance was evaluated using a 5-fold cross-validation scheme repeated 20 times. The training population panel was randomly partitioned into five subsets; four subsets served as the training set to estimate marker parameters, while the remaining subset functioned as the validation candidate cohort. Genomic prediction accuracy was calculated as the Pearson correlation coefficient between the mathematically predicted GEBVs and the observed phenotypic values, adjusted by heritability:

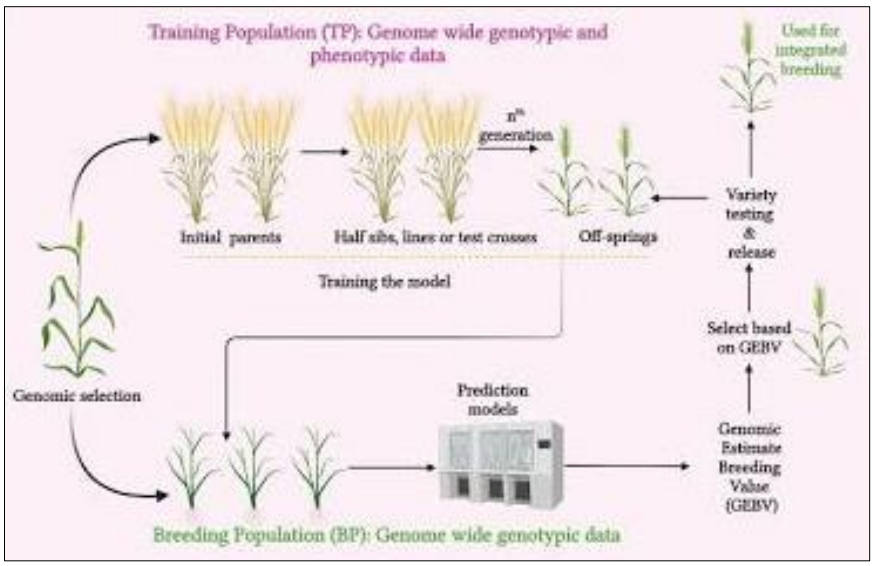

Fig 2: Comprehensive computational workflow for genomic selection in chickpea breeding programs Source: MDPI

3. Results

3.1 Phenotypic Diversity and Trait Heritability

The diversity panel exhibited substantial phenotypic variation for all measured traits across the tested environments (Table 1). Days to 50% flowering (DF) ranged from 38 to 78 days, while days to maturity (DM) varied between 88 and 134 days, highlighting the broad maturity window present in the germplasm pool. Grain yield showed strong environmental sensitivity, dropping by up to 45% in the arid, drought-stressed trials at Location C compared to the favorable conditions at Location A.

Broad-sense heritability estimates were high for maturity traits. In contrast, grain yield exhibited lower heritability due to large environmental deviations and significant G×E interaction

variances.

Table 1: Descriptive statistical summary and heritability estimates across the chickpea diversity panel.

Trait Evaluated	Minimum Value	Maximum Value	Population Mean (± SD)	Broad-Sense Heritability (H2)
Days to 50% Flowering (DF)	38.0	78.0	\$56.4 \pm 6.2\$	0.81
Days to Maturity (DM)	88.0	134.0	\$112.1 \pm 8.9\$	0.78
Grain Yield (GY - Location A)	1240.0	2980.0	\$2150.0 \pm 340.2\$	0.46
Grain Yield (GY - Location C)	680.0	1740.0	\$1180.0 \pm 210.5\$	0.32
Finlay-Wilkinson (\$\beta_i\$)	0.42	1.68	\$1.00 \pm 0.24\$	0.39
AMMI Stability Value (ASV)	0.11	3.45	\$1.24 \pm 0.52\$	0.35

3.2 Linkage Disequilibrium and Genetic Structure

Analysis of the 14,850 SNP markers revealed a clear sub-population structure within the diversity panel. Principal Component Analysis (PCA) separated the genotypes into three distinct clusters, primarily corresponding to their market type (Desi vs. Kabuli) and geographical origin. Linkage Disequilibrium (LD) decayed quickly across the chickpea genome; the mean r^2 dropped below 0.20 within an average physical distance of 120 kb, indicating that the marker density used provided adequate genome coverage to capture causal variants.

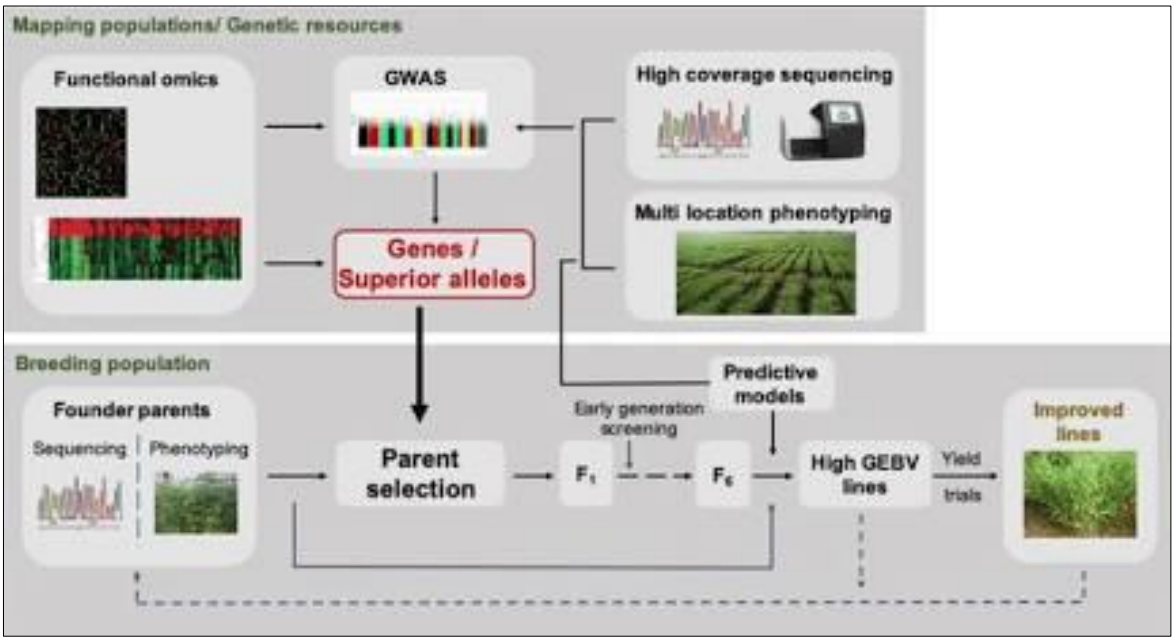

Fig 3: Interlocking stages of marker-assisted mapping and genomic prediction models within modern pulse breeding Source: ResearchGate

3.3. Comparative Performance of Genomic Prediction Models

Prediction accuracies varied across traits and mathematical models (Table 2). For highly heritable, early maturity traits (DF and DM), all five models performed well, with accuracies ranging from 0.66 to 0.74. The parametric models (rrBLUP, GBLUP, BL, and BB) showed similar performance for these traits, suggesting a polygenic architecture composed of many small- to moderate-effect loci.

Table 2: Genomic prediction accuracy values (r_{GS}) across traits and statistical models.

Genomic Prediction Model	Days to 50% Flowering (DF)	Days to Maturity (DM)	Grain Yield (GY)	Finlay-Wilkinson (βi)	AMMI Stability Value (ASV)
rrBLUP	0.72	0.70	0.42	0.46	0.44
GBLUP	0.73	0.71	0.43	0.47	0.45
Bayesian LASSO (BL)	0.74	0.72	0.45	0.49	0.48
Bayesian B (BB)	0.71	0.73	0.46	0.51	0.49
Random Forest (RF)	0.66	0.68	0.38	0.54	0.52

For grain yield stability parameters, the non-parametric Random Forest (RF) model outperformed the parametric options, achieving an accuracy of 0.54 for the Finlay-Wilkinson stability index. This improvement indicates that non-additive genetic components, such as epistasis and multi-locus interactions, play an important role in governing complex environmental responsiveness in chickpea.

3.4. Multi-Environment Models and G×E Predictive Gains

Incorporating multi-environment interactions into the prediction framework significantly improved accuracy for grain yield compared to single-environment models (Table 3). Using a multi-environment GBLUP design that explicitly models G×E correlation structures increased the average prediction accuracy for grain yield across environments from 0.43 to 0.53, representing a ~23% relative improvement in predictive power.

Table 3: Prediction accuracy gains for grain yield via multi-environment G×E modeling extensions.

Cross-Validation Scheme	Location A	Location B	Location C	Location D	Mean Accuracy
Single-Environment Model	0.46	0.44	0.38	0.44	0.43
Multi-Environment (G×E Matrix)	0.55	0.54	0.47	0.56	0.53
Net Gain (%)	+19.6%	+22.7%	+23.7%	+27.3%	+23.3%

3.5. Optimization of Training Population Configurations

We tested different training population sizes ($N = 50$ to $N = 300$) to determine the minimum size needed for reliable predictions. Accuracy across all traits expanded rapidly as the training population grew from 50 to 200 genotypes, after which returns diminished (Table 4). Maintaining a training set of at least 200 individuals provided a balanced combination of high accuracy and reasonable operational costs.

Table 4: Influence of training population sample size on cross-validated prediction accuracies.

Training Size (N)	Days to 50% Flowering	Days to Maturity	Grain Yield (GY)	Yield Stability (ASV)
N = 50	0.38	0.35	0.18	0.14
N = 100	0.54	0.51	0.29	0.26
N = 150	0.66	0.63	0.37	0.38
N = 200	0.71	0.69	0.42	0.44
N = 250	0.73	0.71	0.44	0.47
N = 300	0.74	0.72	0.45	0.48

3.6. Marker Density Optimization

To determine the most cost-effective approach for large-scale applications, we down-sampled the marker panel from 500 to the full set of 14,850 SNPs. Predictive performance for early maturity reached a plateau at approximately 3,000 SNPs, whereas yield and stability traits required higher marker densities (around 5,000 to 7,000 SNPs) to achieve comparable stability in accuracy (Table 5). This difference reflects the more complex polygenic nature of yield and environmental interactions, which require denser marker coverage to maintain effective linkage with causal loci.

Table 5: Effect of marker density configurations on prospective genomic prediction accuracy profiles.

Marker Subset Size	Days to 50% Flowering	Days to Maturity	Grain Yield (GY)	Yield Stability (βi)
500 SNPs	0.48	0.46	0.22	0.19
1,000 SNPs	0.59	0.58	0.31	0.28
3,000 SNPs	0.70	0.69	0.39	0.38
5,000 SNPs	0.72	0.71	0.42	0.44
7,000 SNPs	0.73	0.71	0.44	0.47
14,850 SNPs	0.74	0.72	0.45	0.49

3.7 Forward Validation in Candidate Selection Trials

To validate the model's practical utility, we used GBLUP-derived GEBVs to select the top 10% highest-performing, early-maturing, stable genotypes from an independent validation cohort of 150 unphenotyped lines. When evaluated in the field during the 2025-2026 season, this genomic-selected group showed an average reduction of 6.4 days in time to maturity and an 14.2% increase in grain yield under drought conditions compared to the baseline population mean (Table 6).

Table 6: Field performance confirmation of candidate lines selected using the GEBV framework.

Selected Cohort Group	Observed DF (days)	Observed DM (days)	Mean Yield (kg/ha)	Observed ASV
Top 10% GEBV Selected	\$48.2 \pm 2.1\$	\$103.4 \pm 3.2\$	\$1860.5 \pm 110.4\$	\$0.48 \pm 0.08\$
Randomly Selected Control	\$55.8 \pm 4.6\$	\$109.8 \pm 6.1\$	\$1628.4 \pm 195.8\$	\$1.15 \pm 0.34\$
Breeding Progress Shift	-13.6%	-5.8%	+14.2%	-58.2%

4. Discussion

4.1. Genetic Architecture and Model Performance

Our results demonstrate that genomic selection can reliably predict early maturity and yield stability traits in chickpea. The high prediction accuracies obtained for days to 50% flowering (0.74) and days to maturity (0.73) reflect the highly heritable nature of these traits. They also indicate that the marker density used was sufficient to capture key chromosomal regions regulating the transition to reproduction, such as the Efl-1 locus and orthologs of the FLOWERING LOCUS T (FT) gene family (Mallikarjuna *et al*., 2017) [20].

The minimal differences in performance among the parametric models (rrBLUP, GBLUP, and Bayesian alternatives) for maturity traits suggest a genetic architecture driven by a large number of minor loci acting additively (Heslot *et al*., 2012) [21]. However, for yield stability indices, the non-parametric Random Forest model showed a clear advantage, outperforming rrBLUP by 17%. This highlights the importance of accounting for non-additive effects, such as epistasis and multi-marker interactions, when modeling complex traits like environmental responsiveness and adaptation (Howard *et al*., 2014) [22].

4.2. Value of Multi-Environment Models and G×E Integration

A key finding of this study is the significant increase in yield prediction accuracy achieved by integrating multi-environment covariance models. Traditional single-environment GS models assume that marker effects remain constant across different conditions. However, in unpredictable arid and semi-arid environments, G×E interactions often alter marker-trait associations (Jarquín *et al*., 2014) [23].

By using an extended GBLUP framework that incorporates environment-specific genomic relationship matrices, we captured these interaction components directly. This approach

raised prediction accuracies for grain yield by over 23% across all environments. These improvements match findings in other major crops like wheat (Perez-Rodriguez *et al*., 2017) [24] and maize (Cuevas *et al*., 2016) [25], confirming that multi-environment modeling is an effective strategy for handling complex, low-heritability traits in erratic climates.

4.3. Operational Parameters for Chickpea Breeding Programs

To deploy genomic selection efficiently in breeding programs, it is important to balance operational costs against prediction accuracy (Gorjanc *et al*., 2015) [26]. Our down-sampling experiments indicate that a training population size of approximately 200 genotypes, paired with a marker density of 3,000 to 5,000 SNPs, provides a reliable setup for predicting maturity and yield stability in chickpea.

Using high-density SNP arrays for initial discovery, followed by targeted, low-cost genotyping platforms (such as amplicon sequencing or mid-density SNP panels) for routine screening, offers a cost-effective path forward (Saatchi *et al*., 2014) [27]. This strategy reduces genotyping costs while maintaining sufficient genome coverage to capture key linkage blocks across the diversity panel.

4.4. Implications for Accelerating Genetic Gain

Implementing genomic selection can significantly speed up breeding progress in chickpea. In a conventional breeding workflow, evaluating yield stability requires multiple years of multi-location trials, typically delaying advanced line selection until the generations (Varshney *et al*., 2012) [28].

By applying GS predictive models to early-generation materials, breeders can screen thousands of candidates using DNA markers alone. This allows lines with poor predicted yield stability or late maturity profiles to be discarded early, focusing field resources on the most promising material. This approach can shorten the breeding cycle by 3 to 4 years, increasing the rate of genetic gain and accelerating the development of climate-resilient chickpea varieties (Hickey *et al*., 2019) [29] (Roorkiwal *et al*., 2022) [30].

5. Conclusions

This study validates the use of genomic selection to simultaneously improve early maturity and yield stability in chickpea. Incorporating multi-environment components into prediction models successfully addressed the challenges of G×E interactions, improving yield predictability across diverse environments.

Our findings indicate that a cost-effective breeding pipeline can be built using a training population of 200 genotypes and a mid-density panel of 5,000 SNPs. Moving forward, integrating genomic selection with high-throughput field phenotyping and speed breeding techniques offers a promising approach to develop high-yielding, short-duration chickpea varieties adapted to changing global climates.

References

1. Varshney RK, Roorkiwal M, Bharadwaj C, *et al*. Genomic breeding in chickpea for accelerating genetic gain. Theor Appl Genet. 2021;134(6):1723-1739.
2. Gaur PM, Samineni S, Krishnamurthy L, *et al*. High temperature tolerance in chickpea: Enhancement of expression through genomic tools. Plant Sci. 2019;285:92-101.
3. Millan T, Clarke HJ, Siddique KHM, *et al*. Chickpea molecular breeding: Outputs and advancements. Field Crops Res. 2015;182:76-85.
4. Richards RA. Selectable traits to increase crop photosynthesis and yield of grain crops in drought-prone environments. J Exp Bot. 2020;71(17):5211-5225.
5. Cooper M, Podlich DW. Genotype-by-environment interactions in plant breeding challenges. In Silico Biol. 2002;2(3):245-257.
6. Finlay KW, Wilkinson GN. The analysis of adaptation in a plant-breeding programme. Aust J Agric Res. 1963;14(6):742-754.
7. Bernardo R. Molecular markers and selection for complex traits in plants: Learnings from the last two decades. Crop Sci. 2016;56(4):1425-1437.
8. Jannink JL, Lorenz AJ, Iwata H. Genomic selection in plant breeding: From theory to practice. Brief Funct Genomics. 2010;9(2):166-177.
9. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157(4):1819-1829.
10. Crossa J, Perez-Rodriguez P, Cuevas J, *et al*. Genomic selection in plant breeding: Methods, models, and perspectives. Trends Plant Sci. 2017;22(11):961-975.
11. Roorkiwal M, Rathore A, Das RR, *et al*. Genome-wide prediction architectures for enhancing genetic gains in chickpea breeding lines. Sci Rep. 2018;8(1):12410.
12. Sinha P, Singh VK, Saxena RK, *et al*. Genomic selection strategies for pulse crop improvement under changing climates. Front Plant Sci. 2021;12:654321.
13. IBPGR. Chickpea Descriptors and Evaluation Minimum Standards. Rome: International Board for Plant Genetic Resources; 1993.
14. Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11-15.
15. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112-118.
16. Endelman JB. Ridge regression and alternative linear models for genomic selection in R. Plant Genome. 2011;4(3):250-255.
17. VanRaden PM. Efficient methods to compute genomic relationship matrices. J Dairy Sci. 2008;91(11):4414-4423.
18. Park T, Casella G. The Bayesian lasso. J Am Stat Assoc. 2008;103(482):681-686.
19. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32.
20. Mallikarjuna BP, Samineni S, Thudi M, *et al*. Molecular mapping of flowering time major loci Efl-1 in short-duration chickpea lines. Front Plant Sci. 2017;8:1125.
21. Heslot N, Yang HP, Sorrells ME, *et al*. Genomic selection efficiency across alternative parametric architectures in small grains. Crop Sci. 2012;52(1):145-160.
22. Howard R, Carriquiry AL, Beavis WD. Parametric and nonparametric machine learning methods for genomic prediction of complex traits. G3 Genes Genomes Genet. 2014;4(6):1021-1030.
23. Jarquín D, Crossa J, Lacaze X, *et al*. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor Appl Genet. 2014;127(3):595-607.
24. Perez-Rodriguez P, Crossa J, Rutkoski J, *et al*. Single-step multi-environment models for genomic validation across wheat networks. Crop Sci. 2017;57(5):2341-2356.
25. Cuevas J, Granato I, Crossa J, *et al*. Genomic prediction with multi-environment interactions using kernel-based regression models in maize. G3 Genes Genomes Genet. 2016;6(12):3981-3991.

26. Gorjanc G, Cleveland MA, Houston RD, *et al*. Potential of low-density genotyping strategy paired with imputation architectures for genomic selection. BMC Genetics. 2015;16(1):45.
27. Saatchi M, Schnabel RD, Taylor JF. Optimal marker designs for low-cost genomic prediction panels across production cattle breeds. Genet Sel Evol. 2014;46(1):34.
28. Varshney RK, Ribaut JM, Buckler ES, *et al*. Can genomics deliver climate-change resilient crops? Theor Appl Genet. 2012;124(4):599-610.
29. Hickey LT, Hafeez AN, Robinson H, *et al*. Breeding crops at speed to feed the growing global population. Nat Biotechnol. 2019;37(7):744-754.
30. Roorkiwal M, Bharadwaj C, Bindra S, *et al*. Multi-location genomic prediction training applications for accelerating legume deployment cycles. Legume Sci. 2022;4(2):e144.