



# Empirical Comparison of Tropical Maize Hybrids Selected Through Genomic and Phenotypic Selections

Yoseph Beyene<sup>1\*</sup>, Manje Gowda<sup>1</sup>, Michael Olsen<sup>1</sup>, Kelly R. Robbins<sup>2</sup>, Paulino Pérez-Rodríguez<sup>3</sup>, Gregorio Alvarado<sup>4</sup>, Kate Dreher<sup>4</sup>, Star Yanxin Gao<sup>2</sup>, Stephen Mugo<sup>1</sup>, Boddupalli M. Prasanna<sup>1</sup> and Jose Crossa<sup>4</sup>

<sup>1</sup> Global Maize Program, International Maize and Wheat Improvement Center (CIMMYT), Nairobi, Kenya, <sup>2</sup> School of Integrative Plant Sciences, Cornell University, Ithaca, NY, United States, <sup>3</sup> Colegio de Postgraduados, Montecillos, Mexico, <sup>4</sup> Genetic Resources Program, International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico

## OPEN ACCESS

### Edited by:

Jacqueline Batley,  
University of Western Australia,  
Australia

### Reviewed by:

Wenxin Liu,  
China Agricultural University (CAU),  
China  
Reif Jochen,  
Leibniz Institute of Plant Genetics and  
Crop Plant Research (IPK),  
Germany

### \*Correspondence:

Yoseph Beyene  
y.beyene@cgiar.org

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

**Received:** 12 August 2019

**Accepted:** 29 October 2019

**Published:** 22 November 2019

### Citation:

Beyene Y, Gowda M, Olsen M, Robbins KR, Pérez-Rodríguez P, Alvarado G, Dreher K, Gao SY, Mugo S, Prasanna BM and Crossa J (2019) Empirical Comparison of Tropical Maize Hybrids Selected Through Genomic and Phenotypic Selections. *Front. Plant Sci.* 10:1502. doi: 10.3389/fpls.2019.01502

Genomic selection predicts the genomic estimated breeding values (GEBVs) of individuals not previously phenotyped. Several studies have investigated the accuracy of genomic predictions in maize but there is little empirical evidence on the practical performance of lines selected based on phenotype in comparison with those selected solely on GEBVs in advanced testcross yield trials. The main objectives of this study were to (1) empirically compare the performance of tropical maize hybrids selected through phenotypic selection (PS) and genomic selection (GS) under well-watered (WW) and managed drought stress (WS) conditions in Kenya, and (2) compare the cost-benefit analysis of GS and PS. For this study, we used two experimental maize data sets (stage I and stage II yield trials). The stage I data set consisted of 1492 doubled haploid (DH) lines genotyped with rAmpSeq SNPs. A subset of these lines (855) representing various DH populations within the stage I cohort was crossed with an individual single-cross tester chosen to complement each population. These testcross hybrids were evaluated in replicated trials under WW and WS conditions for grain yield and other agronomic traits, while the remaining 637 DH lines were predicted using the 855 lines as a training set. The second data set (stage II) consists of 348 DH lines from the first data set. Among these 348 best DH lines, 172 lines selected were solely based on GEBVs, and 176 lines were selected based on phenotypic performance. Each of the 348 DH lines were crossed with three common testers from complementary heterotic groups, and the resulting 1042 testcross hybrids and six commercial checks were evaluated in four to five WW locations and one WS condition in Kenya. For stage I trials, the cross-validated prediction accuracy for grain yield was 0.67 and 0.65 under WW and WS conditions, respectively. We found similar responses to selection using PS and GS for grain yield other agronomic traits under WW and WS conditions. The top 15% of hybrids advanced through GS and PS gave 21%–23% higher grain yield under WW and 51%–52% more grain yield under WS than the mean of the checks. The GS reduced the cost by 32% over the PS with similar selection gains. We concluded that the use of GS for yield under WW and WS conditions in maize can produce selection candidates with similar performance as those generated

from conventional PS, but at a lower cost, and therefore, should be incorporated into maize breeding pipelines to increase breeding program efficiency.

**Keywords:** phenotypic selection, genomic selection, genetic gain, maize, well-watered and water stress environments

## INTRODUCTION

With more than 35 million ha harvested each year, maize is the most important staple food crop in sub-Saharan Africa (SSA). In SSA countries, maize is commonly grown by resource-poor farmers and covers large areas with very low average grain yield (1.4 ton/ha) (Smale et al., 2011). The low productivity of maize in SSA is due to several factors including drought and low soil nitrogen stress, foliar diseases, and insect pests among others. The ability to quickly develop germplasm with resistance to important abiotic and biotic stresses will be critical for the resilience of Africa's maize-based cropping systems in the face of climate change. Breeding for drought tolerance and yield stability is an important objective of maize breeding programs in SSA, and a high priority for the International Maize and Wheat Improvement Center (CIMMYT) (Bänziger et al., 2006; Beyene et al., 2017; Cairns and Prasanna, 2018). Over the past decades, CIMMYT and its partners have made significant progress developing maize germplasm that is tolerant to drought, low soil nitrogen, and diseases including maize lethal necrosis (Beyene et al., 2017; Cairns and Prasanna, 2018). To accelerate breeding for drought tolerance, CIMMYT and its partners adopted several breeding approaches including pedigree selection, marker-assisted recurrent selection (MARS), and genomic selection (GS) combined with high-throughput phenotyping, doubled haploids (DH), and year-round nurseries (Beyene et al., 2016; Cairns and Prasanna, 2018).

Genomic prediction is an approach that uses molecular marker data to predict the genetic value of complex traits in progeny for selection and breeding (Meuwissen et al., 2001). When genomic predictions are used to make selections, the process is referred to as GS. The primary difference between GS and traditional forms of marker-assisted selection (MAS) is the simultaneous use of many markers distributed genome-wide, as opposed to a small set of markers linked to quantitative trait loci (Heffner et al., 2009). The objective of GS is to determine the genetic potential of an individual instead of identifying the specific quantitative trait loci. GS could be used in plant breeding programs in rapid recombination cycles and predict the breeding value of untested parents (genomic estimated breeding value, GEBV). Another way of using GS is with sparse testing where some lines are tested in some environments but predicted in others. Implementing genomic prediction and selection requires development of appropriate training sets consisting of individuals that have been both phenotyped and genotyped, followed by model calibration. Bernardo and Yu (2007) were the first to report the use of GS in maize breeding using simulation data. Massman et al. (2013) used real data to compare GS and MARS in a bi-parental maize population derived from temperate lines and reported that GS gave a 14 to 50% advantage over MARS for grain yield and stover

quality. Beyene et al. (2015) reported genetic gains through GS in eight CIMMYT tropical bi-parental maize populations evaluated under managed drought conditions in SSA. The authors showed that (i) the average gain per cycle from GS was 0.086 t/ha under managed drought conditions, (ii) the average grain yield of cycle 3-GS-derived hybrids was significantly higher than that of hybrids derived from  $C_0$ , and (iii) three GS cycles can be achieved in one year. On the other hand, the average gain per cycle using MARS across 10 populations was 0.051 t/ha per cycle under managed drought stress (Beyene et al., 2016). Vivek et al. (2017) reported that the realized genetic gain per year was higher for GS than for phenotypic selection (PS) in two bi-parental populations.

As pointed out by Crossa et al. (2017), the main advantages of GS as compared to phenotype-based selection in breeding are: (i) GS reduces the cost per cycle, and (ii) it increases time efficiency of variety development. For example, in terms of cost reduction in maize breeding, the breeder can testcross 50% of all available lines, evaluate them in first-stage multi-location trials, and then use the phenotypic data to predict the remaining 50% by GS. The time efficiency advantage over PS could come from the second selection cycle, which uses the training population from the previous cycle to predict the new inbred lines, thus excluding testcross formation and first-stage multi-location evaluation trials. As more robust, multi-year training sets are developed, GS can be used to advance the best selection candidates directly to the second stage of multi-location evaluations. This significantly reduces the cost of testcross formation and evaluation in the earliest stage of multi-location multi-year evaluation.

Although testing predictive ability is critical for gathering information for GS, there is a large gap between the findings of these studies and their application in breeding programs (Bernardo, 2016). In maize breeding, the potential of GS was empirically evaluated (Crossa et al., 2010; Windhausen et al., 2012; Massman et al., 2013; Beyene et al., 2015). CIMMYT maize breeding programs have evaluated several GS-related methods with varying levels of success over the past 8 years (Crossa et al., 2010; Burgueño et al., 2012; Windhausen et al., 2012; Beyene et al., 2015; Edriss et al., 2017; Zhang et al., 2017). More recently, some authors have considered using genomic models for predicting hybrid performance (Kadam et al., 2016; Cantelmo et al., 2017; Acosta-Pech et al., 2017; Vélez Torres et al., 2018); these studies have shown that genomic models can give reasonably accurate predictions of the agronomic performance of hybrids. Several of those studies have investigated the accuracy of genomic predictions but there is little empirical evidence on the practical performance of lines selected based on phenotype and GS (untested lines selected solely based on GEBV) in advanced yield trials.

The current study compares the performance of maize DH line testcrosses selected based on GS versus PS in second stage

multi-location yield trials of the CIMMYT maize breeding program in SSA. For this study, we used two experimental maize data sets: first-stage multi-location yield trials (hereafter referred to as stage I) and second-stage multi-location yield trials (hereafter referred to as stage II). The stage I data set consisted of 1492 DH lines genotyped with rAmpSeq (repeat amplification and sequencing) dominant sequence tag markers (<https://doi.org/10.1101/096628>). A subset of these lines (855) was crossed with an individual single-cross tester chosen to complement each specific population and evaluated in replicated trials under optimum and drought conditions for grain yield and other agronomic traits, while the remaining 637 DH lines were predicted using the 855 lines as a training set. The second data set (stage II) consisted of 348 DH lines from the first data set (stage I), of which 172 lines were selected solely based on GEBVs, and 176 lines were selected based on phenotypic performance. In this second data set, each of the 348 DH lines was crossed with three common testers from complementary heterotic groups, and the resulting 1042 testcross hybrids and six commercial checks were evaluated in 4-5 optimum locations and one location with managed drought conditions in Kenya. The objectives of this research were to (1) empirically compare the performance of tropical maize hybrids selected through PS and GS under stress and non-stress conditions, and (2) compare the cost-benefit of genomic and PS in tropical maize.

## MATERIALS AND METHODS

### Plant Materials Used in the Study

The first data set (stage I) comprised a total of 1492 DH lines derived from 12 bi-parental DH populations developed at CIMMYT's Maize DH facility in Kiboko, Kenya. The 12 source populations were obtained by crossing elite CIMMYT maize lines (CMLs) with La Posta Seq C7, a drought tolerant population developed at CIMMYT, Mexico, through recurrent selection among full sib/S1 families (Edmeades et al., 1999). The selected CMLs were drought tolerant lines that have good combining abilities and are adapted across several environments in SSA (Beyene et al., 2013). The DH lines were grown at the Kenya Agricultural and Livestock Research Organization Kiboko Research Station during the 2015/16 short rainy season. Based on the results of *per se* evaluation (germination and good stand establishment, plant type, low ear placement, and well-filled ears), 1492 DH lines were selected for stage I multi-location yield trials (Table 1). The smallest DH family comprised 34 lines, while the largest had 240 lines.

### Field Evaluation of Stage I Hybrid Yield Trials

To implement GS in CIMMYT's maize breeding program, nearly half (855) of 1492 selected DH lines were crossed with a single-cross tester from complementary heterotic group and phenotyped across locations. The 855 hybrids were divided into 14 trials connected by common checks. In each trial, three to six commercial checks were included and planted in an alpha-lattice

**TABLE 1** | List of 12 bi-parental maize populations used in this study.

No.	Population name	# Doubled haploid (DH) lines genotyped	# DH lines phenotyped
1	CML440/LPS-F64	34	34
2	CML445/LPS-F64	181	91
3	CML312/LPS-F64	185	93
4	CML442/LPS-F64	240	126
5	CML505/LPS-F64	162	81
6	CZL04003/LPS-F64	134	67
7	CML536/LPS-F64	180	86
8	CML537/LPS-F64	110	55
9	CML538/LPS-F64	40	40
10	CML540/LPS-F64	51	51
11	ZEWAc1F2-134-4-1-B-1-B*4-1-2-B-B/LPS-F64	75	75
12	CML312/CML540	100	52
	<b>Total</b>	<b>1492</b>	<b>851</b>

design with two replications and phenotyped in three well-watered (WW) environments and one managed drought stress (WS) environment in Kenya during the 2017 growing season. The WS experiment was conducted during the dry (rain-free) season by suspending irrigation starting 2 weeks before flowering until harvest, whereas the WW experiments were conducted during the rainy season, applying supplemental irrigation as needed. Entries were planted in two-row plots, 5 m long, with 0.75 m spacing between rows and 0.25 m between hills. Two seeds per hill were initially planted and then thinned down to one plant per hill three weeks after emergence to obtain a final plant population density of 53,333 plants per hectare. Fertilizers were applied at the rate of 60 kg N and 60 kg P<sub>2</sub>O<sub>5</sub> per ha, as recommended for the area. Nitrogen was applied twice: at planting and 6 weeks after emergence. Fields were kept free of weeds by hand weeding. The following traits were measured: grain yield (GY, tons ha<sup>-1</sup>), anthesis date (AD, days), plant height (PH, cm), grain moisture (MOI, %), gray leaf spot (GLS, 1–5 rating score), and turicum leaf blight (TLB, 1–5 rating score). Plots were manually harvested and GY was corrected to 12.5% moisture. AD was measured from planting to when 50% of the plants shed pollen, and PH was measured from the soil surface to the flag leaf collar on five representative plants within each plot.

### Genotyping 1492 DH Lines Using rAmpSeq

Leaf samples were taken from each of the 1492 DH lines and sent to Intertek, Sweden, for DNA extraction. The DNA sample plates were forwarded to the Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA, for genotyping with repetitive sequences (rAmpSeq markers) as per the procedure described by Buckler et al. (2016). Each sample was first amplified with PCR, and DNAs within each batch were pooled and multiplexed for rAmpSeq sequencing, a new genotyping technology which is used to amplify repetitive (LTR/retroelements) regions of the genome. A K-mer based approach was used to design the primer pairs which range between 150-bp to 200-bp in length and target ~1,500–2,000 loci in the genome (Buckler et al., 2016, <http://www.biorxiv.org/content/early/2016/12/24/096628>). With the

availability of thousands of adaptors, this technology makes it possible to genotype 3,000 samples in a single sequencing run and dramatically reduces the genotyping cost per sample. A total of 4657 markers that passed quality control were used for GS.

## Field Evaluation of Advanced (Stage II) Hybrid Trials Selected Through Phenotypic and GS

From stage I analyses, the top performing 348 (23%) DH lines were chosen for stage II evaluation. Among these 348 DH lines, 172 lines represented selection from the 637 genomic predicted lines that had above average GEBVs and 176 lines were selected from the 855 phenotyped lines that had above average Best Linear Unbiased Estimates (BLUE). Each of these DH lines were crossed with three common testers from complementary heterotic groups. The resultant 1042 testcross hybrids were evaluated in eight connected trials. Six commercial checks were included in each trial and planted in an alpha-lattice design with two replications and phenotyped in 4-5 WW environments and one WS environment in Kenya in 2018. The WS experiment was conducted during the dry (rain-free) season by suspending irrigation starting 2 weeks before flowering until harvest, whereas the WW experiments were conducted during the rainy season, applying supplemental irrigation as needed. Planting and agronomic managements were similar as explained for stage I trials. The following traits were measured: grain yield (GY, tons ha<sup>-1</sup>), anthesis date (AD, days), plant height (PH, cm), grain moisture (MOI, %), gray leaf spot (GLS, 1-5 rating score), and turcicum leaf blight (TLB, 1-5 rating score). Plots were manually harvested and GY was corrected to 12.5% moisture.

## Phenotypic Data Analysis

There were two sets of phenotypic field trials; the first set included 855 hybrids used to predict the performance of unobserved 637 lines (stage I), and the second set (stage II) was made up of 1042 hybrids from 348 DH lines (172 lines selected from GEBV alone and 176 lines selected based on phenotypic data) crossed with three testers. Note that the second set of field trials was used to compare the performance of the GS vs PS of hybrids. All the phenotypic analyses were done to obtain the variance components and BLUEs for the lines under WW and WS. All testcrosses were evaluated in different trials but adjacent to each other and connected by common checks in the same field. Phenotypic data was analyzed first within trials and then across trials.

The BLUEs across WW and WS locations for each trial and each trait were generated using the following linear mixed model carried out using the META-R software (Alvarado et al., 2017):

$$Y_{ijk} = \mu + L_j + R_r(L_j) + B_k[R_r(L_j)] + G_i + GL_{ij} + \epsilon_{ijk}$$

where  $Y_{ijk}$  is the grain yield of genotype  $i$  at location  $j$  in replicate  $r$  within block  $k$ ;  $\mu$  is the general mean;  $L_j$  is the fixed effect of location  $j$ ;  $R_r(L_j)$  is the fixed effect of replicate  $r$  within location  $j$ ;  $B_k[R_r(L_j)]$  is the random effect of incomplete block  $k$  within replicate  $r$  and

location  $j$  is assumed to be independently and identically normal distributed with mean zero and variance  $\sigma_{B(RL)}^2$ ;  $G_i$  is the effect of line  $i$ ;  $GL_{ij}$  is the effect of the line  $\times$  location interaction; and  $\epsilon_{ijk}$  is the random residual error assumed independent and identically normal distributed with mean zero and variance  $\sigma_{\epsilon}^2$ . The variance components and heritability across locations for WW and WS sites were computed. The BLUE of each trait for the single WS location are obtained from the following model

$$Y_{irk} = \mu + R_r + B_k(R_r) + G_i + \epsilon_{irk}$$

The analysis across trials was also performed using similar model as those shown above but including the trial as fixed effect.

## Genomic-Enabled Prediction Models for Stage-I Yield Trial Data

The BLUE of the entries within and across testers were used for genome-based predictions. GEBVs were calculated for GY, AD, MOI and PH using the BGLR statistical R-package (Pérez-Rodríguez and de los Campos, 2014) within and across testers for WW and WS sites. For genome-enabled prediction, a total of 4657 markers that passed quality control were selected. For GS, the Genomic Best Linear Unbiased Predictor (G-BLUP) model was employed. Further, to understand the effect of testers on prediction accuracy, GS was applied to predict unobserved lines within and across testers.

The models described below were used with two purposes: one was to use the 855 lines as a training set to predict the GEBV of 637 lines (testing set) and use the observed and predicted values to select top performing lines. The other objective of the models described below was to study the genome-based prediction accuracy of the 855 lines with phenotypic and genotypic data and determine the prediction accuracy using main effects and main effects plus interaction models for each tester and across testers.

## Environment + Genome Model (E + G + e)

This model can be expressed as

$$y_{ij} = \mu + E_i + g_j + e_{ij}$$

Where  $y_{ij}$  is the response trait for the  $j^{\text{th}}$  hybrid in the  $i^{\text{th}}$  environment,  $\mu$  is the overall mean, and  $E_i$  is the fixed effect of the site (either WW or WS). Here,  $g_j$  corresponds to the genomic breeding value of the  $j^{\text{th}}$  line defined as a linear combination of marker codes and the corresponding marker effects, such that

$g_j = \sum_{m=1}^p x_{jm} b_m$ , where  $p$  is the number of markers,  $x_{jm}$  is the marker code for the  $j^{\text{th}}$  line at the  $m^{\text{th}}$  marker position ( $m = 1, \dots, p$ ), and  $b_m$  is the corresponding marker effect. The marker effects are assumed identically and independently distributed

(IID) such that  $b_m \sim N^{\text{IID}}(\mathbf{0}, \sigma_b^2)$ , with  $\sigma_b^2$  being the variance component of the marker effects. The covariance matrix of the vector of genomic values  $g = \{g_j\}$  can be written as  $\text{Cov}(g) = \mathbf{G}\sigma_g^2$



where  $G$  is the genomic relationship matrix computed as  $G = \frac{XX'}{p}$  where  $X$  is the standardized genotype matrix (by columns),  $p$  is the number of markers, and  $\sigma_g^2 = p \times \sigma_b^2$  denotes the genomic variance component. Hence,  $\mathbf{g} = \{g_j\} \sim N(\mathbf{0}, \mathbf{G}\sigma_g^2)$ . Since molecular marker information varies across individuals (even within the same family), the estimated breeding values are unique for each genotype. Finally,  $e_{ij}$  is the residual assumed to have an identical and independent distribution (IID) such that  $e_{ij} \sim N(\mathbf{0}, \sigma_e^2)$  where  $\sigma_e^2$  is the residual variance.

Note that this model was used for the genomic prediction computed for the WW sites. The predictions for the unique managed sites had only the  $G + e$  component because these trials were established in only one managed drought site.

### Environment + Genome + Genome $\times$ Environment Model (E + G + GE + e)

This is the same as the previous model but includes the interaction term based on marker and environment interaction data. The model (Jarquín et al., 2014) is written as:

$$y_{ij} = \mu + E_i + g_j + gE_{ij} + e_{ij},$$

Where  $gE_{ij}$  is the component representing the interactions between molecular markers of the  $j^{\text{th}}$  line and the  $i^{\text{th}}$  environment. The distributional assumption for this term is such that  $\mathbf{gE} = \{gE_{ij}\} \sim N(\mathbf{0}, (\mathbf{Z}_g \mathbf{G} \mathbf{Z}_g') \otimes (\mathbf{Z}_E \mathbf{Z}_E') \sigma_{gE}^2)$  and  $\sigma_{gE}^2$  is the variance component of the random interaction component  $gE$ . The other terms were already defined in the previously defined E + G + e model. As already mentioned, this model was used for the genomic prediction computed for the WW sites. The prediction for the unique managed sites includes only the  $G + e$  component.

### Random Cross-Validation for Determining the Prediction Accuracy of the Models

The performance of the models when predicting the five traits was evaluated using the average Pearson's correlation coefficient between observed and predicted values. The random cross-validation scheme mimics real plant breeding situations and is a scheme where the performance of 20% of the maize testcrosses was not observed in any of the environments and the rest of the lines (80%) were already observed in the same target environments. For this scheme, a five-fold random partitioning (80% of the data used as the training set, and the remaining 20% as the testing set) was employed. Four folds were used for training the models and for predicting the remaining fold. This procedure was repeated over the five folds and the predictions from the testing fold were joined in a single vector. Then, Pearson's correlations between predicted and observed values within the same environment were computed. The partitioning was repeated 100 times. The cost benefits of PS vs GS were analyzed using spreadsheet-based budgeting tools.

## RESULTS

### Test Cross Hybrid Performance in Stage I and II Trials

For stage I, mean GY averaged across WW locations ranged from 3.49 to 9.14 t/ha with an overall mean of 6.03 t/ha, whereas at stage II it improved further, ranging from 5.1 to 11.6 t/ha with an average of 7.59 t/ha (Table 2). Under WS, GY ranged from 1.08 to 5.76 t/ha with an overall mean of 3.25 t/ha at stage I, whereas at stage II, the range varied from 0.77 to 6.33 t/ha with an average of 3.23 t/ha. The average GY, PH, AD, and MOI were higher under WW conditions than under WS in stage I trials (Table 2). Interestingly, the average performance of hybrids at stage II was also higher for GY, AD, PH, and MOI compared to stage I trials. The magnitude of genotypic variances was higher than the genotype by environment interaction variances for all traits at stage I and only for GY and AD at stage II under WW conditions. For stage I trials, the heritability under WW conditions varied from 0.31 to 0.77, while under WS, it ranged from 0.84 to 0.95. Whereas at Stage II, the heritability under WW conditions varied from 0.18 to 0.82, while under WS it ranged from 0.32 to 0.61. A total 176 lines that had above average phenotypic value for grain yield and other agronomic traits were advanced to stage II evaluation (Supplementary Table S1). These superior lines were derived from all 12 populations, suggesting that the donor parents used to develop the DH lines are excellent sources of germplasm for combining ability with good adaptation to eastern Africa.

### Genomic-Enabled Predictions of GY and Other Agronomic Traits At Stage I

The cross-validation analyses yielded moderately high prediction correlations among optimum and drought conditions for GY and other agronomic traits. The prediction correlations ranged from 0.65–0.67 for GY, 0.57–0.65 for MOI, 0.67–0.75 for AD, and 0.70–0.72 for PH (Table 3). In general, trait–tester combination with higher training set (CML395/CML444) had a higher prediction accuracy than the other two testers (CML312/CML395 and CML312/CML442) that had lower training set (Table 3). The prediction accuracy was lowest for testers CML312/CML442 and CML312/CML395 for all traits and highest for tester CML395/CML444 both under WW and WS conditions.

A total 172 lines that had above average GEBVs were advanced to stage II evaluations (Supplementary Table S1).

### Comparing Hybrids Developed Through PS and GS Under WW Conditions At Stage II

At stage II, 1042 hybrids were evaluated among them 526 were developed from lines selected based on PS and the remaining were derived from lines selected based on GEBVs. The GY of 526 testcross hybrids advanced through PS evaluated across five WW locations (hereafter referred to as PS-WW) ranged from 5.54 to 11.67 t/ha (Supplementary Table S2). In the PS-WW, the top 15% of hybrids produced an average grain yield of 9.4 t ha<sup>-1</sup>, which represents an increase of 1.0 t ha<sup>-1</sup> compared to



**TABLE 4** | Comparison of hybrids advanced through genomic and phenotypic selections and commercial checks evaluated at stage II or advanced yield trials under optimum and managed drought stress across Kenya in 2018.

	Phenotypic selection (PS)		Genomic selection (GS)	
	Well-watered (GY, t/ha)	Water stress (GY, t/ha)	Well-watered (GY, t/ha)	Water stress (GY, t/ha)
All hybrids	7.7	3.2	7.5	3.2
Top 15% of hybrids	9.4	4.7	9.1	4.8
Best hybrid	11.7	6.2	10.4	6.3
Mean of commercial checks	7.2	2.3	7.2	2.3
Best check	8.4	3.3	8.4	3.3
<b>Yield Improvement</b>				
Top 15% of hybrids over commercial checks	23%	51%	21%	52%
Top 15% of hybrids over the best commercial check	12%	42%	8%	47%
The best hybrid over commercial checks	63%	169%	46%	175%
The best hybrid over the best commercial check	39%	88%	24%	92%

hybrid advanced through GS had 24 and 46% higher GY than the best commercial check and the mean of the commercial checks, respectively.

The top 15% of hybrids advanced through PS on average had an increase of 8.5 cm in PH and 4.7 days in AD and a 1% increase in grain moisture content compared to the mean of the commercial checks (**Supplementary Table S2**). However, there was no difference between the top 15% of hybrids and the commercial checks in their responses to the two main foliar diseases, GLS and TLB (**Supplementary Table S2**). The top 15% of hybrids advanced through GS had a 6 cm increase in PH and 4.7 days in flowering and a 1% increase in grain moisture content compared to the mean of the commercial checks (**Supplementary Table S2**). There was no difference among the top 15% of hybrids advanced through PS and GS for AD, MOI and their responses to the two main foliar diseases, GLS and TLB (**Supplementary Table S2**).

## Comparing Hybrids Developed Through PS and GS Under Managed Drought Stress

A total of 526 hybrids advanced through PS were also evaluated under managed drought stress (hereafter referred to as PS-WS); their mean GY ranged from 0.99 to 6.19 t ha<sup>-1</sup> (**Supplementary Table S3**). The top 15% of hybrids (79 hybrids) produced 42% and 51% higher mean GY than the mean GY of the best commercial check and the mean of the checks, respectively (**Table 4**). The best hybrid advanced through PS produced 88% and 169% higher GY than the best check and the mean of the commercial checks, respectively, while for 516 hybrids that were advanced through GS, their GY performance under WS ranged from 1.14 to 6.33 t/ha (**Supplementary Table S3**). The top 15% of hybrids produced 47% and 52% higher mean GY than the best commercial check and the mean of checks, respectively (**Table 4**).

The best hybrid advanced through GS produced 92% and 175% higher GY than the best check and mean of commercial checks, respectively (**Table 4**). Compared to the mean of commercial checks, the top 15% of hybrids advanced through PS

and GS had a 3% increase in MOI, 2.8 days in AD, and a 10 cm increase in PH (**Supplementary Table S3**).

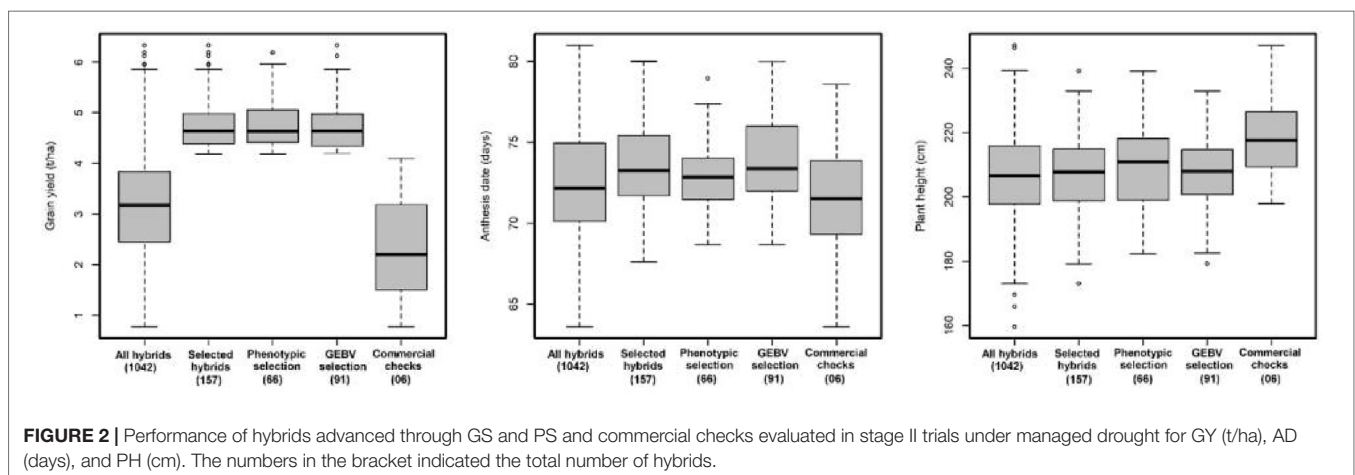
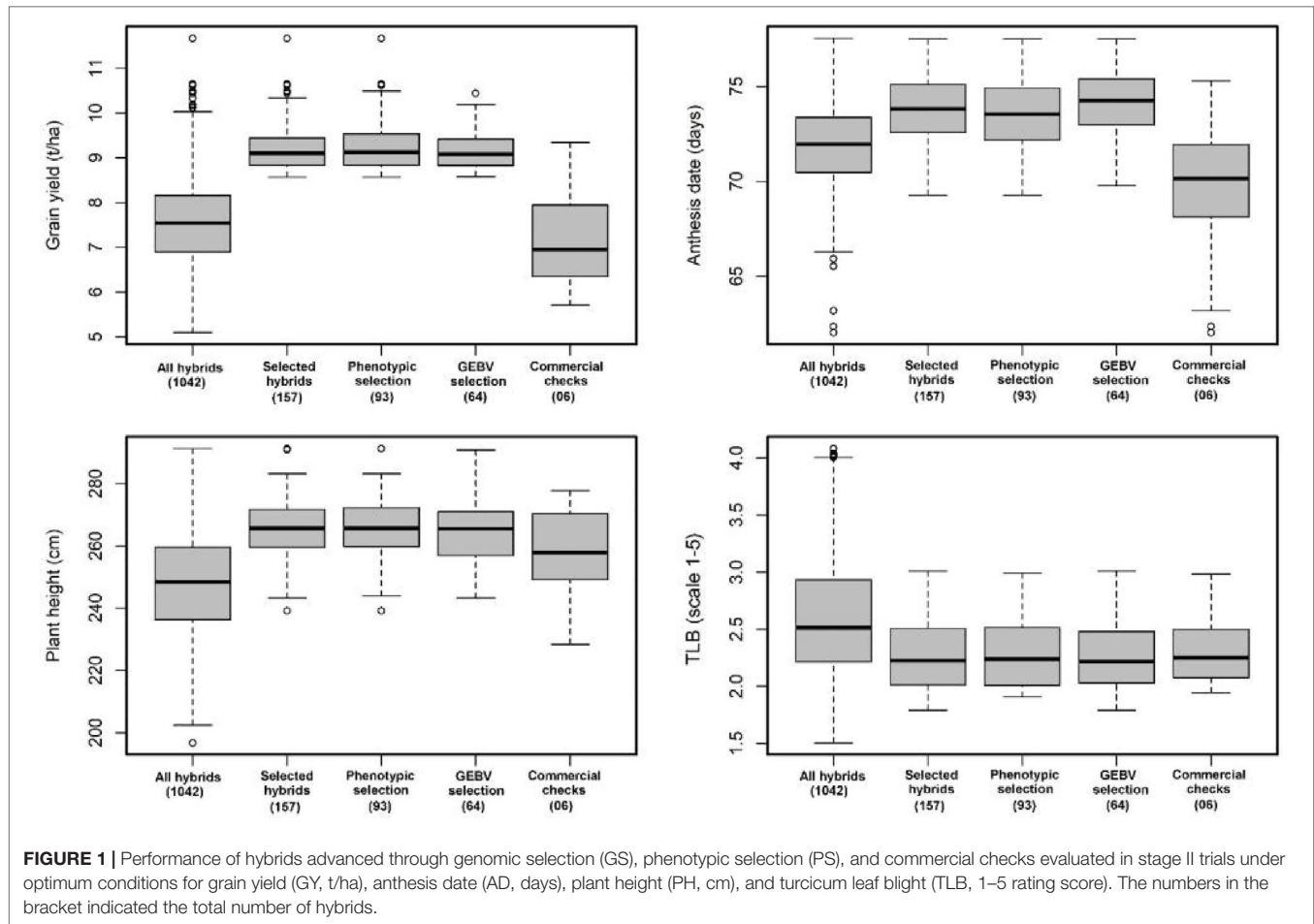
## Advancement Rate of Stage II Candidates to Stage III

An additional metric of interest when considering the overall efficacy of GS as a substitute for conventional PS only schemes is the advancement rate of the GS stage II cohort compared with the advancement rate of the PS stage II cohort. The actual advancement rate of the two methods is a useful means of comparing the overall value of the two groups of advanced lines since it captures all information that the breeder uses to make the decision whether or not to move a given DH line/s into advanced testing (**Figures 1 and 2**). Under WW condition the top 15% (157 of 1042 hybrids) advanced to stage III trials, 93 hybrids were developed through PS, and 64 hybrids were advanced through GS (**Figure 1**).

While under WS condition the top 15% (157 of 1042 hybrids) advanced to stage III trials, 91 hybrids were developed through GS, and 66 hybrids were advanced through PS. There was no significant difference among the top 15% of hybrids advanced through PS and GS for grain yield, AD, and PH (**Figure 2**). The overall advancement rate of GS stage II candidates was 41% and 58% compared with an advancement rate of 59% and 42% for PS stage II candidates under WW and WS condition, respectively, indicating that the two groups of selection candidates had functionally equivalent potential in terms of producing new stage III candidates.

## Cost-Benefit Analysis, PS Versus GS

We compared the costs involved in PS and GS using spreadsheet-based budgeting tools (**Table 5**). At present, for a single entry to make testcross and conduct two row plot yield trial costs US\$ 15 in Kenya. This value represented lower boundary cost because it was mainly based on operational costs excluding personnel and other costs. The cost of genotyping an entry is US\$ 10. Based on this rough estimate, developing and evaluating 1492 testcrosses in stage I trials at four locations with two replications per location would have cost US\$



134,480, while it costs US\$ 91,870 using a combination of phenotypic and GS. Therefore, using the current method (phenotyping half of the materials and predicting the remaining half), the same outcome was achieved with 68% of the phenotyping costs (Table 5). These costs are likely to vary in different breeding programs (primarily due to differences in labor costs), but the results indicated that GS was relatively cheaper than PS.

## DISCUSSION

With the advent of DH technology, thousands of fixed lines are generated each year in maize. However, identifying the best genotypes requires extensive field evaluations with several hybrid combinations, and all DH lines cannot be evaluated because of limited space and resources. One method



**TABLE 5 |** Cost–benefit analysis of phenotypic selection and genomic selection in International Maize and Wheat Improvement Center's (CIMMYT's) maize breeding program in Kenya.

Methods	Cost/entry (US\$)	No. of entries	No. of reps/sites	No. of rows/sites	No. of sites	Total cost (US\$)
PS (making testcrosses)	10	1492	1	1	1	14,920
PS (stage I multi-location yield trials)	5	1492	2	2	4	119,360
GS (making testcrosses)	10	855	1	1	1	8,550
GS (phenotyping training set in stage I multi-location yield trials)	5	855	2	2	4	68,400
GS (genotyping all lines)	10	1492	1	1	1	14,920
Total cost of GS						91,870
Total cost of PS						134,280
GS:PS cost ratio						0.68

for reducing the number of hybrids for field evaluation is crossing all DH lines with a common tester in the early stages of a breeding cycle. Another method is to use a genetic similarity matrix derived from pedigree or molecular markers for predicting performance of untested crosses (Kadam et al., 2016; Cantelmo et al., 2017). GS can be used in plant breeding to improve selection in the early stages of a breeding program without testing all available lines in yield trials. Recently several genomic prediction models (reviewed by Crossa et al., 2017) were used to predict the performance of tested and untested hybrids, and those predictions can be used to decide which hybrids should be selected for further evaluation in field trials.

In this study, we compared the performance of maize DH lines selected from stage I multi-location yield trials based on BLUEs and GEBVs by evaluating the hybrids in common stage II multi-location yield trials of the CIMMYT maize breeding program. We evaluated a total of 855 hybrids under optimum and drought conditions and used BLUEs data to predict the remaining 637 hybrids which were genotyped but have never been phenotyped. In our study, the prediction accuracy for GY under WW conditions was 0.67, and under WS, it was 0.65 (Table 3). Our results agree with previous maize studies by Crossa et al. (2010) and Zhang et al. (2017), who reported medium to high prediction accuracy. In our study, higher marker density, higher heritabilities, and similarity in training and prediction data sets gave higher accuracies of estimated breeding values. Lorenz and Smith (2015) reported that prediction accuracies for GS decreased as the similarity of individuals in the training and test populations decreased. In our case, the training populations used for GS are highly related and purposefully designed (genotyping all and phenotyping half) for reducing field phenotyping, which is costly and logistically complex. Our results agree with those of previous studies by Riedelsheimer et al. (2012) and Laripe et al. (2017), who concluded that prediction accuracies are enough to make GS more efficient than PS.

Identification of optimum size as training and prediction set is crucial for implementing GS in maize breeding program. Cao et al. (2017) reported that high prediction accuracy was observed when 50% of the total genotypes were used as a training population. Zhang et al. (2019) obtained moderate-to high

prediction accuracy trait–environment combinations, when half of the population is used to build the prediction model. In this study we have implemented genotyping all and phenotyping half prediction scheme to reduce the cost of phenotyping all DH line generated each year in stage I yield trials. The ultimate objective is to select untested lines based on GEBV from previous years that improve accuracy and go directly to Stage II yield trials by skipping stage I yield trials. This will require to build multi-year estimation set for specific germplasm groups and targeted growing regions.

## Comparing Performance of Hybrids Advanced to Stage II Under WW Conditions

The mean GY of hybrids advanced through GS and PS methods was significantly higher than the mean of the commercial checks (Figure 1). The top 15% of hybrids advanced through PS were slightly better than hybrids advanced through GS. The top 79 hybrids (15% SI) advanced through PS had mean GY of 9.4 t/ha, while the top 15% of hybrids (77 hybrids) advanced through GS had mean GY of 9.1 t/ha under WW conditions. When historical data from the same breeding program is available, there is the potential to bypass stage I trial evaluation and move material directly into stage II. This approach would reduce both the costs and cycle time but will require accurate predictions from training sets composed of historical data. GS has the potential for increasing the genetic gain per year by accelerating the breeding cycles (Crossa et al., 2017). Beyene et al. (2015) reported that GS can save the time over PS because three rapid cycles of recombination were possible to complete in a year. Gaynor et al. (2016) using simulation data proposed a two-part strategy for GS in plant breeding: namely, population improvement and product development. The population improvement strategy uses GS to perform rapid cycles of recurrent selection to minimize breeding cycle time and maximize the genetic gain per year, while the product development component focuses on developing inbred lines as hybrid parents. The authors concluded that implementing GS in breeding programs increases breeding program efficiency by reducing the cost of achieving a similar outcome.

## Comparing Performance of Hybrids Advanced to Stage II Under WS Conditions

Comparison of hybrids advanced through PS and GS under drought stress conditions revealed that GS did slightly better (4.68 t/ha was the mean of the top 15% of hybrids) than PS (4.48 t/ha, mean of the top 15% of hybrids). There was no significant difference among the top 15% of hybrids advanced through PS and GS for other traits. Longin et al. (2015) found an increased genetic gain when selecting parents based entirely on GEBV for highly heritable traits in wheat. The ultimate objective is to select untested lines based on GEBV from previous years that improve accuracy and go directly to advanced yield trials by skipping preliminary stage I yield trials. In CIMMYT maize breeding program, GS could be implemented to predict untested lines at the early stage of testing by selecting DH with good GEBV and going directly to stage II trials. This could reduce each breeding cycle to less than 2 years. The greatest benefit of GS for achieving genetic gains in crops will come from decreased cycle time (Heffner et al., 2009), as has been predicted and observed in GS of dairy cattle (Schaeffer, 2006; García-Ruiz et al., 2016).

## Cost-Benefit Analysis of GS vs. PS

GS was found to outperform MAS using the same financial investment, even at low prediction accuracies (Bernardo and Yu, 2007; Heffner et al., 2010). GS can be used to identify promising lines much sooner than PS, thereby reducing cycle time and increasing the genetic gain per year (Heffner et al., 2009). Several studies considered the economic aspects of plant breeding while comparing the evaluation of selection strategies. Abalo et al. (2009) found that compared to PS, MAS had 26% lower total operating costs for maize streak virus resistance. For GS, Heffner et al. (2010) found similar results in maize and winter wheat (*Triticum aestivum* L.). Comparison of the genetic gain per unit time and per unit cost for oil palm (*Elaeis guineensis* Jacq.) breeding under PS and GS also revealed higher genetic gain per unit cost for GS (Wong and Bernardo, 2008). In our study, GS reduced the cost by 32% over PS with similar selection gains. Currently, we are testing another set of lines with the aim of using historical data to predict new lines and bypass the first stage of testing, and also using both pedigree and marker data to improve the prediction accuracy, which significantly reduces costs and shortens the breeding cycle.

## CONCLUSIONS

The largest potential advantage of GS is predicting the breeding value of genotyped parents that were never phenotyped. We found similar responses to selection using PS and GS for grain yield

under WW and WS conditions. The top 15% of hybrids advanced through GS and PS produced 21% to 23% higher GY under WW and 52% to 51% under WS than the mean of the commercial checks. The GS reduced the cost by 32% over the PS with similar selection gains. We conclude that the use of GS for yield under optimum and drought conditions in tropical maize can produce selection candidates with similar performance as those generated from conventional PS, but at a lower cost; therefore, this strategy should be effectively incorporated into maize breeding pipelines to enhance breeding program efficiency.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

YB, MG, MO, BP, JC, KR, and SM contributed in the project planning and overall coordination. YB performed and coordinated the field experiments. MG, SG, and KD were responsible for coordinating sample management and genotyping. JC, PP, KR, GA, MG, and YB carried out the analysis. YB, JC, and MG wrote the manuscript. All authors have made their contribution in editing the manuscript and approved the final version.

## ACKNOWLEDGMENTS

This research was supported by the Bill and Melinda Gates Foundation, and the United States Agency for International Development (USAID) through Stress Tolerant Maize for Africa (STMA, Grant # OPP1134248), and the CGIAR Research Program MAIZE. The CGIAR Research Program MAIZE receives W1&W2 support from the Governments of Australia, Belgium, Canada, China, France, India, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Sweden, Switzerland, United Kingdom, United States, and the World Bank. We are grateful to CIMMYT Field Technicians at different stations in Kenya for data collection; CIMMYT Laboratory Technicians in Kenya for sample preparation for genotyping; and Cornell Genotyping facility and Buckler's group for genotyping services and data turn-around.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.01502/full#supplementary-material>

## REFERENCES

Abalo, G., Edema, R., Derera, J., and Tongoona, P. (2009). A comparative analysis of conventional and marker-assisted selection methods in breeding maize streak virus resistance in maize. *Crop Sci.* 49, 509–520. doi: 10.2135/cropsci2008.03.0162

Acosta-Pech, R., Crossa, J., de los Campos, G., Teyssedre, S., Claustres, B., Pérez-Elizalde, S., et al. (2017). Genomic models with genotype × environment interaction for predicting hybrid performance: an application in maize hybrids. *Theor. Appl. Genet.* 130, 1431–1440. doi: 10.1007/s00122-017-2898-0

Alvarado, G., López, M., Vargas, M., Pacheco, Á., Rodríguez, F., Burgueño, J., et al. (2017). META-R (Multi Environment Trial Analysis with R for Windows)

- Version 6.01, hdl:11529/10201, CIMMYT Research Data & Software Repository Network, V20. .
- Bänziger, M., Setimela, P. S., Hodson, D., and Vivek, B. (2006). Breeding for improved drought tolerance in maize adapted to southern Africa. *Agric. Water Manage.* 80, 212–224. doi: 10.1016/j.agwat.2005.07.014
- Bernardo, R. (2016). Bandwagons I, too, have known. *Theor. Appl. Genet.* 129, 2323–2332. doi: 10.1007/s00122-016-2772-5
- Bernardo, R., and Yu, J. M. (2007). Prospects for genome-wide selection for quantitative traits in maize. *Crop Sci.* 47, 1082–1090. doi: 10.2135/cropsci2006.11.0690
- Beyene, Y., Gowda, M., Suresh, L. M., Mugo, S., Olsen, M., Oikeh, S. O., et al. (2017). Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. *Euphytica.* 224, 1–13. doi: 10.1007/s10681-017-2012-3
- Beyene, Y., Mugo, S., Semagn, K., Asea, G., Trevisan, W., Tarekegne, A., et al. (2013). Genetic distance among doubled haploid maize lines and their testcross performance under drought stress and non-stress conditions. *Euphytica.* 192, 379–392. doi: 10.1007/s10681-013-0867-5
- Beyene, Y., Semagn, K., Crossa, J., Mugo, S., Atlin, G. N., Tarekegne, T., et al. (2016). Improving maize grain yield under drought stress and non-stress environments in Sub-Saharan Africa using marker-assisted recurrent selection. *Crop Sci.* 56, 344–353. doi: 10.2135/cropsci2015.02.0135
- Beyene, Y., Semagn, K., Mugo, S., Tarekegne, A., Babu, R., Meisel, B., et al. (2015). Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. *Crop Sci.* 55, 154–163. doi: 10.2135/cropsci2014.07.0460
- Buckler, E. S., Ilut, D. C., Wang, X., Kretschmar, T., Gore, M. A., and Mitchell, S. E. (2016). rAmpSeq: Using repetitive sequences for robust genotyping. bioRxiv. doi: 10.1101/096628
- Burguño, J., de los Campos, G., Weigel, K., and Crossa, J. (2012). Genomic prediction of breeding values when modeling genotype  $\times$  environment interaction using pedigree and dense molecular markers. *Crop Sci.* 52, 707–719. doi: 10.2135/cropsci2011.06.0299
- Cairns, J. E., and Prasanna, B. M. (2018). Developing and deploying climate-resilient maize varieties in the developing world. *Curr. Opin. In Plant Biol.* 45, 226–230. doi: 10.1016/j.pbi.2018.05.004
- Cantelmo, N. F., Von Pinho, R. G., Balestre, M., et al. (2017). Genome-wide prediction for maize single-cross hybrids using the GBLUP model and validation in different crop seasons. *Mol. Breed.* 37, 51. doi: 10.1007/s11032-017-0651-7
- Cao, S., Loladze, A., Yuan, Y., Wu, Y., Zhang, A., and Chen, J. (2017). Genome-wide analysis of tar spot complex resistance in maize using genotyping-by-sequencing SNPs and whole-genome prediction. *Plant Genome* 10, 1–14. doi: 10.3835/plantgenome2016.10.0099
- Crossa, J., de los Campos, G., Pérez-Rodríguez, P., Gianola, D., Burguño, J., Araus, J. L., et al. (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186, 713–724. doi: 10.1534/genetics.110.118521
- Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., et al. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trend Plant Sci.* 22, 961–975. doi: 10.1016/j.tplants.2017.08.011
- Edmeades, G. O., Bolan, J., Chapman, S. C., Lafitte, H. R., and Bänziger, M. (1999). Selection improves drought tolerance in tropical maize populations. *Crop Sci.* 39, 1306. doi: 10.2135/cropsci1999.3951306x
- Edriss, V., Gao, Y., Zhang, X., Jumbo, M. B., Makumbi, D., Olsen, M. S., et al. (2017). Genomic prediction in a large African maize population. *Crop Sci.* 57, 2361–2371. doi: 10.2135/cropsci2016.08.0715
- García-Ruiz, A., Cole, J. B., VanRaden, P. M., Wiggans, G. R., Ruiz-López, F. J., and Tassell, C. P. V. (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci. U.S.A.* 113 (28), E3995–E4004. doi: 10.1073/pnas.1519061113
- Gaynor, R. C., Gorjanc, G., Bentley, A. R., Ober, E. S., Howell, P., Jackson, R., et al. (2016). A Two-Part Strategy for Using Genomic Selection to Develop Inbred Lines. *Crop Sci.* 56, 2372–2386. doi: 10.2135/cropsci2016.09.0742
- Heffner, E. L., Lorenz, A. J., Jannink, J. L., and Sorrells, M. E. (2010). Plant breeding with genomic selection: Gain per unit time and cost. *Crop Sci.* 50, 1681–1690. doi: 10.2135/cropsci2009.11.0662
- Heffner, E. L., Sorrells, M. E., and Jannink, J. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512
- Jarquín, D., Crossa, J., Lacaze, X., Cheyron, P. D., Daucourt, J., Lorgeou, J., et al. (2014). A reaction norm model for genomic selection using high dimensional genomic and environmental data. *Theor. Appl. Genet.* 127, 595–607. doi: 10.1007/s00122-013-2243-1
- Kadam, D. C., Potts, S. M., Bohn, M. O., Lipka, A. E., and Lorenz, A. J. (2016). Genomic prediction of single crosses in the early stages of a maize hybrid breeding pipeline. *G3 Genes Genomes Genet.* 6, 3443–3453. doi: 10.1534/g3.116.031286
- Lariepe, A., Moreau, L., Laborde, J., Bauland, C., Mezouk, S., Decousset, L., et al. (2017). General and specific combining abilities in a maize (*Zea mays* L.) test-cross hybrid panel: Relative importance of population structure and genetic divergence between parents. *Theor. Appl. Genet.* 130, 403–417. doi: 10.1007/s00122-016-2822-z
- Longin, C. F. H., Mi, X. F., and Wurschum, T. (2015). Genomic selection in wheat: Optimum allocation of test resources and comparison of breeding strategies for line and hybrid breeding. *Theor. Appl. Genet.* 128, 1297–1306. doi: 10.1007/s00122-015-2505-1
- Lorenz, A. J., and Smith, K. P. (2015). Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. *Crop Sci.* 55, 2657–2667. doi: 10.2135/cropsci2014.12.0827
- Massman, J. M., Gordillo, A., Lorenzana, R. E., and Bernardo, R. (2013). Genome-wide predictions from maize single-cross data. *Theor. Appl. Genet.* 126, 13–22. doi: 10.1007/s00122-012-1955-y
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1534/genetics.110.116590
- Pérez-Rodríguez, P., and de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198, 483–495. doi: 10.1534/genetics.114.164442
- Riedelshimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., et al. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44, 217–220. doi: 10.1038/ng.1033
- Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123, 218–223. doi: 10.1111/j.1439-0388.2006.00595.x
- Smale, M., Byerlee, D., and Jayne, T. (2011). Maize revolutions in sub-Saharan Africa. The world bank development research group, agriculture and rural development team, pp. 34. doi: 10.1596/1813-9450-5659
- Vélez Torres, M., García Zavala, J. J., Hernández Rodríguez, M., Lobato-Ortiz, R., López-Reynoso, J. J., Benítez Riquelme, I., et al. (2018). Genomic prediction of the general combining ability of maize lines (*Zea mays* L.) and the performance of their single crosses. *Plant Breed.* 137, 379–387. doi: 10.1111/pbr.12597
- Vivek, B. S., Krishna, G. K., Vengadesan, V., Babu, R., Zaidi, P. H., Kha, L. Q., et al. (2017). Use of genomic estimated breeding values results in rapid genetic gains for drought tolerance in maize. *Plant Genome* 10, 1–8. doi: 10.3835/plantgenome2016.07.0070
- Windhausen, V. S., Atlin, G. N., Crossa, J., Hickey, J. M., Jannink, J.-L., Sorrells, M., et al. (2012). Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. *G3: Gene Genet. Genom.* 2, 1427–1436. doi: 10.1534/g3.112.003699
- Wong, C. K., and Bernardo, R. (2008). Genome-wide selection in oil palm: increasing selection gain per unit time and cost with small populations. *Theor. Appl. Genet.* 116, 815–824. doi: 10.1007/s00122-008-0715-5
- Zhang, X., Pérez-Rodríguez, P., Burguño, J., Olsen, M., Buckler, E., Atlin, G., et al. (2017). Rapid cycling genomic selection in a multiparental tropical maize population. *G3 Gene Genome Genet.* 7, 2315–2326. doi: 10.1534/g3.117.043141
- Zhang, H., Yin, L., Wang, M., Yuan, X., Liu, X. (2019). Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Front. Genet.* 10, 189. doi: 10.3389/fgene.2019.00189

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Beyene, Gowda, Olsen, Robbins, Pérez-Rodríguez, Alvarado, Dreher, Gao, Mugo, Prasanna and Crossa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.