Genomic-assisted sparse multi-location testing to increase genetic gains in barley

Julián García Maldonado

19/09/20022

Abstract

In recent years, breeders have been incorporating genomic predictions (GP) into their breeding programs as a useful, cost-effective, and versatile tool to increase genetic gain. Coupled with new experimental designs, genomic prediction maximizes the information quality and quantity per dollar invested. This is the case of genomic-assisted sparse multi-location testing, an experimental design that allows breeders to arrange multi-location trials having only a fraction of the genotypes present in all environments. In this design, the observed genotypes at each environment are used to produce genomic predictions of the non-planted ones and the overlapping genotypes across locations enable the genotype x environment (GxE) connectivity and its modeling. This results in an increase in selection accuracy and/or intensity and thereby higher genetic gain. This strategy has been adopted by the Global Barley Breeding Program of ICARDA as the standard approach for preliminary yield trials. To assess GP accuracies within and across locations, 1,000 new 2- and 6-row stage 1 entries of the Feed Barley for Arid and Semi-Arid Environments Mega-Product Line were assembled in a preliminary yield trial series with four locations. These lines were distributed in 340 plot sparse p-rep trials that included 212 unreplicated entries per location, 18 entries replicated within location (p-rep entries), 80 entries replicated across locations and 6 commercial checks replicated both within and across locations. These trials were planted in four diverse locations in Morocco and Lebanon, each identified as representative of a Target Population Environment. The phenotypic correlation among environments were all lower than r = 0.30. In order to maximize the kinship connectivity, the lines were distributed across locations based on the pedigree-based matrix of relationships. In addition, all 1,000 lines were genotyped using a marker diversity set of 96 SNP distributed across the genome and previously selected for their high minor allele frequency (MAF > 0.40) among the parents. Genomic predictions were calculated using GBLUP fitted using the ASReml-R, following a two-stages analysis. A scheme of 10-fold cross-validation was set up to calculate GP accuracies for the multi-environment trials (MET). Different variables were added to the genomic models to improve the accuracy of the predictions. These included, population structure, pedigree information (hybrid matrix) and row-type. These were added both independently or combined into the models. Best results were obtained when only the hybrid matrix was included in the models. Final GxE genomic prediction accuracies ranged between 0.11 to 0.40.

Resumen

En los últimos años, los mejoradores han ido incorporando predicciones genómicas (GP) en sus programas de mejora como una herramienta útil, rentable y versátil para aumentar la ganancia genética. Junto con nuevos diseños experimentales, la predicción genómica maximiza la calidad y la cantidad de información por dólar invertido. Este es el caso de "genetic-assisted sparse multi location testing", un diseño experimental que permite a los mejoradores organizar ensayos en múltiples ambientes con solo una fracción de los genotipos totales presente en cada uno de los ambientes. En este diseño, los genotipos observados en cada ambiente se utilizan para producir predicciones genómicas de los no plantados, y los genotipos que solapan entre ambiente permiten la conectividad genotipo x ambiente (GxE) y su modelización. Esto da como resultado un aumento en la precisión y/o intensidad de la selección y, por lo tanto, una mayor ganancia genética. Esta estrategia ha sido adoptada por el Programa Global de Mejoramiento de Cebada de ICARDA como el enfoque estándar para las pruebas preliminares de rendimiento. Para evaluar las precisiones de GP dentro de cada ambiente y entre ambientes, se testaron 1,000 nuevas entradas de "stage 1" de 2 y 6 hileras del programa "Feed Barley for Arid and Semi-Arid Environments Mega-Product Line" en un ensayo preliminar de rendimiento con cuatro ambientes. Estas líneas se distribuyeron en 340 ensayos "augmented p-rep" que incluyeron 212 entradas no replicadas en cada ambiente, 18 entradas replicadas dentro del ambiente (entradas p-rep), 80 entradas replicadas en todos los ambientes y 6 checks comerciales replicados tanto en cada dentro ambiente y entre ellos. Estos ensayos se plantaron en cuatro ubicaciones en repartidas entre Marruecos y el Líbano. La correlación fenotípica entre ambientes fue inferior a r = 0,30. Con el fin de maximizar la conectividad de parentesco, las líneas se distribuyeron entre ambientes según la matriz de relaciones basada en el pedigrí. Además, todas las 1000 líneas fueron genotipadas usando un conjunto de marcadores de 96 SNP distribuidos a lo largo del genoma y previamente seleccionados por su alta "minor allele frequency" (MAF > 0.40) entre los padres. Las predicciones genómicas se calcularon usando GBLUP ajustado usando ASReml-R, siguiendo un análisis de dos etapas. Se estableció un esquema de 10-fold cross-validation para calcular las precisiones de GP en los ensayos multi-ambiente (MET). Se agregaron diferentes variables a los modelos genómicos para mejorar la precisión de las predicciones. Estos incluían la estructura de la población, la información genealógica (matriz híbrida) y el tipo de carrera. Estos se agregaron de forma independiente o se combinaron en los modelos. Los mejores resultados se obtuvieron cuando solo se incluyó la matriz híbrida en los modelos. Las precisiones finales de predicción genómica GxE oscilaron entre 0,11 y 0,40.

Résumé

Ces dernières années, les éleveurs ont intégré les prédictions génomiques (GP) dans leurs programmes de sélection comme un outil utile, rentable et polyvalent pour augmenter le gain génétique. Couplée à de nouvelles conceptions expérimentales, la prédiction génomique maximise la qualité et la quantité d'informations par dollar investi. C'est le cas des tests multi-localisés assistés par la génomique, une conception expérimentale qui permet aux sélectionneurs d'organiser des essais multi-localisés n'ayant qu'une fraction des génotypes présents dans tous les environnements. Dans cette conception, les génotypes observés dans chaque environnement sont utilisés pour produire des prédictions génomiques de ceux qui ne sont pas plantés et les génotypes qui se chevauchent entre les emplacements permettent la connectivité génotype x environnement (GxE) et sa modélisation. Il en résulte une augmentation de la précision et/ou de l'intensité de la sélection et donc un gain génétique plus élevé. Cette stratégie a été adoptée par le Programme mondial de sélection de l'orge de l'ICARDA comme approche standard pour les essais préliminaires de rendement. Pour évaluer les précisions GP au sein et entre les emplacements, 1 000 nouvelles entrées de stade 1 à 2 et 6 rangs de la gamme de méga-produits d'orge fourragère pour les environnements arides et semi-arides ont été assemblées dans une série d'essais de rendement préliminaire avec quatre emplacements. Ces lignées ont été distribuées dans 340 essais de p-rep clairsemés sur des parcelles qui comprenaient 212 entrées non répliquées par emplacement, 18 entrées répliquées à l'intérieur de l'emplacement (entrées de p-rep), 80 entrées répliquées à travers les emplacements et 6 contrôles commerciaux répliqués à la fois à l'intérieur et à travers les emplacements. Ces essais ont été plantés dans quatre endroits différents au Maroc et au Liban, chacun identifié comme représentatif d'un environnement de population cible. Les corrélations phénotypiques entre les environnements étaient toutes inférieures à r = 0,30. Afin de maximiser la connectivité de parenté, les lignées ont été réparties entre les emplacements en fonction de la matrice des relations basée sur l'arbre généalogique. De plus, les 1 000 lignées ont été génotypées à l'aide d'un ensemble de marqueurs de diversité de 96 SNP répartis dans le génome et préalablement sélectionnés pour leur fréquence élevée d'allèles mineurs (MAF > 0,40) parmi les parents. Les prédictions génomiques ont été calculées à l'aide de GBLUP ajusté à l'aide de l'ASReml-R, après une analyse en deux étapes. Un schéma de validation croisée de 10 fois a été mis en place pour calculer les précisions GP pour les essais multi-environnements (MET). Différentes variables ont été ajoutées aux modèles génomiques pour améliorer la précision des prédictions. Celles-ci comprenaient la structure de la population, les informations sur l'arbre généalogique (matrice hybride) et le type de ligne. Ceux-ci ont été ajoutés indépendamment ou combinés dans les modèles. Les meilleurs résultats ont été obtenus lorsque seule la matrice hybride était incluse dans les modèles. La précision finale des prédictions génomiques GxE variait entre 0,11 et 0,40.

Index

Introduction

Genomic Prediction (GP) appeared first in animal and then in plant breeding programs as a useful tool to obtain phenotypic performance predictions for a given trait based on the genotypic information of the particular individuals (Meuwissen et al., 2001; Lorenzana and Bernardo, 2009; Crossa et al., 2011). Genomic Prediction has been traditionally used to obtain genomic estimated breeding values (GEBV) of no-tested lines increasing selection intensity, and therefore genetic gain (Jannink et al., 2010; Heslot et al., 2015; Crossa et al., 2017). The other most regular practice is to shorten the breeding cycle time by skipping the phenotypic evaluation (Jannink et al., 2010; Heslot et al., 2015; Crossa et al., 2017). Nowadays, new applications have come up for this technology, such as the evaluation of parental crosses or the evaluation of germplasm (Lado et al., 2017; Kehel et al., 2020; Xu et al., 2020).

In these last years, a new approach for the multi-environment evaluations (MET) in which GP plays a key role has appeared. This is the genetic-assisted multi-location sparse testing, tested firstly by Jarquin et al. (2020), taking the ideas described by Burgueño et al. (2012) about the GP strategies. This MET strategy allows the evaluation of just a fraction of the total lines in each environment, and uses these lines are to obtain the GEBV of the non-planted ones in each environment. The design is based in a small fraction of lines being overlapping between environment (OL), and the rest equally spread across the MET locations. Therefore, this last larger group of lines is evaluated just in one location (NOL) (Jarquin et al., 2020; Atanda et al., 2021, 2022). In this case of OL and NOL entries, the overlapping ones connect the environments allowing the genotype by environment modeling (GE). Therefore, in the real scenario, the GP are supported by the lines evaluated at each location and the GE modelling is made based on the OL across locations. This type of evaluation allows breeders to evaluate a higher number of lines at fixed cost, increasing selection intensity. Also, this approach can be scalable to other environment just by planting the OL and adding a new set of NOL. These characteristic makes this strategy a real alternative to the traditional MET, but for the moment it has been proposed for preliminary yield trials in cereals. In this stage of breeding programs (stage 1), the material has not been characterized or at least not deeply for yield or other traits. That is why this strategy has been suggested for this first screening (Jarquin et al., 2020; Atanda et al., 2021, 2022). The inclusion of this strategy at stage 1 enables a MET evaluation where traditionally a single evaluation in one location was carried out. In a Mediterranean context, where the correlation between environments can be low depending on the locations and the year-by-year correlation is unpredictable, added to the climate change effects already present in the region (Giorgi and Lionello, 2008; Mariotti et al., 2015; Abd-Elmabod et al., 2020), the MET sparse testing allows breeders to evaluate the GE patterns of their genotypes at earlier generation, resulting in a higher genetic gain because of the increase in selection accuracy and the assessment of trait stability. Other factor that is obvious but it can be overlooked is the limitation in the seed amount at stage 1. Depending on the previous strategy followed for the plant material advancement we could have more or less seeds available. In the best scenario, in a traditional mass selection scheme the individuals are evaluated only in one-location plot. Even though the seed amount that we can obtain is higher compared to other strategies, for instance compared to a single seed descent scheme (strategy followed for the ICARDA Global Barley Breeding Program), the seeds

obtained could not reach the required number for a MET. The strategy adopted in the stage 1 will determine the exact number of seed needed. And here is where the sparse testing appears as an efficient MET in terms of seed availability. This MET strategy demands just one-location evaluation for the majority of the lines, the seeds needed are considerably lower in number compared to a common MET.

Genomic Selection (GS) has been implemented in breeding programs, especially for complex trait, as the more efficient way of individual evaluation; showing greater results than marker-assisted selection (MAS) (Heslot et al., 2015; Crossa et al., 2017). As we already mentioned GP enables an earlier selection of superior genotypes and can help shortening the breeding cycle. In addition to that, MAS takes into account the QTLs with higher effect, whereas in GS all marker effects are considered, resulting in a more accurate evaluation of the trait when this is polygenic. The main fact that got GS into breeding programs was the availability of higher number of markers thanks to the new genotyping technologies. At the same time this could be the fact limiting the introduction of GS in regular plant breeding scheme because of the cost of this genotyping. Even though the cost is decreasing year by year, low budget programs cannot afford the genotyping of their population with the technologies commonly used, genotyping-by-sequencing (GBS), SNP arrays or using Diversity Array technology (DArTseq) (Heslot et al., 2015; Crossa et al., 2017). Another genotyping technology that has appeared in regular genomic analyses is the Kompetitive Allele Specific PCR (KASP). These KASP markers are relatively low cost and have been used for QTL mapping, MAS, and procedures where number of SNP required is low, up to several hundreds of data point per individual (Semagn et al., 2013). In part this novel SNP technology has been implemented in regular genetic analyses due to its lower cost. Therefore, this technology could be applied in a GP scheme with a low density of marker for the genotyping.

Regarding the phenotypic evaluation, the analysis of the trial data and the construction of accurate models are crucial for GP. For the analysis of trial data, a proper management of field conditions, and also of the trial data are critical. Spatial variation analysis (Oakey et al., 2016) coupled with GP model, offers a more accurate result specially for unreplicated trials (Gilmour et al., 1997). Regarding the construction of GP models, simulated and real data studies have been carried out since GP got into plant breeding in order to optimize the prediction accuracy (PA) (Crossa et al., 2014; Guo et al., 2014; Jarquín et al., 2014; Isidro et al., 2015; Bustos-Korts et al., 2016; Norman et al., 2018; Edwards et al., 2019). In these works, the authors have tried to assess several questions around GP. The more recurrent ones are the number of markers that should be used, the addition of population structure to the model, the incorporation of the pedigree information into the model, the type of statistical model that would provide the best results or the ideal population size. Some of these questions have been assessed or at least now there is consensus about them, others are still without a clear answer. And so often, the customization of the model is dependent on the target population. What it seems clear for any scenario is that the modeling of the GE and the MET evaluation provides better results than single environment testing (Crossa et al., 2011; Smith et al., 2015b; Oakey et al., 2016). Evaluations on MET with unreplicated trials have become more common due to the increase in the number of lines in breeding programs, and also the extended trait evaluations that makes not possible to measure

all the data for replicated entries. In this last remark, traits measured in the laboratory take most of the workload (Smith et al., 2015a; Terraillon et al., 2022). Traditionally replicated trials are always used at stage 2 because of their accurate measures. And for stage 1, only one-location trial is commonly carried out. In terms of cost and workload per environment it is often not possible to perform replicated MET at stage 1 for the majority of the breeding programs. Even partially replicated MET trials may not be an option. Here the sparse testing overcomes the cost issue and reduces the workload per environment.

Nowadays GS is understood as a multidisciplinary approach in which different disciplines play a key role, such as computer science, machine learning, mathematics, physics, statistics, genetics and quantitative genetics, bioinformatics, biostatistics… (Crossa et al., 2017). All these disciplines have to be taken into account for the development and implementation of a GS breeding scheme, and this process has to adapted and specific to the peculiarities of each breeding program. In this work, we present the genetic-assisted sparse testing strategy adopted by the ICARDA Global Barley Breeding Program for the preliminary yield trials. This testing is carried out in 4 location representing different types of target population of environments with 1000 F_6 unique barley genotypes evaluated targeting Feed Barley for Arid and Semi-Arid Areas. The aim of this work was to identify a low-cost and low-density maker set representative of the genetic diversity that could be used for GP in a large set of genotypes, evaluate the accuracy of the GP of our population using the low-density KASP set (>100 SNP). Optimize the GP models by incorporating population structure and row-type information as covariables and also the pedigree information to construct a hybrid matrix by the combination of the 96 KASP marker kinship and the pedigree matrix.

Material and Methods

Plant material

The plant material was composed of 1000 unique F_6 genotypes (Stage 1) and 6 commercial checks. The unique genotypes came from 534 crosses derived from the combination of 96 parents with different origins: ICARDA Global Breeding Program, landraces, lines from the pre-breeding program and cultivars from USA, Canada and Europe. These lines target the Feed Barley for Arid and Semi-Arid Environments Mega-Product Profile, being the 74% of the lines 6-row and 26% 2-row.

Genotypic data

The population was genotyped with 96 KASP markers distributed along the chromosomes, Figure 1. These markers were selected from 50K SNP chip used previously to genotype the parental material of the set (Crossing Block, CB). The markers were carefully selected for being unlinked and situated beyond the expected recombination distance per chromosome region in the CB. Also, their minor allele frequency (MAF) had to be higher than 0.4 in the CB. Finally, some QTLs identified in the CB were included within the marker set of 96 markers.

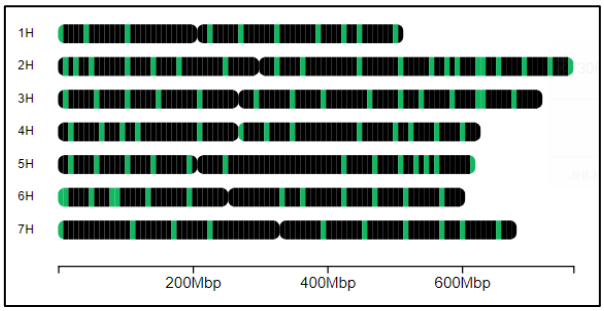


Figure 1. Marker distribution along the chromosomes of the subset of 96 markers selected from the 50K chip SNP (R package, ChromoMap (Anand and Rodriguez Lopez, 2022)).

Field evaluation

The 1006 genotypes were evaluated in 4 locations, 3 of them in Morocco and 1 in Lebanon. These locations were selected as representation of 4 different target population of environments: Sub-optimal rainfed environment (Marchouche, Morocco), North African Highlands (Annoceur, Morocco), Arid environment (Sidi el Aidi, Morocco) and West Asian Highlands (Kafardan, Lebanon)

In each of the 4 locations, 316 genotypes were tested in a spatial p-rep design (Williams et al., 2011) with 340 plots, 1360 plots in total. The allocation across environments of lines overlapping between environments and lines planted only in each of the 4 locations was 86/240. This proportion was selected based on the previous work carried out by Jarquin et al. (2020).

The allocation of the genotypes was done based on the pedigree matrix to ensure an equally genetic distribution of the parental lines. The genotype entries were classified in 4 categories according to the type of replication within and/or across locations. These were: repeated across locations (25% of the lines), repeated within location (6% of the lines), repeated across and within location (checks, 2% of the lines) and unreplicated entries (67% of the lines), Figure 2. At the end, 73% of the lines were planted in any of the 4 environments (unreplicated and repeated within location entries) and the 27% left was overlapping between environments (checks and repeated across location entries).

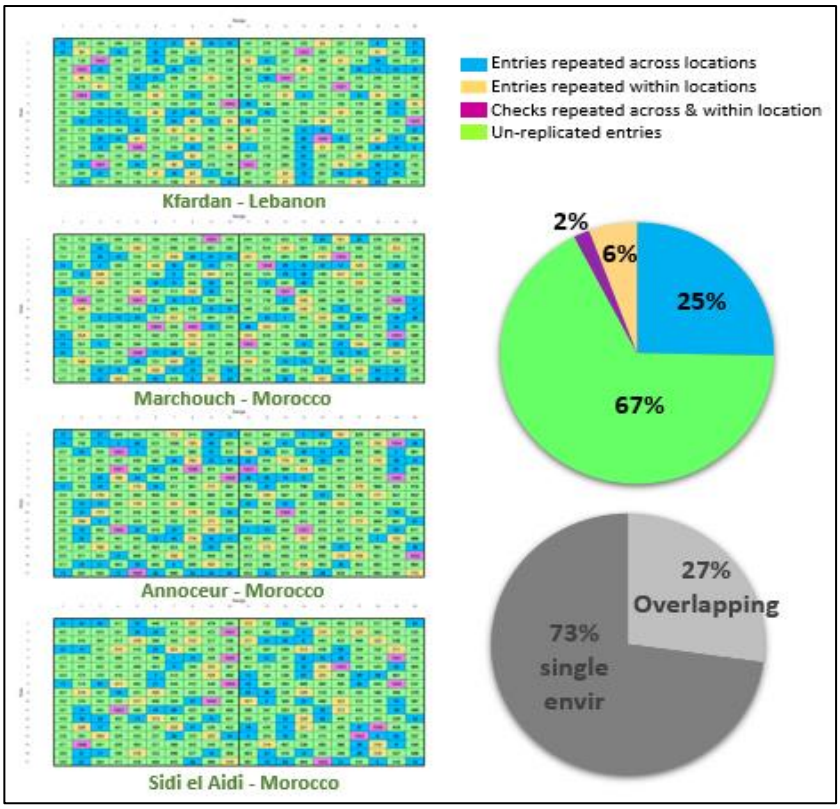


Figure 2. Experimental design and genotype allocation of the stage 1. On the right part of the image there is a representation of experimental design layout for the 4 environments. The lines are arranged in augmented p-rep designs. They appear colored in 4 different colors. The categories for the entries are describe in the upper right corner (in blue entries repeated across locations, in yellow entries repeated within locations, in purple entries repeated across and within locations and in green the unreplicated entries) with their percentages represented in the pile graph. The bottom pile graph represents the percentage of lines that are overlapping and the percentage of the ones that are not.

Genomic predictions

To obtain and validate our GP we set up genomic prediction pipeline specific for spatial p-rep designs in two steps (Piepho et al., 2012; Damesa et al., 2019), the first step consisted in obtaining the BLUE values (and their weights) for all the lines calculated within location and without using MET models. For this part of the analysis, we fitted a list of linear mixed models with the package ASReml-R (Butler et al., 2018). This R script contains a list of 24 models accounting for the different covariables present in the experimental designs. These 24 models are repeated twice, in one case the models are accounting for the spatial analysis (as first-order autoregressive spatial structure errors across columns and rows, AR1 ⊗ AR1) and in the other case do not account for it. At the end, we have a list of 48 models, in which genotypes are set as random effect and the checks as fixed effect. The output of this scrip is a table with relevant genetic information from each of the models (heritability, genetic variance, variance components…), and also, several statistical parameters of the models (BIC, AIC, REML Likelihood ratio tests) to facilitate the selection of the best model. In this case, the decision to select one model or another from the list was made taking into account genetic and statistical parameters, plus information coming from residuals plots or specific environment features. Once the best model for each environment was selected, the genotypes were set as fixed effect and the BLUE values were obtained together with their weights.

For the second step a GBLUP (Habier et al., 2007) model was fitted, also in ASReml-R with the BLUEs and the weights from the first step. As we already did a deep analysis in each of the location, the resulting model in this step was a simplified GBLUP model where only the genetic additive effect of each line is included in the model as random effect. The kinship matrix was calculated by the Yang method (Yang et al., 2010). The variance-covariance matrix for the additive effect of the lines in the different environment was calculated by Factor Analytic model (FA). This approach has been described and evaluated by several authors (Piepho, 1999; Smith et al., 2015a; b; Smith and Cullis, 2018), and it allows a parsimonious evaluation of GE effects. The genetic correlation between environment was also calculated through FA.

Different covariables were added to the model with the goal of evaluating their effects in the GP accuracies. These covariables were the row-type and the population structure (R package for population structure: Adegenet (Jombart, 2008)), incorporating them as fixed effects into the models. And finally, the pedigree information was used to build a hybrid matrix together with the kinship matrix. The construction of this hybrid matrix is made by giving to the pedigree matrix a weight (p) between 0 and 1. Then the pedigree matrix is multiplied by this p-weight and summed to the kinship matrix, which has been multiplied by 1 minus the p-weight. As regular procedure we could evaluate all the possible pedigree weights between 0 and 1 (by increments of 0.01, for instance), but in this work we just presented the values ranging between 0.2 and 0.8 by increments of 0.1.

A 10-folds cross-validation scheme with 25 iterations per analysis was set up to evaluate the GP accuracies. The prediction accuracies was calculated as the pearson‘s correlation between the GP values and the BLUEs. The cross-validation was already evaluated in the MET, where each environment contributed with 340 data point (316 unique genotypes). Due to the random assignment of the genotypes to different folds in the cross-validation scheme, we found three different type of GP validation. The CV1 as described by Burgueño et al. (2012), in which the lines predicted have never been evaluated. In our case all the unreplicated lines get into this category (67% of the lines). When one of the overlapping lines was predicted (25% of the lines), the category corresponded to CV2 (Burgueño et al., 2012), where the line had been evaluated at least in one different environment. And finally, when one of the lines from the other two categories (checks and repeated within location, 8% of the lines) was predicted, the GP was borne by the other replicate in the same environment (in the check case, data points from other environments were also involve in the GP). Even though the majority of the GP in the cross-validation were in the CV1, in the real scenario all the GP would be within the CV2 category, since our goal is to obtain a GEBV for the unreplicated ones in the environments where they were not planted.

The GP values for each line that were obtained in the bare model’s cross-validation were stored and used to make a ranking comparison between them and the ranking done with BLUEs. One ranking comparison was obtained after each of the 25 random iterations, the results presented in this work are the mean value of the 25 iterations. The ranking values were calculated as the number of genotypes overlapping between the two ranking, the one coming from the BLUEs, and the one coming after each iteration. Just the top 30 lines were compared.

Results

Genotyping

Before genotyping our population, we carried out an analysis to see the representativity of the 96 markers selected from the CB. For that, we looked at the correlation between the kinships coming from the two different marker sets, the 50K SNP chip (50K) and the 96 markers selected from the 50K (Subset) in the CB, Figure 3. This correlation was done with off-diagonal values of the two kinship matrices, and the correlation obtained was of 0.78.

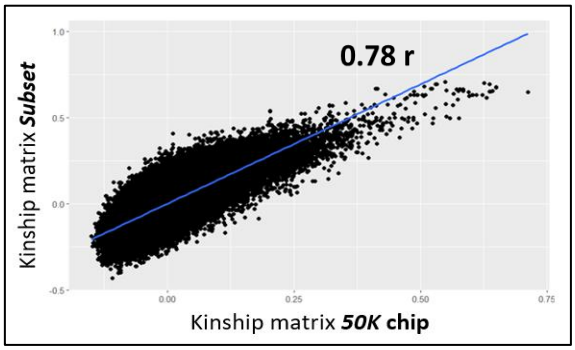


Figure 3. Correlation plot of the two kinship matrices, the one coming from the 50K and the one coming from the subset of 96 markers (genotypes from the CB). The correlation value was of 0.78, only off-diagonal values.

Once our population was genotyped, we checked whether the marker characteristics seen in the CB were still present in our population genotyped with the KASP markers. And as we can see in the Figure 4, the markers were generally unliked and their MAF was generally around 0.4.

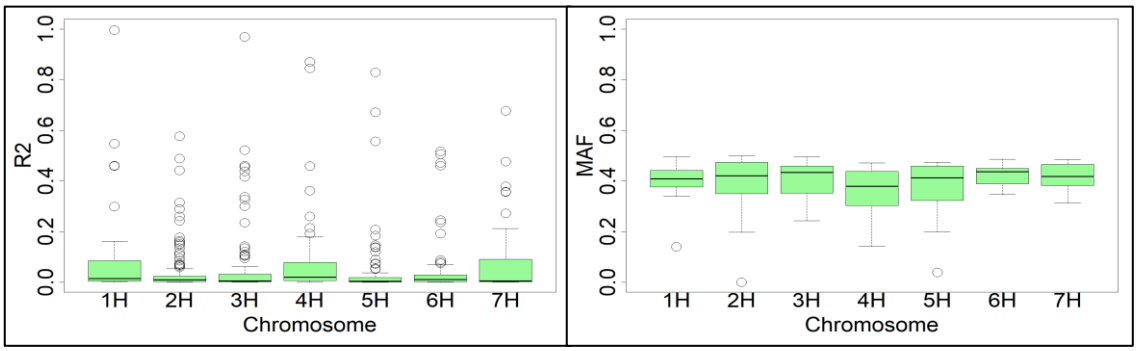
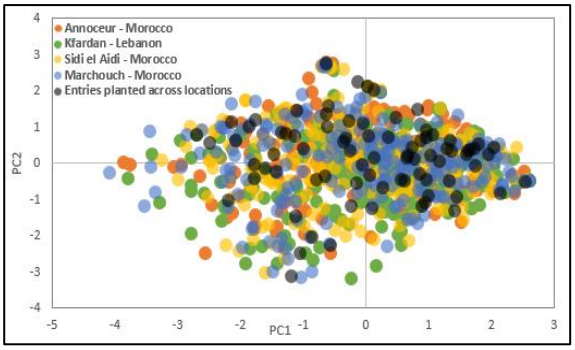


Figure 4. On the left part, boxplot representing the marker disequilibrium in R2 terms per chromosome. On the right part, the boxplot indicating the markers' MAF per chromosome. Both graphs are referring to the KASP genotyping (96 markers) of the target population, not to the CB markers.

Once the KASP genotyping was available for our population we checked if the genotype allocation to the 4 locations that we performed in the experimental design based on the pedigree matrix was corroborated by the marker information. To test that we did a principal component analysis (PCA) with the 96 KASP markers. The entries in the PCA were colored according to their locations, also the overlapping ones were colored in a black, Figure 5. This analysis confirmed that the pedigree allocation ensured the representativity of the genetic diversity in all locations

Figure 5. Principal component analysis of the stage 1 with the 96 KASP markers (PC1=3.21%, PC2=5.50%). The entries are colored according to the different locations. They are named with the name of the location and country. Loc 1 (Marchouch) is colored in bleu, Loc 2 (Annoceur) in orange, Loc 3 (Kfardan) in green and Loc 4 (Sidi el Aidi) in yellow. The entries that are overlapping between environments are colored in black.



Genetic correlation and heritability

The genetic correlations for grain yield between environments ranged from low (and even slightly negative) to moderate values, Figure 6. Higher values were obtained in the pairs Loc 2-Loc 3 and Loc 1-Loc 4. The grain yield heritabilities for the 4 environments were 0.48, 0.49, 0.60 and 0.53 for Loc 1, Loc 2, Loc 3 and Loc 4 respectively.

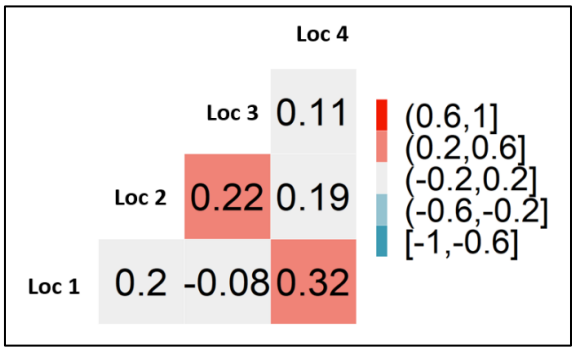


Figure 6. Genetic correlation plot between locations. Higher values are colored in red, lower in blue and middle values in gray

Bare Model

With the bare model the mean PA across the 4 locations was 0.26, and the values ranged between 0.11 and 0.37. Considerable differences were observed between environment (Figure 7), Loc 1 (0.37) and 3 (0.35) got higher results than Loc 2 (0.11) and 4 (0.19).

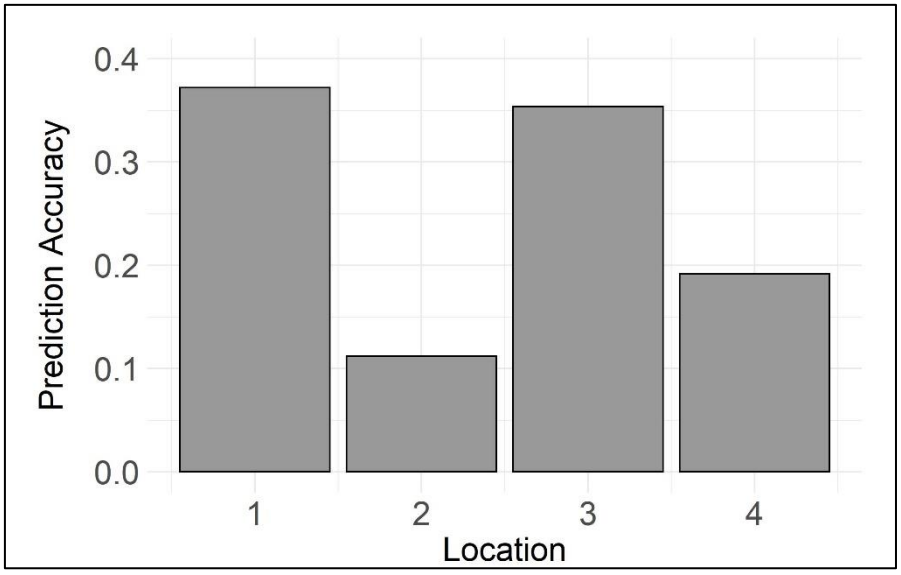


Figure 7. Barplot representing the mean PA for the 4 locations after the 25 iterations. These results correspond to the bare model.

Use of covariables

The addition of row-type or population structure to the model didn't improve the PA compared to the bare model, Figure 8. After adding row-type, the results obtain were practically the same or even slightly worse in one location (Loc 2). The same trend was observed once population structure was added, no clear visible improvement was obtained. The addition of population structure with k-cluster higher than 3 dropped the PA values considerably in Loc 3, and also in Loc 4. Loc 2 was not considerably modified with the incorporation of this covariable, and in Loc 1 some of the k-cluster dropped the PA values.

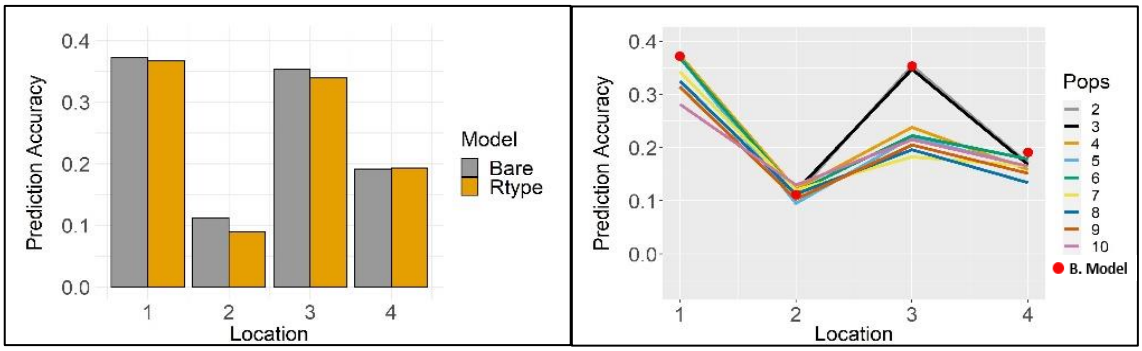


Figure 8. On left the side, barplot comparing bare model's PA (Bare in the plot's legend) with ones coming from the model incorporating row-type (Rtype in the plot's legend). On the right part, PA comparison between the bare model (red dots in the plot) and the model incorporating the different k-cluster (each k-cluster is represented with a different colored line)

Hybrid Matrix

The incorporation of the pedigree matrix into the kinship matrix improved the PA values in all locations except in Loc 3, Figure 9. For the other 3 locations, the higher was weight assigned to the pedigree matrix the better was the improvement in the PA. As Loc 3 was demanding a pedigree negatively correlated to the other 3 locations, it was not possible to obtain any pedigree weight value that maximized the PA results in all locations at once. Here the best result should be the one that maximizes the mean PA value for the 4 locations, other approaches could be applied depending on breeders' criteria about the relevance or interest of each location. The pedigree weight that maximized the PA mean value was 0.6, with a mean value of 0.28. The result in each of the 4 locations were 0.40, 0.18, 0.32 and 0.23 for Loc 1, Loc 2, Loc 3 and Loc 4 respectively.

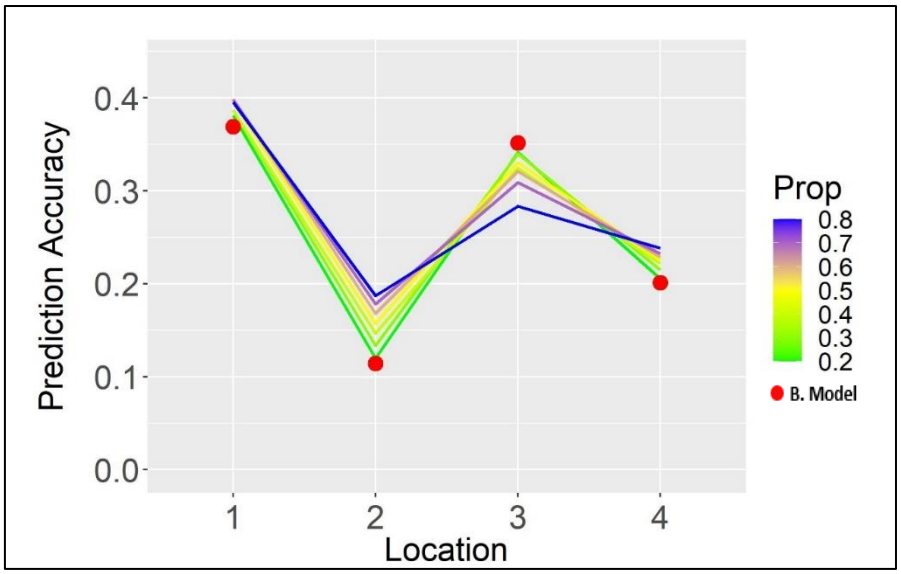


Figure 9. Graph showing the results of the hybrid matrix with different weights for the pedigree matrix (different gradient line). Also, the results obtained in the bare model are represented with red dots

Rank selection

The bare model's GP were able to identify between the 50% and 36% of the top 30 yielding lines across location. In this case, Loc 1 got again the best value with a

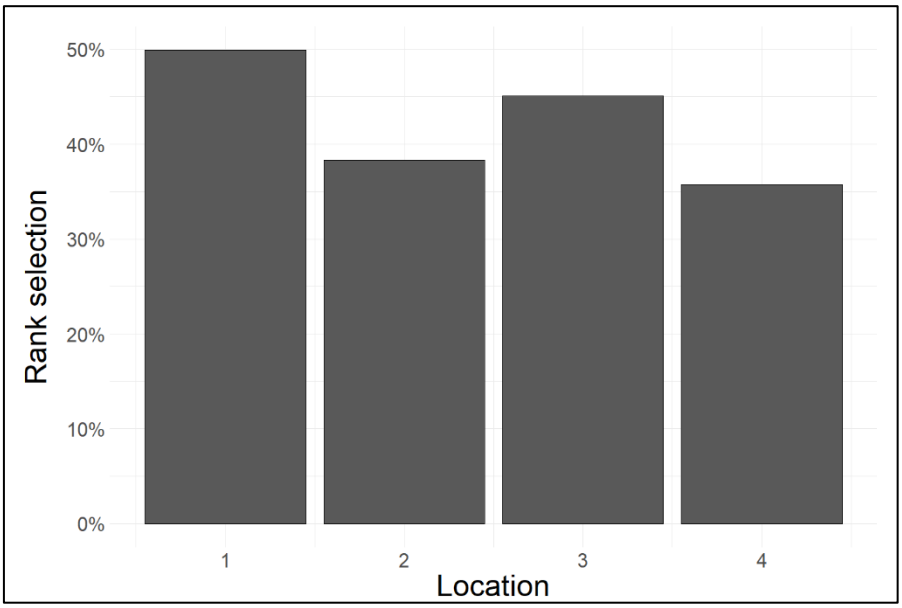


Figure 10. Barplot indicating the percentage of lines overlapping between the two top 30, the one coming from the BLUEs and the one from the cross-validation. The value represented is the average value for the 25 iterations in each location

percentage of genotypes overlapping between the two top 30 of 50%. The lowest value was obtained in Loc 4. Loc 2 and 3 achieved a percentage of 0.38% and 0.45% respectively (Figure 10). On average, the top 30 yielding lines were well identified by the GP in a 42% of the cases.

Discussion

Genomic-assisted sparse multilocation testing evaluation has been already addressed by other authors (Jarquin et al., 2020; Atanda et al., 2021, 2022). However, as far as the author knows, this approach has not been incorporated in a realistic plant breeding scenario. Here, we present the results of the genetic-sparse testing as a strategy already implemented in the ICARDA Global Barley Breeding Program for preliminary yield trials.

As we had our MET placed in Mediterranean conditions, the genetic correlation presented in this work between environments could change year by year. Even though, we expect them to be coupled as they were in this case. As it has been assessed for other authors, the GP will beneficiate from the correlation between environments (Atanda et al., 2021, 2022). In these works, the genetic correlation between environments ranged from low values to high values (-0.22 to 0.67). In our study, the correlation plot showed low correlations between the 4 environments, but some moderate higher values were paired between Loc 1 and 4 (0.32), and between Loc 2 and 3 (0.22) for instance. Therefore, the GP can borrow and model the GE interaction. These results are more similar to those obtained by Jarquin et al. (2020), where their correlations ranged between 0.07 and 0.37. But in this case, they reported phenotypic correlation.

The results gotten in the cross-validation with the bare model show how we were able to obtain PA values higher than 0.3 in two locations (1 and 3), and a mean value across location of 0.26. The differences between environment could be explained by the peculiarities of the single environments where extreme conditions can take place, such as sudden frost (Loc 1) or high temperature days (Loc 4). Atanda et al. (2022) reported PA ranging between 0.09 to 0.46 across locations, with a higher proportion of 0.66 in the ratio OL/NOL, compared to the 0.37 proportion used in this work. They also tested the extreme case where no overlapping lines were present and the results were just slightly lower to those with minimum ratio OL/NOL. In both cases, the results are similar to the ones obtained in this work. Jarquin et al. (2020) reported higher PA with similar ratio of OL/NOL to the one used in this work, the mean values across location were 0.35 and 0.52 for the two population that they worked with. But in their case, the two population were derived from the crosses of one same tester with unique genotypes. Therefore, these half-siblings population were more genetically correlated, which could explain the higher PA. Different crops were used in the two sparse testing works mentioned before, that is, wheat and maize. This could also be one of the reasons that explains the differences between the works. But we observed the main difference was in the number of makers. In both cases, more than 20,000 filtered markers were used. The number of environments and their correlation and especially the number of markers used can affect GP and their PA. Previous studies have tried to assess the effect of the number of markers in barley GP accuracies (Abed et al., 2018), and they concluded that below 2000, loses in PA become apparent.

The addition of the covariables didn't improve the results. As it has been assessed, the incorporation of additional information into the model in order to identify the structure population or specific genotype categories behind our target population does not increase the prediction accuracies (Guo et al., 2014). One of the reasons that makes the incorporation of these covariable useless was the pedigree-based allocation of the

genotypes across locations. This pedigree-based allocation ensured the genetic balance between the training and the test set population. The other factor is that this information is already captured in the kinship matrix.

What did improve the PA values was the addition of the pedigree matrix into the kinship matrix (hybrid matrix). The mean value just increased from 0.26 to 0.28, but the fact that makes this improvement a promising result is that it pushed up prediction accuracies in the locations where the bare model obtained lower results, Loc 2 and 4. The increments in these locations where from 0.11 to 0.18 and from 0.19 to 0.23 for Loc 2 and 4 respectively. Loc 1 also improved its result from 0.37 to 0.40, and Loc 3 was the only one that decreased its results from 0.35 to 0.32. Loc 3 was the only location that didn't beneficiate from the addition of the pedigree matrix, whereas the other three locations were demanding high pedigree weights to obtain the best PA results.

The rank selection got higher results if we compared it to PA values. Also, the trend was the same for the 4 locations in both measurements (Loc 1>Loc 3>Loc 4>Loc 2). Moreover, the differences between environments were less significant, their values ranged between 36% and 50% of lines overlapping between tops. Taking into account that breeders are more interested in making sure that they carry the top yielding lines to the next stage, we should focus more on this kind of indicators than in the prediction accuracies themselves.

The fact that breeding programs target different environments and the improvement in selection accuracies makes MET the best approach for trait evaluation (Crossa et al., 2011; Smith et al., 2015b; Oakey et al., 2016). The genetic-assisted sparse testing enables us to have MET at stage 1, capturing already genotype-by-environment and therefore giving more accurate and earlier selection of the best performing genotypes. That is why the modelling of the GE is crucial in this design. Jarquin et al. (2020) showed how the modelling of the GE improved clearly the PA compared to the model accounting for genotype and environment main effects, as well as for model with genetic main effects. For the works of Atanda et al. (2021 and 2022) there was no discussion about the modeling of the GE, and they only referred to the factor analytic models as a proper option for the GE modelling. The GE modelling makes this strategy a better alternative to one-location testing at stage 1 traditionally performed in breeding programs.

In summary, higher accuracies can be obtained when a large number of markers is used (Jarquin et al., 2020; Atanda et al., 2021, 2022; Abed et al., 2018), however, the present case demonstrates that an extremely low number of very well selected markers and the appropriate modelling can yield breeder-relevant results, reaching PA of up to 0.40.

Conclusion

Considering the special conditions of the Mediterranean area where the lines' performance correlation year-by-year are very low, we consider these results, where the mean PA value was close to 0.3 (0.26), as a good starting point for the genetic-assisted sparse testing at preliminary yield trials.

The incorporation of covariable didn't improve the results obtained for the bare model. By contrast, the incorporation of the hybrid matrix increased slightly the mean PA value from 0.26 to 0.28. But we arrived to in increased the PA values in those locations that

ranged lower in the bare model. This fact makes the MET selection more robust, considering that the hybrid matrix can balance the result across environments. As we saw, the different locations could demand different weights for the pedigree matrix and values ranging in the different extreme cases (toward 0 or toward 1). Therefore, in this point the breeder's criteria should be applied, and the prioritization should be given to the trials of more interest. Other causalities could play a role in this point, for instance, the already commented balanced PA values. But other year-specific condition could appear, for example one severe drought in one location-year. And now, it could be a more interesting environment to select for, in the context of climate change. Therefore, the hybrid matrix improving the PA value in this location should be applied when it comes to the genotype selection.

What is really important for breeders at preliminary yield trials is to select the top yielding lines. Therefore, with the result obtained in the rank selection the population can be already pushed towards the top yielding lines in farther generation, and it allows variety selection with these GP.

With these results we show how genetic sparse testing can be already implemented in breeding programs where the correlation between the stage 1's location and the multiple stage 2's locations is low correlated. The sparse testing overcomes the problem of evaluating only in one location and the genomic prediction allows us to have the performance of all GE combinations. Farther improvement can be implemented in this pipeline. The one that will clear improve these results would the increase in the number of markers. The fact that the hybrid matrix was demanding high values for the pedigree weights in the majority of the locations could be indicating that the marker number was not enough to characterize the genetic relationship between genotypes. Also, the addition of historical data has been used in other works (Atanda et al., 2021), but still the results are very dependent on the target population (and also on the method used: number of markers, population size…). Other options could be applied such as the incorporation of high correlated traits to the GP model or the addition of environmental covariables. But the key point is that in order for these to work we need to have an accurate genetic matrix connecting the genotypes, if it is not the case the GP will be misleading. Therefore, the increment in the number of markers would be a promising next step.

References

Abed, A., P. Pérez, R. José, and C. François. 2018. When less can be better : How can we make genomic selection more cost - effective and accurate in barley ? Theor. Appl. Genet. doi: 10.1007/s00122-018-3120-8.

Atanda, S.A., V. Govindan, R. Singh, K.R. Robbins, J. Crossa, et al. 2022. Sparse testing using genomic prediction improves selection for breeding targets in elite spring wheat. Theor. Appl. Genet. (0123456789). doi: 10.1007/s00122-022-04085-0.

Atanda, S.A., M. Olsen, J. Crossa, J. Burgueño, R. Rincent, et al. 2021. Scalable Sparse Testing Genomic Selection Strategy for Early Yield Testing Stage. Front. Plant Sci. 12(June): 1–10. doi: 10.3389/fpls.2021.658978.

Jarquin, D., R. Howard, J. Crossa, Y. Beyene, M. Gowda, et al. 2020. Genomic prediction enhanced sparse testing for multi-environment trials. G3 Genes, Genomes, Genet. 10(8): 2725–2739. doi: 10.1534/g3.120.401349.

Anand, L., and C.M. Rodriguez Lopez. 2022. ChromoMap: an R package for interactive visualization of multi-omics data and annotation of chromosomes. BMC Bioinformatics 23(1): 1–9. doi: 10.1186/s12859-021-04556-z.

Atanda, S.A., V. Govindan, R. Singh, K.R. Robbins, J. Crossa, et al. 2022. Sparse testing using genomic prediction improves selection for breeding targets in elite spring wheat. Theor. Appl. Genet. (0123456789). doi: 10.1007/s00122-022-04085-0.

Atanda, S.A., M. Olsen, J. Crossa, J. Burgueño, R. Rincent, et al. 2021. Scalable Sparse Testing Genomic Selection Strategy for Early Yield Testing Stage. Front. Plant Sci. 12(June): 1–10. doi: 10.3389/fpls.2021.658978.

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci. 52(2): 707–719. doi: 10.2135/cropsci2011.06.0299.

Bustos-Korts, D., M. Malosetti, S. Chapman, B. Biddulph, and F. van Eeuwijk. 2016. Improvement of predictive ability by uniform coverage of the target genetic space. G3 Genes, Genomes, Genet. 6(11): 3733–3747. doi: 10.1534/g3.116.035410.

Butler, D.G., B.R. Cullis, A.R. Gilmour, B.J. Gogel, and R. Thompson. 2018. ASReml-R Reference Manual Version 4. ASReml-R Ref. Man.: 176. http://www.homepages.ed.ac.uk/iwhite/asreml/uop.

Crossa, J., P. Pérez-Rodríguez, J. Cuevas, O. Montesinos-López, D. Jarquín, et al. 2017. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends Plant Sci. 22(11): 961–975. doi: 10.1016/j.tplants.2017.08.011.

Crossa, J., P. Pérez, J. Hickey, J. Burgueño, L. Ornella, et al. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity (Edinb). 112(1): 48–60. doi: 10.1038/hdy.2013.16.

Crossa, J., P. Pérez, G. de los Campos, G. Mahuku, S. Dreisigacker, et al. 2011. Genomic selection and prediction in plant breeding. J. Crop Improv. 25(3): 239–261. doi:

10.1080/15427528.2011.558767.

Damesa, T.M., J. Hartung, M. Gowda, Y. Beyene, B. Das, et al. 2019. Comparison of weighted and unweighted stage-wise analysis for genome-wide association studies and genomic selection. Crop Sci. 59(6): 2572–2584. doi: 10.2135/cropsci2019.04.0209.

Edwards, S.M.K., J.B. Buntjer, R. Jackson, A.R. Bentley, J. Lage, et al. 2019. The effects of training population design on genomic prediction accuracy in wheat. Theor. Appl. Genet. 132(7): 1943–1952. doi: 10.1007/s00122-019-03327-y.

Gilmour, A.R., B.R. Cullis, and A.P. Verbyla. 1997. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. J. Agric. Biol. Environ. Stat. 2(3): 269–293. doi: 10.2307/1400446.

Giorgi, F., and P. Lionello. 2008. Climate change projections for the Mediterranean region. Glob. Planet. Change 63(2–3): 90–104. doi: 10.1016/j.gloplacha.2007.09.005.

Guo, Z., D.M. Tucker, C.J. Basten, H. Gandhi, E. Ersoz, et al. 2014. The impact of population structure on genomic prediction in stratified populations. Theor. Appl. Genet. 127(3): 749–762. doi: 10.1007/s00122-013-2255-x.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177(4): 2389–2397. doi: 10.1534/genetics.107.081190.

Heslot, N., J.L. Jannink, and M.E. Sorrells. 2015. Perspectives for Genomic Selection Applications and Research in Plants. Crop Sci. 55(1): 1–12. doi: 10.2135/cropsci2014.03.0249.

Isidro, J., J.L. Jannink, D. Akdemir, J. Poland, N. Heslot, et al. 2015. Training set optimization under population structure in genomic selection. Theor. Appl. Genet. 128(1): 145–158. doi: 10.1007/s00122-014-2418-4.

Jannink, J.L., A.J. Lorenz, and H. Iwata. 2010. Genomic selection in plant breeding: From theory to practice. Briefings Funct. Genomics Proteomics 9(2): 166–177. doi: 10.1093/bfgp/elq001.

Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, et al. 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 127(3): 595–607. doi: 10.1007/s00122-013-2243-1.

Jarquin, D., R. Howard, J. Crossa, Y. Beyene, M. Gowda, et al. 2020. Genomic prediction enhanced sparse testing for multi-environment trials. G3 Genes, Genomes, Genet. 10(8): 2725–2739. doi: 10.1534/g3.120.401349.

Jombart, T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24(11): 1403–1405. doi: 10.1093/bioinformatics/btn129.

Kehel, Z., M. Sanchez-Garcia, A. El Baouchi, H. Aberkane, A. Tsivelikas, et al. 2020. Predictive Characterization for Seed Morphometric Traits for Genebank Accessions Using Genomic Selection. Front. Ecol. Evol. 8(March): 1–11. doi:

10.3389/fevo.2020.00032.

Lado, B., S. Battenfield, C. Guzmán, M. Quincke, R.P. Singh, et al. 2017. Strategies for Selecting Crosses Using Genomic Prediction in Two Wheat Breeding Programs. Plant Genome 10(2). doi: 10.3835/plantgenome2016.12.0128.

Lorenzana, R.E., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120(1): 151–161. doi: 10.1007/s00122-009-1166-3.

Mariotti, A., Y. Pan, N. Zeng, and A. Alessandri. 2015. Long-term climate change in the Mediterranean region in the midst of decadal variability. Clim. Dyn. 44(5–6): 1437–1456. doi: 10.1007/s00382-015-2487-3.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4): 1819–1829. doi: 10.1093/genetics/157.4.1819.

Norman, A., J. Taylor, J. Edwards, and H. Kuchel. 2018. Optimising genomic selection in wheat: Effect of marker density, population size and population structure on prediction accuracy. G3 Genes, Genomes, Genet. 8(9): 2889–2899. doi: 10.1534/g3.118.200311.

Oakey, H., B. Cullis, R. Thompson, J. Comadran, C. Halpin, et al. 2016. Genomic selection in multi-environment crop trials. G3 Genes, Genomes, Genet. 6(5): 1313–1326. doi: 10.1534/g3.116.027524.

Piepho, H.P. 1999. Stability analysis using the SAS system. Agron. J. 91(1): 154–160. doi: 10.2134/agronj1999.00021962009100010024x.

Piepho, H.P., J. Möhring, T. Schulz-Streeck, and J.O. Ogutu. 2012. A stage-wise approach for the analysis of multi-environment trials. Biometrical J. 54(6): 844–860. doi: 10.1002/bimj.201100219.

Semagn, K., R. Babu, S. Hearne, and M. Olsen. 2013. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. doi: 10.1007/s11032-013-9917-x.

Smith, A.B., D.G. Butler, C.R. Cavanagh, and B.R. Cullis. 2015a. Multi-phase variety trials using both composite and individual replicate samples: A model-based design approach. J. Agric. Sci. 153(6): 1017–1029. doi: 10.1017/S0021859614000707.

Smith, A.B., and B.R. Cullis. 2018. Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. Euphytica 214(8): 1–19. doi: 10.1007/s10681-018-2220-5.

Smith, A.B., A. Ganesalingam, H. Kuchel, and B.R. Cullis. 2015b. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. Theor. Appl. Genet. 128(1): 55–72. doi: 10.1007/s00122-014-2412-x.

Terraillon, J., M. Frisch, K.C. Falke, H. Jaiser, M. Spiller, et al. 2022. Genomic Prediction Can Provide Precise Estimates of the Genotypic Value of Barley Lines Evaluated in

Unreplicated Trials. Front. Plant Sci. 13(April): 1–10. doi: 10.3389/fpls.2022.735256.

Williams, E., H.P. Piepho, and D. Whitaker. 2011. Augmented p-rep designs. Biometrical J. 53(1): 19–27. doi: 10.1002/bimj.201000102.

Xu, Y., X. Liu, J. Fu, H. Wang, J. Wang, et al. 2020. Enhancing Genetic Gain through Genomic Selection: From Livestock to Plants. Plant Commun. 1(1): 100005. doi: 10.1016/j.xplc.2019.100005.