UNIVERSITÀ degli STUDI di CATANIA

DOTTORATO DI RICERCA IN BIOTECNOLOGIE
UNIVERSITÀ DI CATANIA

Biometec

Department of Biomedical and Biotechnological Sciences

Ph.D. in Biotechnology

Curriculum in Agro-food science

XXXIII Cycle

---

DAMIANO PUGLISI

# Development and validation of barley genetic improvement methodologies based on genomic prediction to underpin varietal innovation and the agriculture of the future

*Ph.D. Thesis*

**Tutors**: *Dr. Agostino Fricano* & *Dr. Luigi Cattivelli*

**Co-tutor**: *Prof.ssa Angela Roberta Lo Piero*

**Coordinator**: *Prof. Vito De Pinto*

---

ACADEMIC YEARS 2017/2020

# Table of Contents

## Riassunto

Il miglioramento genetico delle piante coltivate, iniziato circa 10.000 anni fa, è basato su un complesso processo di selezione della variabilità genetica al fine di creare nuove varietà più produttive, resistenti a malattie, capaci di adattarsi a condizioni climatiche sfavorevoli ed in generale capaci di soddisfare i bisogni delle società umane. Sostenere l'innovazione varietale, significa supportare l'agricoltura del futuro per far fronte ai cambiamenti climatici e alla crescita esponenziale della popolazione mondiale, e rendere più sostenibile l'agricoltura.

Tra gli strumenti a disposizione del miglioratore genetico (*breeder*), la selezione genomica o predizione genomica ha recentemente cominciato a trovare ampie applicazioni nel mondo vegetale. Questa metodologia combina dati fenotipici e genotipici per creare modelli predittivi in grado di stimare gli indici genomici (*Genomic Estimated Breeding Values*), invece di identificare singoli marcatori associati ai caratteri di interesse da utilizzare nei programmi di selezione assistita da marcatori. Per un dato carattere, gli indici genomici rappresentano una stima del valore genetico di un individuo basata sui marcatori molecolari. Con la predizione genomica, famiglie di individui o collezioni di accessioni (*training population*) sono utilizzati per costruire il modello predittivo, utilizzando sia dati fenotipici sia dati genotipici. Successivamente il modello predittivo è applicato ad individui dei quali si conoscono solamente i dati genotipici (*breeding population*) per stimare gli indici genomici usando solamente i marcatori molecolari (**Capitolo 1**).

Nel presente progetto di ricerca, questa metodologia è stata applicata e valutata per la prima volta su una popolazione *Multi Parent Advanced Generation Inter-crosses* sviluppata da 8 parentali di orzo invernale ripetutamente inter-incrociati e autofecondati seguendo uno schema basato su incroci diallelici. Questa popolazione *Multi Parent Advanced Generation Inter-crosses* è stata genotipizzata mediante il "50K SNP chip" e fenotipizzata per la resa in granella e data di fioritura in ambienti temperati e semi-aridi rappresentativi dei diversi areali di coltivazione dell'orzo in prove di campo organizzate in diversi anni ed in condizioni azotate standard ed a basso input. Combinando le informazioni genotipiche e fenotipiche, in questo progetto di ricerca è stata ottimizzata e validata la dimensione ottimale della *training population* al fine di sviluppare diversi modelli di predizione genomica per singolo ambiente (SE-GP) e multi-ambiente (ME-GP) (**Capitolo 2**).

Successivamente, la stessa *training population*, è stata fenotipizzata per l'angolo delle radici seminali, il numero delle radici seminali e l'andamento dell'indice di traspirazione in funzione dell'umidità relativa. Questi caratteri sono stati scelti perché, sulla base di altri studi condotti in altri cereali (frumenti, mais e sorgo), risultano essere particolarmente rilevanti in quanto correlati all'architettura delle radici nella pianta matura e connessi alla resa ed alla tolleranza alla siccità. Dopo aver correlato l'angolo delle radici seminali, il numero delle radici seminali e l'andamento dell'indice di traspirazione in funzione dell'umidità relativa con la resa in granella ottenuta in diversi ambienti, questi dati sono stati utilizzati per creare modelli di predizione genomica implementati con il metodo GBLUP standard e con "*threshold* GBLUP" (**Capitolo 3**).

Gli obiettivi del presente progetto di ricerca sono:

1) Testare e verificare la *performance* della predizione genomica nelle linee *Multi Parent Advanced Generation Inter-crosses* di orzo per selezionare linee con una maggiore resa in granella usando modelli "*single and multi-environment*";
2) Verificare la variabilità delle linee *Multi Parent Advanced Generation Inter-crosses* per l'angolo delle radici seminali, il numero delle radici seminali e l'andamento dell'indice di traspirazione in funzione dell'umidità relativa;
3) Creare modelli di predizione genomica per l'angolo delle radici seminali, il numero delle radici seminali e l'andamento dell'indice di traspirazione.

## Abstract

The genetic improvement of crops, started circa 10.000 years ago, is based on a complex process of selection of genetic variation to create varieties that are 1) resistant to diseases, 2) capable of adapting to unfavorable climatic conditions, 3) high-yielding and 4) fit for purpose for the needs of human society. Varietal innovation is pivotal to underpin the agriculture of the future and cope with climate change and the exponential growth of the world population.

Among the tools in the breeders's toolbox, genomic selection or genomic prediction is gaining momentum and is becoming popular for genetic improvement of crops. This methodology aims to regress genome-wide single nucleotide polymorphisms or other types of DNA markers on phenotypes of individuals to simultaneously predict their effects. The population of individuals having both phenotypic and genotypic information is named training population and is used for constructing predictive models, which allow to compute "Genomic Estimated Breeding Values" in individuals for which only genotyping information is available (breeding population). Typically, the predictive models used in GP require to regress a large number of predictors (DNA markers) that greatly exceeds the number of observations or phenotypes and several parametric and non-parametric models have been proposed to deal with overfitting and the 'large *p*, small *n*' problem as in these conditions the estimation of marker effects using ordinary least squares method is not practicable (**Chapter 1**).

In the present work, genomic prediction has been implemented and investigated on a panel of Multi Parent Advanced Generation Inter-crosses population of barley. This population was created crossing eight winter genotypes following "half-diallel" crosses. The resulting panel of Multi Parent Advanced Generation Inter-crosses lines was genotyped using the barley 50K SNP chip and was phenotyped in different site-by-season and site-by-season-by-management combinations to examine grain yield and heading date. Using phenotypic and genotypic information, models for grain yield predictions have been fitted and cross-validated using single-environment- and multi-environment-genomic prediction models (**Chapter 2**).

Subsequently, the same panel of barley Multi Parent Advanced Generation Inter-crosses was phenotyped for belowground and physiological traits related to drought tolerance and grain yield. Particularly, these lines were phenotyped for seminal root number, seminal root angle and transpiration rate response to increasing evaporative demand. Standard and threshold models were subsequently fitted and cross-validated to predict these traits, which might support ideotype breeding for dry environments (**Chapter 3**).

The aims of this project are:

1) Testing and assessing the performance of genomic prediction on Multi Parent Advanced Generation Inter-crosses to select high-yielding barley lines.
2) Examining the variability of Multi Parent Advanced Generation Inter-crosses lines for seminal root number, seminal root angle and transpiration rate to increasing evaporative demand.
3) Developing, fitting and cross-validating genomic prediction models for seminal root number, seminal root angle and transpiration rate to increasing evaporative demand to underpin ideotype breeding for crop improvement.

## List of abbreviations and keywords

| | |
|---|---|
| **AFLP** | *Amplified Fragment Length Polymorphism* |
| **AIC** | *Advanced Intercrossing* |
| **BAC** | *Bacterial artificial chromosome* |
| **BGLR** | *Bayesian Generalized Linear Regression* |
| **BGTs** | *Belowground traits* |
| **BLUE** | *Best Linear Unbiased Estimator* |
| **BLUP** | *Best Linear Unbiased Prediction* |
| **BP** | *Breeding population* |
| **BRR** | *Bayesian ridge regression* |
| **BV** | *Breeding Value* |
| **CIMMYT** | *International Maize and Wheat Improvement Center* |
| **CRAN** | *Comprehensive R Archive Network* |
| **CSMs** | *Crop simulation models* |
| **CV** | *Cross validation* |
| **DH** | *Doubled Haploid* |
| **Gb** | *Gigabases* |
| **GBLUP** | *Genomic Best Linear Unbiased Prediction* |
| **GCM** | *Global Circulation Model* |
| **GEBV** | *Genomic Estimated Breeding Value* |
| **GK** | *Gaussian Kernel* |
| **GH** | *Grain hardness* |
| **GLM** | *Generalized linear model* |
| **GLMM** | *Generalized linear mixed model* |
| **GP** | *Genomic prediction* |
| **GS** | *Genomic selection* |
| **GY** | *Grain yield* |
| **ha** | *Hectare* |
| **HD** | *Days-to Heading* |
| **ICARDA** | *International Center for Agricultural Research in the Dry Areas* |
| **ISSR** | *Inter Simple Sequence Repeats* |
| **LASSO** | *Least absolute shrinkage and selection operator* |
| **LD** | *Linkage Disequilibrium* |
| **LM** | *Linear models* |
| **LMM** | *Linear mixed models* |
| **LOO** | *Leave-one-out* |
| **LTRs** | *Long terminal repeats* |
| **MAGIC** | *Multi-parent Advanced Generation Inter-Crosses* |
| **MAS** | *Marker Assisted Selection* |
| **MDe** | *METs, environment-specific variance GxE deviation model* |
| **MDs** | *METs, single variance GxE deviation model* |
| **ME-GP** | *Multi Environment Genomic Prediction* |
| **MET** | *Multi-environmental trial* |
| **MM** | *METs, main genotypic effect* |
| **NGS** | *Next Generation Sequencing* |
| **NR** | *Nodal roots* |

| | |
|---|---|
| **PCA** | *Principal Component Analysis* |
| **PH** | *Plant Height* |
| **PS** | *Phenotypic selection* |
| **QTL** | *Quantitative Trait Locus* |
| **RAPD** | *Random Amplification of Polymorphic DNA* |
| **RCBD** | *Randomized Complete Block Design* |
| **REML** | *Restricted Maximum Likelihood* |
| **RFLP** | *Restriction Fragment Length Polymorphism* |
| **RKHS** | *Reproducing Kernel Hilbert space* |
| **RRBLUP** | *Ridge Regression Best Linear Unbiased Prediction* |
| **RR-BLUP** | *Ridge Regression and Other Kernels for Genomic Selection* |
| **SE** | *Single Environment Genomic Prediction* |
| **SNP** | *Single Nucleotide Polymorphism* |
| **SR** | *Seminal root* |
| **SRA** | *Seminal root angle* |
| **SRN** | *Seminal root number* |
| **SSR** | *Simple Sequence Repeats or Microsatellites* |
| **t** | *tonnes* |
| **TASSEL** | *Trait Analysis by aSSociation, Evolution and Linkage* |
| **TGBLUP** | *Threshold Genomic Best Linear Unbiased Prediction* |
| **TP** | *Training population* |
| **TR** | *Transpiration Rate* |
| **VP** | *Validation population* |
| **VPD** | *Vapor Pressure Deficit* |

## Ph.D. affiliations

Dr. Damiano Puglisi during his Industrial Ph.D. course has successfully completed his six-month training at ICARDA and six-month training at ISEA Agroservice SpA - CREA GB. Dr. Damiano Puglisi is jointly conducting his research project in collaboration with:

1) Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A), Università di Catania, Via S. Sofia 98 (CT), Italy: under the supervision of Prof.ssa Angela Roberta Lo Piero as Ph.D. Co-Tutor.
2) Council for Agricultural Research and Economics – Research Centre for Genomics and Bioinformatics (CREA-GB), Via San Protaso 302 Fiorenzuola d'Arda (PC), Italy: under the supervision of Dr. Agostino Fricano and Dr. Luigi Cattivelli as Ph.D. Tutors.
3) Biodiversity and Crop Improvement Program, International Center for Agricultural Research in the Dry Areas (ICARDA) (https://www.icarda.org/), Avenue Hafiane Cherkaoui, Rabat, Morocco: under the supervision of Dr. Andrea Visioni as Research Ph.D. Supervisor.
4) ISEA Agroservice SpA (https://www.agroservicespa.it/), San Severino Marche (MC), Italy: under the supervision of Dr. Eugenio Tassinari as Factory Supervisor, Industrial Ph.D. partner.

# Chapter 1: General Introduction and aim of the Ph.D. research thesis

## 1 Barley (*Hordeum vulgare* ssp. *vulgare* L.)

### 1.1 Use of barley for food and feed production and its economic importance

To date, barley (*Hordeum vulgare* ssp. *vulgare* L.) ranks fourth among the most important worldwide cereal crops in terms of total grain production and acreage after wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), and maize (*Zea mays* L.). Overall, the global barley grain production (**Figure 1A**) in 2019 was around 158.9 million tonnes (Mt), among them 95.6 Mt were produced in Europe, followed by Asia (~25.5 Mt), America (~21.6 Mt), Oceania (~9.2 Mt) and Africa (~6.8 Mt). The global barley acreage (**Figure 1B**) in 2019 was around 51.1 million hectares (Mha), of which around 50% were located in Europe (~24.2 Mha), followed by Asia (~12.5 Mha), America (~5.7 Mha), Oceania (~4.4 Mha) and Africa (~4.1 Mha), while the acreage devoted to barley cultivation in the Mediterranean area (South Europe and North Africa) was around 3 Mha, 261 Kha of which were located in Italy (http://faostat.fao.org).

**Figure 1: World trend of annual barley production and cultivation areas from 1960 to 2019.** **(A)** Annual barley production in different world areas and **(B)** areas devoted to barley cultivation in Africa, Americas, Asia, Europe, Oceania and Mediterranean region (Northern Africa and Southern Europe), respectively.


Barley is one of the oldest crops cultivated by mankind as it was domesticated 10.000 years ago in the Fertile Crescent (Salamini et al. 2002). Since its domestication, farmers and breeders have succeeded to adapt and cultivate this crop in different environmental conditions and currently, barley farming is widespread from extreme Northern latitudes to desertic areas of North Africa. Unlike wheat, barley is more tolerant to drought, which is one of the most important environmental stress in agriculture, particularly in the Mediterranean region (Cattivelli et al. 2008; Giraldo et al. 2019; Ullrich 2014). Along with durum wheat (*Triticum durum* Desf.), barley is a staple food for several rural populations of North Africa (Grando and Macpherson 2005).

Barley is mainly used as animal feed and circa 75% of its global production is devoted to this purpose (Langridge 2018). Compared to other cereals (e.g. maize and sorghum), barley has excellent nutritional characteristics such as higher contents of neutral (18%) and acidic (from 6 to 7%) detergent fibers in kernels, for being used as feed as these compounds allow to maximize the animal digestion (Bleidere and Gaile 2012).

Circa 20% of its global production (Langridge 2018) is used to supply the malting and brewery industry: beer is obtained from the alcoholic fermentation of sugars deriving from starchy sources, the most used of which is malt, that is germinated cereal grains that have been appropriately dried. Barley plays a key role in the brewing industry since barley malt is an excellent substrate for yeast fermentation and beer production (Gupta, Abu-Ghannam, and Gallaghar 2010).

The remaining part of the barley global production (circa 5%) is used as staple food (Langridge 2018; Tricase et al. 2018). Barley is considered as a functional food since a regular barley consumption provides health benefits reducing the risk of chronic diseases (Donato-Capel et al. 2014; Lee 2017). These health benefits have been attributed to the high content of β-glucans (from 4 to 9%) and non-starch polysaccharides present in the endosperm, which are significantly higher than in other cereals (Gupta et al. 2010). Several studies have shown that diets based on barley grain help to prevent insulin resistance and type 2 diabetes, excessive cholesterol blood level (Ramakrishna et al. 2017), coronary heart diseases and colorectal cancer (Ahmad et al. 2012; Din et al. 2018).

In Italy, there is mounting economic interest for this crop and consequently the number of seeds companies engaged in barley breeding programs is increasing (http://www.sementi.it): ISEA Agroservice SpA (https://www.agroservicespa.it) is one of these Italian seeds companies devoted to the improvement of barley and is directly involved in this industrial Ph.D. project.

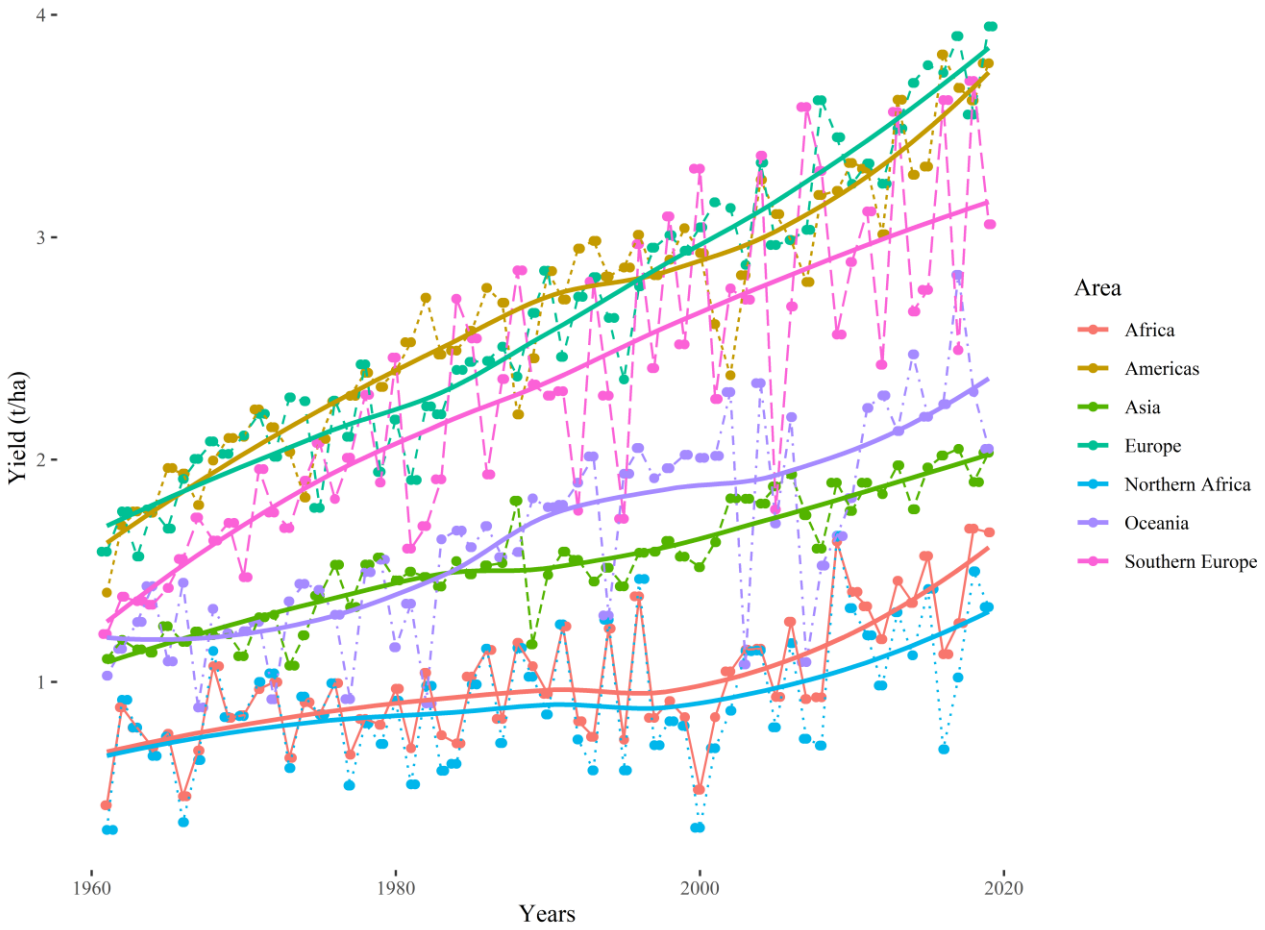
## 1.2 Trend of barley yield in Europe and North-Africa

Globally, barley yield per hectare has grown linearly from 1.3 t/ha in 1960 to 3.1 t/ha in 2019 (http://faostat.fao.org). Particularly, in 2019 Europe has recorded the highest value (~3.9 t/ha) followed by Americas (~3.7 t/ha), Oceania (~2 t/ha), Asia (~2 t/ha) and Africa (~1.6 t/ha) (**Figure 2**). Recent estimates point out that 28% of the yield increase observed in the last decades was obtained targeting barley genetics, corroborating the importance of barley breeding to increase grain yield (GY). In north-Italy, barley yield was circa 4.2 t/ha in 2019 (http://faostat.fao.org).

Although these spectacular achievements obtained at the global level, barley yield in certain marginal and stressful areas and in North Africa is still stagnating and was on average 1.3 t/ha in 2019 (**Figure 2**). Overall, in the Mediterranean region, the growth of the barley yield is not sufficient to guarantee food security for next years, especially for rural human populations of North Africa. Moreover, in North Africa and Southern Europe, barley yield follows a fluctuating trend, which might potentially threaten rural populations (**Figure 2**).

The main reasons that underlie the ample difference in barley yield between North Africa and Europe is due to environmental factors that impose abiotic stresses (e.g., drought and water availability), inefficient agronomic practices due to the lack of means and equipment to implement high-input agricultural practices, the lack of affordable crop protection chemicals to manage pests and diseases and the insufficient genetic progress due to the lack of adapted and high-yielding barley varieties (Pswarayi et al. 2008).

**Figure 2: World trend of barley yield per hectare (t/ha) from 1960 to 2019.** Different points and regression lines indicate barley yield in Africa, Americas, Asia, Europe, Oceania and Mediterranean region (Northern Africa and Southern Europe).




## 2 Targeting barley genetics to improve grain yield: a short list of popular methods

### 2.1 Empirical breeding and phenotypic selection to improve barley

The success of barley breeding programs relies on effective selection of individuals carrying favorable alleles, which has been historically addressed in crops using phenotypic selection (PS), that is the selection of the individuals exhibiting the greatest phenotypic values as parents to mate for generating following progeny. PS relies on the genetic variability of traits in natural or segregating populations, that is populations artificially created by crossing two or more different genotypes. Depending on the type of fertilization, schemes for crop improvement have been historically classified in breeding for predominantly self-fertilizing species and breeding for predominantly outcrossing species. Barley has evolved floral structures that ensure a high percentage of self-pollination and consequently, breeding concepts and schemes aimed at improving this crop are those developed and applied to predominantly self-fertilizing (autogamous) species (Barcaccia and Falcinelli 2006; Lynch and Walsh 1997).

The two key components that enable breeding for crop improvement are 1) a source of genetic variability and 2) the phenotypic evaluation of the traits of interest for selecting the best individuals that fit for purpose. The most common approach to create *ex novo* variability in autogamous species, like barley, consists in crossing two or more inbred parents to generate highly diverse and heterozygous offspring. Crossing of inbred lines allows to shuffle parental genes, which in turn generate new allele combinations and genetic variability. In barley, as well as in other predominantly

autogamous crops, continuous cycles of self-fertilizations of $F_1$ offspring create homogeneous descendant as the percentage of homozygous loci progressively increases with the number of generations (**Table 1**).

**Table 1: Expected trend of the percentage of heterozygous loci as function of self-fertilization cycles.** For each generation of self-fertilizations, the percentage of heterozygosity decreases as the percentage of homozygosity increases.

| Homozygous loci | Heterozygous loci | Number of self-fertilizations |
|---|---|---|
| 0% | 100% | 0 |
| 50% | 50% | 1 |
| 75% | 25% | 2 |
| 87.5% | 12.5% | 3 |
| 93.75% | 6.25% | 4 |
| 96.875% | 3.15% | 5 |

In $F_1$ offspring of predominantly autogamous plants, heterozygosity decreases at each generation of self-pollination, while the fraction of homozygotes loci increases over generations (**Table 1**). Considering a hybrid genotype with $n$ allelic pairs in heterozygous condition, the fraction of homozygous loci after $m$ segregating generations is equal to the following formulas (Barcaccia and Falcinelli 2006):

$$F_{homz} = \left[\frac{(2^m - 1)}{2^m}\right]^n \qquad \textbf{Equation 1}$$

$$F_{heter} = 2\,p\,q\left(\frac{2\,t}{2-s}\right) \qquad \textbf{Equation 2}$$

where $p$ and $q$ indicate the allele frequencies for A and a, $s$ and $t$ indicate the self-fertilization rate and the crossing rate. Consequently, self-pollinating species quickly bring individuals to complete homozygosity, namely obtaining a 'pure line', simply defined as set of individuals derived by self-pollination from a homozygous progenitor. **Equation 1** and **Equation 2** (Barcaccia and Falcinelli 2006) allow to compute the level of homozygosity and heterozygosity maintained in the population after $n$ generations of self-fertilizations, respectively.

Several breeding schemes have been proposed and developed for predominantly autogamous species in combination with PS to exploit the variability of both natural (e.g., mass selection and modified mass selection) or segregating populations artificially created by crossing inbred lines (e.g., pedigree method).

### 2.1.1 Cereal traits that might be targeted to cope with climate change

Climate change is the result of the greenhouse effect caused by gases developed through various activities, which include livestock farming and fossil fuel consumption. It is a global phenomenon that has repercussions in all fields of agriculture, resulting from innumerable anthropic activities. The future effects of climate change have been predicted by the Global Circulation Model (GCM), which projects a rise in the global average temperatures between 0.9 and 2.16 °C, and a variation of rainfall between -24 and +24% (Arneth et al. 2019; Cammarano et al. 2019). Furthermore, the report published in 2018 by the Intergovernmental Panel on Climate Change showed that exceeding the threshold limit of 1.5 °C foreseen in 2040, would have catastrophic effects in all aspects of human societies (Masson-Delmotte et al. 2018). Climate change is a global phenomenon that has a different impact at local level and specifically in the Mediterranean is expected to increase the frequency of drought episodes and heat waves, which in turn will be causing important yield losses in cereal crops

during the next years (Xie et al. 2018). These two phenomena have been already observed during the last years and represent the main concerns for rainfed Mediterranean agriculture.

This dangerous trend along with the increase of world population, which will reach almost 10 billion people by 2050 (http://faostat.fao.org), is increasing worldwide interest on genetic improvement of crops to cope with climate change and underpin food security (Lobell et al. 2008). For these reasons, beyond PS and classical breeding schemes, new breeding strategies and tools are needed to sustain barley yield since this crop plays a relevant role for ensuring food security, especially in North Africa. In fact, in these regions, which are characterized by harsh living conditions and dry environments, barley is one of the major staple food along with durum wheat and plays a relevant role for ensuring food security intended as 'risk of adequate food not being available' (Chakraborty and Newton 2011). In this context, direct phenotypic selection for improving GY is ineffective for two main reasons: 1) this trait shows, in general, low levels of heritability when tested in field trials organized in stressful conditions and 2) the confounding environmental effects are high, and it is difficult to obtain a strong selection response.

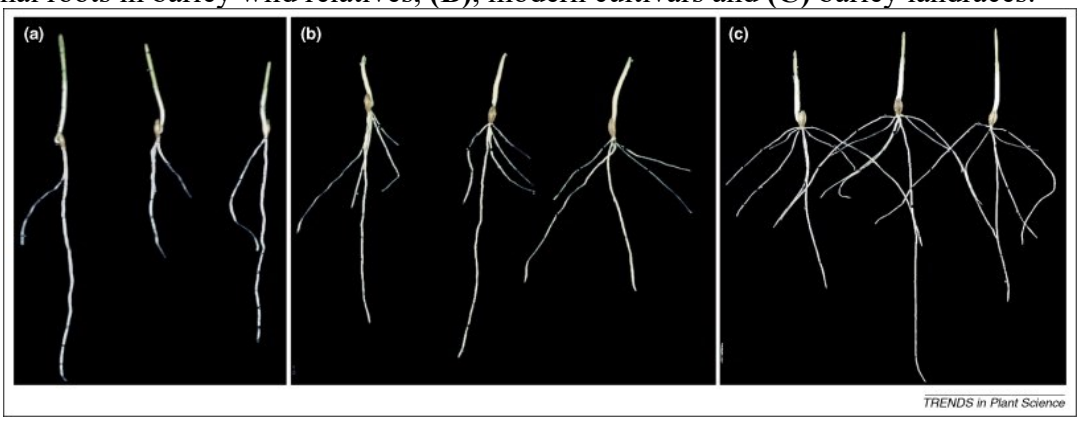
To overcome these limitations, breeding programs often adopt indirect selection to achieve the expected genetic gain of a primary trait. Unlike direct selection, indirect selection allows the selection of a primary trait by selecting one or more directly correlated secondary traits. This methodology is useful when the primary trait to be selected has a low heritability. Some barley traits that might be targeted for indirect selection of GY, coping with climate change and improving drought tolerance are belowground traits (BGTs) and plant transpiration rate (TR).

Improvement of BGTs has the potential to increase water extraction from soil layers for efficiently capturing residual moisture (El Hassouni et al. 2018). Moreover, these traits have important implications in the breeding programs for their relationship with GY (Robinson et al. 2018). In barley and in other cereals, seminal roots (SR) are the first roots differentiated from the embryonic root that allows the development of the primary root system. Differently, nodal roots (NR) begins at the tillering stage from the basal nodes of the crown (Wahbi and Gregory 1995).

The number of radical axes obtained from the primary roots is known as seminal root number (SRN) (**Figure 3**). Several studies carried out in cereal crops have pointed out that modern and high-yielding cultivars show lower values of SRN and inefficient root architectures for water-limiting environments (de Dorlodot et al. 2007), which are probably the result of selection for high input agriculture. The different root architecture and SRN of modern barley cultivars compared to barley landraces corroborate this latter hypothesis (**Figure 4**).

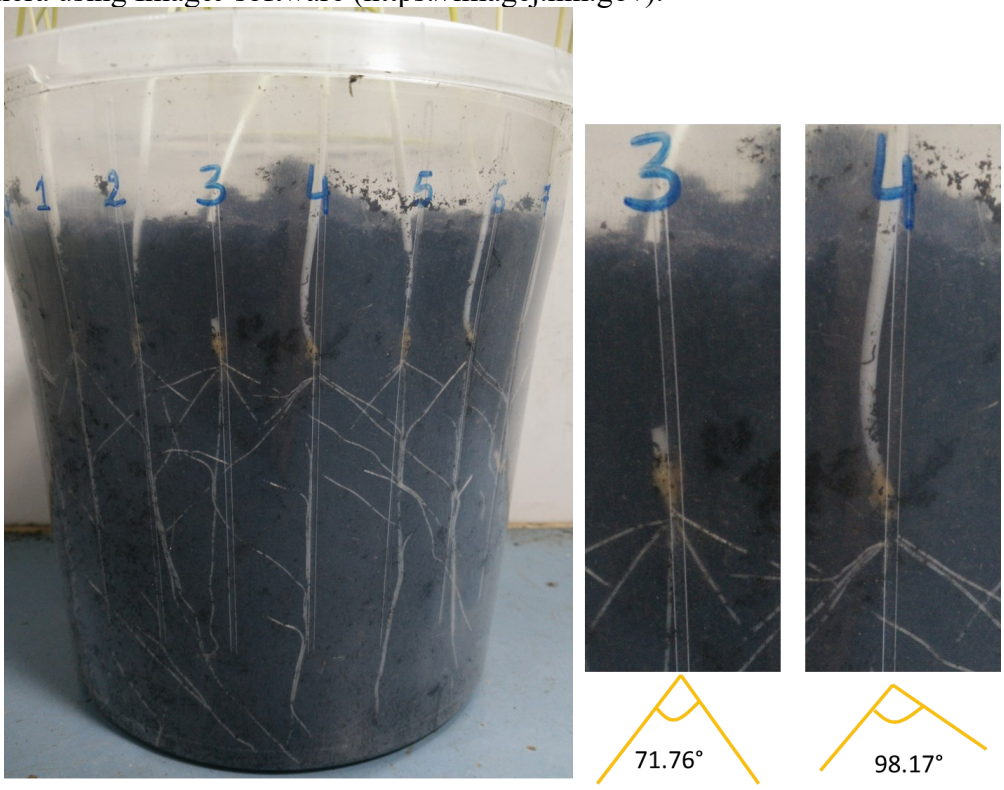
**Figure 3: Image of seminal roots collected in barley using the clear pot method analyzed in this study**. This image was taken eleven-day days after sowing in barley seedlings, removing from the substrate.

**Figure 4: Development and differentiation of seminal roots in barley** (de Dorlodot et al. 2007). **(A)** Seminal roots in barley wild relatives, **(B)**, modern cultivars and **(C)** barley landraces.




The angle obtained by measuring the axes of the SR primary roots is known as seminal root angle (SRA) (**Figure 5**). The modulation of SRA has beneficial effects in several crops: in beans wide SRA allows the development of varieties that better absorb the phosphorus found in the surface soil (Lynch and Brown 2001); in bread wheat it was shown that less productive cultivars tend to have a wide root angle while the more productive and drought tolerant genotypes have a narrow root angle (Manschadi et al. 2008). This experimental evidence indicates that SRA is, in certain crops, a proxy trait for the root architecture system at the mature stage of development and has a fundamental role in the absorption of nutrients and water from the soil and in regulating plant growth (Robinson et al. 2018). As there is a significant genetic correlation between GY and the aforementioned BGTs (El Hassouni et al. 2018), SRA and SRN might be used for indirect selection of high-yielding genotypes in dry environments.

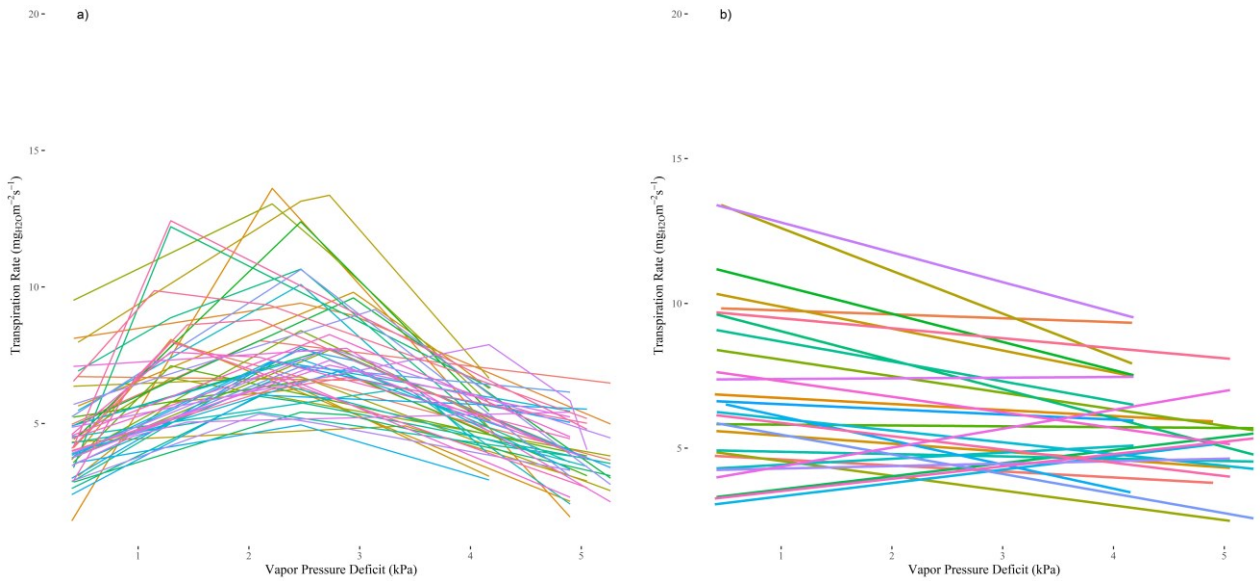
**Figure 5: Image of seminal root angle measured in barley genotypes using the clear pot method analyzed in this study**. In this image, seminal root angle was measured five days after sowing with a digital camera using ImageJ software (https://imagej.nih.gov).



Several non-invasive methods have been developed to phenotype BGTs including the clear pot methodology (Richard et al. 2015), space-planted field pasta strainer (basket) method (El Hassouni et al. 2018), 3D magnetic resonance imaging (MRI) (Van As 2007), 2D-imaging techniques exploiting X-ray absorption and light transmission (de Dorlodot et al. 2007) or, for example, in substrate-filled rhizoboxes (Jia et al. 2019). To date, BGTs (SRA and SRN) have been phenotyped in different barley populations such as advanced backcross DH population of a cross between spring and wild barley (Sayed et al. 2017), for panels of two-rowed spring barley genotypes (Robinson et al. 2018), for wild barley introgression lines (Naz et al. 2014) and for panels of mostly unrelated landraces (Hamada et al. 2012; Jia et al. 2019) but they have never been measured in MAGIC populations of barley.

Another proxy trait for drought tolerance is the transpiration rate (TR), which is evaluated against increasing values of vapor pressure deficit (VPD), that is the difference between the vapor pressure in ambient conditions and the vapor pressure at the point of saturation for the same temperature (Rashed 2016). The regulation of TR under increasing VPD conditions is a relevant factor affecting the agronomic adaptation of wheat and barley to Mediterranean environments limited by water. Medina et al. (2019), showed how the different response can be correlated with different performances in terms of GY and biomass. In barley the analysis of TR in response to increasing VPD values is lagging behind but there is evidence that in durum wheat and in other cereal crops TR varies according to two main trends: 1) genotypes that show a linear TR trend to increasing VPD values and 2) genotypes showing a segmented TR trend to increasing VPD values. In durum wheat, genotypes showing this latter trend have been correlated with increasing drought tolerance in field conditions (Medina et al. 2019). Like durum wheat, other research works have shown that barley genotypes can exhibit linear and segmented TR response to increasing VPD conditions (Sadok and Tamang 2019) and the analyses carried out in the panel of barley lines examined in this research work have shown variability for this trait (**Figure 6**).

**Figure 6: Trend of transpiration rate under increasing vapor pressure deficit in the panel of barley population analyzed in this study. (A)** lines exhibiting a segmented transpiration rate to increasing values of vapor pressure deficit; **(B)** lines showing a linear transpiration rate in response to increasing vapor pressure deficit values.



## 2.2 Genomic tools for enabling molecular breeding in barley

### 2.2.1 Barley reference sequence

The knowledge of barley genomics has the potential to accelerate barley breeding as it allows to investigate the genetic bases of agronomically important traits to enable marker assisted selection (MAS) and to better implement genomic prediction (GP) in breeding programs. Barley is a diploid plant (2n) and has a genome size of 5.3 gigabases (Gb) distributed across seven chromosomes named from 1H to 7H (Monat et al. 2019). The barley genome structure is characterized by approximately 80.2% of repetitive elements, of which 74% of long terminal repeats (LTRs) retrotransposon and 5.39% of DNA Transposon Superfamily (Monat et al. 2019; Wicker et al. 2017). Barley genes are unevenly distributed across the seven chromosomes as the telomeric ends have a higher gene content than the centromeric regions.

To date, several reference sequences of barley are currently available in GenBank (http://plants.ensembl.org). The first reference version of barley cv. Morex was constructed using bacterial artificial chromosome (BAC), and BAC-by-BAC sequencing (Mascher et al. 2017). This reference sequence was later improved using the TRITEX pipeline (Monat et al. 2019), which uses a completely different approach compared to BAC-by-BAC sequencing. Beyond these genomic resources based on barley cv. Morex, Schreiber *et al.* (2020) have recently assembled the reference sequence of barley cv. Golden Promise, a Scotland spring two-row malting and whisky barley, integrating data from several technologies. The cv. Golden Promise is the most efficient genotype for genetic transformations (e.g. Cisgenesis, Genome editing), and for this reason it is considered as "Transformation Reference" in barley (Schreiber et al. 2020). Recently, other additional 19 reference sequences of barley obtained from a set of genetically distant genotypes were recently assembled to investigate the barley pan-genome (Jayakodi et al. 2020). In this research work, the SNPs used to develop GP models refer to the sequence of barley cv. Morex published in 2017 (Mascher et al. 2017).

### 2.2.2 Tools for rapid SNP fingerprinting

Single Nucleotide Polymorphisms (SNPs) are the most popular markers used in crop improvement programs as they have several advantages compared to other type of DNA markers like Restriction Fragment Length Polymorphism (RFLP), Simple Sequence Repeats or Microsatellites (SSR), Random Amplification of Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Inter Simple Sequence Repeats (ISSR) (Khlestkina and Salina 2006). Unlike RAPD, ISSR and AFLP, SNPs mark a single position on the genome and in addition they are codominant markers, that is they allow to distinguish heterozygous and homozygous genotypes. SNPs are prone to be analyzed using high-throughput technologies like chip arrays, which uses sets of probes that have the capability to interrogate specific genome positions. The 50k Illumina Infinium iSelect SNP chip array is a specific barley chip that interrogates 44.040 known SNP positions on the reference genome of barley and compared to other SNP typing technologies, this chip provides faster and more robust results (Bayer et al. 2017).

## 3 Regression models

A statistical model is a mathematical representation with which the data obtained from a given experiment are analyzed. In general, a model has the objective of evaluating the relationship between a response variable and a set of predictor variables as follows:

**VARIABLE = PREDICTORS + ERROR**

In this research work, four types of statistical models were extensively used to analyze experimental data (e.g., field trials, experiments carried out in controlled conditions) and for implementing genomic prediction. These four types of statistical models are known as linear models (LM), linear mixed models (LMM), generalized linear model (GLM) and generalized linear mixed model (GLMM).

### 3.1 Linear model and linear mixed model

In this research work, linear models were used to regress a set of $n$ independent variables ($X_1$, $X_2$, $X_3$, ... $X_n$) with dependent variable ($Y$) for estimating the effects obtained from the fixed values. In LM $y$ is continuous variable while $X_1$, $X_2$, $X_3$, ... $X_n$ can be either continuous or discrete variables. A typical equation describing a LMs is:

$$\begin{aligned} \boldsymbol{y_i} &= \boldsymbol{\mu} && \} \textit{ mean} \\ &+ \boldsymbol{a_1 X_{1i} + a_2 X_{2i} + a_3 X_{3i} + \ldots a_n X_{ni}} && \} \textit{ coefficients + fixed variables} \\ &+ \boldsymbol{e_i} && \} \textit{ error terms} \end{aligned} \qquad \textbf{Equation 3}$$

**Equation 3** is a system of several equations that can be re-written as:

$$\begin{aligned} y_1 &= \mu + a_1 X_{11} + a_2 X_{21} + a_3 X_{31} + \ldots a_n X_{n1} + e_1 \\ y_2 &= \mu + a_1 X_{12} + a_2 X_{22} + a_3 X_{32} + \ldots a_n X_{n2} + e_2 \\ y_3 &= \mu + a_1 X_{13} + a_2 X_{23} + a_3 X_{33} + \ldots a_n X_{n3} + e_3 \\ &\ldots \\ y_i &= \mu + a_1 X_{1i} + a_2 X_{2i} + a_3 X_{3i} + \ldots a_n X_{ni} + e_i \end{aligned}$$

This system of equations can be re-written using a *matrix notation* that is $\boldsymbol{y = aX + e}$:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_i \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu + \begin{bmatrix} X_{11} & X_{21} & X_{31} & \cdots & X_{n1} \\ X_{12} & X_{22} & X_{32} & \cdots & X_{n2} \\ X_{13} & X_{23} & X_{33} & \cdots & X_{n3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ X_{1i} & X_{2i} & X_{3i} & \cdots & X_{ni} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_i \end{bmatrix} \qquad \textbf{Equation 4}$$

where $y_i$ is the response value, $a_n$ are the set of coefficients associated to the independent variables ($X_1$, $X_2$, $X_3$, ... $X_n$), also known as fixed effects and $e_i$ is the error associated to each response. In general, the main aim of LM is to estimate the fixed parameters $a_1$, $a_2$, $a_3$, ... $a_n$ using Best Linear Unbiased Estimators (BLUEs) (Legendre 1805). The main assumption of LM are that: 1) errors are independent and normally distributed *~N(0, $\sigma_e^2$)* with mean of zero and constant variance $\sigma_e^2$ (homoscedasticity); 2) experimental data are obtained from a random extract from a population of X, $y$ units for which the linear relationship applies; 3) for each X value we have a $y$ values normally distributed *~N(0, $\sigma_a^2$)* with the mean located on the regression line.

BLUEs of model parameters are usually obtained using the least squares method, first introduced by Legendre (1805)**,** which allows to choose the best regression line that minimizes the sum of the differences as much as possible. Therefore, the formula for computing $a_n$ is:

$$\boldsymbol{a_n = \frac{Cov\,(X_{ni}, y_i)}{Var(X_{ni})}} \qquad \textbf{Equation 5}$$

where $Cov\,(X_{ni}, y_i)$ is the covariance between dependent and independent variables and $Var(X_{ni})$ is the variance of the dependent variable.

Linear mixed models were used to regress: 1) a set of $n$ independent variables ($X_1$, $X_2$, $X_3$, ... $X_n$) with dependent variable ($Y$) for estimating the effects obtained from the fixed values; 2) a set of $n$

independent variables ($Z_1$, $Z_2$, $Z_3$, ... $Z_n$) with dependent variable ($Y$) for estimating the effects obtained from the random values. In LMM $y$ is continuous variable while $X_1$, $X_2$, $X_3$, ... $X_n$ and $Z_1$, $Z_2$, $Z_3$, ... $Z_n$ can be either continuous or discrete variables.
A typical equation describing a LMMs is:

$$\begin{aligned} \boldsymbol{y_{ij}} &= \boldsymbol{\mu} && \} \textit{ mean} \\ &+\boldsymbol{a_1X_{1i} + a_2X_{2i} + a_3X_{3i} + \ldots a_nX_{ni}} && \} \textit{ coefficients + fixed variables} \\ &+ \boldsymbol{b_1Z_{1j} + b_2Z_{2j} + b_3Z_{3j} + \ldots b_nZ_{nj}} && \} \textit{ coefficients + random variables} \\ &+ \boldsymbol{e_{ij}} && \} \textit{ error terms} \end{aligned} \qquad \textbf{Equation 6}$$

Therefore, the **Equation 6** is a system of several equations that can be re-written as:

$$\begin{aligned} y_{11} &= \mu + a_1X_{11} + a_2X_{21} + a_3X_{31} + \ldots a_nX_{n1} + b_1Z_{11} + b_2Z_{21} + b_3Z_{31} + \ldots b_nZ_{n1} + e_{11} \\ y_{22} &= \mu + a_1X_{12} + a_2X_{22} + a_3X_{32} + \ldots a_nX_{n2} + b_1Z_{12} + b_2Z_{22} + b_3Z_{32} + \ldots b_nZ_{n2} + e_{22} \\ y_{33} &= \mu + a_1X_{13} + a_2X_{23} + a_3X_{33} + \ldots a_nX_{n3} + b_1Z_{13} + b_2Z_{23} + b_3Z_{33} + \ldots b_nZ_{n3} + e_{33} \\ &\ldots \\ y_{ij} &= \mu + a_1X_{1i} + a_2X_{2i} + a_3X_{3i} + \ldots a_nX_{ni} + b_1Z_{1j} + b_2Z_{2j} + b_3Z_{3j} + \ldots b_nZ_{nj} + e_{ij} \end{aligned}$$

This system of equations can be re-written using a *matrix notation* that is $\boldsymbol{y = aX + bZ + e}$

$$\begin{bmatrix} y_{11} \\ y_{22} \\ y_{33} \\ \vdots \\ y_{ij} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mu + \begin{bmatrix} X_{11} & X_{21} & X_{31} & \cdots & X_{n1} \\ X_{12} & X_{22} & X_{32} & \cdots & X_{n2} \\ X_{13} & X_{23} & X_{33} & \cdots & X_{n3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ X_{1i} & X_{2i} & X_{3i} & \cdots & X_{ni} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} Z_{11} & Z_{21} & Z_{31} & \cdots & Z_{n1} \\ Z_{12} & Z_{22} & Z_{32} & \cdots & Z_{n2} \\ Z_{13} & Z_{23} & Z_{33} & \cdots & Z_{n3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ Z_{1j} & Z_{2j} & Z_{3j} & \cdots & Z_{nj} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ \vdots \\ e_{ij} \end{bmatrix}$$

**Equation 7**

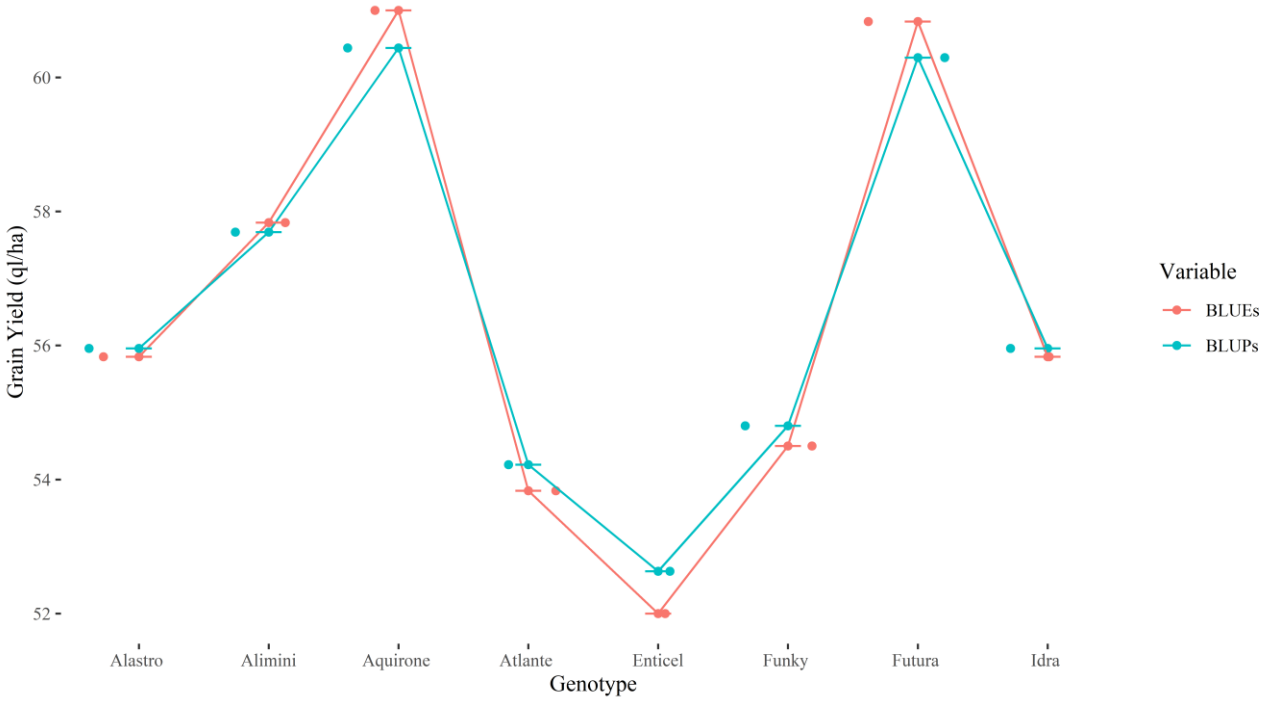
where $y_{ij}$ is the response value, $a_n$ are the set of coefficients associated to the independent variables ($X_1$, $X_2$, $X_3$, ... $X_n$), also known as fixed effects, $b_n$ are the set of coefficients associated to the independent variables ($Z_1$, $Z_2$, $Z_3$, ... $Z_n$) also known as random effects and $e_{ij}$ is the error associated to each response. In general, the main aim of LMM is to estimate the fixed parameters $a_1$, $a_2$, $a_3$, ... $a_n$ using BLUE and random parameters $b_1$, $b_2$, $b_3$, ... $b_n$ using Best Linear Unbiased Prediction (BLUP). Briefly, least squares (Legendre 1805) and shrinkage or "linear unbiased prediction" (Robinson 1991) are used to estimate fixed and random effects, respectively.

In order to highlight this difference, a practical example is presented using the top eight Italian barley varieties (Alastro, Alimini, Aquirone, Atlante, Enticel, Funky, Futura, Idra) evaluated in 6 replicas for GY (ql/ha) in a single environment. The adjusted means, BLUP and BLUE (**Figure 7**), were estimated considering the genotype as a fixed or random effect, respectively, following the formula shown below:

$$\boldsymbol{y_{ij}} = \mu + \boldsymbol{Rep_i + Gen_j + \varepsilon_{ij}} \qquad \textbf{Equation 8}$$

where $y_{ij}$ is the response value GY, μ is the mean, $Rep_i$ are the 6 replicates, $Gen_j$ are the top eight best Italian barley varieties and $\varepsilon_{ij}$ is the errors independent and normally distributed $\sim N(0, \sigma^2_{\varepsilon})$. As shown in **Figure 7**, the random effect causes a shift towards the mean of the genotypes, highlighting its difference with the fixed effect. Indeed, while a fixed effect assumes genotypes as independent and separate, a random effect assumes that they are interrelated and interdependent representative of a large set of levels, which may not even be observed.

**Figure 7: BLUEs and BLUPs estimation of GY (ql/ha) in 6 replicas** of the top eight best Italian barley varieties through LMs and LMMs. Different colored points and its line of barley production Alastro, Alimini, Aquirone, Atlante, Enticel, Funky, Futura and Idra, respectively.




### 3.2 Generalized linear model and generalized linear mixed model

Generalized linear models extend the linear model to variables that are not normally distributed and are characterized by three components: 1) response variable $Y$, 2) a link function connecting the response variable to the set of linear predictors and 3) a set of linear predictors $\eta_i$ having the following form:

$$\boldsymbol{\eta_i = a_1 X_{1i} + a_2 X_{2i} + a_3 X_{3i} + \ldots a_n X_{ni}} \qquad \textbf{Equation 9}$$

**Equation 9** is a system of several equations that can be re-written as:

$$\eta_1 = a_1 X_{11} + a_2 X_{21} + a_3 X_{31} + \ldots a_n X_{n1}$$
$$\eta_2 = a_1 X_{12} + a_2 X_{22} + a_3 X_{32} + \ldots a_n X_{n2}$$
$$\eta_3 = a_1 X_{13} + a_2 X_{23} + a_3 X_{33} + \ldots a_n X_{n3}$$
$$\ldots$$
$$\eta_i = a_1 X_{1i} + a_2 X_{2i} + a_3 X_{3i} + \ldots a_n X_{ni}$$

This system of equations can be re-written using a *matrix notation* that is $\boldsymbol{\eta = aX}$**:**

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \vdots \\ \eta_i \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & X_{31} & \cdots & X_{n1} \\ X_{12} & X_{22} & X_{32} & \cdots & X_{n2} \\ X_{13} & X_{23} & X_{33} & \cdots & X_{n3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ X_{1i} & X_{2i} & X_{3i} & \cdots & X_{ni} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} \qquad \textbf{Equation 10}$$

where $\eta_i$ is the linear predictor, $a_n$ are the set of coefficients associated to the independent variables ($X_1$, $X_2$, $X_3$, ... $X_n$). The link function $g()$ connects the response variable to the set of linear predictors $\eta_i$ as follows:

$$g(\mu_i) = \eta_i = a_n X_{ni}$$ **Equation 11**

**Equation 11** can be re-written as an inverse link function as follows:

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(a_n X_{ni})$$ **Equation 12**

For each probability distribution, there are different link function such as identity, log, inverse, inverse-square, logit and probit. For instance, the identity function is associated with a normal distribution and the expected value of the dependent variable is just the linear predictor.

The main aim of GLM is to estimate the fixed parameters $a_1$, $a_2$, $a_3$, ... $a_n$. Finally, like LMMs, generalized linear mixed models are an extension of GLMs that includes random effects. Therefore, the main aim of GLMM is to estimate the fixed parameters $a_1$, $a_2$, $a_3$, ... $a_n$ using BLUE and random parameters $b_1$, $b_2$, $b_3$, ... $b_n$ using BLUP.

## 4 Field trial analysis

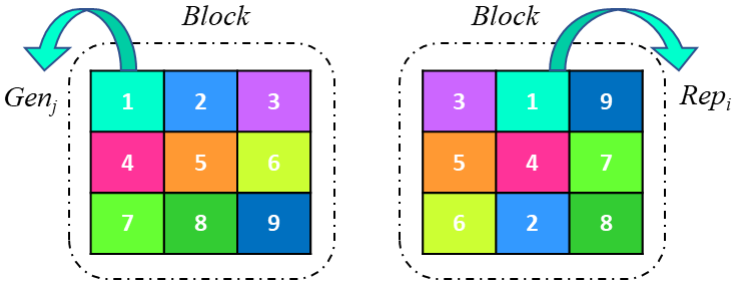
Field variability originates from soil heterogeneity and non-uniformity and is one of the most important confounding factors that must be considered in trial analyses carried out during breeding programs as it can increase the variances of errors and masks the genetic variation that underlies the traits of interest (Araus and Cairns 2014). There are plenty of factors that cause soil heterogeneity and non-uniformity and generally the larger is the field trial, the greater is the internal variability: it is generally observed that in field trials the same genotype grown in adjacent plots tends to have similar yield or trait values than the same genotype grown in distant plots.

There are different experimental designs that allow to control field variability and correct the phenotypic values measured in field trials considering these environmental factors. The most popular experimental designs adopted for analyzing field trials are the Randomized Complete-Block Designs (RCBD) and the alpha-lattice. These experimental designs are conveniently analyzed using LMs or LMMs and depending on whether the genotypes are considered as random or fixed variables, allow to compute the adjusted phenotypic values using BLUPs or BLUEs, respectively. Therefore, the adjusted means are pivotal to analyze field trial data.

### 4.1 Randomized complete block design

RCBD is a widely used experimental design used for field trial analyses, particularly when the number of treatments (e.g., the number of lines to test) is limited. The experimental unit of the RCBD is the block and RCBD assumes that these blocks are homogeneous, that is there is no variability within each block and that each block contains all treatments (e.g., all the lines that must be examined) (**Figure 8**).

**Figure 8: Graphical representation of a RCBD.** This experimental design consists of two blocks or replicates and nine treatments (e.g., different genotypes). Each block contains all treatments and RCBD assumes that these blocks are homogeneous.




Breeders uses RCBD to estimate the effect of each treatment (e.g., plant genotype) and assume that the raw phenotypic values are the sum of the block effects, the treatment effect plus a random error.

Using LMs or LMMs, a RCBD can be formalized as follows:

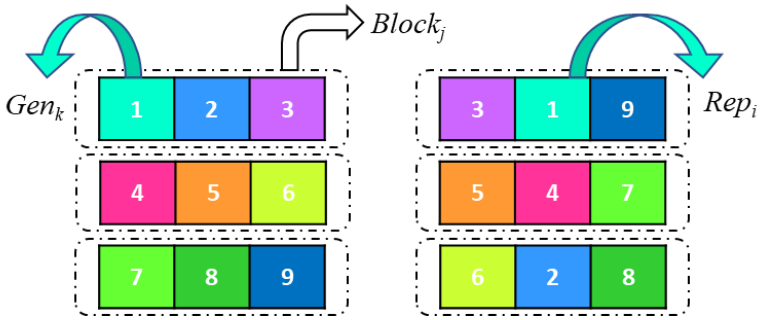
$$\boldsymbol{y_{ij}} = \mu + \boldsymbol{Rep_i} + \boldsymbol{Gen_j} + \boldsymbol{Cov} + \boldsymbol{\varepsilon_{ij}} \qquad \textbf{Equation 13}$$

where $y_{ij}$ is the raw phenotypic value of the trait of interest measured in the replicate $i^{th}$ for the genotype $j^{th}$, μ is the overall mean, $Rep$ is the effect of replicate $i^{th}$, $Gen$ is the effect of the genotype $j^{th}$, $Cov$ might indicate another covariate that can exert its effect on trait $y$ while $\varepsilon$ is the error associated with the phenotypic value measured in the replicate $i^{th}$ for genotype $j^{th}$. Moreover, this model assumes $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$, that is, the error terms $\varepsilon$ is independent and normally distributed with mean 0 and variance $\sigma_\varepsilon^2$. With RCBD, breeders are interested to estimate the effect of treatments, in our case the effect of genotype $Gen$. Depending on the purpose, genotypes might be treated as random or fixed effects. Regardless whether genotypes are considered as random or fixed effects, their effects computed with **Equation 13** are usually indicated as adjusted means.

### 4.2 Alpha-lattice

Like RCBD, the alpha-lattice design is widely used in field trial analyses when the number of treatments (e.g., the number of lines to test) is large. The experimental units of the alpha-lattice design are the blocks, which are assumed to be homogeneous, that is there is no variability within each block and contain only a part of all treatments (e.g., all the lines that must be examined) (**Figure 9**).

**Figure 9: Graphical representation of an alpha-lattice design.** In this design there are six blocks (three blocks per replicate) and nine treatments (e.g., different genotypes). Each block contains only a subset (3) of all treatments.




Breeders use alpha-lattice to estimate the effects of each treatment (e.g., plant genotype) and assume that the phenotypic values are the sum of the block effects on each replication, the replicates, the treatment effect plus a random error. Using LMs or LMMs, an alpha-lattice design can be formalized as follows:

$$\boldsymbol{y_{ijk}} = \mu + \boldsymbol{Rep_i} + \boldsymbol{Block_j(Rep_i)} + \boldsymbol{Gen_k} + \boldsymbol{Cov} + \boldsymbol{\varepsilon_{ijk}} \qquad \textbf{Equation 14}$$

where $y_{ijk}$ is the phenotypic value of the trait of interest measured in the replicate $i^{th}$ of the block $j^{th}$ for the genotype $k^{th}$, μ is the overall mean, $Rep$ is the effect of replicate $i^{th}$, $Block$ is the effect of block $j^{th}$ on replication $i^{th}$, $Gen$ is the effect of the genotype $k^{th}$, $Cov$ might indicate another covariate that can exert its effect on trait $y$ and $\varepsilon$ is the error associated with the phenotypic value measured in the replicate $i^{th}$ of the block $j^{th}$ for genotype $k^{th}$. Moreover, this model assumes $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$, that is the error terms $\varepsilon$ are independent and normally distributed with mean 0 and variance $\sigma_\varepsilon^2$.

Using the alpha-lattice design, breeders are interested to estimate the effect of treatments, in our case the effect of genotype $Gen$. Regardless whether genotypes considered as random or fixed effects, the effects of genotypes computed with **Equation 14** are usually indicated as adjusted means.

### 4.3 Multi environmental Trials

The estimation of genotypic effects from **Equation 13** (or **Equation 14**) does not allow to further dissect them in additive and interaction effects. Consequently, the estimation of the genotypic effects conducted in a single environment might not be reliable owing to the confounding effects of genotype x environment interaction. Multi-environmental trials (METs) are field trials organized in different site-by-season combinations using pre-defined experimental designs (e.g., RCBD or alpha-lattice) and allow to better estimate the genotypic values of the lines tested. This can be demonstrated comparing the models of single and multi-environmental trials organized using RCBD. A model describing a multi-environmental trial organized using RCBD can be described as follows:

$$\boldsymbol{y_{ijk} = \mu + Env_i + Rep_j(Env_i) + Gen_k + Gen_k \text{ x } Env_i + Cov + \varepsilon_{ijk}} \qquad \textbf{Equation 15}$$

where $y_{ijk}$ is the phenotypic value of the trait of interest measured in the replicate $j^{th}$ for the genotype $k^{th}$ in environment $i^{th}$, μ is the overall mean, $Env$ is the effect of the environment $i^{th}$, $Rep$ is the effect of replicate $j^{th}$ on each environment $i^{th}$, $Gen$ is the effect of the genotype $k^{th}$, $Gen\ x\ Env$ is the interaction effect between genotype $k^{th}$ and environment $i^{th}$, $Cov$ is the effect of covariate and $\varepsilon$ is the error associated with the phenotypic value measured in the replicate $j^{th}$ for genotype $k^{th}$ in environment $i^{th}$. Moreover, this model assumes $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$, that is the error terms $\varepsilon$ are independent and normally distributed with mean 0 and variance $\sigma_\varepsilon^2$.

The comparison of **Equation 13** with **Equation 15** shows that the effects of genotypes $Gen$ in single-environmental trial is further dissected in $Gen_k$ and $Gen_k$ x $Env_i$ terms in **Equation 15**. In other words, METs allow to estimate genotype x environment interaction effects and provide better estimate of the additive values. Similarly, METs organized with alpha-lattice designs provide better estimates of genetic effects. A model describing a MET organized using alpha-lattice can be described as follows:

$$\boldsymbol{y_{ijkl} = \mu + Env_i + Rep_j(Env_i) + Block_k(Env_i\, Rep_j) + Gen_l + Gen_l\, x\, Env_i + Cov + \varepsilon_{ijkl}} \qquad \textbf{Equation 16}$$

where $y_{ijkl}$ is the phenotypic value of the trait of interest measured in the replicate $j^{th}$ of the block $k^{th}$ for the genotype $l^{th}$ and in environment $i^{th}$, μ is the overall mean, $Env$ is the effect of the environment $i^{th}$, $Rep$ is the effect of replicate $j^{th}$ on environment $i^{th}$, $Block$ is the effect of block on replication $j^{th}$ and environment $i^{th}$, $Gen$ is the effect of the genotype $l^{th}$, $Gen\ x\ Env$ is the interaction effect between genotype $l^{th}$ and environment $i^{th}$, $Cov$ is the effect of covariate and $\varepsilon$ is the error associated with the phenotypic value measured in the replicate $j^{th}$ of the block $k^{th}$ for genotype $l^{th}$ in environment $i^{th}$. Moreover, this model assumes $\varepsilon_{ijkl} \sim N(0, \sigma_\varepsilon^2)$, that is, the error terms $\varepsilon$ are independent and normally distributed with mean 0 and variance $\sigma_\varepsilon^2$.

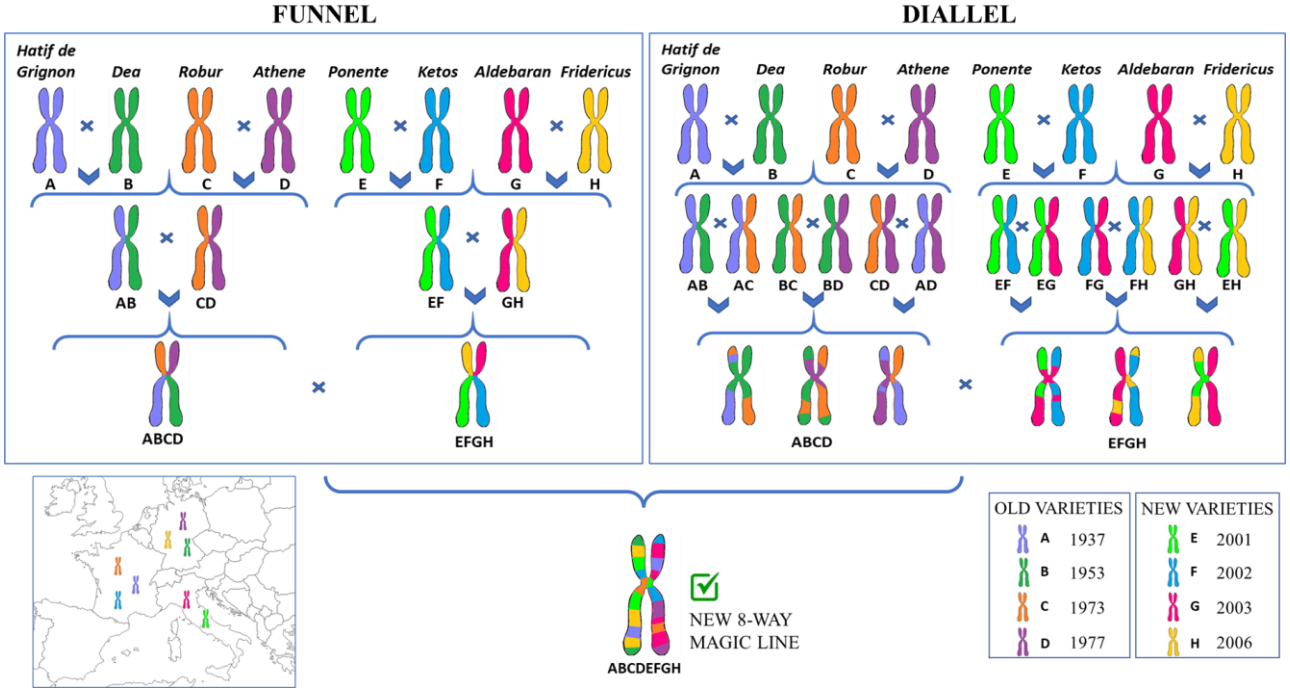
Like RCBD, the comparison of **Equation 14** with **Equation 16** shows that the effects of genotypes $Gen$ computed in single-environmental trial can be further dissected in $Gen_k$ and $Gen_k$ x $Env_i$ using **Equation 16** and then METs allow to estimate genotype x environment interaction effects and provide better estimate of the genetic values.

## 5 Multi-parent Advanced Generation Inter-Crosses population

A Multi-parent Advanced Generation Inter-Crosses (MAGIC) population is a collection of related plants obtained inter-crossing a set of founder lines for several generations in each single line (Huang et al. 2015). MAGIC population are an extension of the advanced inter-crossing (AIC) lines created for the first time in mice (Yalcin, Flint, and Mott 2005) and later in plants.

In cereal crops, MAGIC populations have been developed and established for rice (Chen et al. 2013; Han et al. 2020; Ogawa et al. 2018), maize (Dell'Acqua et al. 2015; Jiménez-Galindo et al. 2019), bread wheat (Huang et al. 2012; Mackay et al. 2014; Milner et al. 2016; Rebetzke et al. 2014; Sannemann et al. 2018; Shah et al. 2019; Stadlmeier, Hartl, and Mohler 2018; Thépot et al. 2014), sorghum (Ongom and Ejeta 2018) and barley including an eight-way population of 5000 DH (Sannemann et al. 2015) and another of 324 MAGIC lines $F_6$ 32-way (Bülow, Nachtigall, and Frese 2019). MAGIC populations were explicitly developed for genetic research purpose as they allow to increase power and precision for detecting and mapping quantitative trait loci (QTLs) (Cavanagh et al. 2008; Huang et al. 2015). In cereal crops, MAGIC populations were used to map QTLs that underlie several key traits (Arrones et al. 2020) such as site-related to agronomic adaptation (Bülow et al. 2019), flowering time (Sannemann et al. 2015), plant height (Huang et al. 2012; Milner et al. 2016; Sannemann et al. 2018), test weight (Huang et al. 2012), coleoptile length, thickness, shoot length (Rebetzke et al. 2014), grain yield (Milner et al. 2016) and powdery mildew resistance (Stadlmeier et al. 2018). Usually, the set of founder lines used to construct MAGIC populations, is composed by 8 or 16 different parents, although MAGIC population can be constructed using more founders. As shown in **Figure 10**, two are the commonly used schemes to construct MAGIC populations. In the first scheme, namely funnel scheme (Arrones et al. 2020; Huang et al. 2015), the eight parents are crossed to obtain 4 $F_1$ hybrid populations. In the next stage, individuals from pairs of these $F_1$ hybrid populations are inter-mated to generate 2 double-cross $F_2$ populations. Similarly, individuals of these two $F_2$ populations are intermated to create a population of individuals, which is self-fertilized for several generations to create inbred lines. The second approach for creating MAGIC populations is based on diallel cross (Arrones et al. 2020; Hayman 1954) where the eight parents are crossed in all possible combination both as male and female (*full-diallel*) or only as male or female (*half-diallel*).

**Figure 10: Graphical representation of schemes to construct 8-way MAGIC.** The eight parents, (Hatif de Grignon, Dea, Robur and Athene, Ponente, Ketos, Aldebaran and Fridericus) of 6-rowed barley varieties are crossed as biparental population (*funnel*) or in all possible combination (*diallel*) for several generations to obtain a mosaic genome of MAGIC.



The development of MAGIC lines is time consuming, nevertheless Doubled Haploid (DH) technology allows to shorten their creations. DHs are obtained using *in vitro* culture of anthers or pollen or ovaries or egg cells, to which phytohormones are applied for stimulating cell divisions and forming the callus, which is the undifferentiated mass of meristematic cells from which a plant develops. This technology is often used to speed up the creation of MAGIC population instead of applying self-fertilizing cycles (Sannemann et al. 2015).

## 6 Genomic selection (GS)

Genomic selection (GS), also known as genomic prediction (GP), is a promising tool that can be used for accelerating the genetic gain of barley breeding and for coping with climate change (Abberton et al. 2016). GP aims to regress genome-wide single nucleotide polymorphisms (SNPs) or other types of DNA markers on phenotypes of individuals to simultaneously predict their effects (Meuwissen, Hayes, and Goddard 2001) and combines phenotypic and genotypic data using a training population (TP) to create predictive models for computing the "*Genomic Estimated Breeding Values*" (GEBVs). In GP, this predictive model is subsequently applied to individuals for which only genotyping information is available for ranking individuals based on their predicted performance or GEBVs. This latter collection of individuals having only genotypic information is known as breeding population (BP). In GP, a third set of individuals, defined as validation population (VP), might be optionally genotyped and phenotyped to validate the predictive ability of GP models, that is the accuracy of GP models, without using complex cross validation schemes (**Figure 11**).

Like the breeding values (BV) introduced in the animal models (Henderson 1977), GEBVs point out the genetic merit of individuals that might be mated for the following generation, but are computed using molecular markers. In GP, the prediction of GEBVs is carried out for a TP of plants using different types of statistical models, which aim to fit the observed phenotypic values of the TP along with their genomic profiles.

GP is a cutting-edge methodology that allows reducing selection cycles and efficiently capturing both major and minor gene effects using whole-genome marker regression without knowing the QTLs that underlie target traits (Desta and Ortiz 2014; Meuwissen et al. 2001; Xu et al. 2020). Although this methodology was described almost 20 years ago (Meuwissen et al. 2001), it is now gaining momentum in plant breeding although its application requires genotyping of many loci.

Unlike MAS, GP allows efficiently capturing both major and minor gene effects using whole-genome marker regression without knowing the QTLs that underlie target traits (Cavanagh et al. 2008; Huang et al. 2015). GP exceeds the MAS limits for complex traits (e.g., grain yield, drought tolerance), based on the interaction of many genes and low heritability.

MAGIC population has never been used along with a GP methodology and, potentially, this approach is a revolutionary method that combines good prediction accuracy and training population model recycling for several breeding populations. This approach allows to obtain a GP model that can theoretically be used for different breeding programs.

### 6.1 Genomic predictive ability and cross validation schemes

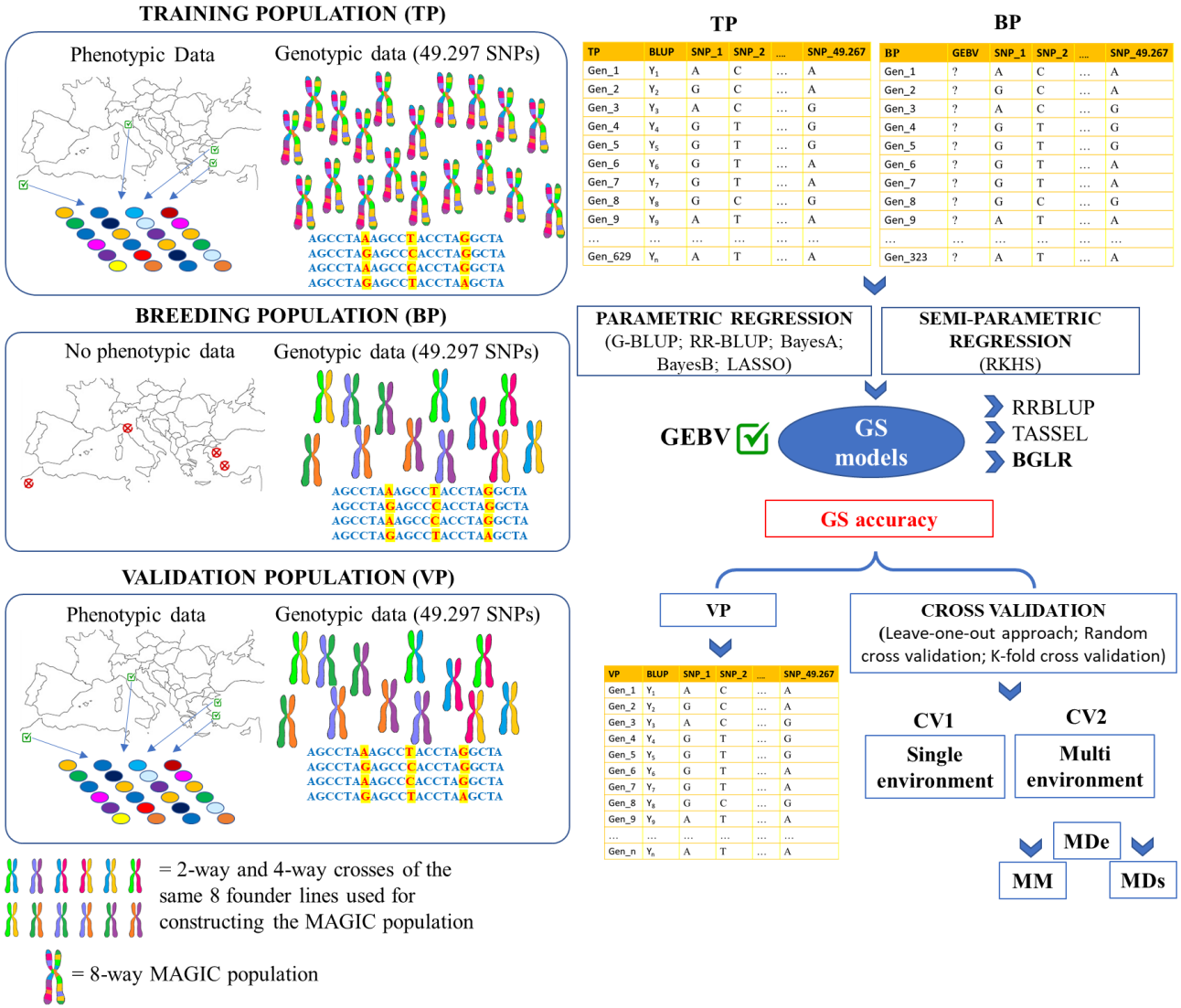
One of the parameters for evaluating GP models is the accuracy. Model accuracy ($r_A$), or predictive ability points out the Pearson's correlation coefficient between the unknown true BV and the GEBVs. In practice, the true BV of plants are unknown quantities and model accuracy is usually computed between GEBVs and the phenotypic values or other estimators (e.g., adjusted means). In GP, model accuracy depends on several genetic factors such as LD, size and genetic diversity of the TP, number

of markers and genetic relationship between training and breeding populations (Crossa et al. 2017) according to the following formula:

$$r_A = \sqrt{\frac{h^2}{h^2 + \frac{M_e}{N_p}}} \qquad \textbf{Equation 17}$$

where $h^2$ is the narrow sense heritability, $N_p$ is the number of individuals in a TP, and $M_e$ is the number of independent chromosome segments, which in turn depend on and consider the extent of LD. There are different approaches for estimating the accuracy of GP models: 1) use of a VPs and 2) cross-validation (CV) schemes, which do not require examining additional individuals for computing model accuracy (**Figure 11**).

**Figure 11: Example of GP applied to MAGIC lines phenotyped for GY and genotyped using 50K SNP array.** The TP is used to predict GEBVs for individuals that have only genotypic information (BP) using parametric and semi-parametric regressions. Finally, the GP accuracy is achieved using either the VP approach or the Cross-Validation approach (CV1 and CV2).




The first approach for estimating the accuracy combines phenotypic and genotypic data using a TP to create predictive models to compute GEBVs of VPs. Subsequently, Pearson's correlation coefficient between GEBVs and the phenotypic values or other estimators (e.g., adjusted means) of VPs are used to estimate the prediction accuracy of GP model. The use of VP implies that additional

resources for genotyping these individuals are available along with phenotypic data. Unlike the use of VPs, CV schemes allow estimating GP model accuracy without investing resources for fingerprinting other individuals. Using CV the different data used as TP are recursively divided into training sets and test sets, which in turn are used to calculate the average accuracy of the models (Jia 2017).

The CV approach can be applying for single environment or multi-environments GP models and two main CV schemes can be pursued. Cross validation 1 (CV1) can be applied to either single-environment or multi-environment GP models and allows predicting GEBVs of individuals that have not been evaluated in any environment (Bandeira e Sousa et al. 2017). Cross validation 2 (CV2) can be applied to multi-environment GP models and allows predicting GEBVs for individuals that have been evaluated in some environments but not included in other environments. Either CV1 or CV2 can be implemented pursuing different strategies (**Figure 11**).

In the leave-one-out (LOO) cross-validation *n* GP models using the *n* lines of the TP are created (Gianola and Schon 2016). These models are created recursively using *n*-1 lines of the TP and computing the GEBV for the line excluded from the model. At the end of this procedure, the Pearson's correlation coefficient between GEBVs and phenotypic values or other estimators (e.g., adjusted means) is used to estimate the accuracy of GP model (**Figure 12**).

Using random cross-validation *n* lines of TP were randomly selected for *x* times (Roorkiwal et al. 2016). Therefore, the accuracy values are obtained from each randomization and estimated through the Pearson's correlation between the phenotypic values or other estimators (e.g. adjusted means) and GEBVs. Then, the mean and standard deviation GP models are estimated (**Figure 13**).

Finally, K-fold cross-validation consists in randomly partitioning *k* portions of population with genotypic and phenotypic information (Jia 2017). At each cycle, a GP model is created using *k* -1 portions while the remaining individuals are used as a test to compute the Pearson's correlation coefficient between GEBVs and phenotypic values or other estimators (e.g., adjusted means) for estimating the accuracy of GP model.

**Figure 12: Graphical representation of leave-one out CV approach.** The prediction accuracy was estimated recursively creating n predictive models, using n-1 lines of the training population and calculating the GEBV for the line excluded from the model.

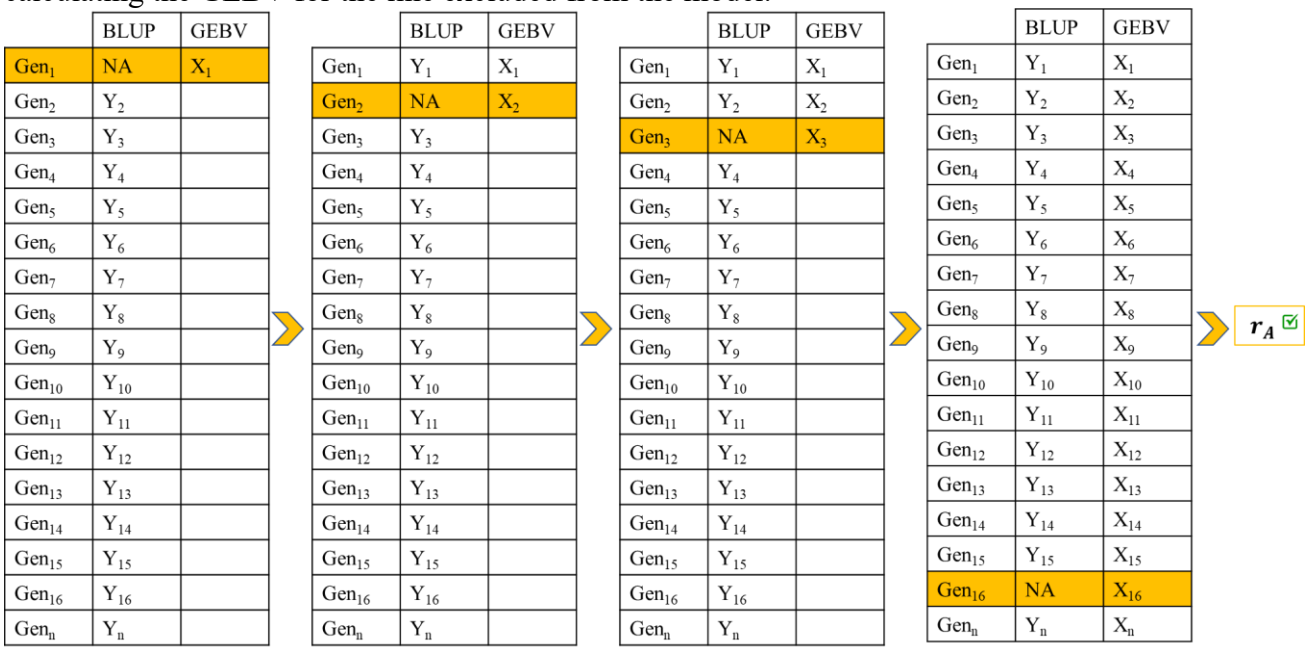
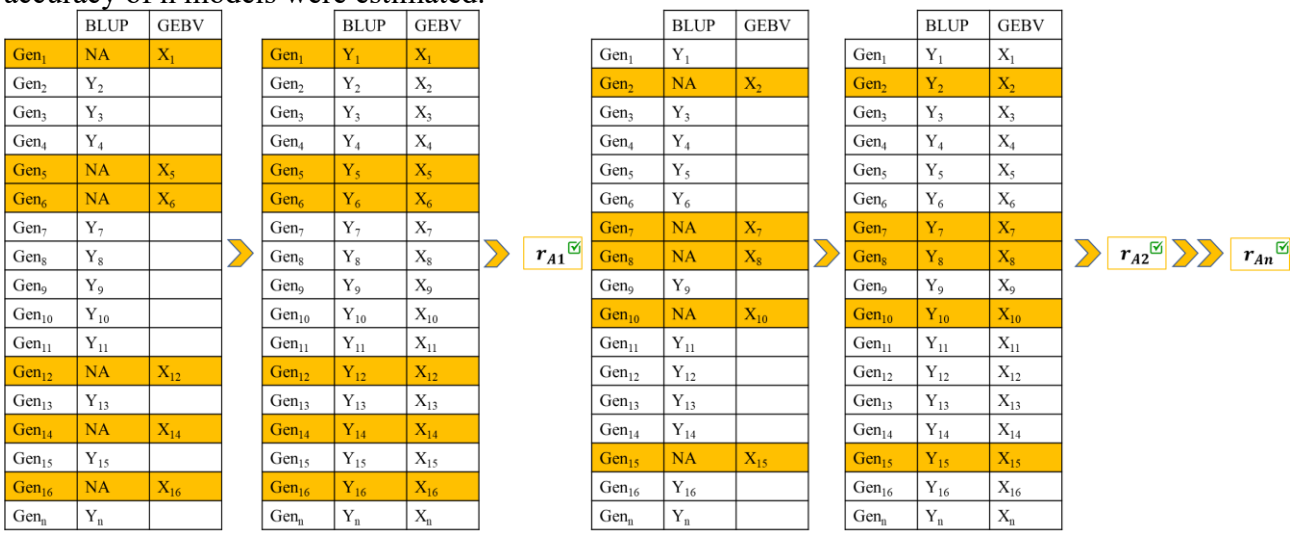


**Figure 13: Graphical representation of random CV approach.** Through the random assignment *n* times of n individuals in the training population, the mean and standard deviation of the prediction accuracy of n models were estimated.




## 6.2 A deeper review of the statistical models that underlie GP

The models used for genomic selection can be summarized as follows:

$$\boldsymbol{y_i = f\,(x_i) + \varepsilon_i} \qquad \textbf{Equation 18}$$

where $\boldsymbol{f\,(x_i)}$ represents the parametric or semi-parametric functions and $\boldsymbol{\varepsilon_i}$ is the residual error with normal distribution.
Alternatively, using matrix notation, the vector of phenotypic observations $y$ for a given trait can be written as an LMM:

$$\boldsymbol{y = X\beta + Zu + e} \qquad \textbf{Equation 19}$$

where $X$ is the incidence matrix of fixed effects, $\beta$ is the vector of fixed effects, $Z$ is the incidence matrix of random effects that in GP indicates molecular markers, $u$ is the vector of random effects that in GP points out the predicted effects of markers and $e$ is the vector of residual error terms. **Equation 19** assumes that $u \sim N(0, G)$, that is the vector of random genetic effects follows a multi-variate normal distribution with mean 0 and covariance structure equals to $G$ matrix. Similarly, it assumes that $e \sim (0, I\sigma_e^2)$, that is the terms of the error vector are independently normally distributed with mean 0 and variance structure equals to $\sigma_e^2$. Moreover, it supposes that $cov(u, e) = 0$, that is that the vectors of random effects and of model errors are independent and not correlated. Generally, $y$, $X$ and $Z$ are known, while $\beta$ and $u$ must be estimated and predicted, respectively. Henderson's mixed model equations (Henderson 1984) provide the mathematical solutions to compute the estimators of $\beta$ and $u$:

$$\boldsymbol{\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} y} \qquad \textbf{Equation 20}$$

$$\boldsymbol{\hat{u} = G Z^T V^{-1} (y - X\hat{\beta})} \qquad \textbf{Equation 21}$$

where $\boldsymbol{V = ZGZ^T + R}$. In the standard animal model (Henderson 1984), the covariance matrix $\boldsymbol{G}$ of additive effects is modelled as $G = \sigma_A^2 A$, where $\sigma_A^2$ is the variance of additive effects and $A$ is the additive genetic relationship matrix, which is usually based on pedigree, that is using expected relatedness between individuals. This $A$ matrix does not consider the actual genetic relationships between individuals as mendelian sampling causes deviation from the expected relatedness. In GP,

the matrix $A$ is replaced with the matrix of genomic relationship $K$, which is constructed using molecular markers. Unlike matrix $A$, $K$ measures the relatedness between individuals considering mendelian sampling and allows to better estimate $\hat{u}$ according to **Equation 20** and **Equation 21** (VanRaden 2008). The incorporation of genomic information using matrix $K$ leads to estimate marker effects using genomic BLUP (GBLUP) which has become of the most popular method used for implementing GP in plants and animals (VanRaden 2008). GP models implemented using Ridge Regression BLUP (RRBLUP) (Piepho et al. 2012) have been demonstrated to be equivalent to GBLUP (Habier et al. 2010).

Referring to **Equation 18,** parametric and non-parametric methods are based on different hypotheses and algorithms and for this reason the accuracies and variances that can produce are different. Beyond the aforementioned GBLUP method described above, the most widely used parametric methods are BRR (Bayesian ridge regression), BayesA, BayesB and Bayesian LASSO (least absolute shrinkage and selection operator). Instead, the most popular non-parametric method is RKHS (Reproducing Kernel Hilbert space) regression. These methods, parametric and non-parametric, aim to regress whole genome markers on phenotypes.

## 7 Aim of the Ph.D. research thesis

The main aim of this research project was to develop and validate genetic improvement methodologies based on genomic prediction to underpin barley breeding in the Mediterranean environments.

The first objective (**Chapter 2**) was to characterize a new barley MAGIC population constructed with a diverse founder set of winter cultivars showing contrasting GY, which was examined across an ample range of site-by-season combinations characterized by different temperature and precipitation patterns.

The second objective (**Chapter 2**) was to combine the phenotypic data of MAGIC collected across these field trials with genotypic information to fit different single-environment genomic prediction and multi environment genomic prediction models for empirically assessing the predictive ability in multi-parent populations. Moreover, we applied different untargeted optimization methods to this MAGIC population for assembling and benchmarking the performance of optimized TPs. Fitting SE-GP and ME-GP models to MAGIC lines, we aimed at broadening the use of these experimental populations beyond classical QTL mapping and analysis of epistatic effects for sustaining and accelerating barley breeding.

The third objective (**Chapter 3**) was to investigate the diversity and distribution of morphological and physiological traits in the same MAGIC barley population. We correlated phenotypic data of SRA, SRN and TR along with GY obtained in different site-by-season combinations to re-assess the relevance of belowground and physiological traits of seedlings for predicting drought-tolerant and drought-intolerant lines. Leveraging on phenotypic and genotypic information we fitted and cross-validated different GP models using standard and threshold GBLUP and models for count data including additive interaction. We showed that these models can successfully predict SRA, SRN and TR and might pave the way for characterizing large plant collections while minimizing resources for phenotyping.

# 8 Reference


Abberton, Michael, Jacqueline Batley, Alison Bentley, John Bryant, Hongwei Cai, James Cockram, Antonio Costa de Oliveira, Leland J. Cseke, Hannes Dempewolf, Ciro De Pace, David Edwards, Paul Gepts, Andy Greenland, Anthony E. Hall, Robert Henry, Kiyosumi Hori, Glenn Thomas Howe, Stephen Hughes, Mike Humphreys, David Lightfoot, Athole Marshall, Sean Mayes, Henry T. Nguyen, Francis C. Ogbonnaya, Rodomiro Ortiz, Andrew H. Paterson, Roberto Tuberosa, Babu Valliyodan, Rajeev K. Varshney, and Masahiro Yano. 2016. "Global Agricultural Intensification during Climate Change: A Role for Genomics." *Plant Biotechnology Journal* 14(4):1095–98.

Ahmad, Asif, Faqir Muhammad Anjum, Tahir Zahoor, Haq Nawaz, and Syed Muhammad Raihan Dilshad. 2012. "Beta Glucan: A Valuable Functional Ingredient in Foods." *Critical Reviews in Food Science and Nutrition* 52(3):201–12.

Araus, José Luis, and Jill E. Cairns. 2014. "Field High-Throughput Phenotyping: The New Crop Breeding Frontier." *Trends in Plant Science* 19(1):52–61.

Arneth, A., F. Denton, F. Agus, A. Elbehri, K. Erb, Elasha, B. Osman, M. Rahimi, M. Rounsevell, A. Spence, and R. Valentini. 2019. "Framing and Context. In: Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems [P.R. Shukla, J. Skea, E. Calvo ." *IPCC Intergovernmental Panel on Climate Change* 46–54.

Arrones, Andrea, Santiago Vilanova, Mariola Plazas, Giulio Mangino, Laura Pascual, María José Díez, Jaime Prohens, and Pietro Gramazio. 2020. "The Dawn of the Age of Multi-Parent Magic Populations in Plant Breeding: Novel Powerful next-Generation Resources for Genetic Analysis and Selection of Recombinant Elite Material." *Biology* 9(8):1–25.

Van As, Henk. 2007. "Intact Plant MRI for the Study of Cell Water Relations, Membrane Permeability, Cell-to-Cell and Long Distance Water Transport." *Journal of Experimental Botany* 58(4):1–14.

Bandeira e Sousa, Massaine, Jaime Cuevas, Evellyn Giselly de Oliveira Couto, Paulino Pérez-Rodríguez, Diego Jarquín, Roberto Fritsche-Neto, Juan Burgueño, and Jose Crossa. 2017. "Genomic-Enabled Prediction in Maize Using Kernel Models with Genotype 3 Environment Interaction." *G3: Genes, Genomes, Genetics* 7(6):1995–2014.

Barcaccia, Gianni, and Mario Falcinelli. 2006. "Genetica e Genomica. Vol. 2: Miglioramento Genetico." P. 440 in *Genetica e Genomica*, edited by Liguori.

Bayer, Micha M., Paulo Rapazote-Flores, Martin Ganal, Pete E. Hedley, Malcolm Macaulay, Jörg Plieske, Luke Ramsay, Joanne Russell, Paul D. Shaw, William Thomas, and Robbie Waugh. 2017. "Development and Evaluation of a Barley 50k ISelect SNP Array." *Frontiers in Plant Science* 8:1–10.

Bleidere, Mara, and Zinta Gaile. 2012. "Grain Quality Traits Important in Feed Barley." *Proceedings of the Latvian Academy of Sciences, Section B: Natural, Exact, and Applied Sciences* 66(1–2):1–9.

Bülow, Lorenz, Marion Nachtigall, and Lothar Frese. 2019. "A MAGIC Population as an Approach to the Conservation and Development of Genetic Diversity of Winter Barley for Breeding Purposes by On-Farm Management." *Journal Fur Kulturpflanzen* 71(11):286–98.

Cammarano, Davide, Salvatore Ceccarelli, Stefania Grando, Ignacio Romagosa, Abdelkader Benbelkacem, Tanek Akar, Adnan Al-Yassin, Nicola Pecchioni, Enrico Francia, and Domenico Ronga. 2019. "The Impact of Climate Change on Barley Yield in the Mediterranean Basin." *European Journal of Agronomy* 106(February):1–11.

Cattivelli, Luigi, Fulvia Rizza, Franz W. Badeck, Elisabetta Mazzucotelli, Anna M. Mastrangelo, Enrico Francia, Caterina Marè, Alessandro Tondelli, and A. Michele Stanca. 2008. "Drought Tolerance Improvement in Crop Plants: An Integrated View from Breeding to Genomics." *Field Crops Research* 105(1–2):1–14.

Cavanagh, Colin, Matthew Morell, Ian Mackay, and Wayne Powell. 2008. "From Mutations to MAGIC: Resources for Gene Discovery, Validation and Delivery in Crop Plants." *Current Opinion in Plant Biology* 11(2):215–21.

Chakraborty, S., and A. C. Newton. 2011. "Climate Change, Plant Diseases and Food Security: An Overview." *Plant Pathology* 60(1):2–14.

Chen, Xifeng, Shufang Fu, Pinghua Zhang, Zhimin Gu, Jianzhong Liu, Qian Qian, and Bojun Ma. 2013. "Proteomic Analysis of a Disease-Resistance-Enhanced Lesion Mimic Mutant Spotted Leaf 5 in Rice." *Rice* 6(1):1–10.

Crossa, José, Paulino Pérez-Rodríguez, Jaime Cuevas, Osval Montesinos-López, Diego Jarquín, Gustavo de los Campos, Juan Burgueño, Juan M. González-Camacho, Sergio Pérez-Elizalde, Yoseph Beyene, Susanne Dreisigacker, Ravi Singh, Xuecai Zhang, Manje Gowda, Manish Roorkiwal, Jessica Rutkoski, and Rajeev K. Varshney. 2017. "Genomic Selection in Plant Breeding: Methods, Models, and Perspectives." *Trends in Plant Science* 22(11):961–75.

Dell'Acqua, Matteo, Daniel M. Gatti, Giorgio Pea, Federica Cattonaro, Frederik Coppens, Gabriele Magris, Aye L. Hlaing, Htay H. Aung, Hilde Nelissen, Joke Baute, Elisabetta Frascaroli, Gary A. Churchill, Dirk Inzé, Michele Morgante, and Mario Enrico Pè. 2015. "Genetic Properties of the MAGIC Maize Population: A New Platform for High Definition QTL Mapping in Zea Mays." *Genome Biology* 16(1):1–23.

Desta, Zeratsion Abera, and Rodomiro Ortiz. 2014. "Genomic Selection: Genome-Wide Prediction in Plant Improvement." *Trends in Plant Science* 19(9):592–601.

Din, A., M. F. J. Chughtai, M. R. K. Khan, A. Shahzad, A. Khaliq, and M. A. Nasir. 2018. "Nutritional and Functional Perspectives of Barley β-Glucan." *International Food Research Journal* 25(5):1773–84.

Donato-Capel, L., C. L. Garcia-Rodenas, E. Pouteau, U. Lehmann, S. Srichuwong, A. Erkner, E. Kolodziejczyk, E. Hughes, T. J. Wooster, and L. Sagalowicz. 2014. *Technological Means to Modulate Food Digestion and Physiological Response*. Elsevier Inc.

de Dorlodot, Sophie, Brian Forster, Loïc Pagès, Adam Price, Roberto Tuberosa, and Xavier Draye. 2007. "Root System Architecture: Opportunities and Constraints for Genetic Improvement of Crops." *Trends in Plant Science* 12(10):474–81.

Gianola, Daniel, and Chris Carolin Schon. 2016. "Cross-Validation without Doing Cross-Validation in Genome-Enabled Prediction." *G3: Genes, Genomes, Genetics* 6(10):3107–28.

Giraldo, Patricia, Elena Benavente, Francisco Manzano-Agugliaro, and Estela Gimenez. 2019. "Worldwide Research Trends on Wheat and Barley: A Bibliometric Comparative Analysis." *Agronomy* 9(7).

Grando, Stefania, and Helena Gomez Macpherson. 2005. *Food Barley: Importance, Uses and Local Knowledge*. Aleppo, Syria,.

Gupta, Mahesh, Nissreen Abu-Ghannam, and Eimear Gallaghar. 2010. "Barley for Brewing: Characteristic Changes during Malting, Brewing and Applications of Its by-Products." *Comprehensive Reviews in Food Science and Food Safety* 9(3):318–28.

Habier, David, Jens Tetens, Franz Reinhold Seefried, Peter Lichtner, and Georg Thaller. 2010. "The Impact of Genetic Relationship Information on Genomic Breeding Values in German Holstein Cattle." *Genetics Selection Evolution* 42(1):1–12.

Hamada, Alhosein, Miyuki Nitta, Shuhei Nasuda, Kenji Kato, Masaya Fujita, Hitoshi Matsunaka, and Yutaka Okumoto. 2012. "Novel QTLs for Growth Angle of Seminal Roots in Wheat (Triticum Aestivum L.)." *Plant and Soil* 354(1–2):395–405.

Han, Zhongmin, Gang Hu, Hua Liu, Famao Liang, Lin Yang, Hu Zhao, Qinghua Zhang, Zhixin Li, Qifa Zhang, and Yongzhong Xing. 2020. "Bin-Based Genome-Wide Association Analyses Improve Power and Resolution in QTL Mapping and Identify Favorable Alleles from Multiple Parents in a Four-Way MAGIC Rice Population." *Theoretical and Applied Genetics* 133(1):59–71.

El Hassouni, K., S. Alahmad, B. Belkadi, A. Filali-Maltouf, L. T. Hickey, and F. M. Bassi. 2018.

"Root System Architecture and Its Association with Yield under Different Water Regimes in Durum Wheat." *Crop Science* 58(6):2331–46.

Hayman, B. I. 1954. "The Analysis of Variance of Diallel Tables." *International Biometric Society* 10(2):235–44.

Henderson, C. R. 1977. "Best Linear Unbiased Prediction of Breeding Values Not in the Model for Records." *Journal of Dairy Science*.

Henderson, C. R. 1984. *Applications of Linear Models in Animal Breeding*.

Huang, B. Emma, Klara L. Verbyla, Arunas P. Verbyla, Chitra Raghavan, Vikas K. Singh, Pooran Gaur, Hei Leung, Rajeev K. Varshney, and Colin R. Cavanagh. 2015. "MAGIC Populations in Crops: Current Status and Future Prospects." *Theoretical and Applied Genetics* 128(6):999–1017.

Huang, Bevan E., Andrew W. George, Kerrie L. Forrest, Andrzej Kilian, Matthew J. Hayden, Matthew K. Morell, and Colin R. Cavanagh. 2012. "A Multiparent Advanced Generation Inter-Cross Population for Genetic Analysis in Wheat." *Plant Biotechnology Journal* 10(7):826–39.

Jayakodi, Murukarthick, Sudharsan Padmarasu, Georg Haberer, Venkata Suresh Bonthala, Heidrun Gundlach, Cécile Monat, Thomas Lux, Nadia Kamal, Daniel Lang, Axel Himmelbach, Jennifer Ens, Xiao Qi Zhang, Tefera T. Angessa, Gaofeng Zhou, Cong Tan, Camilla Hill, Penghao Wang, Miriam Schreiber, Lori B. Boston, Christopher Plott, Jerry Jenkins, Yu Guo, Anne Fiebig, Hikmet Budak, Dongdong Xu, Jing Zhang, Chunchao Wang, Jane Grimwood, Jeremy Schmutz, Ganggang Guo, Guoping Zhang, Keiichi Mochida, Takashi Hirayama, Kazuhiro Sato, Kenneth J. Chalmers, Peter Langridge, Robbie Waugh, Curtis J. Pozniak, Uwe Scholz, Klaus F. X. Mayer, Manuel Spannagl, Chengdao Li, Martin Mascher, and Nils Stein. 2020. "The Barley Pan-Genome Reveals the Hidden Legacy of Mutation Breeding." *Nature* 588(7837):284–89.

Jia, Zhenyu. 2017. "Controlling the Overfitting of Heritability in Genomic Selection through Cross Validation." *Scientific Reports* 7(1):1–9.

Jia, Zhongtao, Ying Liu, Benjamin D. Gruber, Kerstin Neumann, Benjamin Kilian, Andreas Graner, and Nicolaus von Wirén. 2019. "Genetic Dissection of Root System Architectural Traits in Spring Barley." *Frontiers in Plant Science* 10(April):1–14.

Jiménez-Galindo, José Cruz, Rosa Ana Malvar, Ana Butrón, Rogelio Santiago, Luis Fernando Samayoa, Marlon Caicedo, and Bernardo Ordás. 2019. "Mapping of Resistance to Corn Borers in a MAGIC Population of Maize." *BMC Plant Biology*.

Khlestkina, E. K., and E. A. Salina. 2006. "SNP Markers: Methods of Analysis, Ways of Development, and Comparison on an Example of Common Wheat." *Russian Journal of Genetics* 42(6):585–94.

Langridge, Peter. 2018. "Economic and Academic Importance of Barley." Pp. 345–61 in *The Barley Genome*. Springer International Publishing.

Lee, S. 2017. "Strategic Design of Delivery Systems for Nutraceuticals." *Nanotechnology Applications in Food: Flavor, Stability, Nutrition and Safety* 65–86.

Legendre, Adrien Marie. 1805. *Nouvelles Méthodes Pour La Détermination Des Orbites Des Comètes [New Methods for the Determination of the Orbits of Comets]*. edited by F. Didot. Paris (French).

Lobell, David B., Marshall B. Burke, Claudia Tebaldi, Michael D. Mastrandrea, Walter P. Falcon, and Rosamond L. Naylor. 2008. "Prioritizing Climate Change Adaptation Needs for Food Security in 2030." *Science* 319(5863):607–10.

Lynch, Jonathan P., and Kathleen M. Brown. 2001. "Topsoil Foraging - An Architectural Adaptation of Plants to Low Phosphorus Availability." *Plant and Soil* 237(2):225–37.

Lynch, Michael, and Bruce Walsh. 1997. *Genetics and Analysis of Quantitative Traits*. Sunderland, MA.

Mackay, Ian J., Pauline Bansept-Basler, Alison R. Bentley, James Cockram, Nick Gosman, Andy J. Greenland, Richard Horsnell, Rhian Howells, Donal M. O'Sullivan, Gemma A. Rose, and Phil J. Howell. 2014. "An Eight-Parent Multiparent Advanced Generation Inter-Cross Population for

Winter-Sown Wheat: Creation, Properties, and Validation." *G3: Genes, Genomes, Genetics* 4(9):1603–10.

Manschadi, Ahmad M., Graeme L. Hammer, John T. Christopher, and Peter DeVoil. 2008. "Genotypic Variation in Seedling Root Architectural Traits and Implications for Drought Adaptation in Wheat (Triticum Aestivum L.)." *Plant and Soil* 303(1–2):115–29.

Mascher, Martin, Heidrun Gundlach, Axel Himmelbach, Sebastian Beier, Sven O. Twardziok, Thomas Wicker, Volodymyr Radchuk, Christoph Dockter, Pete E. Hedley, Joanne Russell, Micha Bayer, Luke Ramsay, Hui Liu, Georg Haberer, Xiao Qi Zhang, Qisen Zhang, Roberto A. Barrero, Lin Li, Stefan Taudien, Marco Groth, Marius Felder, Alex Hastie, Hana Šimková, Helena Stanková, Jan Vrána, Saki Chan, Mariá Munõz-Amatriaín, Rachid Ounit, Steve Wanamaker, Daniel Bolser, Christian Colmsee, Thomas Schmutzer, Lala Aliyeva-Schnorr, Stefano Grasso, Jaakko Tanskanen, Anna Chailyan, Dharanya Sampath, Darren Heavens, Leah Clissold, Sujie Cao, Brett Chapman, Fei Dai, Yong Han, Hua Li, Xuan Li, Chongyun Lin, John K. McCooke, Cong Tan, Penghao Wang, Songbo Wang, Shuya Yin, Gaofeng Zhou, Jesse A. Poland, Matthew I. Bellgard, Ljudmilla Borisjuk, Andreas Houben, Jaroslav Doleael, Sarah Ayling, Stefano Lonardi, Paul Kersey, Peter Langridge, Gary J. Muehlbauer, Matthew D. Clark, Mario Caccamo, Alan H. Schulman, Klaus F. X. Mayer, Matthias Platzer, Timothy J. Close, Uwe Scholz, Mats Hansson, Guoping Zhang, Ilka Braumann, Manuel Spannagl, Chengdao Li, Robbie Waugh, and Nils Stein. 2017. "A Chromosome Conformation Capture Ordered Sequence of the Barley Genome." *Nature* 544(7651):427–33.

Masson-Delmotte, V., P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield. 2018. "IPCC, 2018: Summary for Policymakers." *Global Warming of 1.5°C. An IPCC Special Report on the Impacts of Global Warming of 1.5°C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change,*.

Medina, Susan, Rubén Vicente, Maria Teresa Nieto-Taladriz, Nieves Aparicio, Fadia Chairi, Omar Vergara-Diaz, and José Luis Araus. 2019. "The Plant-Transpiration Response to Vapor Pressure Deficit (VPD) in Durum Wheat Is Associated with Differential Yield Performance and Specific Expression of Genes Involved in Primary Metabolism and Water Transport." *Frontiers in Plant Science* 9(January):1–19.

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):1819–29.

Milner, Sara Giulia, Marco Maccaferri, Bevan Emma Huang, Paola Mantovani, Andrea Massi, Elisabetta Frascaroli, Roberto Tuberosa, and Silvio Salvi. 2016. "A Multiparental Cross Population for Mapping QTL for Agronomic Traits in Durum Wheat (Triticum Turgidum Ssp. Durum)." *Plant Biotechnology Journal* 14(2):735–48.

Monat, Cécile, Sudharsan Padmarasu, Thomas Lux, Thomas Wicker, Heidrun Gundlach, Axel Himmelbach, Jennifer Ens, Chengdao Li, Gary J. Muehlbauer, Alan H. Schulman, Robbie Waugh, Ilka Braumann, Curtis Pozniak, Uwe Scholz, Klaus F. X. Mayer, Manuel Spannagl, Nils Stein, and Martin Mascher. 2019. "TRITEX : Chromosome-Scale Sequence Assembly of Triticeae Genomes with Open- Source Tools." *Genome Biology* 20(1):1–35.

Naz, Ali Ahmad, Md Arifuzzaman, Shumaila Muzammil, Klaus Pillen, and Jens Léon. 2014. "Wild Barley Introgression Lines Revealed Novel QTL Alleles for Root and Related Shoot Traits in the Cultivated Barley (Hordeum Vulgare L.)." *BMC Genetics* 15(1):1–12.

Ogawa, Daisuke, Eiji Yamamoto, Toshikazu Ohtani, Noriko Kanno, Hiroshi Tsunematsu, Yasunori Nonoue, Masahiro Yano, Toshio Yamamoto, and Jun Ichi Yonemaru. 2018. "Haplotype-Based Allele Mining in the Japan-MAGIC Rice Population." *Scientific Reports* 8(1):1–11.

Ongom, Patrick O., and Gebisa Ejeta. 2018. "Mating Design and Genetic Structure of a Multi-Parent Advanced Generation Intercross (MAGIC) Population of Sorghum (Sorghum Bicolor (L.)

Moench).” *G3: Genes, Genomes, Genetics* 8(1):331–41.

Piepho, H. P., J. O. Ogutu, T. Schulz-Streeck, B. Estaghvirou, A. Gordillo, and F. Technow. 2012. “Efficient Computation of Ridge-Regression Best Linear Unbiased Prediction in Genomic Selection in Plant Breeding.” *Crop Science* 52(3):1093–1104.

Pswarayi, A., F. A. Van Eeuwijk, S. Ceccarelli, S. Grando, J. Comadran, J. R. Russell, E. Francia, N. Pecchioni, O. Li Destri, T. Akar, A. Al-Yassin, A. Benbelkacem, W. Choumane, M. Karrou, H. Ouabbou, J. Bort, J. L. Araus, J. L. Molina-Cano, W. T. B. Thomas, and I. Romagosa. 2008. “Barley Adaptation and Improvement in the Mediterranean Basin.” *Plant Breeding* 127(6):554–60.

Ramakrishna, Ramnarain, Dipayan Sarkar, Paul Schwarz, and Kalidas Shetty. 2017. “Phenolic Linked Anti-Hyperglycemic Bioactives of Barley (Hordeum Vulgare L.) Cultivars as Nutraceuticals Targeting Type 2 Diabetes.” *Industrial Crops and Products* 107:509–17.

Rashed, Md Rais Uddin. 2016. “Substrate Effects on Plant Transpiration Rate under Several Vapour Pressure Deficit (VPD) Levels.” *Journal of Plant Pathology & Microbiology* 7(7).

Rebetzke, Greg J., Arunas P. Verbyla, Klara L. Verbyla, Matthew K. Morell, and Colin R. Cavanagh. 2014. “Use of a Large Multiparent Wheat Mapping Population in Genomic Dissection of Coleoptile and Seedling Growth.” *Plant Biotechnology Journal* 12(2):219–30.

Richard, Cecile Al, Lee T. Hickey, Susan Fletcher, Raeleen Jennings, Karine Chenu, and Jack T. Christopher. 2015. “High-Throughput Phenotyping of Seminal Root Traits in Wheat.” *Plant Methods* 11(1):1–11.

Robinson, G. K. 1991. “That BLUP Is a Good Thing: The Estimation of Random Effects.” *Statistical Science* 6(1):15–51.

Robinson, Hannah, Alison Kelly, Glen Fox, Jerome Franckowiak, Andrew Borrell, and Lee Hickey. 2018. “Root Architectural Traits and Yield: Exploring the Relationship in Barley Breeding Trials.” *Euphytica* 214(151):1–14.

Roorkiwal, Manish, Abhishek Rathore, Roma R. Das, Muneendra K. Singh, Ankit Jain, Samineni Srinivasan, Pooran M. Gaur, Bharadwaj Chellapilla, Shailesh Tripathi, Yongle Li, John M. Hickey, Aaron Lorenz, Tim Sutton, Jose Crossa, Jean Luc Jannink, and Rajeev K. Varshney. 2016. “Genome-Enabled Prediction Models for Yield Related Traits in Chickpea.” *Frontiers in Plant Science* 7(NOVEMBER2016).

Sadok, Walid, and Bishal G. Tamang. 2019. “Diversity in Daytime and Night-Time Transpiration Dynamics in Barley Indicates Adaptation to Drought Regimes across the Middle-East.” *Journal of Agronomy and Crop Science* 205(4):372–84.

Salamini, Francesco, Hakan Özkan, Andrea Brandolini, Ralf Schäfer-Pregl, and William Martin. 2002. “Genetics and Geography of Wild Cereal Domestication in the near East.” *Nature Reviews Genetics* 3(6):429–41.

Sannemann, Wiebke, Bevan Emma Huang, Boby Mathew, and Jens Léon. 2015. “Multi-Parent Advanced Generation Inter-Cross in Barley: High-Resolution Quantitative Trait Locus Mapping for Flowering Time as a Proof of Concept.” *Molecular Breeding* 35(3).

Sannemann, Wiebke, Antonia Lisker, Andreas Maurer, Jens Léon, Ebrahim Kazman, Hilmar Cöster, Josef Holzapfel, Hubert Kempf, Viktor Korzun, Erhard Ebmeyer, and Klaus Pillen. 2018. “Adaptive Selection of Founder Segments and Epistatic Control of Plant Height in the MAGIC Winter Wheat Population WM-800.” *BMC Genomics* 19(1):1–16.

Sayed, Mohammed A., Alhosein Hamada, Jens Lèon, and Ali Ahmed Naz. 2017. “Genetic Mapping Reveals Novel Exotic QTL Alleles for Seminal Root Architecture in Barley Advanced Backcross Double Haploid Population.” *Euphytica* 213(1).

Schreiber, Miriam, Martin Mascher, Jonathan Wright, Sudharasan Padmarasu, Axel Himmelbach, Darren Heavens, Linda Milne, Bernardo J. Clavijo, Nils Stein, and Robbie Waugh. 2020. “A Genome Assembly of the Barley ‘ Transformation Reference ’ Cultivar Golden Promise.” *Genetics* 10(June):1823–27.

Shah, Rohan, B. Emma Huang, Alex Whan, Marcus Newberry, Klara Verbyla, Matthew Morell, and

Colin Cavanagh. 2019. “The Complex Genetic Architecture of Recombination and Structural Variation in Wheat Uncovered Using a Large 8-Founder MAGIC Population.” *BioRxiv* 594317.

Stadlmeier, Melanie, Lorenz Hartl, and Volker Mohler. 2018. “Usefulness of a Multiparent Advanced Generation Intercross Population with a Greatly Reduced Mating Design for Genetic Studies in Winter Wheat.” *Frontiers in Plant Science* 871(December):1–12.

Thépot, Stéphanie, Gwendal Restoux, Isabelle Goldringer, Frédéric Hospital, David Gouache, Ian Mackay, and Jérôme Enjalbert. 2014. “Efficiently Tracking Selection in a Multiparental Population: The Case of Earliness in Wheat.” *Genetics* 199(2):609–23.

Tricase, Caterina, Vera Amicarelli, Emilia Lamonaca, and Roberto Leonardo Rana. 2018. “Chapter 2: Economic Analysis of the Barley Market and Related Uses.” *IntechOpen* i:38.

Ullrich, Steven E. 2014. *The Barley Crop: Origin and Taxonomy, Production, and End Uses*. 2nd ed. Elsevier Ltd.

VanRaden, P. M. 2008. “Efficient Methods to Compute Genomic Predictions.” *Journal of Dairy Science* 91(11):4414–23.

Wahbi, A., and P. J. Gregory. 1995. “Growth and Development of Young Roots of Barley (Hordeum Vulgare L.) Genotypes.” *Annals of Botany* 75(5):533–39.

Wicker, Thomas, Alan H. Schulman, Jaakko Tanskanen, Manuel Spannagl, Sven Twardziok, Martin Mascher, Nathan M. Springer, Qing Li, Robbie Waugh, Chengdao Li, Guoping Zhang, Nils Stein, Klaus F. X. Mayer, and Heidrun Gundlach. 2017. “The Repetitive Landscape of the 5100 Mbp Barley Genome.” *Mobile DNA* 8(1):1–16.

Xie, Wei, Wei Xiong, Jie Pan, Tariq Ali, CUi Qi, Dabo Guan, Jing Meng, Nathaniel D. Mueller, Erda Lin, and Steven J. Davis. 2018. “Decreases in Global Beer Supply Due to Extreme Drought and Heat.” *Nature Plants* 4(11):964-973.

Xu, Yunbi, Xiaogang Liu, Junjie Fu, Hongwu Wang, Jiankang Wang, Changling Huang, Boddupalli M. Prasanna, Michael S. Olsen, Guoying Wang, and Aimin Zhang. 2020. “Enhancing Genetic Gain through Genomic Selection: From Livestock to Plants.” *Plant Communications* 1:1–21.

Yalcin, B., J. Flint, and R. Mott. 2005. “Using Progenitor Strain Information to Identify Quantitative Trait Nucleotides in Outbred Mice.” *Genetics*.

# Chapter 2: Genomic prediction of grain yield in a barley MAGIC population modelling genotype per environment interaction

**Damiano Puglisi**[1ǂ], Stefano Delbono[2ǂ], Andrea Visioni[3], Hakan Ozkan[4], İbrahim Kara[5], Ana M. Casas[6], Ernesto Igartua[6], Giampiero Valè[7], Angela Roberta Lo Piero[1], Luigi Cattivelli[2], Alessandro Tondelli[2], Fricano Agostino[2*]

[1]Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A), Università di Catania, Via S. Sofia 98, 95123 Catania, Italy
[2]Council for Agricultural Research and Economics – Research Centre for Genomics and Bioinformatics, Via San Protaso 302 Fiorenzuola d'Arda (PC), Italy
[3]Biodiversity and Crop Improvement Program, International Center for Agricultural Research in the Dry Areas, Avenue Hafiane Cherkaoui, Rabat, Morocco
[4]Department of Field Crops, Faculty of Agriculture, University of Cukurova, 01330 Adana, Turkey
[5]Bahri Dagdas International Agricultural Research Institute, Ereğli Yolu 3. Km Karatay / Konya, 42020, Turkey
[6]Aula Dei Experimental Station (EEAD-CSIC), Spanish Research Council, Avda Montañana 1005, 50059 Zaragoza, Spain
[7]Dipartimento di Scienze e Innovazione Tecnologica (DiSIT), Università del Piemonte Orientale, Piazza San Eusebio 5, I-13100 Vercelli, Italy

[ǂ]These authors contributed equally to this work.
***Correspondence:**
Corresponding Author
agostino.fricano@crea.gov.it

**Keywords**: Genomic prediction, MAGIC, barley, GBLUP, Genotype x Environment interaction

**ORIGINAL RESEARCH article**
Front. Plant Sci., 24 May 2021 | https://doi.org/10.3389/fpls.2021.664148

## 1 Abstract

Multi-parent Advanced Generation Inter-crosses (MAGIC) lines have mosaic genomes that are generated shuffling the genetic material of the founder parents following pre-defined crossing schemes. In cereal crops, these experimental populations have been extensively used to investigate the genetic bases of several traits and dissect the genetic bases of epistasis. In plants, genomic prediction models are usually fitted using either diverse panels of mostly unrelated accessions or individuals of biparental families and several empirical analyses have been conducted to evaluate the predictive ability of models fitted to these populations using different traits.

In this paper, we constructed, genotyped and evaluated a barley MAGIC population of 352 individuals developed with a diverse set of eight founder parents showing contrasting phenotypes for grain yield. We combined phenotypic and genotypic information of this MAGIC population to fit several genomic prediction models which were cross-validated to conduct empirical analyses aimed at examining the predictive ability of these models varying the sizes of training populations. Moreover, several methods to optimize the composition of the training population were also applied to this MAGIC population and cross-validated to estimate the resulting predictive ability. Finally, extensive phenotypic data generated in field trials organized across an ample range of water regimes

and climatic conditions in the Mediterranean were used to fit and cross-validate multi-environment genomic prediction models including GxE interaction, using both genomic best linear unbiased prediction and reproducing kernel Hilbert space along with a non-linear Gaussian Kernel.

Overall, our empirical analyses showed that genomic prediction models trained with a limited number of MAGIC lines can be used to predict grain yield with values of predictive ability that vary from 0.25 to 0.60 and that beyond QTL mapping and analysis of epistatic effects, MAGIC population might be used to successfully fit genomic prediction models. We concluded that for grain yield, the single-environment genomic prediction models examined in this study are equivalent in terms of predictive ability while, in general, multi-environment models that explicitly split marker effects in main and environmental-specific effects outperform simpler multi-environment models.

## 2 Introduction

The experimental design that underlies Multi-parent Advanced Generation Intercrosses (MAGIC) populations traces its origins to the advanced inter-cross lines, which were originally developed in animal model species (Yalcin et al. 2005). MAGIC populations are developed crossing multiple inbred parents or founders, which are subsequently inter-mated several times following pre-defined crossing schemes to shuffle founder genomes in each single line (Huang et al. 2015). In plants, MAGIC populations have been explicitly developed for genetic research purposes as they allow to increase power and precision for detecting and mapping quantitative trait loci (QTLs) (Cavanagh et al. 2008; Huang et al. 2015; Scott et al. 2020). Theoretically, MAGIC populations have the potential to dissect the genetic bases of complex traits at sub-centimorgan scale, allowing to overcome common issues related to the use of biparental families for QTL mapping and detection such as low-resolution power, low genetic diversity of parents and limited number of recombination events (Valdar, Flint, and Mott 2006). In cereal crops, MAGIC populations have been developed and established for rice (Bandillo et al. 2013; Ponce, Ye, and Zhao 2018), bread wheat (Mackay et al. 2014; Sannemann et al. 2018; Stadlmeier et al. 2018), maize (Dell'Acqua et al. 2015; Jiménez-Galindo et al. 2019) and barley (Mathew et al. 2018) and to date they have been deployed for unravelling the genetic bases of biotic and abiotic stresses, grain yield (GY) and seed quality traits. Beyond the aforementioned applications, barley MAGIC populations have been recently exploited to disentangle the effect of epistasis on flowering time (Afsharyan et al. 2020; Mathew et al. 2018; Sannemann et al. 2018).

Similarly to MAGIC, the theory underlying genomic prediction (GP) was originally developed and deployed in animal species. The pivotal component of GP is a population of individuals having phenotypic and genotypic information, which is known as training population (TP) and is used to regress genome-wide single nucleotide polymorphisms (SNPs) or other types of DNA markers on phenotypes to simultaneously predict their effects (Meuwissen et al. 2001), that is for training GP models. Trained GP models are subsequently used in combination with the genotypic information of candidate individuals that must be selected for computing their genomic estimated breeding values (GEBVs) and ranking them to apply truncation selection (Heffner, Sorrells, and Jannink 2009; Meuwissen et al. 2001). This latter population of candidate individuals having only genotypic information is known as breeding population (BP) (Heffner et al. 2009; Meuwissen et al. 2001). To date, GP has been largely applied for crop improvement fitting GP models trained with individuals from either biparental families or diversity panels of mostly unrelated accessions. As the genetic relatedness of TP and BP affects the prediction ability of GP models (Ben Hassen et al. 2018; Norman et al. 2018), these two approaches have profound differences in terms of versatility as DNA marker effects estimated on diversity panels have the potential of a broader applicability and might be used in different breeding programs (Bassi et al. 2016), while GP models trained with individuals of

biparental families can allow to accurately predict the performance of offspring produced within the same cross.

Typically, GP models require to regress a number of predictors (DNA markers) that greatly exceeds the number of observations or phenotypes and several parametric and non-parametric models have been proposed to deal with overfitting and the 'large *p*, small *n*' problem (Jannink, Lorenz, and Iwata 2010; Meuwissen et al. 2001; Pérez and de los Campos 2014) as in these conditions the estimation of marker effects using ordinary least squares method is not practicable. A commonly used solution is to estimate marker effects jointly using the Least Absolute Shrinkage and Selection Operator (LASSO) method (Tishbirani 1996) and its Bayesian counterpart (Bayesian Lasso or BL), which uses a penalizing or regularization parameter (λ) that denotes the amount of shrinkage for regressing markers (De Los Campos et al. 2009). Other popular whole genome regression methods based on Bayesian theory are BayesA and BayesB (Meuwissen et al. 2001), which relax the assumption of common variance across marker effects adopted in other models (e.g. ridge regression) and allow each marker to have its own variance. Differently to BayesA, BayesB allows having markers with no effects in the model and theoretically assumes more realistic conditions as it is plausible that a large fraction of genome-wide markers does not contribute to explaining the observed phenotypic variance. Beyond these methods, whole genome regression based on reproducing kernel Hilbert space (RKHS) has been proposed and applied to implement GP models (Gianola and Kaam 2008; Gota and Gianola 2014). In the RKHS regression, a reproducing kernel, that is any positive definite function for mapping from pairs of points in input space to other pairs of points, is used to transform DNA markers of individuals in square distance matrix that are used in a linear model (Gota and Gianola 2014). The Gaussian Kernel (GK) is one of the most common function used as reproducing kernel and depends on the bandwidth (or smoothing) parameter $h$ that controls the decay rate of the kernel as two points step away. Several studies have shown that the use of GK in combination with RKHS improves the prediction of genetic values if the bandwidth parameter $h$ is correctly chosen (Pérez-Elizalde et al. 2015). Moreover as RKHS regression does not assume linearity, this model might allow to better capture nonadditive effects without explicitly including epistatic interactions and dominance in GP models (Gianola and Kaam 2008). Differently from methods based on whole genome regression of markers, the genomic best linear unbiased prediction (GBLUP) method treats genomic values of individuals as random effects in a linear mixed model and uses a genomic relationship matrix based on DNA marker data to compute GEBVs (VanRaden 2008; Wang et al. 2018). Notably, the use of RKHS along with the genomic relationship matrix is equivalent to the mixed linear model of GBLUP, that is GBLUP method represents a special case of RKHS regression (Gota and Gianola 2014).

The effectiveness of GP depends, among other factors, on the degree of correlation between GEBVs and true genetic values, that is the predictive ability of the model. In practice, the predictive ability is evaluated using the Pearson's correlation coefficient between GEBVs and the realized phenotypes or other estimators (e.g. adjusted means). To date several empirical studies have been conducted for fitting GP models on biparental populations and panels of mostly unrelated accessions across different species and traits, which point out that, depending on the genetic architecture of the trait, each statistical model has its own advantages and disadvantages in term of predictive ability and estimation of marker effects (Ben Hassen et al. 2018; Heslot et al. 2012). Other factors that strongly influence the predictive ability are the size of the TP, its structure, and its relatedness with the BP (Desta and Ortiz 2014). Several targeted and untargeted methods have been developed to optimize the composition of TP for maximizing the predictive ability for a given set of individuals (Akdemir, Sanchez, and Jannink 2015; Rincent et al. 2012). Nevertheless, these methods generally generate trait-dependent TPs which might hamper the implementation of these procedures in real breeding programs.

The first objective of the present study was to create a new barley MAGIC population using a diverse founder set of old and new 6-rowed, winter cultivars showing contrasting GY, which was examined across an ample range of site-by-season combinations characterized by different temperature and precipitation patterns. The second objective of this study was to combine data collected across these field trials with genotypic information to fit different single-environment genomic prediction (SE-GP) and multi environment genomic prediction (ME-GP) models for empirically assessing the predictive ability in multi-parent populations. Moreover, we applied different untargeted optimization methods to this MAGIC population for assembling and benchmarking the performance of optimized TPs. Fitting SE-GP and ME-GP models to MAGIC lines, we aimed at broadening the use of these experimental populations beyond classical QTL mapping and analysis of epistatic effects for sustaining and accelerating barley breeding.

## 3 Materials and Methods

### 3.1 Development of the barley MAGIC population

The MAGIC population used in this study was developed using a founder set of eight 6-rowed barley genotypes with a winter growth habit, which were selected on the basis of their pedigrees and similarity in days-to-heading (DH) (**Table 1**). At the first stage of MAGIC development, four $F_1$ populations were created crossing one of the four old 6-rowed barley varieties (Hatif de Grignon, Dea, Robur and Athene) with one of the four 6-rowed modern barley varieties (Ponente, Ketos, Aldebaran and Fridericus). At the second stage of MAGIC development, half-diallel crosses of these four $F_1$ individuals were carried out to generate 6 sets of plants. Finally, these 6 sets of genotypes, each of which contained the alleles of four out eight founder parents, were appropriately crossed in predefined funnel schemes to combine the genome of the eight founders in single lines. Differently from the original crossing schemes developed for constructing MAGIC populations (Cavanagh et al. 2008), instead of recursively self-fertilizing these plants for several generations, seeds of the eight-way inter-crosses were sent to an external lab (SAATEN-UNION GmbH, Germany) to generate 352 inbred MAGIC lines using doubled haploid technology.

**Table 1: Founder set of barley varieties that were intermated for creating the barley MAGIC population.** For each genotype of the founder set, the adjusted means of days to heading (DH), plant height (PH) and grain yield (GY) scored in 8 different trials were reported along with available pedigree information.

| Genotype | Year of release | Country of release | Pedigree | DH (days) | PH (cm) | GY (t/ha) |
|---|---|---|---|---|---|---|
| **Hatif de Grignon** | 1937 | France | Selection from French landraces | 208.3 | 95.9 | 4.1 |
| **Dea** | 1953 | Germany | ((Ragusa x Peragis12) x (Heils Franken x Frw.Berg)) x ((Ragusa x Mahnd.Viktoria) (Ragusa x Bolivia)) | 212.1 | 95.3 | 6.0 |
| **Robur** | 1973 | France | Ager x ( Hatif de Grignon x Ares) | 208.3 | 78.8 | 6.3 |
| **Athene** | 1977 | Germany | (Herfodia x Hord.sp.nigrum H204) x (Madru x Weissenhaus-Stamm) | 211.5 | 94.0 | 6.0 |
| **Ponente** | 2001 | Italy | (Vetulkio x Arma) x Express | 209.7 | 85.0 | 6.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Ketos** | 2002 | France | (Gotic x Orblonde) x (12813 x 91H595) | 208.6 | 81.9 | 6.8 |
| **Aldebaran** | 2003 | Italy | Rebelle x Jaidor | 208.5 | 83.0 | 7.2 |
| **Fridericus** | 2006 | Germany | Carola x LP 6-564 | 211.6 | 89.3 | 7.3 |

### 3.2 Field trials and plant phenotyping

The MAGIC population of 352 inbred individuals and the eight founder parents (**Table 1**) were sown during the fall of two consecutive growing seasons (2015-2016 and 2016-2017) in Fiorenzuola d'Arda (Italy) at CREA-Centro di Genomica e Bioinformatica (44°55'39.0"N 9°53'40.6"E, 78 m above sea level), using an alpha-lattice design with two-replicates. The whole set of MAGIC and the founder parents were also sown during the fall of 2015-2016 growing season in Marchouch (Morocco) at the Experimental station (33°36'43.5"N 6°42'53.0"W, 390 m above sea level) of the "International Center for Agricultural Research in the Dry Areas" using the same experimental design. Similarly, the subset of 82 MAGIC lines included in the optimized TP (TP-Diverse) and the eight founder parents were sown during the fall in 2017-2018 and 2018-2019 growing seasons in Fiorenzuola d'Arda under two different levels of nitrogen fertilization using alpha lattice experimental designs with two replicates. Trials conducted under ideal nitrogen conditions were fertilized with 100 kg/ha of nitrogen applied in two doses: 50 kg/ha were used at the sowing and 50 kg/ha were applied at the stem elongation stage. Field trials conducted under low nitrogen conditions received 50 kg/ha of nitrogen, 25 of which were applied at sowing while the remaining amount was applied at the stem elongation stage. In the growing season 2018-2019, other two field trials were conducted in Konya (Turkey) (37°53'37.9"N 32°37'26.0"E, 1005 m above sea level) and in Adana (Turkey) (36°59'52.9"N 35°20'28.0"E, 24 m above sea level) to phenotype the optimized TP (TP-Diverse) using the same experimental design. For each trial considered in this study, plots of 3 square meters and a sowing density of 350 seeds per square meter were adopted, respectively. Local check cultivars were included as internal checks in all experiments to compare phenotypes with trait observations collected in past seasons. Common protocols were adopted for each trial to phenotype plant genotypes for GY and DH. Phenotyping of MAGIC lines for GY was conducted as follows: from each plot grains were collected using a combine harvester and the total grain weight recorded in each plot was converted in tonnes per hectare. DH was measured as the number of days between sowing date and the date of heading stage, which was defined when 50% of the plants in a plot were at Zadoks' 55 growth stage (Zadoks et al., 1974). For each trial, phenotypic data of GY used in GP models were centred by subtracting the overall mean and standardized dividing by the sample standard deviation.

### 3.3 Statistical models for computing the adjusted means of GY

The adjusted means of GY were computed in each site-by-season combination and across environments including DH as fixed covariate using the approach described in Emrich et al., 2008. The resulting model for computing the adjusted means of GY collected in field trials organized according to alpha-lattice design was:

$$\mathbf{y_{ijk}} = \mathbf{1}\mu + \mathbf{Rep_i} + \mathbf{Block_j(Rep_i)} + \mathbf{Gen_k} + \mathbf{DH_k} + \mathbf{e_{ijk}} \qquad \textbf{Equation 1}$$

where $y_{ijk}$ is the response variable, that is the raw GY, $\mu$ is the general mean, $Rep_i$ is the effect of the $i^{th}$ replicate, $Block_j(Rep_i)$ is the effect of the $j^{th}$ incomplete block within the $i^{th}$ replicate, $Gen_k$ is the random effect of the $k^{th}$ genotype and DH is the effect of "Days-to-heading" covariate measured

in each plot. In this model it is supposed that the random effects of $\mathrm{Gen}_k$ follow a normal distribution with mean 0 and variance $\sigma_g^2$, that is $\mathrm{Gen}_k \sim NIID(0, \sigma_g^2)$, and similarly, the residual terms $\mathrm{e}_{ijk}$ are normally distributed with mean 0 and variance equals to $\sigma^2$, that is $\mathrm{e}_{ijk} \sim NIID(0, \sigma^2)$. The adjusted GY values obtained predicting the random terms $\mathrm{Gen}_k$ from the aforementioned model were used as phenotypes for training GP models. The linear mixed model reported in **Equation 1** was fitted for each site-by-season combination using R 3.6.2 statistical environment and lme4 package (Bates et al. 2015) and variance components of fitted models were used to compute broad sense heritability ($H^2$) of GY.

### 3.4 Genotyping of genetic materials

DNA was extracted from plant leaves using the Macherey Nagel Plant II extraction kit (Macherey Nagel, Dueren,Germany) and analyzed using gel electrophoresis and Quant-iT™ PicoGreen™ dsDNA Assay Kit (ThermoFisher,Grand Island NY, USA) following manufacturer's instructions to assess quality and concentration, respectively. DNA samples were shipped to a propel-certified service provider (Trait Genetics GmbH, Gatersleben, Germany) and fingerprinted using the Illumina Infinium technology along with the Barley 50k iSelect SNP Array (Bayer et al. 2017). To update the physical positions of SNP markers interrogated with the Barley 50k iSelect SNP Array, probe sets used to design this array were mapped against the new reference sequence of barley (Monat et al. 2019). The raw genotyping table was imported in R software using "synbreed" package (Wimmer et al. 2012) to filter out markers with more than 10% of missing data and impute remaining missing data using Beagle 4.1 (Browning and Browning 2016). 20 random leaf samples from field trials organized in Adana and Marchouch were genotyped using Illumina Infinium technology and Barley 50k iSelect SNP Array to assess whether mislabelling of genotypes occurred during phenotyping operations and data collection.

### 3.5 Clustering and linkage disequilibrium analyses of the MAGIC population

Principal component analysis was used to assess the diversity of the whole MAGIC population and was carried on imputed SNP data of the 352 MAGIC lines and the eight founders using ade4 package along with R version 3.6.2 (Team R. Core. 2019; Thioulouse et al. 2018). The first two principal components were used to visualize the dispersion of MAGIC lines in a graph. Linkage disequilibrium between pairs of markers was measured using $r^2$ (Hill and Robertson 2008) in the subset of MAGIC genotypes included in the optimized TP and computed using Plink 1.9 software (Chang et al. 2015; Purcell et al. 2007). $r^2$ values showing *p*-values above 0.001 were filtered out, while the remaining pairwise $r^2$ values were imported and examined with a custom script developed for R 3.6.2 (Team R. Core. 2019) to compute the mean $r^2$ in 100 kb windows, which was plotted in R 3.6.2 using ggplot2 package (Wickham 2016).

### 3.6 Statistical models used for fitting SE-GP

SE-GP models were fitted using BayesA, BayesB and BL models (Meuwissen et al. 2001; Park and Casella 2008; Tishbirani 1996). Moreover, RKHS regression models were fitted using a linear GBLUP kernel (GB) and a non-linear GK (Gianola and Kaam 2008; Gota and Gianola 2014). For the GK, that is $K(x_i, x_{i'}) = e^{-\left(h * d_{ii'}^2\right)}$, where $d_{ii'}^2$ points out the squared Euclidean distance between individuals $i$ and $i'$, the rate of decay imposed by the bandwidth parameter $h$, was estimated using an empirical Bayesian methodology (Pérez-Elizalde et al. 2015) modifying published R codes (Cuevas et al. 2016).

### 3.7 Statistical models used for fitting ME-GP

Beyond SE-GP models, the adjusted means of GY computed across different site-by-season combinations were fitted to three previously described ME-GP models. Following the model nomenclature reported in Bandeira e Sousa et al. ( 2017), these three models were indicated in this study as "multi-environment, main genotypic effect" (MM) model (Bandeira e Sousa et al. 2017; Jarquín et al. 2014; Lopez-Cruz et al. 2015), "multi-environment, single variance GxE deviation model" (MDs) (Bandeira e Sousa et al. 2017; Jarquín et al. 2014) and the "multi-environment, environment-specific variance GxE deviation model" (MDe) (Bandeira e Sousa et al. 2017; Lopez-Cruz et al. 2015). Site-by-season combinations were considered as environments in MM, MDs and MDe regression models, which are briefly defined and summarized as follows. In the MM model, environments were considered as fixed effects while the random genetic effects were considered constant across all environments without modelling marker x environment interactions. Following matrix notation, the MM regression model is defined as follows:

$$\mathbf{y} = \mathbf{1}\boldsymbol{\mu} + \mathbf{Z}_e\boldsymbol{\beta}_e + \mathbf{Z}_u\mathbf{u} + \boldsymbol{\varepsilon} \qquad \textbf{Equation 2}$$

where y is the vector of observations collected in all environments, μ is the overall mean, $Z_e$ is the incidence matrix that connects observed phenotypes to the environments in which they were measured, $\beta_e$ is the vector of environmental fixed effects that must be estimated, $Z_u$ is an incidence matrix connecting genotypes with phenotypes for each environment, u is the vector of random genetic effects that must be predicted while ε is a vector of model residuals. In this model, marker genetic effects are assumed as $u \sim N(0, \sigma^2_{\mu 0} K)$, that is, they follow a multivariate normal distribution with mean and variance-covariance matrix equal to zero and $\sigma^2_{\mu 0} K$, respectively. The term $\sigma^2_{\mu 0}$ of the variance-covariance matrix is the variance of additive genetic effects across environments, while $K$ can be either a genomic relationship matrix (VanRaden 2008) or a kernel function as discussed below. Model residuals of the vector ε are assumed to be independent and normally distributed with null mean and variance equal to $\sigma^2_e$, that is $\varepsilon \sim N(0, I\sigma^2_e)$, where $I$ points out the identity matrix. Overall, the MM regression model estimates marker effects across all environments and does not split them in main marker effects and in environmental-specific effects as in MDs and MDe models. As already substantiated in Lopez-Cruz et al. (2015), for balanced field trial designs, MM is equivalent to fitting a genomic regression model using the average performance of each line across environments as phenotype.

Differently from the MM model, the MDe model allows markers to assume different effects in each $j^{th}$ environment (Bandeira e Sousa et al. 2017; Lopez-Cruz et al. 2015), and consequently allows to account for marker x environment interactions. This model assumes that the effects of the $j^{th}$ environments, and the effects of markers are separated into two components, which are the main effect of markers for all environments, names as $b_{0k}$, and the peculiar random effect $b_{ik}$, of the markers in each $j^{th}$ environment, that is the effects of marker x environment interactions (Lopez-Cruz et al. 2015). Consequently, in MDe models, the effect of the $k^{th}$ marker on the $j^{th}$ environment ($\beta_{jk}$) is described as the sum of an effect common to all environments ($b_{0k}$), plus a random deviation ($b_{ik}$) peculiar to the $j^{th}$ environment, that is $\beta_{jk} = b_{0k} + b_{ik}$. Following matrix notation, the MDe regression model is defined as follows:

$$\mathbf{y} = \mathbf{1}\boldsymbol{\mu} + \mathbf{Z}_e\boldsymbol{\beta}_e + \mathbf{Z}_u\mathbf{u}_o + \mathbf{u}_E + \boldsymbol{\varepsilon} \qquad \textbf{Equation 3}$$

where μ, $Z_e$, $\beta_e$ have the same meaning of the MM regression model, $u_o$ represents the main effect of markers across all environments with a variance–covariance structure similar to MM model, that is, $u \sim N(0, \sigma^2_{\mu 0} K)$. As pointed out by López-Cruz et al. (2015) $\sigma^2_{\mu 0}$ is common to all environments, and the borrowing of information among environments is generated through the kernel matrix $K$. $u_E$ points out the specific effects of marker x environment interactions, which follow a multi-variate

normal distribution with null mean and a variance–covariance matrix $K_E$, that is, $u_E \sim N(0, K_E)$. For $j$ environments, the variance-covariance matrix $K_E$ is defined as follows:

$$K_e = \begin{bmatrix} \sigma^2_{\mu E1} K_1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \sigma^2_{\mu Em} K_m & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & \sigma^2_{\mu Ej} K_j \end{bmatrix}$$

As explained in Bandeira e Sousa et al. (2017), $K_E$ can be discomposed as a sum of $j$ matrices, one for each $j$ environment. Consequently, the interaction term $u_E$ can be decomposed in $j$ environmental specific effects to transform **Equation 3** as follows:

$$\boldsymbol{y = 1\mu + Z_e\beta_e + Z_u u_0 + u_{E1} + u_{E2} + u_{E3} + \ldots + u_{Ej} + \varepsilon} \qquad \textbf{Equation 4}$$

where each interaction effect $u_{Ej}$ has a normal distribution with null mean and a variance-covariance structure $\sigma^2_{\mu Ej} K_j$.

Starting from the MM regression model, the MDs model adds the random interaction effect of the environments with the genetic information of the lines pointed out with ue. Following matrix notation, the MDs modes is described as follows:

$$\boldsymbol{y = 1\mu + Z_e\beta_e + Z_u u + ue + \varepsilon} \qquad \textbf{Equation 5}$$

where μ, $Z_e$, $\beta_e$, $Z_u$, u and ε have the same meaning of the MM regression model. As substantiated in Jarquín et al. (2014) the interaction term ue has a multi-variate normal distribution with null mean and variance-covariance matrix equal to $[Z_u K Z_u']°[Z_E Z_e']$, where the Haddamar product operator denotes the element to element product between the two matrices in the same order.

In the present study, MM, MDs and MDe regression models were fitted using either the linear GB kernel method (VanRaden 2008) or the non-linear GK method (Bandeira e Sousa et al. 2017). For the linear GB kernel method, the matrix $K$ of the aforementioned models was the genomic relationship matrix and was computed as $K = \left(\frac{XX'}{p}\right)$ (VanRaden 2008), where X is the standardized matrix of molecular markers for the individuals, of order $n$ by $p$; where $n$ and $p$ are the number of observations and the number of markers, respectively. For GK method, the matrix $K$ of MM, MDs and MDe regression models was computed as $K_j(x_{ij}, x_{ij'}) = e^{-\left(h_j * d^2_{ii'}\right)}$ where $d^2_{ii'}$ is the squared Euclidean distance of the markers genotypes in individuals $i$ and $i'$ for the $j^{th}$ environment. Similarly to SE-GP models, the bandwidth parameter $h$ was computed using an empirical Bayes method (Cuevas et al. 2016; Pérez-Elizalde et al. 2015).

MM, MDs and MDe regression models used in this study were fitted using BGLR package 1.08 (Pérez and de los Campos 2014) in R 3.6.2 statistical environment, adapting scripts provided in the framework of other studies (Bandeira e Sousa et al. 2017). For each model implemented in this study, predictions were based on 500,000 iterations collected after discarding 10,000 iterations for burn-in period-and using a thinning interval of 5 iterations. Trace plots for each of the variance parameters were created to assess whether the number of burn-in iterations was sufficient.

### 3.8 Optimization of the TPs

In this study three different untargeted optimization criteria based on coefficient of determination (Laloe 1993), predictive error variance (Rincent et al. 2012) and rScore (Ou and Liao 2019) were used to assemble three corresponding TPs, each of which groups a set of 90 MAGIC individuals. The R package TSDFGS (Ou and Liao 2019) was used to assemble these three optimized TPs using the

aforementioned criteria. A fourth empirical untargeted optimization criterion was adopted for assembling another TP from the whole MAGIC population and aimed at maximizing the average distance between each selected accession and the closest other line using the modified Roger's distance (Thachuk et al. 2009). This criterion was implemented in R 3.6.2 using the heuristic algorithm implemented in the package Core Hunter3 (De Beukelaer, Davenport, and Fack 2018) and was used to select a subset of 82 out 352 MAGIC individuals along with the eight MAGIC founder parents.

### 3.9 Cross validation schemes

In this study several cross-validation (CV) schemes were adopted for estimating the predictive ability of GP models along with their standard errors (Burgueño et al. 2012; Gianola and Schon 2016). For estimating the predictive ability of SE-GP models implemented with BayesA, BayesB, Bayesian Lasso, GB and RKHS with GK, cross validation was carried out using 100 repeated random partitioning of MAGIC population into training and validation sets. Using increasingly larger TPs of 80, 90, 100, 110, 120, 130, 140, 150 and 160 individuals, CV schemes were applied to compute mean and standard deviation of predictive ability for each TP size. Totally 4,500 models were fitted to carry out this CV experiment, combining the five statistical models with the aforementioned dimensions of the TP and 100 repeated random partitioning of MAGIC in training and validation sets.

CV of SE-GP models fitted using optimized TPs was carried out using the standard leave-one-out (LOO) strategy to estimate their predictive ability (Gianola and Schon 2016). Basically, using LOO strategy, N GP models are fitted using N-1 individuals excluding recursively one individual from the TP and the GEBV of the excluded line is predicted from a model trained using all other lines. In our LOO experiment, this was carried out separately for each group of 90 lines included in the optimized TPs, and the accuracy of these predictions was calculated as the Pearson's correlation coefficient between GEBVs and the corresponding adjusted means of GY.

The predictive ability of ME-GP models was assessed using cross-validation 1 (CV1) and cross-validation 2 (CV2) schemes (Burgueño et al. 2012), assigning 90% of lines to the training set and the remaining 10% to the validation set. In both CV schemes, all the parameters of the MM, MDs and MDe regression models were recursively re-estimated in each of 100 random partitions. For each random partitioning, models were fitted using genotypes included in the training sets and the predictive ability was computed as the Pearson's correlation coefficient between GEBVs and the corresponding adjusted means of GY. Overall, 100 Pearson's correlations were computed for each model and the mean and standard deviation of these values were computed to estimate the predictive ability of GP models.

## 4 Results

### 4.1 Development of the barley MAGIC population

The barley genotypes included in the founder set of MAGIC were examined in field trials organized in height site-by-season combinations in Italy, Germany and Scotland (Xu et al. 2018) for assessing the diversity of European cultivars for GY, plant height and DH. These field trials showed that the founder set, which includes four elite and four old barley varieties with different genetic background, exhibits limited variation of DH values (**Table 1**). Following a modified version of the standard crossing design (Huang et al. 2015), this founder set was intermated to create an eight-way MAGIC population of 352 individuals, which were subsequently genotyped to assess the contribution of each founder parent to the mosaic genome of each line.

### 4.2 Estimating the predictive ability of GP models as a function of TP size

In GP models, the variation of predictive ability as a function of the TP size has been empirically investigated on segregating families and in collections of mostly unrelated accessions (Norman et al. 2018). Here, we investigated the relationship between TP size and the predictive ability of different GP statistical models fitted to the barley MAGIC population. To carry out this analysis, the whole panel of 352 MAGIC lines and the founder parents were genotyped using the Barley 50k iSelect SNP Array (Bayer et al. 2017). SNPs with more than 10% of missing data were discarded, while the remaining missing genotypes were imputed using the algorithm implemented in BEAGLE (Browning and Browning 2016). This procedure allowed to identify 19,723 polymorphic SNPs, which were combined to the adjusted means (BLUPs) of GY computed in three site-by-season combinations (**Table 2**) to fit and cross-validate SE-GP models. Overall, five different whole genome regression methods based on BayesA, BayesB, BL, GB and RKHS fitted with the non-linear GK (Crossa et al. 2017; Cuevas et al. 2016; Gianola and Kaam 2008; Gota and Gianola 2014) were compared.

**Table 2: Field trials carried out for phenotyping the whole MAGIC population and the founder set for GY.**

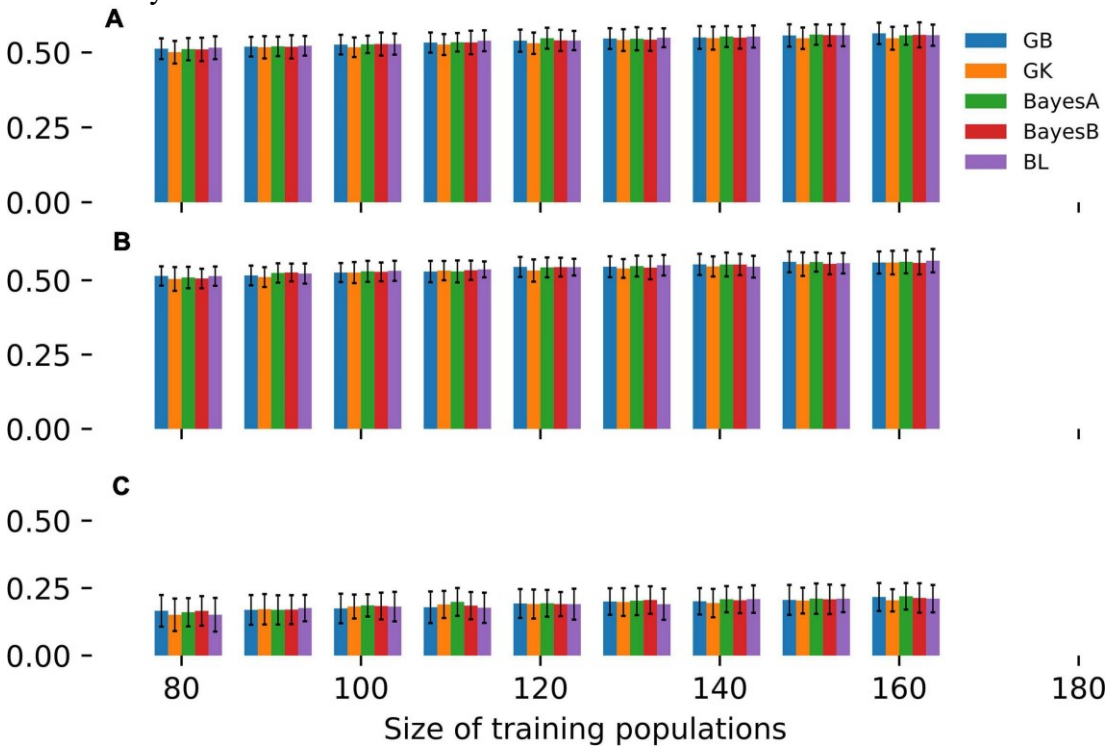
| Acronym | Site | Country | Growing season | Populations | Traits |
|---|---|---|---|---|---|
| **Fio16IN** | Fiorenzuola d'Arda | Italy | 2015-2016 | 352 MAGIC and the founder set | DH, GY |
| **Fio17IN** | Fiorenzuola d'Arda | Italy | 2016-2017 | 352 MAGIC and the founder set | DH, GY |
| **Mar16IN** | Marchouch | Morocco | 2015-2016 | 352 MAGIC and founder set | DH, GY |

These aforementioned SE-GP models were fitted to the MAGIC population and cross-validated for estimating the trend of predictive ability as a function of TP size (**Figure 1**). Specifically, CV was implemented randomly partitioning 100 times the whole panel of MAGIC lines in a TP and in a validating population (VP). Overall, nine different CV experiments were carried out, using TP sizes of 80, 90, 100, 110, 120, 130, 140, 150 and 160 MAGIC lines and the remaining genotypes as VPs (**Figure 1**). The CV of these GP models points out that in the three site-by-season combinations **Table 2**), GB, GK, BayesA, BayesB and BL show comparable predictive abilities across the entire range of TP sizes considered (**Figure 1**).

Moreover, these CV experiments point out that in temperate locations (Fio16IN, Fio17IN, **Table 2**), the predictive ability of SE-GP models exceeds 0.50 even using TPs of 80 or 90 individuals (**Figure 1**), while in the harsh and pre-desertic environment of Mar16IN (**Table 2**), it does not exceed 0.25 and shows larger standard deviation. Varying the size of TPs from 80 to 160 individuals slightly increases the values of predictive ability for GY in the remaining individuals of the MAGIC population (**Figure 1**; **Supplementary Table 1**) as already substantiated in other GP models fitted using collection of mostly unrelated genotypes (Norman et al. 2018).

Overall, this empirical analysis shows that 80 or 90 MAGIC individuals are sufficient to fit SE-GP models yielding high values of predictive ability and that larger TPs do not significantly improve the predictive ability of GP models either in temperate or stressful environments (**Figure 1**; **Supplementary Table 1**).

**Figure 1: CV of different SE-GP models fitted to GY measured in the MAGIC population.** Bars report the values of predictive ability for GY computed in A) Fio16IN, B) Fio17IN and C) Mar16IN. Bars of different colors point out values of predictive ability computed using GB, GK, BayesA, BayesB and BL models as a function of TP sizes, while the error bars point out the standard deviation of predictive ability values.

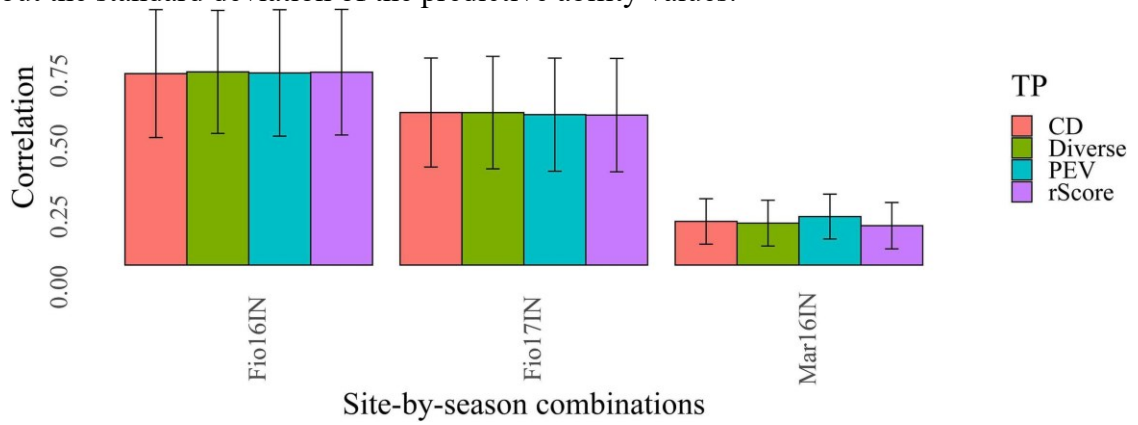



### 4.3 Designing optimized TPs of MAGIC

The predictive ability of GP models fitted in collection of mostly unrelated accessions and in biparental populations depends on the size of TP, the genome distribution and number of molecular markers used for whole genome regression, the genetic composition of TP and its genetic relationship with the BP (Berro et al. 2019; Desta and Ortiz 2014; Heffner et al. 2009; Jannink et al. 2010). Particularly, it was assessed that using a large reference panel of accessions, the predictive ability of GP models can be improved increasing the diversity of the TPs (Norman et al. 2018). Along with these empirical findings, several statistical criteria and algorithms have been proposed to optimize TPs for maximizing predictive ability using reference panels of accessions or sets of advanced lines (Akdemir et al. 2015; Berro et al. 2019; Ou and Liao 2019).

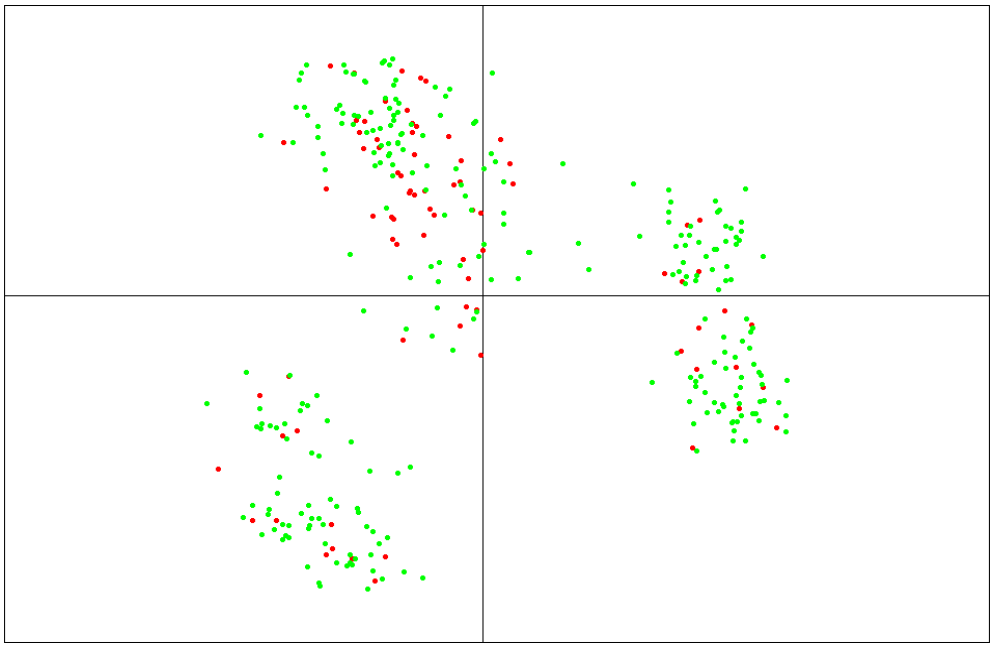
Here, we examined three different untargeted optimization criteria based on the coefficient of determination (CD_mean) (Laloe 1993), prediction error variance (PEV) (Rincent et al. 2012) and rScore (Ou and Liao 2019) and benchmarked them against a method that samples a diverse TP from the whole MAGIC population using SNP markers (**Figure 2**). The rationale of this latter method is to maximize the average distance, computed using the modified Roger's method, between each selected accession and the closest other genotype (Thachuk et al. 2009). This criterion, named entry-to-nearest entry was maximized with a heuristic algorithm to construct a highly diverse TP in which all MAGIC lines are maximally different (De Beukelaer et al. 2018). The TP assembled with this latter untargeted optimization criterion, named "TP-Diverse" (**Figure 2**), was constructed using the panel of 19,723 polymorphic SNPs detected in the whole MAGIC population, and was subsequently used as optimized TP and benchmarked to TPs assembled using CD_mean, PEV and rScore optimization methods (**Figure 2**).

**Figure 2: Benchmarking of different methods for optimizing TPs of MAGIC.** Bars of different colors report the values of predictive ability obtained with GP models fitted using CD_mean (CD), prediction error variance (PEV), rScore and Diverse optimization criteria. The error bars of each plot point out the standard deviation of the predictive ability values.



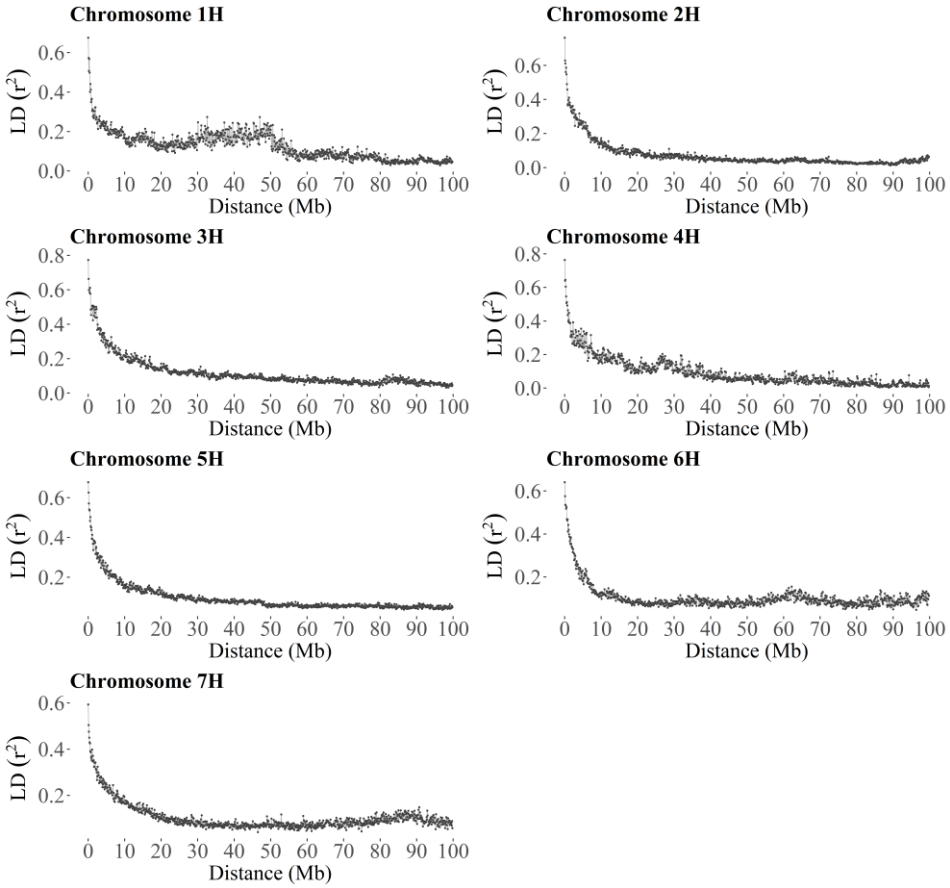

Following this "TP-Diverse" optimization, our procedure led to identify a set of 82 MAGIC lines as the smallest population subset fulfilling the aforementioned criterion, which was used as TP along with the eight founder parents. Overall, when applied to MAGIC populations, the four optimized TPs spawned similar predictive abilities across the three site-by-season combinations (**Figure 2**) and consequently the genetic makeup of this TP was further investigated. The genetic relationships between TP-Diverse and the remaining MAGIC lines were assessed conducting a principal component analysis (PCA) on genetic data, which pointed out that the first two principal components explain 22.3 and 5.5 percent of the total genetic variability of the MAGIC population, respectively (**Figure 3**). PCA shows three main clusters of MAGIC lines and corroborates those individuals included in the TP-Diverse are representative of the whole diversity of MAGIC lines (**Figure 3;** red points).

**Figure 3: Principal component analysis (PCA) of the MAGIC population based on 19.723 SNPs.** The first two axes of PCA explain 22.3% and 5.5% of the total variability, respectively. Red points represent the subset of MAGIC lines included in TP-Diverse, while green points represent the remaining MAGIC lines.



In segregating families and collections of mostly unrelated accessions, a large number of molecular markers is often needed to capture the effects of all QTLs or alternatively, strong linkage disequilibrium (LD) between markers and causative variants that control the traits of interest is desirable to achieve high values of predictive ability in GP (Heffner, Jannink, and Sorrells 2011; Lorenzana and Bernardo 2009; Norman et al. 2018). Consequently, the extent of LD was investigated in TP-Diverse to assess its correlation with the predictive ability values of GP models. Firstly, SNP markers of the barley 50K SNP chip used to fingerprint the whole MAGIC population were lifted over to the new barley reference sequence (Monat et al. 2019) and secondly, the average extent of $r^2$ was computed for each barley chromosome. Overall, a large fraction of the 44,040 SNPs of the barley 50k SNP chip were lifted over and 18,248 out 19,723 polymorphic SNPs unambiguously mapped to the reference sequence of barley (**Supplementary Table 2**) were used to estimate the decay of average LD computed in bins of 100kb (**Figure 4**). This analysis indicated that across the seven barley chromosomes $r^2$ decays relatively slowly as SNPs mapped more than 10Mbp apart show $r^2$ values of circa 0.2, while the average $r^2$ values of markers within 1 MB or less exceed 0.4 (**Figure 4).** Considering the average number of markers per chromosome (**Supplementary Table 2**), the levels of LD measured in TP-Diverse are sufficiently high and higher marker densities might not significantly increase the predictive ability of GP models fitted in our MAGIC population of barley as empirically observed in other crops (Norman et al. 2018). Overall, the predictive ability values obtained with GP models fitted with the three optimization methods are substantially equivalent to the prediction accuracy obtained with TP-Diverse (**Figure 2**) and consequently this latter TP was chosen for fitting further single- and multi-environment GP models.

**Figure 4: Extent of the average linkage disequilibrium in TP-Diverse.** For each barley chromosome, each point shows the average $\boldsymbol{r^2}$ computed in 100 kb windows as a function of marker distance.



### 4.4 Using the optimized TP for fitting SE-GP and ME-GP models

Field trials of TP-Diverse were organized in nine site-by-season combinations and phenotypic data for GY and DH were collected using common phenotyping protocols, while the remaining set of MAGIC lines were used in Fio16IN, Fio17IN and Mar16IN as VP (**Table 3**). Alpha-lattice experimental designs were adopted for all field trials and mixed linear models were used to compute adjusted means of GY and broad sense heritability ($H^2$) for each site-by-season combination considering genotypes as random variables (BLUPs) (**Table 3**). This analysis indicated that $H^2$ varies significantly across the nine field trials and spans from 0.805 in Kon19IN to 0.122 in Mar16IN (**Table 3**). The adjusted means of GY were subsequently used as phenotypes for fitting GP models along with genotypic information.

**Table 3: Summary of field trials carried out for phenotyping TP and VP for GY.** For each site-by-season combination, the estimates of broad sense heritability ($H^2$) of GY were reported. $H^2$ was computed for the whole panel of MAGIC lines for Fio16IN, Fio17IN and Mar16IN.

| Acronym | Site | Country | Growing season | Populations | $H^2$ |
|---|---|---|---|---|---|
| **Fio16IN** | Fiorenzuola d'Arda | Italy | 2015-2016 | TP and VP | 0.660 |
| **Fio17IN** | Fiorenzuola d'Arda | Italy | 2016-2017 | TP and VP | 0.472 |
| **Fio18IN** | Fiorenzuola d'Arda | Italy | 2017-2018 | TP | 0.532 |
| **Fio18LN** | Fiorenzuola d'Arda – Low Nitrogen | Italy | 2017-2018 | TP | 0.395 |
| **Fio19IN** | Fiorenzuola d'Arda | Italy | 2018-2019 | TP | 0.652 |
| **Fio19LN** | Fiorenzuola d'Arda – Low Nitrogen | Italy | 2018-2019 | TP | 0.663 |
| **Mar16IN** | Marchouch | Morocco | 2015-2016 | TP and VP | 0.122 |
| **Ada19IN** | Adana | Turkey | 2018-2019 | TP | 0.737 |
| **Kon19IN** | Konya | Turkey | 2018-2019 | TP | 0.805 |

To assess the performance of MAGIC lines included in TP-Diverse, across different locations and years, a pairwise correlation analysis of the adjusted means of GY computed in the nine site-by-season combinations considered in this study was carried out (**Figure 5**). The correlations of GY across environments spanned from -0.030 to 0.553 and, as expected, values were higher between field trials carried out in the same environments but in different years, while lower values were observed among Mar16IN and other site-by-season combinations, corroborating the hypothesis that the climatic peculiarity of this environment imposes higher levels of stress to MAGIC lines (**Figure 5**). Similarly, the adjusted means of GY computed in Fio18LN exhibited lower correlation values with other site-by-season combinations (**Figure 5**). These adjusted means of GY were used to train SE-GP and ME-GP models using "TP-Diverse". For each site-by-season combination, phenotypic and genotypic data were standardized, and nine different SE-GP models were fitted using GB and GK statistical models (**Table 4**). As expected after standardization, for models fitted using GB, the summation of variance components was circa 1 (**Table 4**), while the distribution of the residuals after fitting all GP models to the nine site-by-season combinations was approximately normal. The analysis of variance components of SE-GP models showed that the values of error variance in GK models are lower than those obtained for the corresponding GB models (**Table 4**), and similarly in GK models the values of genetic component variance are always higher than the corresponding quantities computed for GB models (**Table 4**).

**Figure 5: Pairwise correlations of GY obtained in the nine site-by-season combinations for TP and VP.** Numbers reported in black, red, and blue on the upper graph show pairwise Pearson correlations computed between adjusted means of GY for the whole set of lines tested, TP and VP, respectively. The lower graph shows scatter plots of GY adjusted means computed in pairs of site-by-season combinations.

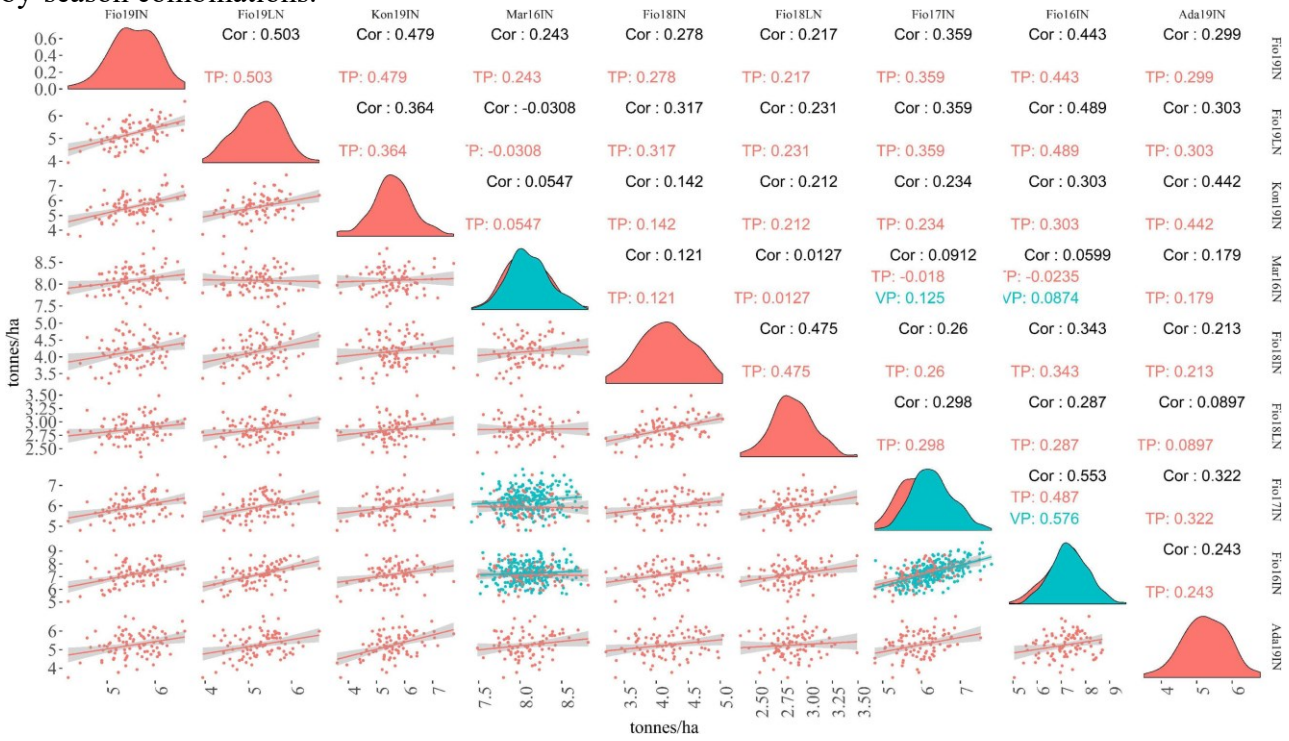



**Table 4: Variance components of SE-GP models fitted using GBLUP (GB) and GK statistical model.** For each site-by-season combination, the estimated variance components of genetic effects and residuals fitted with GB and GK models are reported, while bracketed numbers point out the corresponding standard deviation.

| Site-by-season combination | GB | | GK | |
|---|---|---|---|---|
| | Genetic effect variance | Residual variance | Genetic effect variance | Residual variance |
| **Kon19IN** | 0.586 (0.010) | 0.557 (0.045) | 0.660 (0.016) | 0.489 (0.068) |
| **Mar16IN** | 0.467 (0.089) | 0.719 (0.067) | 0.590 (0.013) | 0.588 (0.078) |
| **Fio18IN** | 0.491 (0.059) | 0.560 (0.029) | 0.632 (0.086) | 0.455 (0.043) |
| **Fio18LN** | 0.412 (0.048) | 0.752 (0.050) | 0.544 (0.000) | 0.611 (0.069) |
| **Fio17IN** | 0.537 (0.072) | 0.480 (0.016) | 0.655 (0.084) | 0.417 (0.041) |
| **Fio16IN** | 0.618 (0.066) | 0.336 (0.094) | 0.680 (0.065) | 0.348 (0.011) |
| **Ada19IN** | 0.561 (0.086) | 0.543 (0.036) | 0.654 (0.019) | 0.498 (0.070) |
| **Fio19IN** | 0.480 (0.079) | 0.651 (0.049) | 0.659 (0.005) | 0.469 (0.054) |
| **Fio19LN** | 0.479 (0.058) | 0.566 (0.024) | 0.632 (0.083) | 0.446 (0.041) |

The adjusted means of GY computed at the nine site-by-season combinations were used to fit ME-GP, particularly three models were fitted, which were named "Multi-environment, main genotypic effect" (MM), "Multi-environment, single variance GxE deviation" (MDs) (Jarquín et al. 2014) and "Multi-environment, environment specific variance GxE deviation" (MDe) (Lopez-Cruz et al. 2015) following recent model nomenclature (Bandeira e Sousa et al. 2017). Similarly to SE-GP models, MM, MDs and MDe models were fitted using GB and GK methods and totally six model method combinations were used to fit multi-environment predictions. The analysis of variance components

showed that for all three models (MM, MDs and MDe), GK methods exhibit lower values of the estimated residual variances pointing out a better model fitting (**Table 5**). Moreover, model comparisons showed that the inclusion of the interaction term (GxE) in MDe model induces a reduction in the estimated residual variance for GY compared to MM models either using GB or GK methods, but MDs models fitted better the data compared to MDe. For the MDe models, the residual variance components of MDe-GK were smaller than those of the MDe-GB, whereas the estimated variance components for the genetic main effect and genetic environment specific effect variances were higher for the GK than for the GB (**Table 5)**.

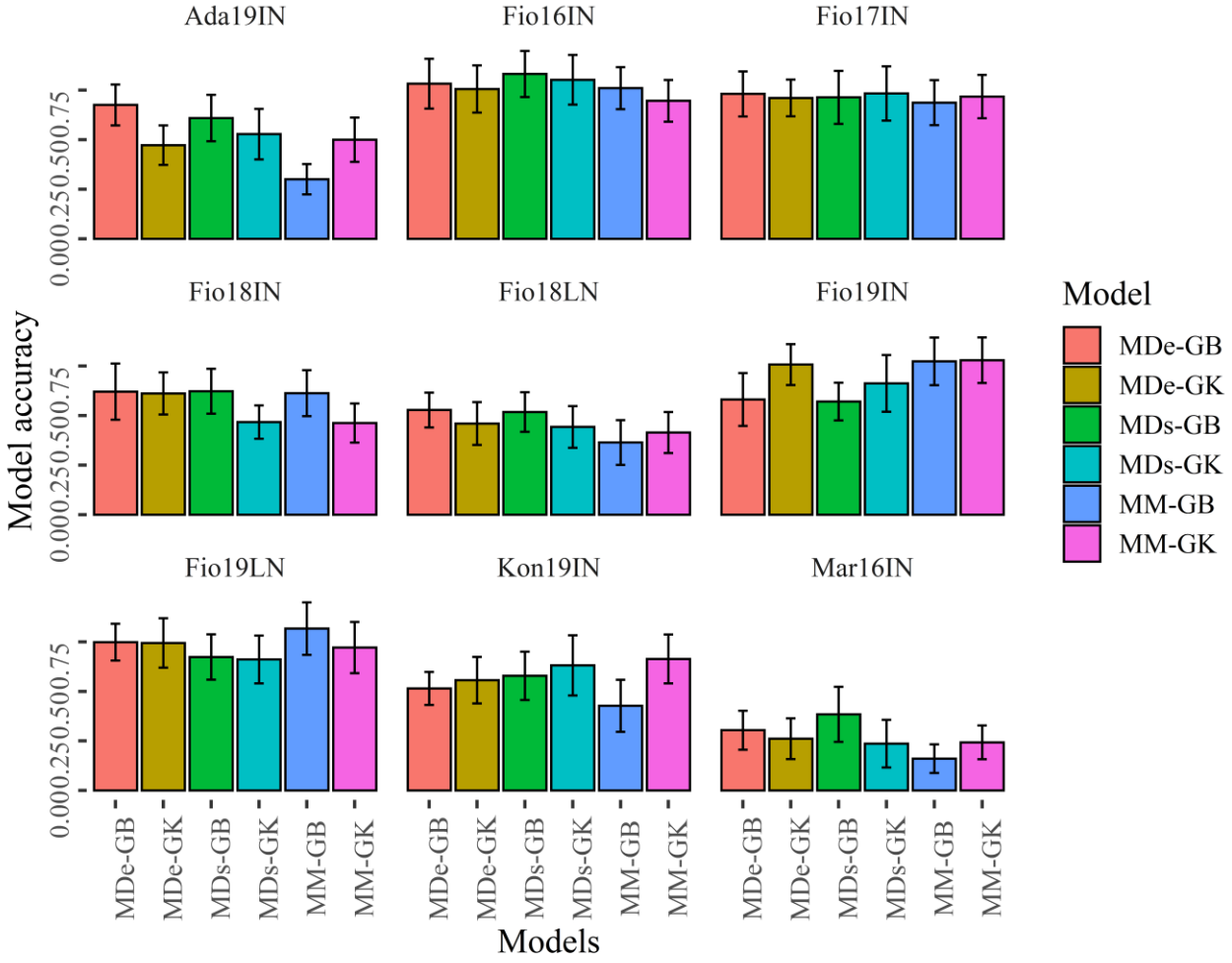
**Table 5: Variance components of ME-GP models fitted using GBLUP (GB) and RKHS along with the Gaussian Kernel (GK) methods.** For each of the three regression models (MM, MDs and MDe), the estimated variance components fitted with GB and GK methods are reported, while bracketed numbers point out the corresponding standard deviation of variance component estimates.

| Component | Environment | GB | GK |
|---|---|---|---|
| **Multi-environment, main genotypic effect (MM) model** | | | |
| Residual ($\sigma_e^2$) | - | 0.758 (0.047) | 0.746 (0.045) |
| Genetic main effect ($\sigma_{\mu 0}^2$) | - | 0.249 (0.069) | 0.373 (0.088) |
| **Multi-environment, single variance GxE deviation (MDs) model** | | | |
| Residual ($\sigma_e^2$) | - | 0.516 (0.056) | 0.389 (0.071) |
| Genetic main effect ($\sigma_{u0}^2$) | - | 0.281 (0.077) | 0.374 (0.089) |
| Genetic interaction effect ($\sigma_{ue}^2$) | - | 0.247 (0.066) | 0.589 (0.140) |
| **Multi-environment, environment specific variance GxE deviation (MDe) model** | | | |
| Residual ($\sigma_e^2$) | - | 0.602 (0.016) | 0.592 (0.018) |
| Genetic main effect ($\sigma_{u0}^2$) | - | 0.292 (0.026) | 0.402 (0.031) |
| Genetic environment specific effect ($\sigma_{uEj}^2$) | Ada19IN | 0.251 (0.054) | 0.353 (0.083) |
| | Fio16IN | 0.035 (0.027) | 0.054 (0.046) |
| | Fio17IN | 0.010 (0.066) | 0.024 (0.023) |
| | Fio18LN | 0.062 (0.054) | 0.116 (0.085) |
| | Fio18IN | 0.007 (0.006) | 0.018 (0.015) |
| | Mar16IN | 0.549 (0.085) | 0.873 (0.122) |
| | Kon19IN | 0.217 (0.050) | 0.312 (0.079) |
| | Fio19LN | 0.008 (0.007) | 0.053 (0.018) |
| | Fio19IN | 0.004 (0.003) | 0.055 (0.011) |

### 4.5 Predictive ability of ME-GP models with GB and GK methods

The predictive ability of MM, MDs and MDe models implemented using GB and GK methods was estimated with cross-validation 1 (CV1) and cross-validation 2 (CV2) schemes using 100 random partitions. For each of the six multi-environment model-method combinations, the values of predictive ability for CV1 and CV2 schemes were obtained for the set of 100 random partitions, which were used to compute the average predictive ability and the associated standard deviation. Overall, CV2 showed that in four site-by-season combinations (Fio16IN, Fio17IN, Fio19IN and Fio19LN) the predictive ability is generally higher and exceed 0.70 for certain ME-GP models, while for Mar16IN the six model-method combinations exhibit, on average, the lowest values of predictive ability as for this site-by-season combination the lowest values of 0.161 and 0.236 were observed for MM-GB and MDs-GK models, respectively (**Figure 6; Supplementary Table 3**).

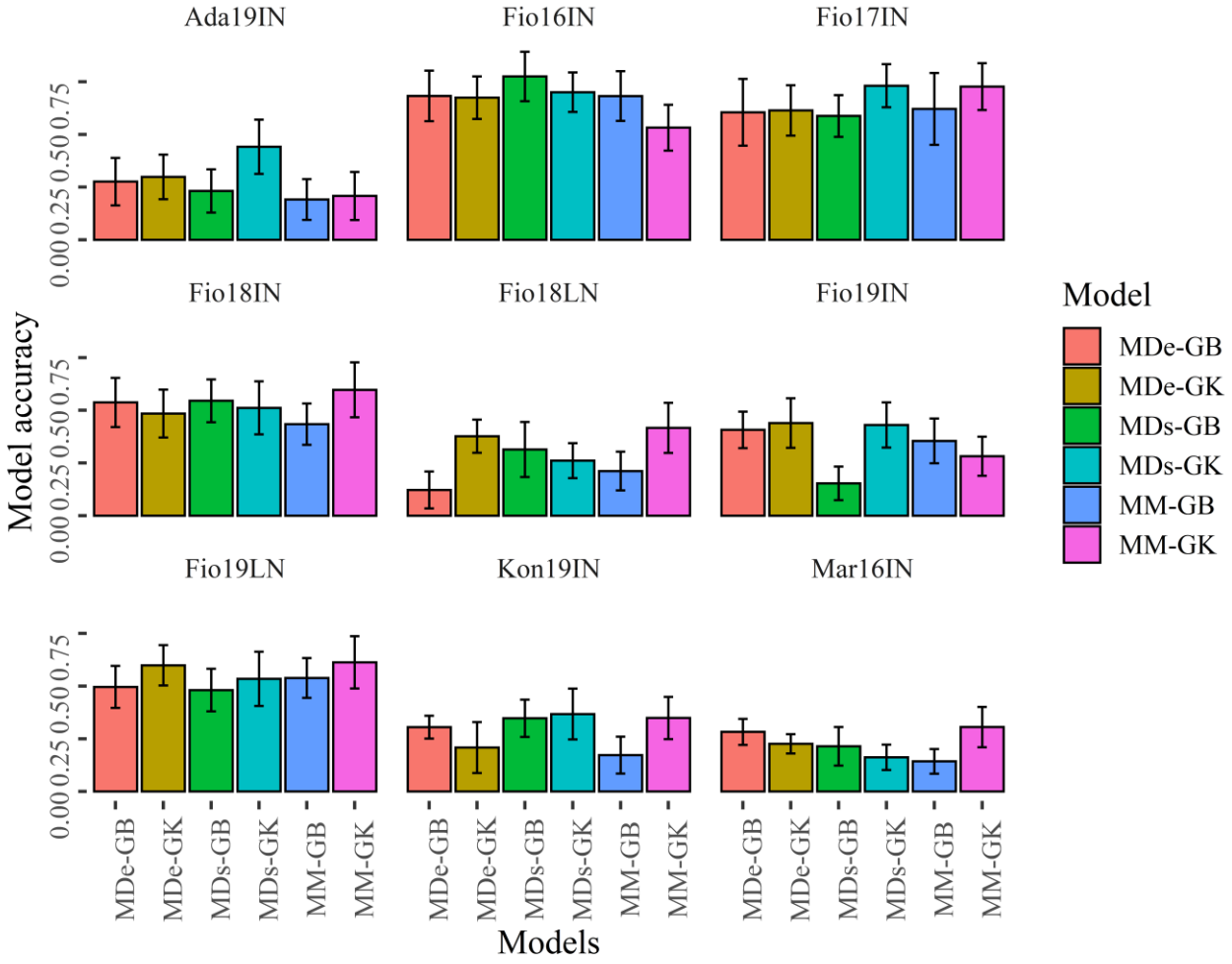
**Figure 6: Bar plots of the predictive ability values obtained with CV2**. Bar plots show the mean correlation between observed and predicted values of GY obtained with 100 random CV2 partitions for MM, MDs and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Error bars point out the standard deviation of predictive ability values.




As in most of the case, the standard deviations associated to the values of predictive ability were overlapping (**Figure 6; Figure 7**), Welch's *t*-tests were applied to determine whether pairwise comparisons of predictive ability values obtained with ME-GP models were statistically different (**Supplementary Figure 1 and 2**). CV2 experiments showed that in Fio17IN the values of predictive ability computed with the six-model method combinations were comparable except for MM-GB, which was significantly lower than the predictive ability of MDs-GK, while in Fio16IN the predictive ability of MM-GK was significantly lower than the predictive ability obtained with the remaining model-method combinations (**Figure 6; Supplementary Figure 2**). In Fio16IN, CV2 showed that MDe-GB and MDe-GK have similar performance and significantly higher values of predictive ability compared to MM models, either implemented with GB or GK statistical methods (**Figure 6; Supplementary Table 3; Supplementary Figure 2**). In Ada19IN the best model predictive ability using CV2 scheme was obtained with MDe-GB, while for Fio18LN the best values of predictive ability were obtained with MDe-GB and MDs-GB models. Overall, CV2 experiments indicated that in four out nine site-by-season combinations (Fio16IN, Fio17IN, Fio18IN and Mar16IN) MDe-GB and MDe-GK models have higher values of predictive ability compared to MM models, either implemented with GB or GK statistical methods (**Figure 6; Supplementary Table 3; Supplementary Figure 2**). Differently, Fio19IN, Fio19LN and Kon19IN deviate from this trend as for these site-by-season combinations the values of predictive ability for MM models were higher

(**Supplementary Table 3**). In Fio19IN, MM-GB and MM-GK had the higher predictive ability values along with MDe-GK, while for Fio19LN the higher value of predictive ability was found for MM-GB.

**Figure 7: Bar plots of the predictive ability values obtained with CV1**. Bar plots show the mean correlation between observed and predicted values of GY obtained with 100 random CV1 partitions for MM, MDs and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Error bars point out the standard deviation of predictive ability values.




The values of predictive ability obtained for random CV1 decreased **(Figure 7; Supplementary Table 4)** as compared with those computed for CV2 for all models. Similarly, to the results obtained for CV2, CV1 experiments indicated that in four site-by-season combinations (Fio16IN, Fio17IN, Fio18IN and Fio19LN) the predictive ability of GP-ME models is generally higher than the values of predictive ability observed in other site-by-season combinations for all models. MDs-GB and MD-GK yielded the higher values of predictive ability in Ada19IN, Fio16IN and Fio17IN, respectively. In Fio18IN, Fio18LN, Mar16IN and Fio19LN, the higher predictive ability values were found for MM-GK, although in this latter site-by-season combination the accuracy of MDe-GK does not differ significantly (**Supplementary Figure 1**). In Fio19IN, the highest values of predictive ability were obtained for MDe-GB and MD-GK models (**Figure 7; Supplementary Figure 1**).

# 5 Discussion

## 5.1 Broadening the use of MAGIC populations for plant breeding

MAGIC populations were conceived to improve precision and efficiency of QTL mapping in plants and animals as they allow overcoming limitations of biparental populations and association mapping panels (Huang et al. 2015). In cereal crops, these experimental populations have been extensively used for research purpose and contributed to dissecting the genetic bases of several traits among which biotic stress resistance (Jiménez-Galindo et al. 2019; Riaz et al. 2020; Stadlmeier et al. 2018), GY, grain quality (Zaw et al. 2019) and DH (Afsharyan et al. 2020). Recently, these genomic resources have been established in barley to investigate the effects of epistasis and environmental interactions on flowering time (Afsharyan et al. 2020; Mathew et al. 2018), further broadening the original scope for which they were devised.

In the present study, we constructed a new MAGIC population shuffling alleles of winter 6-rowed barley varieties, and demonstrated that, along with biparental populations and collections of mostly unrelated accessions, these genomic resources might be used to train GP models with high predictive ability and might speed up barley breeding. Under this point of view, the large number of MAGIC populations developed in the last years in several crops (Kover et al. 2009; Mathew et al. 2018; Rebetzke et al. 2014; Stadlmeier et al. 2018) can be considered as untapped resources that would contribute to further strengthening and stimulating the application of GP in plant breeding. On the other side, de novo creation of MAGIC populations to train GP models for actual breeding purposes is hampered because of their time consuming and costly development, which requires to intermate and self-fertilize the founder parents for several cycles. The results presented in this study show that these limitations might be softened using doubled haploid technology, which allows to short self-fertilization stages to obtain fully homozygous lines. Similarly, speed breeding might contribute to accelerating the development of new MAGIC populations (Watson et al. 2018).

To examine the genetic relationship between the whole set of MAGIC and the subset of lines included in the “TP-Diverse”, a PCA was carried out using 19,723 SNPs, which detected genetic structure in the MAGIC population and three main clusters of individuals. The nature of these clusters is unclear, but it is plausible that they might reflect subgroups of individuals showing segregation distortion for one or more founders. In our eight-way MAGIC population, the expected segregation rate of the eight founder haplotypes is 1:1:1:1:1:1:1:1, but the haplotypes of some founders (e.g. Dea) deviate from the expected ratio (**Data not shown**). Segregation distortion is a common phenomenon that occurs in MAGIC populations as pointed out in other studies (Sannemann et al. 2018). Although this did not hamper our ability to train GP models with this population, this phenomenon might explain the genetic structure pointed out with PCA.

Overall, the use of SE-GP and ME-GP models trained with MAGIC populations might find effective applications when the diversity of BPs originates from the same parents included in the founder set. In this case, GP models based on MAGIC populations might be applied to select the best offspring from crosses obtained with the MAGIC founders.

## 5.2 Benchmarking of different TPs to improve the predictive ability of GP models

The composition of TPs and their genetic relationship with BPs affect the predictive ability of GP models as pointed out in several studies (Desta and Ortiz 2014; Edwards et al. 2019; Norman et al. 2018) and to date several algorithms for optimizing TPs have been developed to increase the predictive ability of GP models (Akdemir et al. 2015). Untargeted and targeted optimization criteria based on GBLUP have been so far developed and tested in biparental populations and panel of mostly

unrelated accessions. Nevertheless, the use of these optimization methods in actual breeding programs is hampered as the optimization process can lead to different optimized TP per each trait of interest. These optimization algorithms require *a priori* information (knowledge of the BP genotypes and traits for which GP models must be developed) and output trait-dependent TPs (Akdemir et al. 2015). Moreover, in real breeding programs, BPs change over time and it might be difficult to implement these optimization procedures. Previous studies have shown that the relatedness between TPs and BPs has a large impact on the predictive ability of GP models, which can be improved increasing the genetic diversity of TPs (Norman et al. 2018). In fact, when the TP exhibits a narrow genetic diversity, low values of the predictive ability are often obtained in GP as it becomes impossible to predict all the marker effects that contribute to determining the phenotypic variations (Norman et al., 2018). Following these empirical findings, in this study we assembled a TP of 90 barley genotypes, which was named "TP-Diverse", maximizing the genetic diversity among MAGIC lines and assessing its predictive ability using random CV schemes. Surprisingly, the predictive ability obtained with TP-Diverse was comparable with the predictive ability of GP models trained with the other three optimized TPs used in this study (**Figure 2**). One of the main advantages of using this approach is that the criterion adopted to assemble "TP-Diverse" depends only on genetic data and does not generate trait-dependent TPs. On the other side, in this study we have not developed mathematical models to demonstrate or justify the rationale of this empirical criterion and consequently its validity should be further validated in other studies.

### 5.3 Fitting SE-GP and ME-GP models using the MAGIC population of barley

Several empirical analyses have been conducted to benchmark the predictive ability of different GP models in barley, maize and wheat panels of mostly unrelated accessions, biparental populations of *A.thaliana* and diallel crosses of maize and wheat to predict GY and other traits (Heslot et al. 2012). In this study, we presented another empirical analysis to assess the most promising GP models for MAGIC populations, implementing CV schemes for estimating the standard deviation of predictive ability values.

Three out five models fitted in this study (BayesA, BayesB and BL) belong to the group of so called "Bayesian alphabet", which denotes Bayesian linear regressions that differ in their prior density distribution (Gianola 2013). In these Bayesian regression models, the prior density distribution assigned to marker effects controls the shrinkage of estimates and then different priors induce different types of shrinkage of marker effects. In the original description both BayesA and BayesB were introduced as hierarchical structures (Meuwissen et al. 2001) and it was later demonstrated that BayesA adopts a scaled t-distribution prior, while BayesB adopts priors that are mixtures of a peak in the vicinity of zero and of a continuous density priors (e.g., t, or normal density distribution) (Gianola et al. 2009). BL adopts a double exponential prior density distribution, which behaves similar to that of BayesA as both priors used in these models do not allow marker effects to be equal to zero and shrink estimates of the remaining marker effects. While the priors adopted in BL and BayesA prevent to have marker effects equal to zero, the prior used in BayesB allows to have null marker effects. The rationale of this prior is that in GP many markers might have a null contribution to the observed phenotypic variation. Although marker effects might be estimated differently, the predictive ability of the Bayesian models fitted in this study does not differ significantly (**Figure 1**). Moreover, our empirical analysis shows that the predictive ability of Bayesian models fitted to MAGIC populations is comparable with that of GB and GK models (**Figure 1**). Several empirical analyses have been carried out in cereal crops to highlight advantages and limits of different whole genome regression methods. In rice, SE-GP models fitted with BayesA, GB and GK for three traits were compared using a reference panel of 284 accessions under different linkage disequilibrium

scenarios (Ben Hassen et al. 2018). These results showed that under high linkage disequilibrium scenarios GK models slightly outperform GB in terms of prediction ability. Differently, when a subset of rice reference panel was used to predict the performance of 97 advanced lined derived from biparental crosses, GK and GB prediction ability showed comparable results for the three traits considered (Ben Hassen et al. 2018). Anyway, the results obtained in this study are limited to one (complex) trait and it might plausible that for simpler traits GP models fitted in MAGIC might have different trend of the predictive ability.

Beyond SE-GP models, in this study we used the MAGIC population of barley to fit three different ME-GP models, two of which (MDs and MDe models) include terms for incorporating GxE interaction. In plant breeding, multi-environment field trials are routinely carried out to evaluate and exploit GxE interaction as it contributes to creating high-yielding genotypes. Consequently, modelling GxE interaction in GP has the potential to differentiate marker effects. MDe models used in this study (Bandeira e Sousa et al. 2017; Lopez-Cruz et al. 2015) partition marker effects in main effects, that is effects that are stable across environments and environment-specific effects, that is interaction effects between markers and specific genotypes. As pointed out in other studies, MDe models are known to be more efficient when used along with sets of environments that have positive correlations. This limit arises as the pairwise correlation between environments is represented by the variance of the main marker effects, which in turn forces the co-variance between a pair of environments to be positive (Bandeira e Sousa et al. 2017; Lopez-Cruz et al. 2015). This requirement is not trivial and might not allow to fit correctly MDe models. In our study, the adjusted means of GY in Mar16IN showed low or negative correlation with the other site-by-season combinations tested in this study and this might be the reason for which we have found that MDs models fit better the data, particularly when used in combination with the non-linear GK.

GP models based on reproducing kernel Hilbert Space along with the non-linear GK have the potential to capture non-additive genetic effects and potentially might outperform GB in terms of model fitting and predictive ability. In maize and wheat, comparison between the same GP models fitted with GB and the nonlinear GK for GY, unveiled that the latter method outperforms GB in terms of predictive ability in both single environment and multi-environment models (Bandeira e Sousa et al. 2017; Cuevas et al. 2016). In cereal crops, GY is a complex trait controlled by nonlinearity effects between genotypes and phenotypes owing to epistasis, environmental interactions (Bandeira e Sousa et al. 2017; Cuevas et al. 2018) and other interactions that are not considered in standard quantitative genetic models (Gianola, Fernando, and Stella 2006). GK models have the potential to capture small and complex interactions, which are more evident in quantitative traits and this can explain the higher prediction ability of GK for GY. The empirical analysis presented in this study using barley MAGIC population corroborates that, for complex traits like GY, the predictive ability of GK outperforms that of GB. Overall, considering the number of models and methods fitted and the extensive field trials carried out across the Mediterranean, this study has delivered the most comprehensive empirical analysis of GP models fitted with MAGIC populations.

## 6 Reference

Afsharyan, Nazanin P., Wiebke Sannemann, Jens Léon, and Agim Ballvora. 2020. "Effect of Epistasis and Environment on Flowering Time in Barley Reveals a Novel Flowering-Delaying QTL Allele." *Journal of Experimental Botany* 71(3):893–906.

Akdemir, Deniz, Julio I. Sanchez, and Jean Luc Jannink. 2015. "Optimization of Genomic Selection Training Populations with a Genetic Algorithm." *Genetics Selection Evolution* 47(1):1–10.

Bandeira e Sousa, Massaine, Jaime Cuevas, Evellyn Giselly de Oliveira Couto, Paulino Pérez-Rodríguez, Diego Jarquín, Roberto Fritsche-Neto, Juan Burgueño, and Jose Crossa. 2017. "Genomic-Enabled Prediction in Maize Using Kernel Models with Genotype 3 Environment Interaction." *G3: Genes, Genomes, Genetics* 7(6):1995–2014.

Bandillo, Nonoy, Chitra Raghavan, Pauline Andrea Muyco, Ma Anna Lynn Sevilla, Irish T. Lobina, Christine Jade Dilla-Ermita, Chih Wei Tung, Susan McCouch, Michael Thomson, Ramil Mauleon, Rakesh Kumar Singh, Glenn Gregorio, Edilberto Redoña, and Hei Leung. 2013. "Multi-Parent Advanced Generation Inter-Cross (MAGIC) Populations in Rice: Progress and Potential for Genetics Research and Breeding." *Rice*.

Bassi, Filippo M., Alison R. Bentley, Gilles Charmet, Rodomiro Ortiz, and Jose Crossa. 2016. "Breeding Schemes for the Implementation of Genomic Selection in Wheat (Triticum Spp.)." *Plant Science* 242:23–36.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67(1).

Bayer, Micha M., Paulo Rapazote-Flores, Martin Ganal, Pete E. Hedley, Malcolm Macaulay, Jörg Plieske, Luke Ramsay, Joanne Russell, Paul D. Shaw, William Thomas, and Robbie Waugh. 2017. "Development and Evaluation of a Barley 50k ISelect SNP Array." *Frontiers in Plant Science* 8:1–10.

Berro, Inés, Bettina Lado, Rafael S. Nalin, Martin Quincke, and Lucía Gutiérrez. 2019. "Training Population Optimization for Genomic Selection." *The Plant Genome* 12(3):190028.

De Beukelaer, Herman, Guy F. Davenport, and Veerle Fack. 2018. "Core Hunter 3: Flexible Core Subset Selection." *BMC Bioinformatics*.

Browning, Brian L., and Sharon R. Browning. 2016. "Genotype Imputation with Millions of Reference Samples." *American Journal of Human Genetics*.

Burgueño, Juan, Gustavo de los Campos, Kent Weigel, and José Crossa. 2012. "Genomic Prediction of Breeding Values When Modeling Genotype × Environment Interaction Using Pedigree and Dense Molecular Markers." *Crop Science* 52(2):707–19.

Cavanagh, Colin, Matthew Morell, Ian Mackay, and Wayne Powell. 2008. "From Mutations to MAGIC: Resources for Gene Discovery, Validation and Delivery in Crop Plants." *Current Opinion in Plant Biology* 11(2):215–21.

Chang, Christopher C., Carson C. Chow, Laurent C. A. M. Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience*.

Crossa, José, Paulino Pérez-Rodríguez, Jaime Cuevas, Osval Montesinos-López, Diego Jarquín, Gustavo de los Campos, Juan Burgueño, Juan M. González-Camacho, Sergio Pérez-Elizalde, Yoseph Beyene, Susanne Dreisigacker, Ravi Singh, Xuecai Zhang, Manje Gowda, Manish Roorkiwal, Jessica Rutkoski, and Rajeev K. Varshney. 2017. "Genomic Selection in Plant Breeding: Methods, Models, and Perspectives." *Trends in Plant Science* 22(11):961–75.

Cuevas, Jaime, José Crossa, Víctor Soberanis, Sergio Pérez-Elizalde, Paulino Pérez-Rodríguez, Gustavo de los Campos, O. A. Montesinos-López, and Juan Burgueño. 2016. "Genomic Prediction of Genotype × Environment Interaction Kernel Regression Models." *The Plant Genome* 9(3):plantgenome2016.03.0024.

Cuevas, Jaime, Italo Granato, Roberto Fritsche-Neto, Osval A. Montesinos-Lopez, Juan Burgueño, Massaine Bandeira Sousa, and José Crossa. 2018. "Genomic-Enabled Prediction Kernel Models

with Random Intercepts for Multi-Environment Trials." *G3: Genes, Genomes, Genetics* 8(4):1347–65.

Dell'Acqua, Matteo, Daniel M. Gatti, Giorgio Pea, Federica Cattonaro, Frederik Coppens, Gabriele Magris, Aye L. Hlaing, Htay H. Aung, Hilde Nelissen, Joke Baute, Elisabetta Frascaroli, Gary A. Churchill, Dirk Inzé, Michele Morgante, and Mario Enrico Pè. 2015. "Genetic Properties of the MAGIC Maize Population: A New Platform for High Definition QTL Mapping in Zea Mays." *Genome Biology* 16(1):1–23.

Desta, Zeratsion Abera, and Rodomiro Ortiz. 2014. "Genomic Selection: Genome-Wide Prediction in Plant Improvement." *Trends in Plant Science* 19(9):592–601.

Edwards, Stefan Mc Kinnon, Jaap B. Buntjer, Robert Jackson, Alison R. Bentley, Jacob Lage, Ed Byrne, Chris Burt, Peter Jack, Simon Berry, Edward Flatman, Bruno Poupard, Stephen Smith, Charlotte Hayes, R. Chris Gaynor, Gregor Gorjanc, Phil Howell, Eric Ober, Ian J. Mackay, and John M. Hickey. 2019. "The Effects of Training Population Design on Genomic Prediction Accuracy in Wheat." *Theoretical and Applied Genetics* 132(7):1943–52.

Emrich, K., F. Wilde, T. Miedaner, and H. P. Piepho. 2008. "REML Approach for Adjusting the Fusarium Head Blight Rating to a Phenological Date in Inoculated Selection Experiments of Wheat." *Theoretical and Applied Genetics* 117(1):65–73.

Gianola, Daniel. 2013. "Priors in Whole-Genome Regression: The Bayesian Alphabet Returns." *Genetics* 194(3):573–96.

Gianola, Daniel, Rohan L. Fernando, and Alessandra Stella. 2006. "Genomic-Assisted Prediction of Genetic Value with Semiparametric Procedures." *Genetics* 173(3):1761–76.

Gianola, Daniel, and Johannes B. C. H. M. Van Kaam. 2008. "Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits." *Genetics Society of America* 178(April):2289–2303.

Gianola, Daniel, Gustavo De Los Campos, William G. Hill, Eduardo Manfredi, and Rohan Fernando. 2009. "Additive Genetic Variability and the Bayesian Alphabet." *Genetics* 183(1):347–63.

Gianola, Daniel, and Chris Carolin Schon. 2016. "Cross-Validation without Doing Cross-Validation in Genome-Enabled Prediction." *G3: Genes, Genomes, Genetics* 6(10):3107–28.

Gota, Morota, and Daniel Gianola. 2014. "Kernel-Based Whole-Genome Prediction of Complex Traits: A Review." *Frontiers in Genetics* 5(OCT):1–13.

Ben Hassen, M., T. V. Cao, J. Bartholomé, G. Orasen, C. Colombi, J. Rakotomalala, L. Razafinimpiasa, C. Bertone, C. Biselli, A. Volante, F. Desiderio, L. Jacquin, G. Valè, and N. Ahmadi. 2018. "Rice Diversity Panel Provides Accurate Genomic Predictions for Complex Traits in the Progenies of Biparental Crosses Involving Members of the Panel." *Theoretical and Applied Genetics* 131(2):417–35.

Heffner, Elliot L., Jean-Luc Jannink, and Mark E. Sorrells. 2011. "Genomic Selection Accuracy Using Multifamily Prediction Models in a Wheat Breeding Program." *The Plant Genome* 4(1):65–75.

Heffner, Elliot L., Mark E. Sorrells, and Jean Luc Jannink. 2009. "Genomic Selection for Crop Improvement." *Crop Science* 49(1):1–12.

Heslot, Nicolas, Hsiao Pei Yang, Mark E. Sorrells, and Jean Luc Jannink. 2012. "Genomic Selection in Plant Breeding: A Comparison of Models." *Crop Science* 52(1):146–60.

Hill, W. G., and Alan Robertson. 2008. "The Effect of Linkage on Limits to Artificial Selection." *Genetics Research*.

Huang, B. Emma, Klara L. Verbyla, Arunas P. Verbyla, Chitra Raghavan, Vikas K. Singh, Pooran Gaur, Hei Leung, Rajeev K. Varshney, and Colin R. Cavanagh. 2015. "MAGIC Populations in Crops: Current Status and Future Prospects." *Theoretical and Applied Genetics* 128(6):999–1017.

Jannink, Jean Luc, Aaron J. Lorenz, and Hiroyoshi Iwata. 2010. "Genomic Selection in Plant Breeding: From Theory to Practice." *Briefings in Functional Genomics and Proteomics* 9(2):166–77.

Jarquín, Diego, José Crossa, Xavier Lacaze, Philippe Du Cheyron, Joëlle Daucourt, Josiane Lorgeou, François Piraux, Laurent Guerreiro, Paulino Pérez, Mario Calus, Juan Burgueño, and Gustavo de los Campos. 2014. "A Reaction Norm Model for Genomic Selection Using High-Dimensional Genomic and Environmental Data." *Theoretical and Applied Genetics* 127(3):595–607.

Jiménez-Galindo, José Cruz, Rosa Ana Malvar, Ana Butrón, Rogelio Santiago, Luis Fernando Samayoa, Marlon Caicedo, and Bernardo Ordás. 2019. "Mapping of Resistance to Corn Borers in a MAGIC Population of Maize." *BMC Plant Biology*.

Kover, Paula X., William Valdar, Joseph Trakalo, Nora Scarcelli, Ian M. Ehrenreich, Michael D. Purugganan, Caroline Durrant, and Richard Mott. 2009. "A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in Arabidopsis Thaliana." *PLoS Genetics* 5(7).

Laloe, D. 1993. "Precision and Information in Linear Models of Genetic Evaluation." *Genetics Selection Evolution* 25(6):557–76.

Lopez-Cruz, Marco, Jose Crossa, David Bonnett, Susanne Dreisigacker, Jesse Poland, Jean Luc Jannink, Ravi P. Singh, Enrique Autrique, and Gustavo de los Campos. 2015. "Increased Prediction Accuracy in Wheat Breeding Trials Using a Marker × Environment Interaction Genomic Selection Model." *G3: Genes, Genomes, Genetics* 5(4):569–82.

Lorenzana, Robenzon E., and Rex Bernardo. 2009. "Accuracy of Genotypic Value Predictions for Marker-Based Selection in Biparental Plant Populations." *Theoretical and Applied Genetics* 120(1):151–61.

De Los Campos, Gustavo, Hugo Naya, Daniel Gianola, José Crossa, Andrés Legarra, Eduardo Manfredi, Kent Weigel, and José Miguel Cotes. 2009. "Predicting Quantitative Traits with Regression Models for Dense Molecular Markers and Pedigree." *Genetics* 182(1):375–85.

Mackay, Ian J., Pauline Bansept-Basler, Alison R. Bentley, James Cockram, Nick Gosman, Andy J. Greenland, Richard Horsnell, Rhian Howells, Donal M. O'Sullivan, Gemma A. Rose, and Phil J. Howell. 2014. "An Eight-Parent Multiparent Advanced Generation Inter-Cross Population for Winter-Sown Wheat: Creation, Properties, and Validation." *G3: Genes, Genomes, Genetics* 4(9):1603–10.

Mathew, Boby, Jens Léon, Wiebke Sannemann, and Mikko J. Sillanpää. 2018. "Detection of Epistasis for Flowering Time Using Bayesian Multilocus Estimation in a Barley MAGIC Population." *Genetics* 208(2):525–36.

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):1819–29.

Monat, Cécile, Sudharsan Padmarasu, Thomas Lux, Thomas Wicker, Heidrun Gundlach, Axel Himmelbach, Jennifer Ens, Chengdao Li, Gary J. Muehlbauer, Alan H. Schulman, Robbie Waugh, Ilka Braumann, Curtis Pozniak, Uwe Scholz, Klaus F. X. Mayer, Manuel Spannagl, Nils Stein, and Martin Mascher. 2019. "TRITEX : Chromosome-Scale Sequence Assembly of Triticeae Genomes with Open- Source Tools." *Genome Biology* 20(1):1–35.

Norman, Adam, Julian Taylor, James Edwards, and Haydn Kuchel. 2018. "Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy." *G3: Genes, Genomes, Genetics* 8(9):2889–99.

Ou, Jen Hsiang, and Chen Tuo Liao. 2019. "Training Set Determination for Genomic Selection." *Theoretical and Applied Genetics* 132(10):2781–92.

Park, Trevor, and George Casella. 2008. "The Bayesian Lasso." *Journal of the American Statistical Association* 103(482):681–86.

Pérez-Elizalde, Sergio, Jaime Cuevas, Paulino Pérez-Rodríguez, and José Crossa. 2015. "Selection of the Bandwidth Parameter in a Bayesian Kernel Regression Model for Genomic-Enabled Prediction." *Journal of Agricultural, Biological, and Environmental Statistics* 20(4):512–32.

Pérez, Paulino, and G. de los Campos. 2014. "BGLR : A Statistical Package for Whole Genome Regression and Prediction." *Genetics* 198(2):483–95.

Ponce, Kimberly S., Guoyou Ye, and Xiangqian Zhao. 2018. "QTL Identification for Cooking and

Eating Quality in Indica Rice Using Multi-Parent Advanced Generation Intercross (MAGIC) Population." *Frontiers in Plant Science* 9(July):1–9.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul I. W. De Bakker, Mark J. Daly, and Pak C. Sham. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics*.

Rebetzke, Greg J., Arunas P. Verbyla, Klara L. Verbyla, Matthew K. Morell, and Colin R. Cavanagh. 2014. "Use of a Large Multiparent Wheat Mapping Population in Genomic Dissection of Coleoptile and Seedling Growth." *Plant Biotechnology Journal* 12(2):219–30.

Riaz, Adnan, Petra KockAppelgren, James Gerard Hehir, Jie Kang, Fergus Meade, James Cockram, Dan Milbourne, John Spink, Ewen Mullins, and Stephen Byrne. 2020. "Genetic Analysis Using a Multi-Parent Wheat Population Identifies Novel Sources of Septoria Tritici Blotch Resistance." *Genes* 11(8):1–26.

Rincent, R., D. Laloë, S. Nicolas, T. Altmann, D. Brunel, P. Revilla, V. M. Rodríguez, J. Moreno-Gonzalez, A. Melchinger, E. Bauer, C. C. Schoen, N. Meyer, C. Giauffret, C. Bauland, P. Jamin, J. Laborde, H. Monod, P. Flament, A. Charcosset, and L. Moreau. 2012. "Maximizing the Reliability of Genomic Selection by Optimizing the Calibration Set of Reference Individuals: Comparison of Methods in Two Diverse Groups of Maize Inbreds (Zea Mays L.)." *Genetics* 192(2):715–28.

Sannemann, Wiebke, Antonia Lisker, Andreas Maurer, Jens Léon, Ebrahim Kazman, Hilmar Cöster, Josef Holzapfel, Hubert Kempf, Viktor Korzun, Erhard Ebmeyer, and Klaus Pillen. 2018. "Adaptive Selection of Founder Segments and Epistatic Control of Plant Height in the MAGIC Winter Wheat Population WM-800." *BMC Genomics* 19(1):1–16.

Scott, Michael F., Olufunmilayo Ladejobi, Samer Amer, Alison R. Bentley, Jay Biernaskie, Scott A. Boden, Matt Clark, Matteo Dell'Acqua, Laura E. Dixon, Carla V. Filippi, Nick Fradgley, Keith A. Gardner, Ian J. Mackay, Donal O'Sullivan, Lawrence Percival-Alwyn, Manish Roorkiwal, Rakesh Kumar Singh, Mahendar Thudi, Rajeev Kumar Varshney, Luca Venturini, Alex Whan, James Cockram, and Richard Mott. 2020. "Multi-Parent Populations in Crops: A Toolbox Integrating Genomics and Genetic Mapping with Breeding." *Heredity*.

Stadlmeier, Melanie, Lorenz Hartl, and Volker Mohler. 2018. "Usefulness of a Multiparent Advanced Generation Intercross Population with a Greatly Reduced Mating Design for Genetic Studies in Winter Wheat." *Frontiers in Plant Science* 871(December):1–12.

Team R. Core. 2019. "R: A Language and Environment for Statistical Computing." *Industrial and Commercial Training*.

Thachuk, Chris, José Crossa, Jorge Franco, Susanne Dreisigacker, Marilyn Warburton, and Guy F. Davenport. 2009. "Core Hunter: An Algorithm for Sampling Genetic Resources Based on Multiple Genetic Measures." *BMC Bioinformatics*.

Thioulouse, Jean, Anne Béatrice Dufour, Thibaut Jombart, Stéphane Dray, Aurélie Siberchicot, and Sandrine Pavoine. 2018. *Multivariate Analysis of Ecological Data with Ade4*.

Tishbirani, R. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* 58(1):267–88.

Valdar, William, Jonathan Flint, and Richard Mott. 2006. "Simulating the Collaborative Cross: Power of Quantitative Trait Loci Detection and Mapping Resolution in Large Sets of Recombinant Inbred Strains of Mice." *Genetics* 172(3):1783–97.

VanRaden, P. M. 2008. "Efficient Methods to Compute Genomic Predictions." *Journal of Dairy Science* 91(11):4414–23.

Wang, Xin, Yang Xu, Zhongli Hu, and Chenwu Xu. 2018. "Genomic Selection Methods for Crop Improvement: Current Status and Prospects." *Crop Journal* 6(4):330–40.

Watson, Amy, Sreya Ghosh, Matthew J. Williams, William S. Cuddy, James Simmonds, María Dolores Rey, M. Asyraf Md Hatta, Alison Hinchliffe, Andrew Steed, Daniel Reynolds, Nikolai M. Adamski, Andy Breakspear, Andrey Korolev, Tracey Rayner, Laura E. Dixon, Adnan Riaz,

William Martin, Merrill Ryan, David Edwards, Jacqueline Batley, Harsh Raman, Jeremy Carter, Christian Rogers, Claire Domoney, Graham Moore, Wendy Harwood, Paul Nicholson, Mark J. Dieters, Ian H. Delacy, Ji Zhou, Cristobal Uauy, Scott A. Boden, Robert F. Park, Brande B. H. Wulff, and Lee T. Hickey. 2018. "Speed Breeding Is a Powerful Tool to Accelerate Crop Research and Breeding." *Nature Plants* 4(1):23–29.

Wickham, Hadley. 2016. *Ggplot 2: Elagant Graphics for Data Analysis*. Vol. 35.

Wimmer, Valentin, Theresa Albrecht, Hans Jürgen Auinger, and Chris Carolin Schön. 2012. "Synbreed: A Framework for the Analysis of Genomic Prediction Data Using R." *Bioinformatics*.

Xu, Xin, Rajiv Sharma, Alessandro Tondelli, Joanne Russell, Jordi Comadran, Florian Schnaithmann, Klaus Pillen, Benjamin Kilian, Luigi Cattivelli, William T. B. Thomas, and Andrew J. Flavell. 2018. "Genome-Wide Association Analysis of Grain Yield-Associated Traits in a Pan-European Barley Cultivar Collection." *The Plant Genome* 0(0):0.

Yalcin, B., J. Flint, and R. Mott. 2005. "Using Progenitor Strain Information to Identify Quantitative Trait Nucleotides in Outbred Mice." *Genetics*.

Zadoks, J. C., T. T. Chang, and C. F. Konzak. 1974. "A Decimal Code for the Growth Stages of Cereals." *World Pumps* 14:415–21.

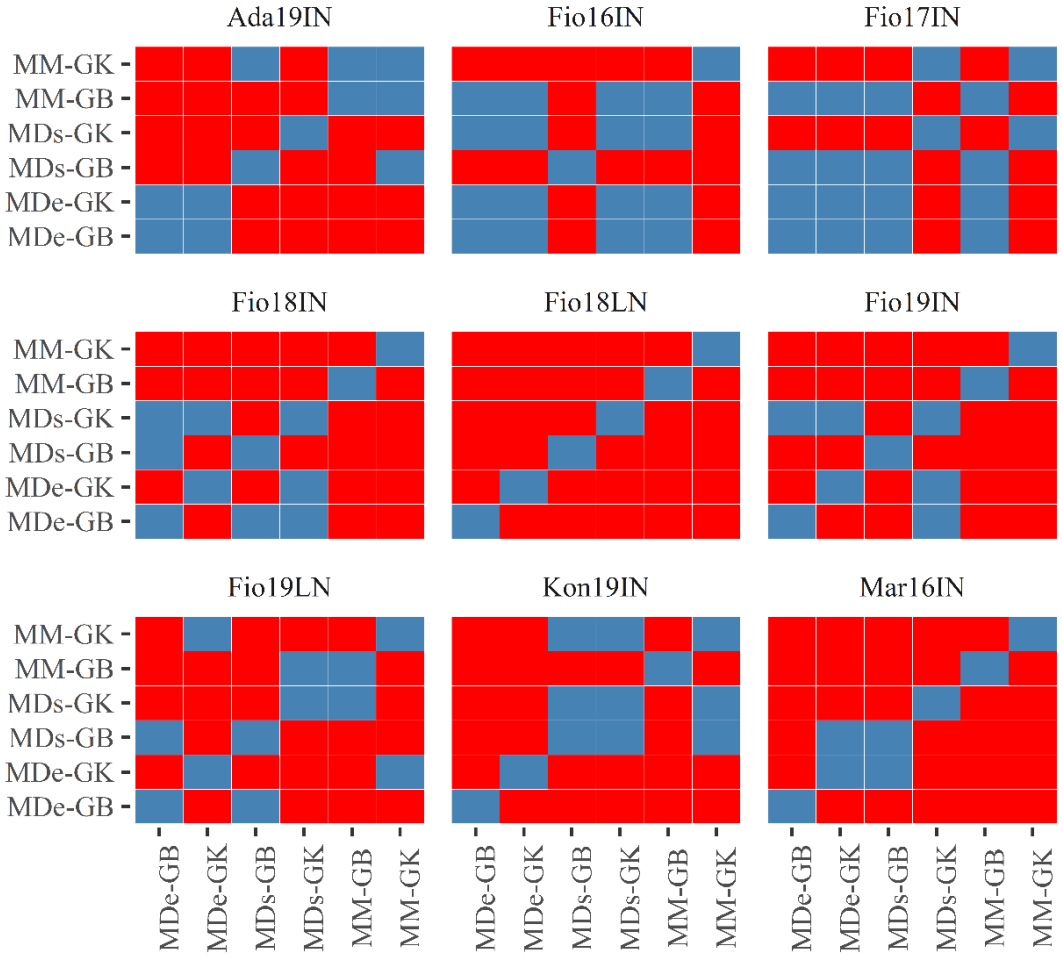
Zaw, Hein, Chitra Raghavan, Arnel Pocsedio, B. P. Mallikarjun. Swamy, Mona Liza Jubay, Rakesh Kumar Singh, Justine Bonifacio, Ramil Mauleon, Jose E. Hernandez, Merlyn S. Mendioro, Glenn B. Gregorio, and Hei Leung. 2019. "Exploring Genetic Architecture of Grain Yield and Quality Traits in a 16-Way Indica by Japonica Rice MAGIC Global Population." *Scientific Reports* 9(1):1–11.

## 7 Supplementary Material

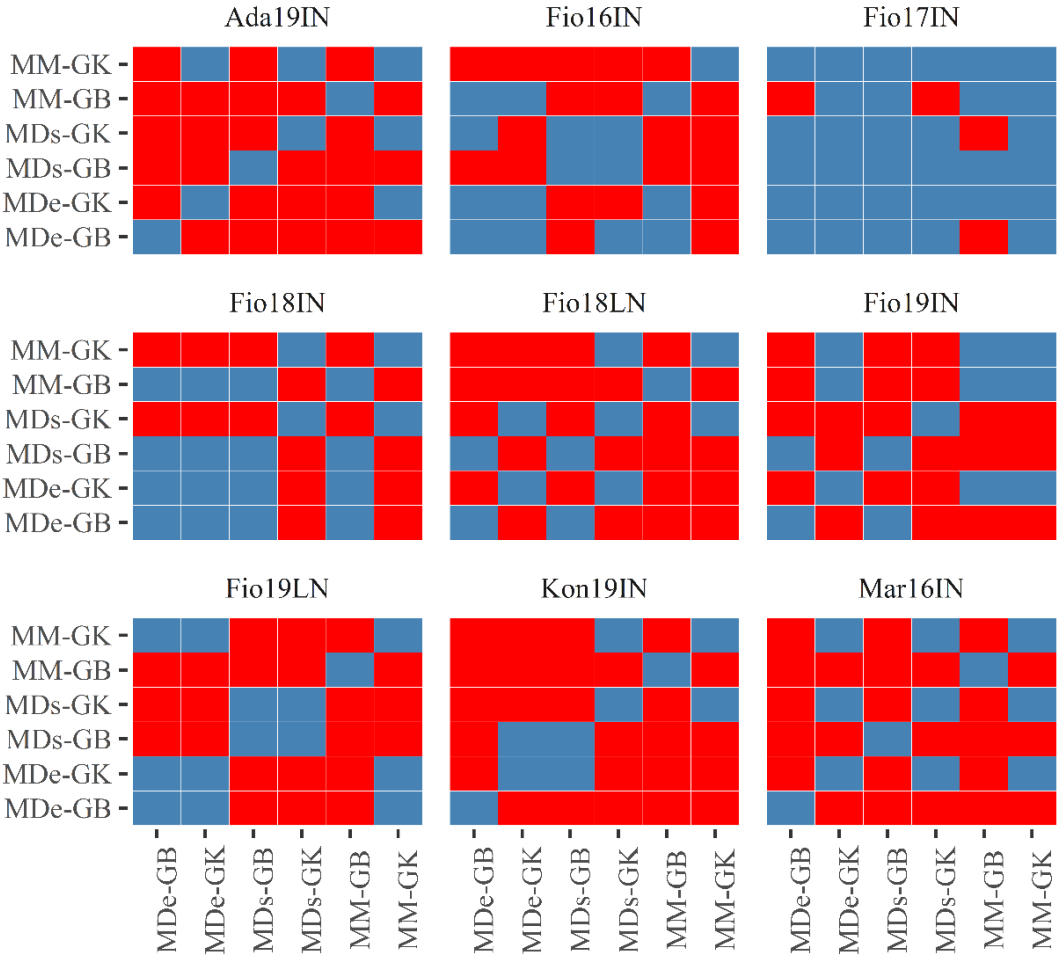
The Supplementary Material, figures and tables, can be found online at https://www.frontiersin.org/articles/10.3389/fpls.2021.664148/full#supplementary-material and are shown below:

### 7.1 Supplementary Figures

**Supplementary Figure 1.** Heatmap of *P*-values obtained applying Welch's *t*-test to accuracy parameters (means and standard deviations) obtained from pairs of ME-GP models using CV1 schemes. Red tiles in the heatmap point out pairs of models that exhibit significantly different accuracy (*P*-values<0.05), while blue tiles point out pairs of models that do not show different model accuracy.



**Supplementary Figure 2.** Heatmap of *P*-values obtained applying Welch's *t*-test between model accuracy values obtained using CV2 schemes for each ME-GP model. Red tiles in the heatmap point out pairs of models that exhibit significantly different model accuracy values (*P*-values<0.05), while blue tiles point out pairs of models that do not show different model accuracy.




## 7.2 Supplementary Tables

**Supplementary Table 1: CV of different SE-GP models fitted using MAGIC population.** Model accuracies and associated standard deviation of SE-GP models, reported within brackets, computed in Fio16IN, Fio17IN and Mar16IN using GB, GK, BayesA, BayesB and BL models as a function of TP sizes.

| Env | TP size | BayesA | BayesB | BL | GB | GK |
|---|---|---|---|---|---|---|
| **Fio16IN** | 80 | 0.511 (0.037) | 0.51 (0.039) | 0.516 (0.038) | 0.512 (0.035) | 0.501 (0.038) |
| **Fio16IN** | 90 | 0.52 (0.033) | 0.519 (0.039) | 0.522 (0.033) | 0.519 (0.033) | 0.517 (0.037) |
| **Fio16IN** | 100 | 0.527 (0.029) | 0.528 (0.038) | 0.528 (0.035) | 0.526 (0.033) | 0.517 (0.033) |
| **Fio16IN** | 110 | 0.534 (0.031) | 0.533 (0.039) | 0.539 (0.035) | 0.533 (0.034) | 0.526 (0.035) |
| **Fio16IN** | 120 | 0.547 (0.035) | 0.54 (0.036) | 0.54 (0.032) | 0.539 (0.037) | 0.531 (0.036) |
| **Fio16IN** | 130 | 0.546 (0.039) | 0.543 (0.037) | 0.549 (0.031) | 0.546 (0.035) | 0.541 (0.036) |
| **Fio16IN** | 140 | 0.553 (0.035) | 0.55 (0.037) | 0.553 (0.037) | 0.55 (0.038) | 0.548 (0.038) |
| **Fio16IN** | 150 | 0.56 (0.034) | 0.558 (0.035) | 0.558 (0.037) | 0.557 (0.037) | 0.547 (0.036) |
| **Fio16IN** | 160 | 0.557 (0.031) | 0.559 (0.042) | 0.558 (0.035) | 0.564 (0.036) | 0.547 (0.038) |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Fio17IN** | 80 | 0.509 (0.036) | 0.506 (0.033) | 0.514 (0.032) | 0.514 (0.032) | 0.504 (0.04) |
| **Fio17IN** | 90 | 0.525 (0.032) | 0.526 (0.03) | 0.523 (0.034) | 0.516 (0.033) | 0.511 (0.033) |
| **Fio17IN** | 100 | 0.53 (0.035) | 0.528 (0.032) | 0.532 (0.034) | 0.526 (0.032) | 0.526 (0.035) |
| **Fio17IN** | 110 | 0.53 (0.037) | 0.534 (0.033) | 0.537 (0.027) | 0.529 (0.036) | 0.532 (0.032) |
| **Fio17IN** | 120 | 0.543 (0.033) | 0.544 (0.032) | 0.544 (0.028) | 0.545 (0.034) | 0.533 (0.037) |
| **Fio17IN** | 130 | 0.548 (0.035) | 0.542 (0.039) | 0.551 (0.035) | 0.545 (0.035) | 0.54 (0.032) |
| **Fio17IN** | 140 | 0.553 (0.04) | 0.553 (0.036) | 0.546 (0.037) | 0.553 (0.036) | 0.546 (0.034) |
| **Fio17IN** | 150 | 0.561 (0.032) | 0.555 (0.035) | 0.557 (0.034) | 0.562 (0.035) | 0.554 (0.04) |
| **Fio17IN** | 160 | 0.562 (0.039) | 0.558 (0.039) | 0.566 (0.039) | 0.56 (0.037) | 0.559 (0.04) |
| **Mar16IN** | 80 | 0.16 (0.052) | 0.165 (0.055) | 0.151 (0.062) | 0.166 (0.059) | 0.151 (0.06) |
| **Mar16IN** | 90 | 0.169 (0.054) | 0.17 (0.054) | 0.176 (0.049) | 0.169 (0.055) | 0.171 (0.056) |
| **Mar16IN** | 100 | 0.186 (0.041) | 0.183 (0.05) | 0.181 (0.055) | 0.174 (0.055) | 0.182 (0.044) |
| **Mar16IN** | 110 | 0.199 (0.052) | 0.185 (0.05) | 0.177 (0.056) | 0.179 (0.059) | 0.189 (0.05) |
| **Mar16IN** | 120 | 0.194 (0.049) | 0.191 (0.045) | 0.190 (0.057) | 0.193 (0.054) | 0.191 (0.054) |
| **Mar16IN** | 130 | 0.203 (0.054) | 0.206 (0.051) | 0.190 (0.058) | 0.2 (0.049) | 0.198 (0.051) |
| **Mar16IN** | 140 | 0.209 (0.049) | 0.205 (0.048) | 0.209 (0.051) | 0.201 (0.049) | 0.194 (0.053) |
| **Mar16IN** | 150 | 0.211 (0.056) | 0.208 (0.055) | 0.211 (0.05) | 0.206 (0.056) | 0.204 (0.048) |
| **Mar16IN** | 160 | 0.22 (0.049) | 0.213 (0.054) | 0.211 (0.051) | 0.217 (0.052) | 0.205 (0.041) |

**Supplementary Table 2: Distribution of polymorphic SNPs mapped in the seven barley chromosomes.** 18,248 out 19,723 polymorphic SNPs were unambiguously mapped to the barley reference sequence and used to compute $r^2$ to estimate the extent of LD.

| Chromosome | Number of SNPs | Average SNPs per Mb |
|---|---|---|
| **Chromosome 1H** | 2095 | 4.00 |
| **Chromosome 2H** | 2281 | 3.38 |
| **Chromosome 3H** | 2855 | 4.54 |
| **Chromosome 4H** | 1444 | 2.31 |
| **Chromosome 5H** | 3589 | 5.99 |
| **Chromosome 6H** | 2615 | 4.56 |
| **Chromosome 7H** | 3369 | 5.31 |
| **Total** | 18248 | 4.36 |

**Supplementary Table 3:** Mean correlation between observed and predicted values of GY (average of 100 random CV partitions) computed using CV2 scheme for MM, MDs and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Numbers between brackets point out the standard deviation of model accuracy values.

| Model | MDe-GB | MDe-GK | MDs-GB | MDs-GK | MM-GB | MM-GK |
|---|---|---|---|---|---|---|
| **Ada19IN** | 0.676 (0.103) | 0.472 (0.1) | 0.609 (0.117) | 0.528 (0.128) | 0.301 (0.076) | 0.500 (0.112) |
| **Fio16IN** | 0.783 (0.125) | 0.756 (0.119) | 0.832 (0.116) | 0.802 (0.125) | 0.760 (0.106) | 0.696 (0.105) |
| **Fio17IN** | 0.731 (0.114) | 0.711 (0.093) | 0.714 (0.134) | 0.734 (0.137) | 0.687 (0.113) | 0.718 (0.109) |
| **Fio18IN** | 0.621 (0.142) | 0.612 (0.106) | 0.623 (0.113) | 0.467 (0.084) | 0.613 (0.116) | 0.462 (0.099) |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Fio18LN** | 0.528 (0.088) | 0.46 (0.108) | 0.518 (0.1) | 0.443 (0.105) | 0.364 (0.113) | 0.414 (0.104) |
| **Fio19IN** | 0.581 (0.133) | 0.757 (0.103) | 0.571 (0.095) | 0.662 (0.143) | 0.774 (0.12) | 0.779 (0.115) |
| **Fio19LN** | 0.749 (0.093) | 0.744 (0.125) | 0.674 (0.114) | 0.661 (0.121) | 0.817 (0.132) | 0.721 (0.129) |
| **Kon19IN** | 0.515 (0.083) | 0.557 (0.118) | 0.579 (0.122) | 0.632 (0.152) | 0.427 (0.131) | 0.664 (0.123) |
| **Mar16IN** | 0.304 (0.098) | 0.261 (0.103) | 0.384 (0.139) | 0.236 (0.12) | 0.161 (0.072) | 0.243 (0.085) |

**Supplementary Table 4:** Mean correlation between observed and predicted values of GY (average of 100 random CV partitions) computed using CV1 scheme for MM, MDs and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Numbers between brackets point out the standard deviation of predictive ability values.

| Model | MDe-GB | MDe-GK | MDs-GB | MDs-GK | MM-GB | MM-GK |
|---|---|---|---|---|---|---|
| **Ada19IN** | 0.276 (0.113) | 0.298 (0.106) | 0.232 (0.102) | 0.442 (0.129) | 0.192 (0.097) | 0.208 (0.114) |
| **Fio16IN** | 0.683 (0.120) | 0.675 (0.101) | 0.775 (0.117) | 0.701 (0.094) | 0.682 (0.118) | 0.532 (0.109) |
| **Fio17IN** | 0.605 (0.158) | 0.614 (0.119) | 0.588 (0.099) | 0.731 (0.102) | 0.621 (0.17) | 0.727 (0.111) |
| **Fio18IN** | 0.537 (0.117) | 0.484 (0.113) | 0.545 (0.102) | 0.512 (0.126) | 0.434 (0.098) | 0.597 (0.13) |
| **Fio18LN** | 0.122 (0.088) | 0.377 (0.079) | 0.314 (0.13) | 0.261 (0.083) | 0.211 (0.092) | 0.417 (0.119) |
| **Fio19IN** | 0.407 (0.087) | 0.439 (0.118) | 0.153 (0.080) | 0.43 (0.107) | 0.355 (0.106) | 0.282 (0.093) |
| **Fio19LN** | 0.496 (0.1) | 0.599 (0.096) | 0.481 (0.101) | 0.535 (0.129) | 0.539 (0.095) | 0.613 (0.124) |
| **Kon19IN** | 0.305 (0.054) | 0.208 (0.121) | 0.348 (0.089) | 0.367 (0.121) | 0.173 (0.087) | 0.349 (0.1) |
| **Mar16IN** | 0.283 (0.062) | 0.226 (0.046) | 0.215 (0.091) | 0.162 (0.06) | 0.143 (0.059) | 0.306 (0.096) |

## 8 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.

## 9 Author contribution

AF designed and supervised the research along with the help of EI and LC. AF wrote the manuscript along with DP and significant contributions from AV, EI, HO, AC, AT, GP, LC and ARLP. DP performed the research, while GP, AT and SD developed the MAGIC population. AV, HO, EI, AC, SD, AT, IK, ARLP and DP carried out field trials and plant phenotyping. All authors contributed to the article and approved the submitted version.

## 10 Funding

This research was carried out in the framework of the iBarMed project, which has been funded through the ARIMNet2 initiative and the Italian "Ministry of Agricultural, Food and Forestry Policies" under grant agreement "DM n. 20120". ARIMNet2 has received funding from the EU 7th Framework Programme for research, technological development and demonstration under grant agreement no. 618127. The work was also supported by SYSTEMIC_1063 (An integrated approach to the challenge of sustainable food systems: adaptive and mitigatory strategies to address climate change and malnutrition: From cereal diversity to plant breeding), a research project funded by Italian "Ministry of Agricultural, Food and Forestry Policies" in the frame of the Knowledge Hub on Food and Nutrition Security.

## 11 Abbreviations

Days-to Heading (DH), Plant Height (PH), Grain Yield (GY), Multi-parent Advanced Generation Inter-crosses (MAGIC), Training Population (TP), Breeding Population (BP), Genomic Prediction (GP), Genomic Best Linear Unbiased Prediction (GBLUP), Reproducing Kernel Hilbert Space (RKHS), Gaussian Kernel (GK), Linkage Disequilibrium (LD), Principal Component Analysis (PCA), Quantitative Trait Locus (QTL), Genomic Estimated Breeding Value (GEBV), Single Nucleotide Polymorphism (SNP), Single Environment Genomic Prediction (SE-GP), Multi Environment Genomic Prediction (ME-GP)

## 12 Acknowledgements

This paper is dedicated to the memory of Prof. Antonio Michele Stanca, who passed away on March 19th, 2020. Beyond supporting this study, Prof. Stanca's research encompassed a wide range of plant science topics and contributed to exploiting morphological barley mutants to unveil the mechanisms that underlie plant developmental biology.

## 13 Data Availability Statement

The data underpinning the results presented in this manuscript are available from the authors upon request.

# Chapter 3: Genome enabled prediction of belowground and physiological traits using a MAGIC population of barley

**Damiano Puglisi**[1ǂ], Andrea Visioni[2ǂ], Hakan Ozkan[3], İbrahim Kara[4], Rosa Marchetti[5], Angela Roberta Lo Piero[1], Alessandro Tondelli[6], Luigi Cattivelli[6], Fricano Agostino[6*]

[1]Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A), Università di Catania, Via S. Sofia 98, 95123 Catania, Italy.
[2]Biodiversity and Crop Improvement Program, International Center for Agricultural Research in the Dry Areas, Avenue Hafiane Cherkaoui, Rabat, Morocco
[3]Department of Field Crops, Faculty of Agriculture, University of Cukurova, 01330 Adana, Turkey
[4]Bahri Dagdas International Agricultural Research Institute, Ereğli Yolu 3. Km Karatay / Konya, 42020, Turkey
[5]Council for Agricultural Research and Economics – Research Centre for Animal Production and Aquaculture, Via Beccastecca 345, 41018, San Cesario sul Panaro (MO), Italy
[6]Council for Agricultural Research and Economics – Research Centre for Genomics and Bioinformatics, Via San Protaso 302 Fiorenzuola d'Arda (PC), Italy

[ǂ]These authors contributed equally to this work.
***Correspondence:**
Corresponding Author
agostino.fricano@crea.gov.it

**Keywords:** seminal root angle, seminal root number, transpiration rate, MAGIC, barley, genomic prediction, threshold GBLUP.

## 1 Abstract

With the introduction of simpler and faster phenotyping methods, the study of belowground traits is gaining momentum due to their importance in yield formation and water and nutrient uptake. Among belowground traits, in cereal crops seminal root number and seminal root angle have shown to be correlated and predict the root system architecture at the mature stages. Along with seminal root number and seminal root angle, experimental evidence has indicated that the transpiration rate response to evaporative demand or vapor pressure deficit is a key physiological trait that might be targeted to cope with drought tolerance as the reduction of the water flux to leaves for limiting transpiration rate at high levels of vapor pressure deficit allows to better manage soil moisture.

In the present study, we examined the phenotypic diversity of seminal root number, seminal root angle and transpiration rate in a panel of 90 eight-way Multi-parent Advanced Generation Inter-crosses (MAGIC) lines of winter barley and correlated these traits with grain yield measured in different site-by-season combinations. Phenotypic and genotypic information of MAGIC population were combined to fit and cross-validate different genomic prediction models. Genomic prediction for seminal root number were fitted using threshold and log-normal models, considering these data as ordinal discrete variable and as count data, respectively, while for seminal root angle and transpiration rate genomic prediction models were fitted using extended GBLUP.

Our empirical analyses showed that genomic prediction models can be used to predict seminal root number, seminal root angle and transpiration rate with high predictive ability and that the best models investigated included first order additive x additive epistatic interaction effects. Overall, this study

showed that beyond grain yield, genomic prediction models might be used to predict belowground and physiological traits.

## 2 Introduction

Despite their importance for water use and nutrient uptake, belowground traits have been largely neglected for crop improvement as breeding efforts have predominantly targeted aboveground traits related to yield formation. Nevertheless, there is evidence that thousands of years of empirical selection have also indirectly reshaped the root system architecture of domesticated species, corroborating the importance of belowground traits for crop yield (de Dorlodot et al. 2007; Herder et al. 2010; Jia et al. 2019). In cereals, experimental results and crop simulation models (CSMs) have pointed out that genotypes with deeper root system architecture can cope with drought and heat stresses and increase grain yield (GY) in dry environments (Manschadi et al. 2008; Mu et al. 2015; Liu et al. 2017; Tao et al. 2017). For instance in durum wheat, contrasting root system architectures correlate with drought-intolerant and drought-tolerant genotypes showing higher GY under sub-optimal water regimes (El Hassouni et al. 2018). In this species, it has been shown that deeper root system architectures can increase GY from 16 to 35% in environments with limited soil moisture and from 9 to 24% in irrigated sites (El Hassouni et al. 2018). Similarly, in bread wheat narrower and deeper root system architectures with more branching at depth allow to provide greater access to soil moisture in environments experiencing terminal drought (Manschadi et al. 2008). In maize, it has been shown that the increase of root size improves nitrogen absorption and GY (Mu et al. 2015) and that a more efficient root system is more important than canopy architecture for determining plant growth rate and biomass accumulation (Hammer et al. 2009).

In cereals, the root system architecture of seedlings can be dissected into primary or seminal roots and nodal or secondary roots. While seminal roots develop first from the primordia of the embryo and grow out from the coleorhizae, the development of nodal roots begins at the tillering stage from the basal nodes of the crown (Wahbi and Gregory 1995). In bread and durum wheat, it has been shown that the seminal root number (SRN) and the seminal root angle (SRA), that is the angle measured between the first pair of seminal roots or between the two outmost seminal roots at the seedling stage, are two proxy traits that can predict the root system architecture at the adult stages (Manschadi et al. 2008; El Hassouni et al. 2018; Alahmad et al. 2019). For instance, reduced SRA and higher SRN in bread wheat seedlings correlate with drought-tolerant genotypes (e.g. Baxter, Babax, and Dharwar Dry, SeriM82), which exhibit a deeper and more compact root system architecture at the adult stages (Manschadi et al. 2008). In barley the assumption that SRA and SRN measured in seedlings are proxies of the root system architecture of mature plants has not been directly assessed, although in spring genotypes moderate correlations between these belowground traits and GY have been observed in field trials organized in 20 rainfed and irrigated site-by-season combinations (Robinson et al. 2018). Recently, phenotypic variation for SRA and SRN has been assessed in a large panel of spring barley and exploited to map loci that underlie these traits using genome wide association studies (Jia et al. 2019).

The evaporative demand or vapor pressure deficit (VPD) indicates the difference between the saturated and the actual vapor pressure of air at a given temperature and drives the transpiration rate (TR) of crops (Kholová et al. 2012). In field conditions, either soil drought or atmospheric drought, that is the combination of high temperatures and low humidity, does not allow crops to satisfy the required evaporative demand and climate change is expected to exacerbate these phenomena (Lobell and Gourdji 2012; Medina et al. 2019). CSMs have pointed out that, beyond the root system architecture, the TR response to VPD is an important physiological trait that might be targeted to cope with high evaporative demand and increase GY (Tao et al. 2009, 2017) as the reduction of water

flux to leaves for limiting TR at high levels of VPD is a water-saving strategy that imposes physiological trade-offs in leaf dehydration and senescence and allows crops to better manage soil moisture to overcome drought stress. While this water saving strategy might cause yield penalty when soil moisture is not a limiting factor, in sorghum and maize experimental evidence has shown that limiting TR at high evaporative demand can allow to increase GY in dry environments (Sinclair et al. 2005). As substantiated for belowground traits, TR response to VPD is a key physiological trait that can serve as proxy trait for drought tolerance (Schoppach and Sadok 2012, 2013; Schoppach et al. 2016). In durum wheat, the variation of TR response has allowed to identify at least two different sets of genotypes showing linear and segmented trends of TR in response to VPD and interestingly, these different responses have been correlated with different performances in terms of GY and biomass production in rainfed and irrigated field trials (Medina et al. 2019). Similarly, in sorghum and chickpeas, closing of stomata for limiting TR has been correlated with genotypes that have a better ability to retain soil moisture and contribute to yield formation under drought stress (Devi et al. 2015; Sivasakthi et al. 2017; Medina et al. 2019). While phenotypic diversity in TR response to VPD has been widely assessed in bread and durum wheat (Schoppach and Sadok 2012; Schoppach et al. 2017), the knowledge of this trait in barley has lagged behind and to date its natural variation has been assessed in a limited panel of 25 wild barley and in one cultivar, which corroborates the existence of diversity for this trait in barley germplasm (Sadok and Tamang 2019).

Genomic prediction (GP) aims to regress genome-wide single nucleotide polymorphisms (SNPs) or other types of DNA markers on phenotypes of individuals to simultaneously predict their effects (Meuwissen et al. 2001). The population of individuals having both phenotypic and genotypic information is named training population (TP) and is used for constructing predictive models, which allows to compute "Genomic Estimated Breeding Values" (GEBVs) in individuals for which only genotyping information is available (Desta and Ortiz 2014). Typically, the predictive models used in GP require to regress a number of predictors (DNA markers) that greatly exceeds the number of observations or phenotypes and several parametric and non-parametric models have been proposed to deal with overfitting and the 'large *p*, small *n*' problem (Meuwissen et al. 2001; Jannink et al. 2010; Pérez and de los Campos 2014) as in these conditions the estimation of marker effects using ordinary least squares method is not practicable. Unlike methods based on whole genome regression of markers, the genomic best linear unbiased prediction (GBLUP) method treats genomic values of individuals as random effects in a linear mixed model and uses a genomic relationship matrix based on DNA marker data to compute GEBVs (VanRaden 2008; Wang et al. 2018). To date, in plant breeding GP has been mainly applied for improving GY (Crossa et al. 2017), but this methodology offers the possibility to predict other traits of agricultural interest that cannot be easily scored (e.g. belowground and physiological traits correlated with abiotic stress tolerances). For instance, GP models have been fitted to seed size (Nielsen et al. 2016), thousand grain weight, number of grains per $m^2$, grain plumpness (Bhatta et al. 2020), straw breaking and lodging (Tsai et al. 2020) beta-glucan and grain protein content (Bhatta et al. 2020) and starch (Tsai et al. 2020).

In plant and animal breeding, categorical data often arise when traits are scored either as dichotomous phenotypes (e.g., presence or absence) or ordered categorical data varying on discrete values (e.g., ordered scales indicating increasing resistance to a pathogen). Similarly, count data arise when the trait of interest is the sum of discrete quantities that can take only integer values (e.g., the number of tillers per plant or the number of seminal roots in seedlings). These data are often analyzed assuming that the number of categories is sufficiently large, and that consequently they can be approximate to a normal distribution, but several studies have shown that significant bias is observed with low number of samples and discrete categories (Montesinos-López et al. 2015a). Treating count or ordinal categorical data as a random variable that follows a continuous distribution violates several

assumptions of standard GP linear models as already substantiated in other studies (Montesinos-López et al. 2015b, a, 2016). To overcome these limitations, GP based on generalized linear mixed models (GLMM) have been developed and fitted to dichotomous and ordinal categorical traits. While in standard linear models the response variable is modelled as the sum of the explanatory variables and a probability assumption about the residual is made, in GLMM the response variable is connected to the explanatory variables via a link function, that is a function that maps explanatory variables to the observations. Compared to standard GP linear models, it has been empirically shown that GP models based on GLMM allow to increase the predictive ability for count data and ordinal categorical traits especially when the trait heritability is low (González-Recio and Forni 2011; Villanueva et al. 2011; Wang et al. 2013; Montesinos-López et al. 2015a, 2016).

In the present study, we examined the diversity and distribution of belowground and physiological traits in an eight-way MAGIC population of winter barley and in its founder parents (Puglisi et al. 2021). We correlated phenotypic data of SRA, SRN and TR along with GY obtained in different site-by-season combinations to re-assess the relevance of belowground and physiological traits of seedlings for predicting drought-tolerant and drought-intolerant lines. Leveraging on phenotypic and genotypic information we fitted and cross-validated different GP models including different sets of linear predictors and showed that these models can successfully predict SRA, SRN and TR and might pave the way for underpinning ideotype breeding and mining large plant collections.

## 3 Material and methods

### 3.1 Plant materials and genotyping

The MAGIC population examined in this study along with the set of 90 MAGIC lines used for genome-enabled predictions has been extensively described elsewhere (Puglisi et al. 2021). Particularly, this latter set of 90 MAGIC lines used in the present study (**Supplementary Table 1 and 2**) corresponds to the TP previously assembled for fitting multi-environment GP models and genotyped using the barley 50K SNP chip (Puglisi et al. 2021).

### 3.2 Phenotyping of plant material

The set of MAGIC lines was phenotyped in the following site-by-season combinations to examine GY and heading date (HD): Fiorenzuola d'Arda (Italy, 2016, 2017, 2018, 2019) at CREA-Centro di Genomica e Bioinformatica (44°55'39.0"N 9°53'40.6"E, 78 $m$ above sea level), Marchouch (Morocco, 2016, 2019) at ICARDA Experimental station (33°36'43.5"N 6°42'53.0"W, 390 $m$ above sea level), Adana (36°59'52.9"N 35°20'28.0"E, 24 $m$ above sea level) and Konya (37°53'37.9"N 32°37'26.0"E, 1005 $m$ above sea level) (Turkey, 2019). These data, excluding phenotypic data collected in Marchouch during the growing season 2018-2019, are part of the data set previously analyzed (Puglisi et al. 2021). All these experiments were conducted following local management practices, except for field trials organized in Fiorenzuola d'Arda during 2017-2018 and 2018-2019 growing seasons as they were conducted using two different levels of nitrogen fertilization as previously described (Puglisi et al. 2021). All field trials were analyzed using an alpha-lattice experimental design with two replicates and adjusted means of GY were computed using a mixed linear model considering genotypes as random effects (Puglisi et al. 2021). Best linear unbiased predictions (BLUPs) of genotypic effects were computed per environment and per site-by-season-by-management combination including HD and both HD and year as covariates, respectively. Mixed linear models for analyzing field trial data were implemented in lme4 package (Bates et al. 2015) along with R 4.0.3 (Core R Team 2019). The adjusted means of GY were used to study genotype × environment (GE) interaction for delineating clusters of mega-environments using genotype plus

genotype × environment (GGE) biplot analyses (Yan et al. 2000; Yan and Holland 2010) implemented in GGEBiplots package and for assessing correlation with physiological and belowground traits.

SRA and SRN were phenotyped using the clear pot method (Richard et al. 2015; Robinson et al. 2016) at ICARDA's physiology laboratory under controlled temperature and humidity according to the original protocol (Richard et al. 2015) and using transparent ANOVApot pots (Anovapot Pty Ltd, Brisbane, QLD, Australia, www.anovapot.com/php/anovapot.php) with a diameter of 200 $mm$, height of 190 $mm$ and a volume of 4 $L$. To phenotype SRA and SRN, the set of 90 MAGIC genotypes was randomized in 45 pots using 12 replicates per genotypes. In each pot 24 random genotypes were sowed at a distance of 2.5 $cm$ from each other and at a depth of circa 1.5-2 $cm$, positioning the embryo towards the bottom of the pot to support the correct root development. After sowing, pots were irrigated and placed inside black container pots, in a lightless growing chamber with controlled climate conditions with data logger at 19 °C and 70% of relative humidity (Type TGU-4550, Gemini Data Loggers, UK). Five days after sowing twelve images of SRA per genotype were acquired with a digital camera, which were subsequently analyzed using ImageJ software (https://imagej.nih.gov) (Schneider et al. 2012) to measure the angle between the first pairs of seminal roots according to published protocols (Richard et al. 2015). Eleven-day days after sowing seedlings were carefully removed from soil and SRNs were manually counted in the same seedlings used for measuring SRA. The incomplete block design used for phenotyping SRA was analyzed with the following linear model:

$$y_{ijk} = 1\mu + Rep_i + Pot_j + Gen_k + e_{ijk} \qquad \textbf{Equation 1}$$

where $y_{ijk}$ is the response variable, that is the raw SRA data, μ is the general mean, $Rep_i$ is the fixed effect of the $i^{th}$ replicate, $Pot_j$ is the random effect of the $j^{th}$ pot, $Gen_k$ is the fixed effect of the $k^{th}$ genotype and $e_{ijk}$ is the error associated to each response. The adjusted means of SRA and their 95% confidence interval were estimated in R 4.0.3 statistical (Core R Team 2019) using lme4 package (Bates et al. 2015) and used to fit censored GP models and compute broad sense heritability ($H^2$) and for seeking correlations with other belowground and physiological traits.

The TR response to VPD was examined at ICARDA's physiology laboratory under controlled temperature and humidity. This experiment was designed randomizing the 90 MAGIC lines using three biological replicates per pot, using 2 $L$ pots with diameter and height of 104 and 200 $mm$, respectively. In each pot, plants were sown at a depth of circa 2 $cm$ and were uniformly irrigated every 2 days. At Zadoks stage 14 (ZADOKS et al. 1974), which was reached after 4-5 weeks after sowing, depending on the genotype, pots were irrigated until reaching the maximum water holding capacity of the substrate and subsequently closed with plastic bags and balls to limit evaporation. TR was measured under increasing VPD ranging from 0.4 to 5.2 $KPa$ in a greenhouse with controlled conditions (temperature and humidity) accurately monitored with data loggers (Type TGU-4550, Gemini Data Loggers, UK). Phenotyping of plants and the computation of TR under increasing VPD conditions was carried out following published protocols (Fletcher et al. 2007; Sadok and Sinclair 2009a, b; Schoppach and Sadok 2012; Schoppach et al. 2017; Tamang and Sadok 2018; Sadok and Tamang 2019). Briefly, to determine the amount of water loss by transpiration, the weight of pots was measured using an electronic balance with a resolution of 0.1 $g$ (KB Kern 573, Kern & Sohn GmbH, Balingen, Germany), at seven values of VPD. For calculating evapotranspiration, the weight of five empty pots was measured seven times and at the end of the experiment, plants leaf area, temperature and humidity were immediately collected using leaf area meters (LI-3000C, LiCor Biosciences, Lincoln, Nebraska, USA) and data loggers (Type TGU-4550, Gemini Data Loggers, UK). The VPD applied to plants was computed as a function of temperature and humidity and TR

measures in $mg_{H2O}m^{-2}s^{-1}$ was normalized as function of total leaf area and time in which pot weight was collected. TR data were regressed on VPD values using linear (**Equation 2**) and segmented models (**Equations 2 and 3**) (Medina et al. 2019).

$$y_1 = S_1x + I_1 \quad \textbf{Equation 2}$$

$$y_2 = S_2x + I_2 \quad \textbf{Equation 3}$$

where $S$ and $I$ point out slope and intercept, respectively. These models were computed using "segmented" package (V. R. M. Muggeo 2008; Muggeo 2017) implemented in R 4.0.3 statistical (Core R Team 2019). Goodness-of-fit based on R-squared values was used as main criterion to select the best model and classify genotypes in a first group of lines showing linear trend of TR and a second group showing a segmented trend.

Raw data of TR measured at 2.7 $KPa$ of VPD were analyzed using the following linear model:

$$y_{ik} = 1\mu + Rep_i + Gen_k + e_{ik} \quad \textbf{Equation 4}$$

where $y_{ik}$ are the raw TR values, $\mu$ is the general mean, $Rep_i$ is the effect of the $i^{th}$ replicate, $Gen_k$ is the fixed effect of the $k^{th}$ genotype and $e_{ik}$ is the error associated to each response. The adjusted means of TR measured at low levels of VPD along with their 95% confidence intervals were estimated in R 4.0.3 statistical (Core R Team 2019) using lme4 package (Bates et al. 2015) and used to fit censored GP models and compute broad sense heritability ($H^2$).

### 3.3 Descriptive and correlation analysis

Variation in GY and SRA in the panel MAGIC lines was investigated computing maximum and minimum values, mean, median and standard deviation. This descriptive analysis was computed using "metan" package (Olivoto and Lúcio 2020) implemented in R 4.0.3 statistical (Core R Team 2019). The adjusted means of GY and SRA, along with SRN and TR measured at a VPD of 2.7 $KPa$ were analyzed and correlated each other. Two different types of correlations were applied on the basis of variable type: Pearson's correlation coefficient was applied to compute correlation between continuous traits (GY, SRA and TR measured at a VPD of 2.7 $KPa$), while polyserial correlations were computed to measure correlations between continuous and categorical variables (Drasgow 2006). These correlation analyses were computed using "polycor" package (Fox 2019) implemented in R 4.0.3 statistical (Core R Team 2019).

### 3.4 GP models fitted to SRN

In the present study three different GP models were fitted to SRN combining phenotypic data with genotypic information obtained with the Barley 50k SNP chip (Puglisi et al. 2021). For this trait, GP models were fitted following two different hypotheses. Firstly, we assumed that SRN varies as an ordinal discrete variable that indicates the performance of plants at the adult stage under nitrogen or water deficiency and for this aim threshold genomic best linear unbiased predictor (TGBLUP) models and extended TGBLUP models were fitted.

Formal presentation of the model theory of GP for ordinal discrete data was disserted elsewhere (Montesinos-López et al. 2015a). Here, we shortly introduce the TGBLUP models used in the present study for implementing GP. For SRN, we assumed that the ordinal response variable $y_{ik}$, that is the number of observed seminal roots, can take $C$=7 mutually exclusive $c$ values, where $i$ indicates the genotype, $k$ points out the number of replicates and $c$ takes values equal to the number of observed seminal roots observed in the MAGIC population, that is $c = 2, 3, 4, 5, 6, 7, 8$. Moreover, we supposed that the ordinal response variable $y_{ik}$ follows a multinomial distribution with parameters $N_{ik}$ and $\pi_{ik(c=2)}, \pi_{ik(c=3)}, \pi_{ik(c=4)}, \pi_{ik(c=5)}, \pi_{ik(c=6)}, \pi_{ik(c=7)}, \pi_{ik(c=8)}$, that is $(y_{ik(c=2)}, y_{ik(c=3)}, y_{ik(c=4)}, y_{ik(c=5)}, y_{ik(c=6)}, y_{ik(c=7)}, y_{ik(c=8)}) \sim MULTINOMIAL(N_{ik}, \pi_{ik(c=2)}, \pi_{ik(c=3)}, \pi_{ik(c=4)}, \pi_{ik(c=5)}, \pi_{ik(c=6)}, \pi_{ik(c=7)}, \pi_{ik(c=8)})$

where $N_{ik}$ points out the number of observation and $\pi_{ik(c=2)}, \pi_{ik(c=3)}, \ldots \pi_{ik(c=8)}$ point out the probabilities of getting values $c = 2, 3, \ldots 8$ in the $i^{th}$genotype in the $k^{th}$replicate. Threshold models assume that $y_{ik}$ is generated from an underlying continuous random variable $l_{ik}$, having a normal distribution, which is called latent "liability" variable (Sorensen et al. 1995; Montesinos-López et al. 2015a) and imply that for $C$ ordinal and mutually exclusive categories the existence of $C - 1 = 6$ unknown $\gamma$ thresholds that must be estimated such as $\gamma_{min} < \gamma_1 < \gamma_2 < \gamma_3 .. < \gamma_{max}$, with $\gamma_{min} = -\infty$ and $\gamma_{max} = +\infty$. In threshold models, values of $l_{ik}$ are mapped to the ordinal categorical response according to the following conditions:

$$y_{ik} = \begin{cases} 2 & if \quad \gamma_{min} < l_{ik} < \gamma_1 \\ 3 & if \quad \gamma_1 < l_{ik} < \gamma_2 \\ 4 & if \quad \gamma_2 < l_{ik} < \gamma_3 \\ & \ldots\ldots . \ldots\ldots\ldots\ldots\ldots \\ 8 & if \quad \gamma_6 < l_{ik} < \gamma_{max} \end{cases}$$

In these models the link function relating linear predictors with the probability of observing data is the cumulative probit $\Phi(.)$, that is the cumulative distribution function of a standard normal distribution and $\Phi^{-1}$ is the corresponding inverse function. Consequently, threshold models are specified with $C - 1$ linear predictors $\eta_{ikc}$ as follows:

$$\eta_{ik(c=2)} = \Phi^{-1}\left(\pi_{ik(c=2)}\right) = \gamma_1 - X_{ik}^T\beta - Z_{ik}^T u$$

$$\eta_{ik(c=3)} = \Phi^{-1}\left(\pi_{ik(c=2)} + \pi_{ik(c=3)}\right) = \gamma_2 - X_{ik}^T\beta - Z_{ik}^T u$$

$$\eta_{ik(c=4)} = \Phi^{-1}\left(\pi_{ik(c=2)} + \pi_{ik(c=3)} + \pi_{ik(c=4)}\right) = \gamma_3 - X_{ik}^T\beta - Z_{ik}^T u$$

$$\eta_{ik(c=5)} = \Phi^{-1}\left(\pi_{ik(c=2)} + \pi_{ik(c=3)} + \pi_{ik(c=4)} + \pi_{ik(c=5)}\right) = \gamma_4 - X_{ik}^T\beta - Z_{ik}^T u$$

$$\eta_{ik(c=6)} = \Phi^{-1}\left(\pi_{ik(c=2)} + \pi_{ik(c=3)} + \pi_{ik(c=4)} + \pi_{ik(c=5)} + \pi_{ik(c=6)}\right) = \gamma_5 - X_{ik}^T\beta - Z_{ik}^T u$$

$$\eta_{ik(c=7)} = \Phi^{-1}\left(\pi_{ik(c=2)} + \pi_{ik(c=3)} + \pi_{ik(c=4)} + \pi_{ik(c=5)} + \pi_{ik(c=6)} + \pi_{ik(c=7)}\right) = \gamma_6 - X_{ik}^T\beta - Z_{ik}^T u$$

where $X_{ik}^T$ is a known row incidence vectors of fixed effects, $Z_{ik}^T$ is a known row incidence vectors of random effects, $\beta$ points out the vector of fixed effects and $b$ is the vector of random effects. The probabilities $\pi_{ikc}$ are linked to the linear predictors $\eta_{ikc}$as follows:

$$\pi_{ik(c=2)} = \Phi\left(\pi_{ik(c=2)}\right), \pi_{ik(c=3)} + \pi_{ik(c=2)} = \Phi\left(\pi_{ik(c=3)}\right), \ldots\ldots, \pi_{ik(c=7)} + \pi_{ik(c=6)} + \pi_{ik(c=5)} + \pi_{ik(c=4)} + \pi_{ik(c=3)} + \pi_{ik(c=2)} = \Phi\left(\pi_{ik(c=7)}\right)$$

As mentioned above, threshold models assume that the latent and normally distributed variable $l_{ik}$ generates the observed $C$ categories as follows:

$$l_{ik} = X_{ik}^T\beta + Z_{ik}^T u + e_{ik} \qquad \textbf{Equation 5}$$

where the error terms $e_{ik}$ are independent and identically distributed and follow a normal distribution with mean 0 and standard deviation equals to 1, that is $e_{ik} \sim N(0,1)$. In the present study, different combinations of linear predictors including replicates, lines, markers and first order epistatic effects, were incorporated in $X_{ik}^T$ and $Z_{ik}^T$ for fitting five extended threshold models (**Table 1**), which were already substantiated and described in other studies (Jarquín et al. 2014; Montesinos-López et al. 2015a).

**Table 1: Summary of the linear predictors incorporated in the GBLUP and TGBLUP models used to analyze SRN, SRA and TR**. R, replicate; L, line; G, marker covariates; GxG, additive x additive epistasis.

| Model | Main Effects | | | Interaction |
|---|---|---|---|---|
| | R | L | G | GxG |
| SRN-Model 1<br>SRA-Model 1<br>TR-Model 1 | x | x | | |
| SRN-Model 2<br>SRA-Model 2<br>TR-Model 2 | x | | x | |
| SRN-Model 3<br>SRA-Model 3<br>TR-Model 3 | x | | x | x |
| SRN-Model 4<br>SRA-Model 4<br>TR-Model 4 | x | x | x | |
| SRN-Model 5<br>SRA-Model 5<br>TR-Model 5 | x | x | x | x |

The resulting five models include the following sets of linear predictors:

**SRN-Model 1**: $l_{ik} = R_k + L_i + \varepsilon_{ik}$ **Equation 6**

**SRN-Model 2**: $l_{ik} = R_k + g_i + \varepsilon_{ik}$ **Equation 7**

**SRN-Model 3**: $l_{ik} = R_k + g_i + g_{Ai} + \varepsilon_{ik}$ **Equation 8**

**SRN-Model 4**: $l_{ik} = R_k + L_i + g_i + \varepsilon_{ik}$ **Equation 9**

**SRN-Model 5**: $l_{ik} = R_k + L_i + g_i + g_{Ai} + \varepsilon_{ik}$ **Equation 10**

where $l_{ik}$ is the latent "liability" variable of $k^{th}$replicates in the $i^{th}$ line. SRN-Model 1 includes $R_k$, which is the fixed effect of $k^{th}$ replicates and $L_i$ that is the random effect of the $i^{th}$ line supposed to be independent and normally distributed as $L_i \sim N(0, \sigma_L^2)$. SRN-Model 2 includes $g_i$, which points out the additive genetic value of the $i^{th}$ line, that is $g_i = \sum_{n=1}^{p} x_{in} b_n$, where $x_{in}$ is the genotype of the $i^{th}$ line at marker $n$ and $b_n$ is the corresponding effect of marker $n$. The vector of additive genetic value $\boldsymbol{g} = (g_1, g_2, g_3 \ldots \ldots g_i)$ is supposed to be normally distributed as $\boldsymbol{g} \sim N(0, G\sigma_g^2)$ with mean 0 and variance-covariance structure $G\sigma_g^2$, where $\sigma_g^2$ points out the additive genetic variance $\sigma_g^2$ and **G** is the genomic marker relationship matrix (VanRaden 2008). SRN-Model 3 extends SRN-Model 2 including first order multiplicative epistatic effects $\boldsymbol{g_A} = (g_{A1}, g_{A2}, \ldots g_{Ai})$, which are modelled as $\boldsymbol{g_A} \sim N(0, G_A\sigma_{g_A}^2)$, that is the vector of epistatic effects follows a normal distribution with mean 0 and epistatic additive x additive genetic variance $\sigma_{gA}^2$ (Montesinos-López et al. 2015a). Finally, SRN-Model 4 includes $R_k, L_i$ and $g_i$ as linear predictors while SRN-Model 5 extends SRN-Model 4 including $g_{Ai}$ effects. In the present study the aforementioned threshold models were implemented in a Bayesian framework using BGLR package (Pérez and De Los Campos 2014) in R 4.0.3 statistical (Core R Team 2019) using default prior distributions and modifying codes published in other studies (Montesinos-López et al. 2015a).

Secondly, we handled SRN as count data for predicting this trait *per se*, fitting five log-normal GP models based on GBLUP (Montesinos-López et al. 2015b). In these five log-normal model the response variable is the logarithm of SRN, that is $\log(y_{kj})$ and were fitted using the same sets of

linear predictors ($R_k, L_j, g_j, g_{Aj}$) described for the five extended TGBLUP models (**Table 1**). In these models $R_k, L_j, g_j, g_{Aj}$ follow the same distributions defined for the extended TGBLUP models except for the error terms $\varepsilon_{\text{ik}}$ of $K^{th}$ replicates in $i^{th}$ line, which in these models is $\varepsilon_{ik} \sim N(0, \sigma_e^2)$, that is the residuals are independent and normally distributed with mean 0 and variance $\sigma_e^2$. Like TGBLUP models, log-normal models were implemented using BGLR package (Pérez and De Los Campos 2014) in R 4.0.3 statistical (Core R Team 2019).

### 3.5 Genomic prediction models fitted for SRA and TR

The five sets of linear predictors used in the extended TGBLUP models (**Table 1**) were used for predicting SRA and TR measured at a VPD of 2.7 KPa, using the following models:

**SRA-Model 1**: $y_i = L_i + \varepsilon_i$ **Equation 11**
**SRA-Model 2**: $y_i = g_i + \varepsilon_i$ **Equation 12**
**SRA-Model 3**: $y_i = g_i + g_{Ai} + \varepsilon_i$ **Equation 13**
**SRA-Model 4**: $y_i = L_i + g_i + \varepsilon_i$ **Equation 14**
**SRA-Model 5**: $y_i = L_i + g_i + g_{Ai} + \varepsilon_i$ **Equation 15**

**TR-Model 1**: $y_i = L_i + \varepsilon_i$ **Equation 16**
**TR-Model 2**: $y_i = g_i + \varepsilon_i$ **Equation 17**
**TR-Model 3**: $y_i = g_i + g_{Ai} + \varepsilon_i$ **Equation 18**
**TR-Model 4**: $y_i = L_i + g_i + \varepsilon_i$ **Equation 19**
**TR-Model 5**: $y_i = L_i + g_i + g_{Ai} + \varepsilon_i$ **Equation 20**

where $y_i$ is the adjusted mean of SRA or TR, $\epsilon_i$ is the error term of the $i^{th}$ measurement with $\epsilon_i \sim N(0, \sigma_\varepsilon^2)$, that is that the errors are independent and identically distributed with mean 0 and variance $\sigma_\varepsilon^2$. In these models, the linear predictors $L_j, g_j, g_{Aj}$ follow the same distribution defined for TGBLUP models. These five extended GBLUP models (**SRA-Model 1-5, TR-Model 1-5**) were implemented using BGLR package (Pérez and De Los Campos 2014) in R 4.0.3 statistical (Core R Team 2019) as censored data described with the following interval

$$a_i < y_i < b_i$$

where $y_i$ is the adjusted mean of SRA or TR computed as BLUE, $a_i$ is the lower bound estimate of $y_i$ computed as the difference between $y_i$ and 2 standard deviations (SD) and $b_i$ is the upper bound estimate of $y_i$ computed as the sum of $y_i$ with 2 SD.

### 3.6 Cross-validation of GP models

For the extended TGBLUP models, leave-one-out cross-validation was carried out and predictive ability was estimated using both Brier Score (BS) and the proportion of cases correctly classified (PCCC) (BRIER 1950; Montesinos-López et al. 2015a, 2020). BS was computed as follows:

$$\boldsymbol{BS = n^{-1} \sum_{i=1}^{n} \sum_{c=1}^{g} (\hat{\pi}_{ic} - d_{ic})^2} \quad \textbf{Equation 21}$$

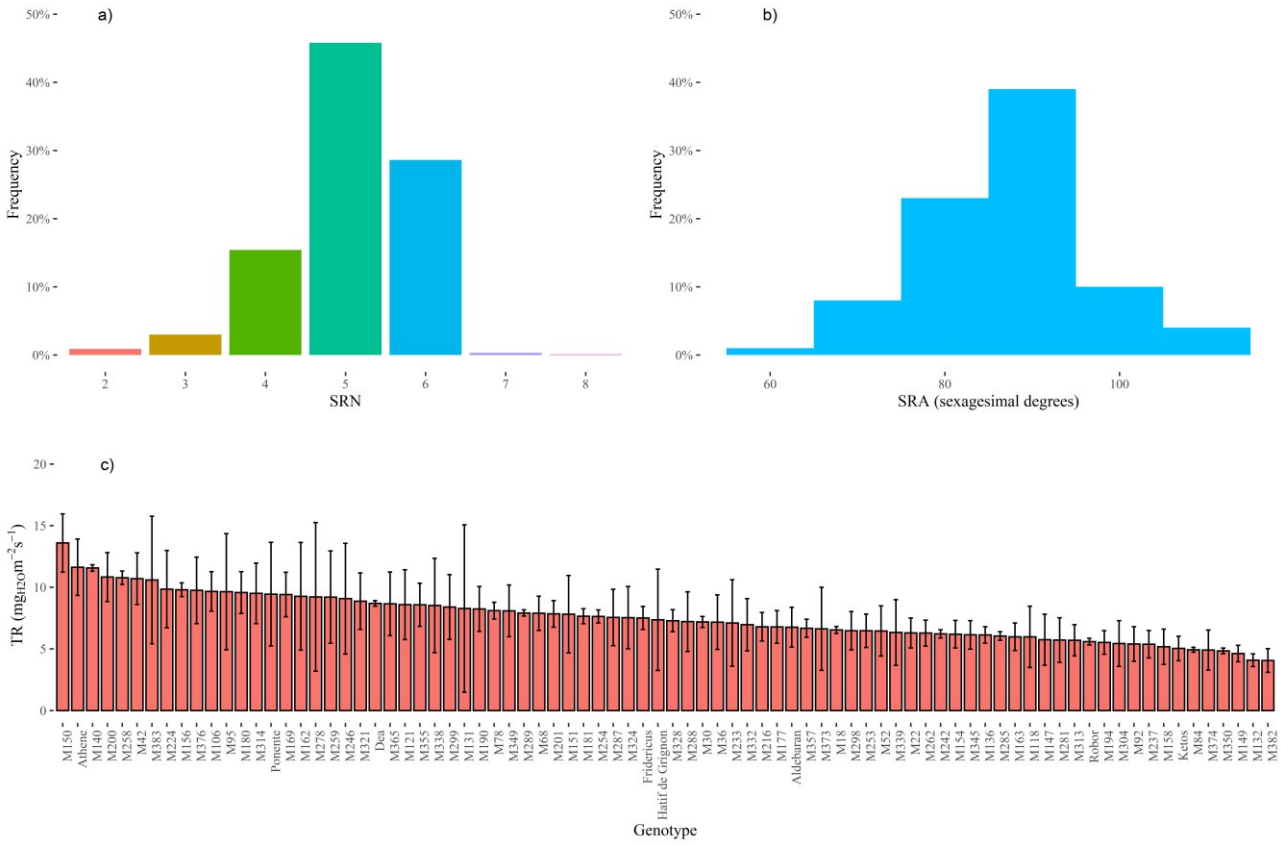
where $\boldsymbol{(\hat{\pi}_{ic} - d_{ic})^2}$ is the average square difference between $\hat{\pi}_{ic}$ predictions and $d_{ic}$ classes for observation *i* into category *c*. BS obtained with **Equation 21** was divided by two in order to have a range that varies from 0 to 1 (Brier, G. W. 1950; Montesinos-López et al. 2015a). For the other models used in the present study (extended GBLUP and log-normal models), the predictive accuracy of GP models was calculated as the Pearson's correlation coefficient between GEBVs and the corresponding adjusted means of the trait (SRA, TR measured at a VPD of 2.7 KPa). Unlike the Pearson's correlation coefficient used for the extended GBLUP models for SRA and TR, lower values of BS point out higher predictive ability of the models, while higher values of BS point out lower predictive ability of models.

# 4 Results

## 4.1 Phenotypic distribution and analysis of belowground and physiological traits in the barley MAGIC population

To assess the phenotypic variability of belowground traits, the panel of MAGIC lines was phenotyped for SRA and SRN at the seedling stage under controlled conditions. This analysis showed that SRN greatly varies in this population as it ranges between 2 and 8 with a mean of 5 roots and a standard deviation of 0.84 (**Figure 1A**; **Supplementary Table 1**). A linear mixed model was fitted to analyze the raw measurements of SRA following the adopted experimental design to compute the adjusted means and the corresponding 95% confidence intervals of SRA using best linear unbiased estimators (BLUEs). This analysis indicated that the phenotypic distribution of SRA ranges from 57.75° (genotype "M18") to 106.40° (genotype "M324") with an average value of 86.51° and SD of 9.6° (**Figure 1B**; **Supplementary Table 2**). Both belowground traits exhibit a bell-shaped distribution (**Figure 1A and Figure 1B**) and particularly, for SRA Shapiro-Wilk normality test shows that the null hypothesis, that is that the adjusted means of SRA follow a normal distribution, cannot be rejected (*P*-values = 0.1844).

**Figure 1: Phenotypic distribution of SRA, SRN and whole-plant TR measured at high evaporative demand in the panel of MAGIC lines:** (a) Histogram of SRN counted in the MAGIC lines; (b) Histogram of the adjusted means of SRA measured in sexagesimal degrees; (c) Bar plot of TR measured at under high evaporative demand (2.7 KPa). Error bars point out the 95% confidence interval of TR values.



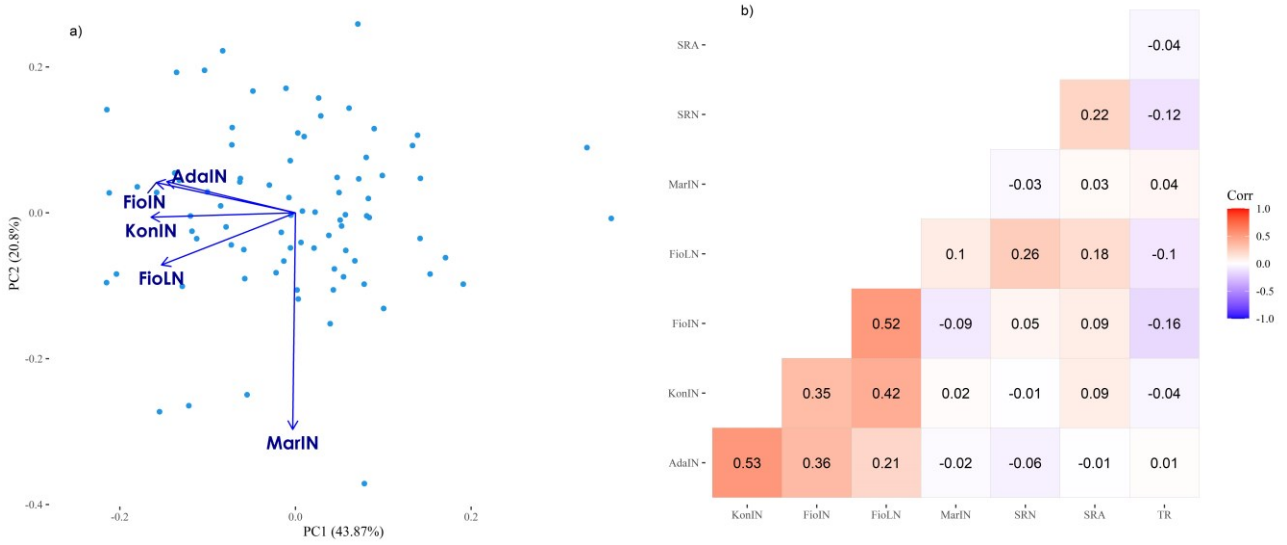

The whole-plant TR was measured in the set of MAGIC lines at the seedling stage using increasing VPD values ranging from 0.4 to 2.7 KPa. Regression of whole-plant TR on VPD values was carried out using linear and segmented models and R-squared was used as goodness-of-fit measure for model selection. Regression of whole-plant TR on VPD values shows a segmented response in a large

fraction of genotypes, while 28 MAGIC lines show a linear response of TR to increasing levels of VPD (**Supplementary Table 2**). Following the results showed in previous studies (Sadok and Tamang 2019), we investigated the variability of TR measured at high evaporative demand (2.7 KPa), which showed that the panel of MAGIC exhibits circa a five-fold variation of TR as measured values range from 4.05 (genotype "M382") to 13.6 (genotype "M150") $mg_{H2O}m^{-2}s^{-1}$ (**Figure 1C, Supplementary Table 2**).

## 4.2 Correlation of belowground and physiological traits with GY

To assess the level of correlation among site-by-season combinations, a genotype × environment (GGE) biplot analysis was carried using the adjusted means of GY (**Figure 2A**). This analysis indicated that AdaIN and FioIN show the highest environmental similarities, while compared to the remaining environments, MarIN is the most dissimilar one as already substantiated in other studies (Puglisi et al. 2021). For assessing whether belowground and physiological traits might contribute to determining yield formation, we correlated GY with SRN, SRA and TR measured at high evaporative demand computing Pearson's correlation coefficient for pairs of continuous traits (GY, SRA and TR) and polyserial correlation coefficient for pairs of continuous and discrete (SRN) traits (**Figure 2B**). This pairwise correlation analysis indicated that SRN shows a positive and moderate correlation with SRA (r = 0.22) and a negative correlation with the TR measured at high evaporative demand (VPD=2.7 KPa) (r = -0.12), that is MAGIC lines with higher SRN tend to transpire less at high VPD conditions (**Figure 2B**). Furthermore, TR under high evaporative demand showed a slight negative correlation with GY in FioIN (r = -0.16). GY in FioLN exhibited positive correlations with SRA and SRN, showing values of 0.18 and 0.26, respectively (**Figure 2B**). Overall, GY was positively correlated among KonIN, AdaIN and FioIN while no significant correlations were observed between MarIN and the remaining sites (KonIN, AdaIN, FioIN and FioLN) (**Figure 2B**).

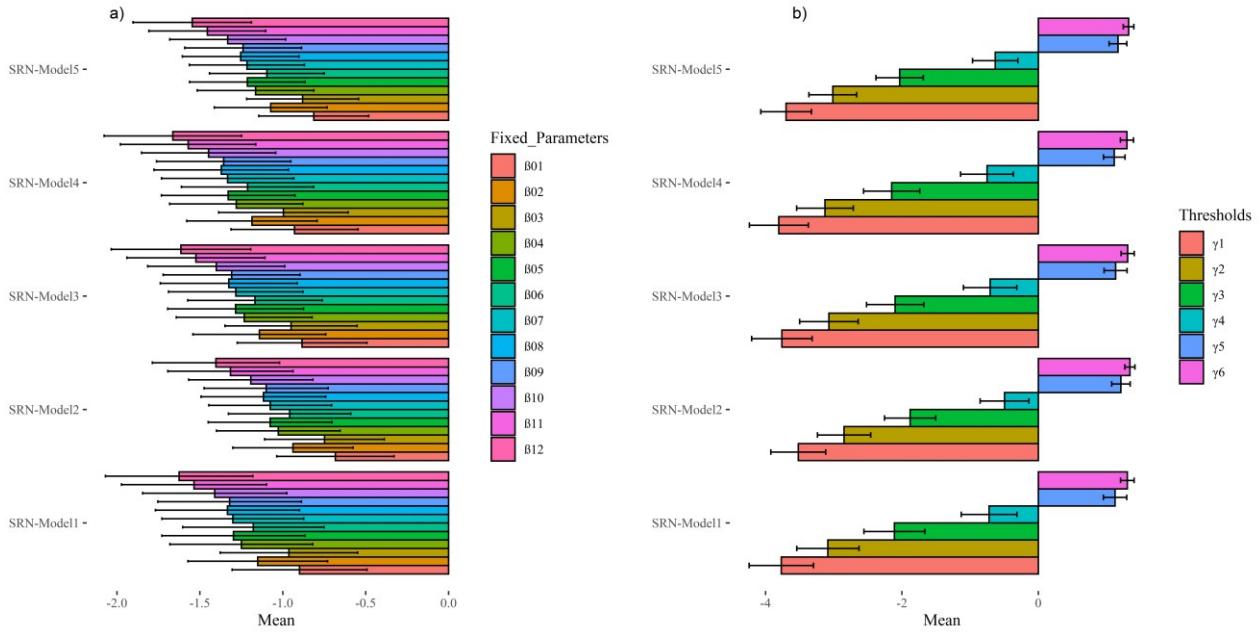
**Figure 2: GGE biplot of GY and pairwise correlations of SRA, SRN, GY and TR at high evaporative demand.** (a) The environment-vector view of the GGE biplot indicates similarities among test environments in discriminating the genotypes. (b) Depending on the trait distribution type (discrete or continuous), pairwise values indicate Pearson's correlation or polyserial correlation between GY, SRA, SRN and TR measured under high evaporative demand of VPD (2.7 KPa).



### 4.3 GP models based on proxy traits for nitrogen limiting conditions

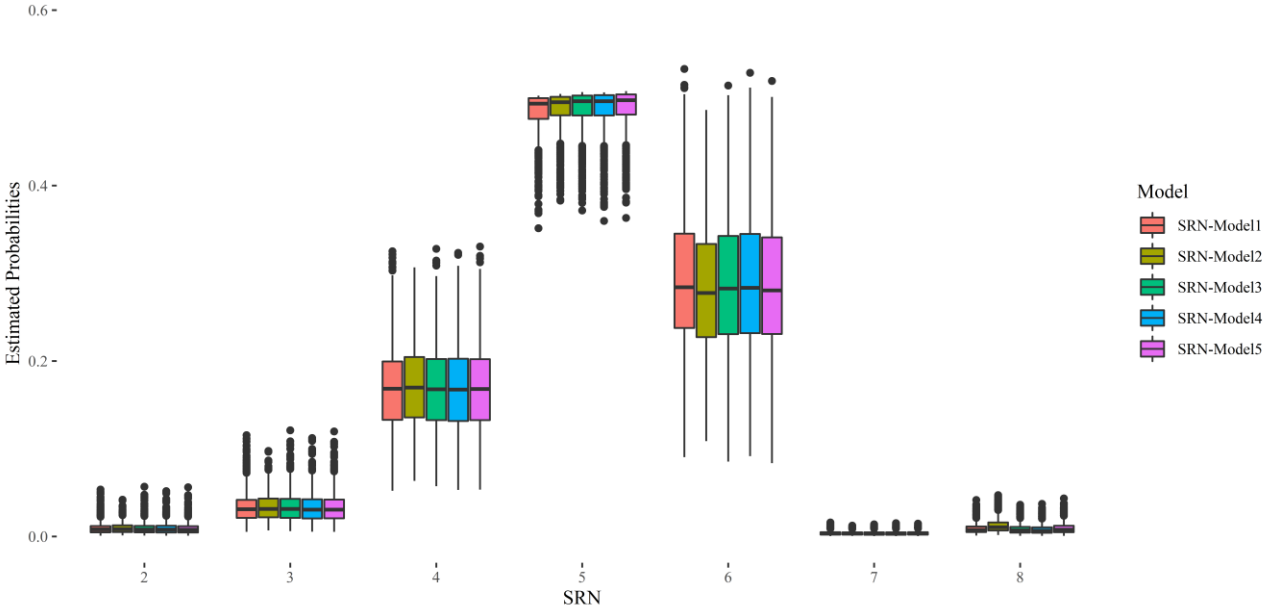
Here, following the moderate correlation observed between GY obtained with limiting nitrogen conditions and the number of seminal root (**Figure 2B**), we assumed that SRN might serve as a proxy trait for predicting the performance of MAGIC lines under nitrogen deficiency and consequently the number of seminal roots was analyzed as an ordinal categorical phenotype, that is we supposed that genotypes exhibiting less seminal roots are more sensitive to nitrogen deficiency and vice versa. Genotyping data of MAGIC (Puglisi et al. 2021) were combined with SRN counted in seedlings to fit five threshold GP models, which include the fixed effect of replicates (all models), the effect of lines (**SRN-Model 1**), the effect of molecular markers (**SRN-Model 2**), the effect of molecular markers and epistasis (**SRN-Model 3**), the effect of lines and molecular markers (**SRN-Model 4**) and the effect of lines, markers and epistasis (**SRN-Model 5**) (**Table 1**) (Montesinos-Lopes et al., 2015). **SRN-Model 2** represents a standard TGBLUP, while **SRN-Model 3**, **SRN-Model 4** and **SRN-Model 5** extend TGBLUP models for including line effects and epistasis (**Table 1**). The estimates of fixed effects and their 95% confidence interval, that is the effects of the twelve replicates used for phenotyping SRN, showed similar values in the five threshold models (**Figure 3A; Supplementary Table 3**). Similarly, the estimates of the six model thresholds ($\gamma_1, \gamma_2, .. \gamma_6$) showed similar values across the five threshold models **(Figure 3B; Supplementary Table 4**). Overall, the five threshold models considered in this study showed similar estimates of fixed effects and thresholds (**Figure 3; Supplementary Table 3; Supplementary Table 4**).

**Figure 3: Bar plots of the estimated parameters and thresholds of SRN in SRN-Models 1-5.** (a) Posterior mean and 95% confidence interval of the twelve fixed effects ($\beta_1, \beta_2, \ldots \beta_{12}$) estimated in the five TGBLUP models (b) Posterior mean and 95% confidence interval of threshold parameters ($\boldsymbol{\gamma_1, \gamma_2, .. \gamma_6}$). In both graphs, error bars point out the posterior 95% confidence interval of parameter values.




The probabilities for each ordinal categorical phenotype estimated in the five TGBLUP models for the whole data set are shown in **Figure 4**. This analysis shows that the average probabilities for category 5 (5 seminal roots) are circa 0.5 in the whole data set for all five models (**Figure 4**) followed by categories 6 and 4. Unlike the distribution estimated from raw data, these probability estimates consider the effect of replicates but, overall, show similar trends from the estimates obtained based on raw frequencies (**Figure 1A**).

**Figure 4: Boxplots of the estimated probabilities of SRN in SRN-Models 1-5.** In each boxplot, upper and lower hinges show the 25$^{th}$ percentiles of the estimated probabilities. Lines across boxes point out the average estimated probability.



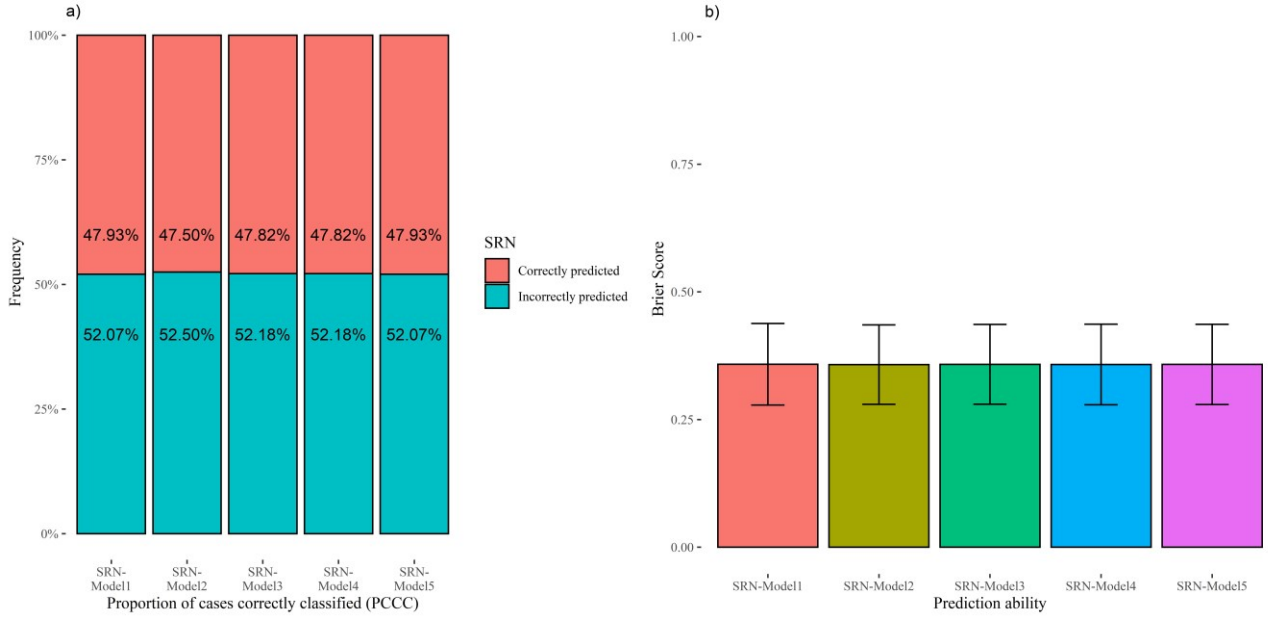

The analysis of the estimated variance components in the five TGBLUP models fitted to SRN data shows that overall, the total variance explained in the five TGBLUP models is 1.07 (**Table 2**). Molecular markers explain a variance of circa 0.05 in **SRN-Model 2**, and of circa 0.02 in **SRN-Model 3**, **4** and **5** (**Table 2**). Similarly, the variance explained by the additive x additive epistatic terms is 0.04 for the **SRN-Model 3** and 0.02 for the **SRN-Model 5**. Overall, the molecular markers explain a small fraction of the total variance observed for SRN in the five TGBLUP models used in the present study.

**Table 2: Estimated variance components of SRN-Model 1-5.** L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG points out the variance of additive x additive epistatic effects. Numbers between brackets point out the standard deviation of the estimated variances.

| Model | L | G | GxG | Error variance | Total Variance |
|---|---|---|---|---|---|
| **SRN-Model 1** | 0.07 (0.03) | | | 1 | 1.07 |
| **SRN-Model 2** | | 0.05 (0.02) | | 1 | 1.05 |
| **SRN-Model 3** | | 0.02 (0.01) | 0.04 (0.02) | 1 | 1.07 |
| **SRN-Model 4** | 0.04 (0.02) | 0.02 (0.01) | | 1 | 1.06 |
| **SRN-Model 5** | 0.02 (0.01) | 0.02 (0.01) | 0.02 (0.01) | 1 | 1.06 |

To assess the predictive ability of **SRN-Model 1-5**, leave-one-out (LOO) cross-validation was implemented to compute BS between predicted and observed categorical values and the proportion of cases correctly classified (PCCC) (Gianola and Schon 2016; Montesinos-López et al. 2020). Cross-validation analysis points out the PCCC is circa 52% for the five models considered in the present study (**Figure 5A; Supplementary Table 5**). Similarly, BS points out a high predictive ability as a value of circa 0.3 was estimated in the five models (**Figure 5B**).

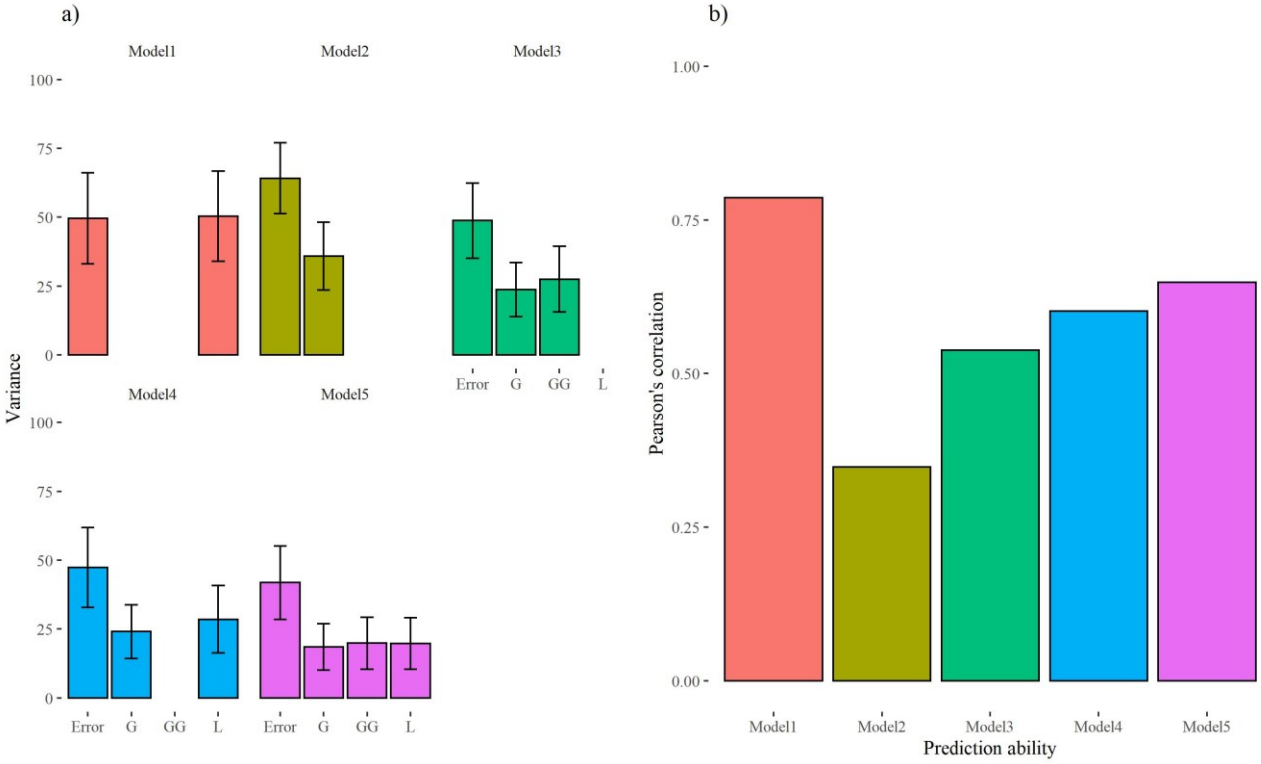
**Figure 5: Predictive ability of SRN-Model 1-5 fitted to SRN:** (a) Proportion of cases correctly classified using the five TGBLUP models; (b) Brier scores obtained from LOO cross-validation of the five TGBLUP models.




## 4.4 Genomic prediction of SRN using log-normal transformation of count data.

Beyond using SRN as a proxy trait for predicting drought tolerance, GP models for count data were fitted to predict this trait independently from its association with drought tolerance and performance under nitrogen deficiency at the mature stage, using log-normal models incorporating the same combinations of linear predictors included in the **Model 1-5** (**Table 1**). The analysis of variance components of these five log-normal models shows that "**Model 5**", which incorporate line (L), marker (G) and additive x additive epistatic interaction (GxG) effects, has a lower error variance compared to the other models considered in the present study and allows to better fit data (**Figure 6A; Supplementary Table 6**). The variance of GxG is 27.52 % for **Model 3** and 19.83 % for **Model 5** (**Figure 6A; Supplementary Table 6**). LOO cross validation points out that the predictive ability values of these models, measured using Pearson's correlation coefficient between predicted and observed data, ranges from 0.35 (**Model 2**) to 0.79 (**Model 1**), while **Model 3**, **Model 4** and **Model 5** show predictive ability values of 0.54, 0.60 and 0.65, respectively (**Figure 6B; Supplementary Table 6; Supplementary Figure 1**). Overall, this analysis indicates that for SRN, models that explicitly incorporate marker and interaction effects have higher efficiency in term of both predictive ability and variance components captured.

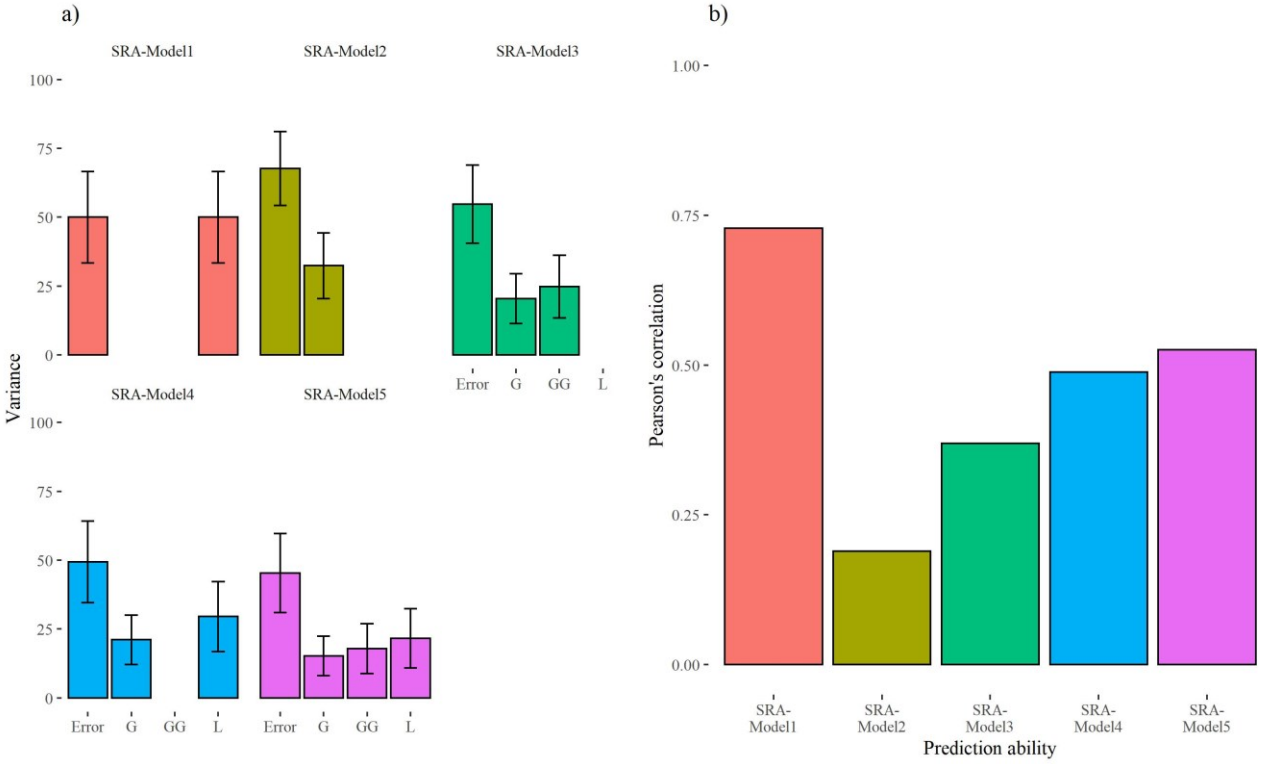
**Figure 6: Estimated variance components and prediction ability of log- normal models fitted to SRN.** (a) Bar plots indicated variances of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG and "Error" point out the variance of additive x additive epistatic effects and the residual variance, respectively. Error bars point out the 95% confidence interval of the estimated variances**;** (b) Bar plots of predictive ability values computed as Pearson's correlation between estimated and observed data using LOO cross validation strategy.




## 4.5 GP models for SRA

As with log-normal models, the GBLUP counterpart of the five threshold models used for predicting SRN were fitted to SRA. These five models incorporate the main effects and interactions used for SRN (**Table 1**) but assume that the response variable, that is SRA, is continuous and follows a normal distribution. The adjusted means of SRA were combined with 19,723 polymorphic SNPs detected in this panel of MAGIC population to fit these extended GBLUP models. The analysis of variance components of these five models showed that "**SRA-Model 5**", which incorporates line (L), marker (G) and additive x additive epistatic interaction (GxG) effects, has a lower error variance compared to the other models considered in the present study and allows to better fit data (**Figure 7A; Supplementary Table 7**). The variance of GxG is 27.79 % for **SRA-Model 3** and 17.81 % for **SRA-Model 5** (**Figure 7A; Supplementary Table 7**). LOO cross validation pointed out that the predictive ability values of these models, measured using Pearson's correlation coefficient between predicted and observed data, range from 0.19 (**SRA-Model 2**) to 0.73 (**SRA-Model 1**), while **SRA-Model 3**, **SRA-Model 4** and **SRA-Model 5** show predictive ability values of 0.37, 0.49 and 0.53, respectively (**Figure 7B; Supplementary Table 7; Supplementary Figure 2**). Like observed for log-normal models, this analysis indicates that for SRA, models that explicitly incorporate marker and interaction effects have higher efficiency in term of both predictive ability and variance components captured.

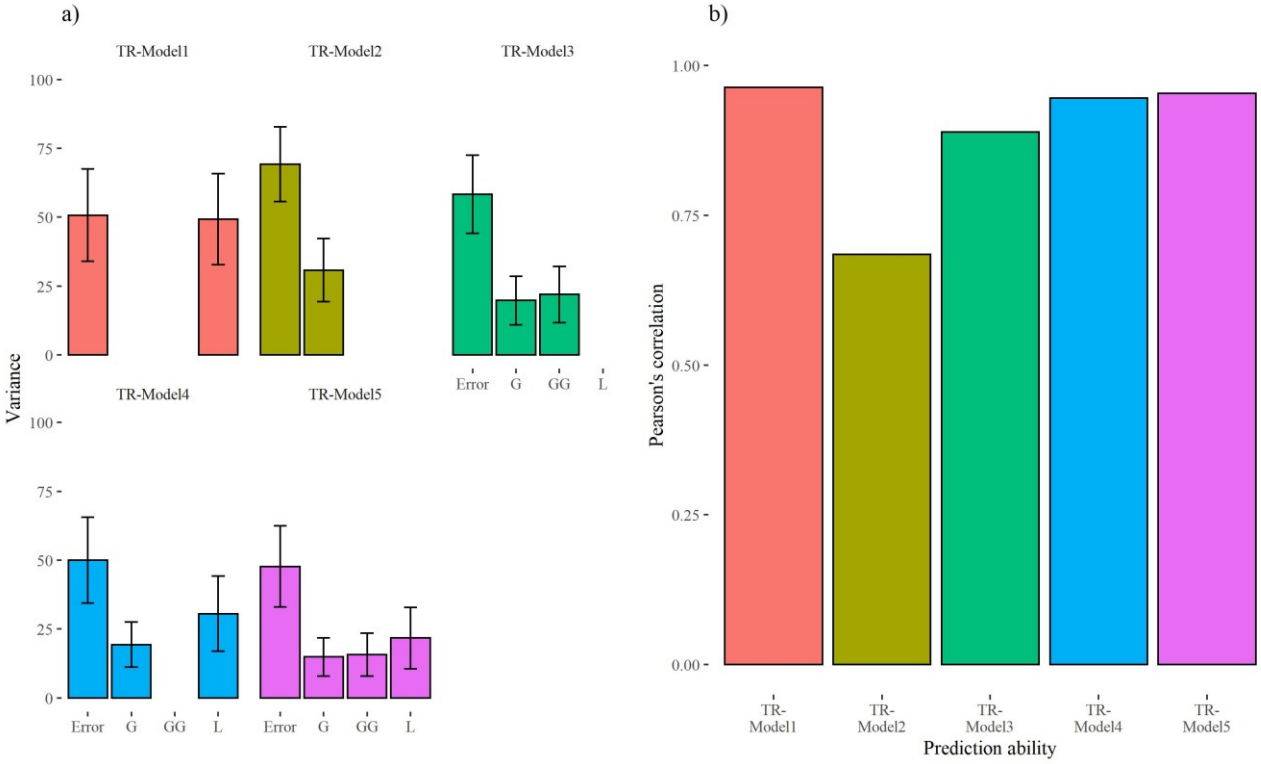
**Figure 7: Estimated variance components and prediction ability of the SRA-Models 1-5 used to predict SRA: (A)** Bar plots indicate the variances of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG and "Error" point out the variance of additive x additive epistatic effects and the residual variance, respectively. Error bars point out the 95% confidence interval of the estimated variances**; (B)** Bar plot of predictive ability values computed using Pearson's correlation between estimated and observed SRA.




## 4.6 Models for predicting TR under high evo-transpiration demand

The five combinations of linear predictors incorporated in GP models for SRA and SRN (**Table 1**) were used to predict TR at a VPD of 2.7 kPa. The analysis of variance components of these five models showed that "**TR-Model 5**", which incorporates line (L), marker (G) and additive x additive epistatic interaction (GxG) effects, has a lower error variance compared to the other models considered in the present study and allows to better fit data (**Figure 8A; Supplementary Table 8**). The variance explained by GxG is 21.93 % for **TR-Model 3** and 15.69 % for **TR-Model 5** (**Figure 8A; Supplementary Table 8**). LOO cross validation points out that the predictive ability values of these models, measured using Pearson's correlation coefficient between predicted and observed data, range from 0.68 (**TR-Model 2**) to 0.96 (**TR-Model 1**), while **TR-Model 3**, **TR-Model 4** and **TR-Model 5** show predictive ability values of 0.89, 0.95 and 0.95, respectively (**Figure 8B; Supplementary Table 8; Supplementary Figure 3**). Like observed for log-normal models, this analysis indicates that for TR, models that explicitly incorporate marker and interaction effects have higher performance in term of both predictive ability and percentage of explained variance.

**Figure 8: Estimated variance components and prediction ability of the TR-Models 1-5 used to predict TR under high evaporative demand: (A)** Bar plots indicate variances of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG and "Error" point out the variance of additive x additive epistatic effects and the residual variance, respectively. Error bars point out the 95% confidence interval of the estimated variances**; (B)** Bar plot of predictive ability values expressed as Pearson's correlation between estimated and observed TR under high evaporative demand (2.7 KPa).




## 5 Discussion

In the present study, the phenotypic variability of SRN, SRA and TR under high evaporative demand was surveyed in a panel of 90 MAGIC lines of barley (**Supplementary Table 1, Supplementary Table 2**). The results showed that in this genetic material, SRN can vary from 2 to 8 (**Figure 1A**), consistently with other barley studies (Robinson et al. 2016; Jia et al. 2019). Anyway, this comparison is not straightforward as in other studies SRN has been handled as a continuous trait and the phenotypic variability was presented using the adjusted means (BLUE or BLUP), while in this research work, SRN was analyzed as a discrete trait.

In the literature, SRA has been measured between the first pair of seminal roots (Robinson et al. 2016) or between the two outmost seminal roots (Jia et al. 2019). Following the original protocol, in the present study SRA between the first pair of seminal roots has been measured and the comparison of the phenotypic distribution of SRA with the results obtained in other work carried out following the same methodology (Robinson et al. 2016, 2018) point out that the SRA observed in MAGIC lines has wider phenotypic distribution. Unlike the findings reported in other studies (Robinson et al. 2016), SRA and SRN show a moderate correlation (r = 0.22). Anyway, analyzing SRN as continuous or discrete phenotypic trait implies different assumptions that hamper the comparisons of correlations as in our analysis SRN was considered as a discrete trait and consequently the connection between SRA and SRN was assessed using the polyserial correlation instead of Pearson's correlation coefficient. Interestingly, the present study indicated a moderate and positive correlation of SRA (r = 0.18) and SRN (r = 0.26) with GY measured in FioLN, which is a field trial fertilized with reduced

amount of nitrogen (**Figure 2B**). This correlation might suggest a link between the root architecture system and the ability of barley to grow in soils with a reduced nitrogen fertilization. Anyway, this correlation emerged in one out two field trials organized with a limited nitrogen fertilization and other studies are needed to support the link between these belowground traits measured at the seedling stage and the ability to promote yield formation in limiting nitrogen conditions.

In recent years, TR has been widely targeted to exploit its correlation to drought tolerance (Schoppach and Sadok 2012, 2013; Schoppach et al. 2016). Nevertheless, in barley the analysis of TR in response to high evaporative demand has lagged behind and to date has been investigated on a limited panel of 25 wild barley and in one cultivar (Sadok and Tamang 2019), which showed that at a VPD of circa 2.7 KPa, TR ranges from circa 25 to 75 $mg_{H2O}m^{-2}s^{-1}$ depending on the genotype. Our study confirms that in barley the TR at high evaporative demands is significantly lower than the values observed in other cereal crops (Schoppach and Sadok 2012; Sinclair et al. 2017; Tamang and Sadok 2018) and that the TR measured in our MAGIC population exhibits lower values of TR compared to other results obtained in barley at the same VPD values (Sadok and Tamang 2019). As with SRA analysis, technical causes and the lack of common genotypes hamper the comparison of different TR studies and absolute TR values, but it is plausible that, in general, our MAGIC population has a lower TR response to high evaporative demand compared to the barley genotypes investigated in other studies (Sadok and Tamang 2019). This trait showed a moderate and negative correlation of SRN (r = -0.12) with TR (**Figure 2B**), indicating that a high number of roots might decrease TR, which in turn could explain how in a stressful environment a high number of seminal roots can allow greater adaptation.

Breeding for crop ideotypes specifically designed for specific target environments and cropping systems has been proposed as an alternative of empirical breeding (Donald 1968) and different ideotypes of future cereal cultivars have been proposed to cope with climate change (Rötter et al. 2015). Designing and breeding for crop ideotypes requires *a priori* knowledge of sets of belowground and physiological traits that can cope with the peculiar environmental conditions of target environments to increase yield. In the present study, GP models for traits that are potentially relevant for barley ideotypes for dry environments were fitted using different methodologies. As substantiated in several reports, in cereal crops these are important traits as they are related to drought and GY but require time-consuming phenotyping operations that cannot be easily automated. Leveraging on prediction models for different types of data (continuous, count and ordinal), we showed that SRA, SRN and TR can be easily predicted paving the way to characterize large plant collection to sustain ideotype breeding while minimizing phenotyping costs.

For standard, log normal and TGBLUP models we used five sets of linear predictors, which differ for the type of effect considered (**Table 1**). Specifically, we considered models with main effects and models taking into account interactions between genetic factors (additive x·additive epistatic effect) and we found that in general the inclusion of interaction effects brings advantages in term of model fitting and predictive ability. For instance, these latter models showed predictive ability values of 0.65, 0.53 and 0.95 for SRN, SRA and TR, respectively (**Figure 6**; **Figure 7**; **Figure 8**). Overall, we observed that the inclusion of additive x additive improves model fitting in MAGIC population as it is plausible that in these lines epistasis might easier to model (Ehrenreich 2017; Mathew et al. 2018).

Several traits that are relevant for plant breeding are not normally distributed and need to be analyzed using special statistical techniques (Montesinos-López et al. 2015a). Traits that fall in this category are proportion of plants that overcome a stress, disease resistance scored using discrete scales and SRN. In the present study we assumed that this latter belowground trait varies as an ordinal discrete variable indicating the level of drought tolerance of plants at the adult stage using TGBLUP GP models (Montesinos-López et al. 2015a). Moreover, we modelled SRN as count data fitting log-

normal GP models. Using TGBLUP models, we observed that including additive x additive epistatic interaction (GxG) in the set of linear predictors (**SRN-Model 3** and **SRN-Model 5**) increases the total variance explained in the models (**Table 2**) pointing out the importance of epistasis in the genetic architecture of this trait. These results are corroborated in log-normal GP models, as they benefit of the inclusion of additive x additive epistatic interaction in the set of linear predictors in term of model fitting and predictive ability (**Figure 6**; **Figure 7**; **Figure 8**). Despite the low values of variance explained by molecular markers (**Table 2**), the TGBLUP models used to fit SRN showed high predictive ability values (Brier Score equals to 0.36) (**Figure 5**), highlighting that GP can be successfully applied to traits showing low heritability (**Supplementary Table 1, Supplementary Table 2**) as already substantiated in other studies carried out in plants (Zhang et al. 2017; Klápště et al. 2020) and animals (Guo et al. 2014; Iheshiulor et al. 2016).

## 6 Reference

Alahmad, Samir, Khaoula El Hassouni, Filippo M. Bassi, Eric Dinglasan, Chvan Youssef, Georgia Quarry, Alpaslan Aksoy, Elisabetta Mazzucotelli, Angéla Juhász, Jason A. Able, Jack Christopher, Kai P. Voss-Fels, and Lee T. Hickey. 2019. "A Major Root Architecture QTL Responding to Water Limitation in Durum Wheat." *Frontiers in Plant Science* 10(April):1–18.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67(1).

Bhatta, Madhav, Lucia Gutierrez, Lorena Cammarota, Fernanda Cardozo, Silvia Germán, Blanca Gómez-Guerrero, María Fernanda Pardo, Valeria Lanaro, Mercedes Sayas, and Ariel J. Castro. 2020. "Multi-Trait Genomic Prediction Model Increased the Predictive Ability for Agronomic and Malting Quality Traits in Barley (Hordeum Vulgare L.)." *G3: Genes, Genomes, Genetics* 10(3):1113–24.

Brier, G. W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78:1–3.

Crossa, José, Paulino Pérez-Rodríguez, Jaime Cuevas, Osval Montesinos-López, Diego Jarquín, Gustavo de los Campos, Juan Burgueño, Juan M. González-Camacho, Sergio Pérez-Elizalde, Yoseph Beyene, Susanne Dreisigacker, Ravi Singh, Xuecai Zhang, Manje Gowda, Manish Roorkiwal, Jessica Rutkoski, and Rajeev K. Varshney. 2017. "Genomic Selection in Plant Breeding: Methods, Models, and Perspectives." *Trends in Plant Science* 22(11):961–75.

Desta, Zeratsion Abera, and Rodomiro Ortiz. 2014. "Genomic Selection: Genome-Wide Prediction in Plant Improvement." *Trends in Plant Science* 19(9):592–601.

Devi, Mura Jyostna, Thomas R. Sinclair, and Earl Taliercio. 2015. "Comparisons of the Effects of Elevated Vapor Pressure Deficit on Gene Expression in Leaves among Two Fast-Wilting and a Slow-Wilting Soybean." *PLoS ONE* 10(10):1–21.

Donald, C. M. 1968. "The Breeding of Crop Ideotypes." *Euphytica* 17(3):385–403.

de Dorlodot, Sophie, Brian Forster, Loïc Pagès, Adam Price, Roberto Tuberosa, and Xavier Draye. 2007. "Root System Architecture: Opportunities and Constraints for Genetic Improvement of Crops." *Trends in Plant Science* 12(10):474–81.

Drasgow, Fritz. 2006. "Polychoric and Polyserial Correlations." *Encyclopedia of Statistical Sciences* 1–6.

Ehrenreich, Ian M. 2017. "Epistasis: Searching for Interacting Genetic Variants Using Crosses." *G3: Genes, Genomes, Genetics* 7(6):1619–22.

Fletcher, Andrew L., Thomas R. Sinclair, and L. Hartwell Allen. 2007. "Transpiration Responses to Vapor Pressure Deficit in Well Watered 'slow-Wilting' and Commercial Soybean." *Environmental and Experimental Botany* 61(2):145–51.

Fox, John. 2019. "Polycor: Polychoric and Polyserial Correlations. R Package Version 0.7-10."

Gianola, Daniel, and Chris Carolin Schon. 2016. "Cross-Validation without Doing Cross-Validation in Genome-Enabled Prediction." *G3: Genes, Genomes, Genetics* 6(10):3107–28.

González-Recio, Oscar, and Selma Forni. 2011. "Genome-Wide Prediction of Discrete Traits Using Bayesian Regressions and Machine Learning." *Genetics Selection Evolution* 43(1):1–12.

Guo, Gang, Fuping Zhao, Yachun Wang, Yuan Zhang, Lixin Du, and Guosheng Su. 2014. "Comparison of Single-Trait and Multiple-Trait Genomic Prediction Models." *BMC Genetics* 1–7.

Hammer, Graeme L., Zhanshan Dong, Greg McLean, Al Doherty, Carlos Messina, Jeff Schussler, Chris Zinselmeier, Steve Paszkiewicz, and Mark Cooper. 2009. "Can Changes in Canopy and/or Root System Architecture Explain Historical Maize Yield Trends in the U.S. Corn Belt?" *Crop Science* 49(1):299–312.

El Hassouni, K., S. Alahmad, B. Belkadi, A. Filali-Maltouf, L. T. Hickey, and F. M. Bassi. 2018. "Root System Architecture and Its Association with Yield under Different Water Regimes in Durum Wheat." *Crop Science* 58(6):2331–46.

Herder, Griet Den, Gert Van Isterdael, Tom Beeckman, and Ive De Smet. 2010. "The Roots of a New Green Revolution." *Trends in Plant Science* 15(11):600–607.

Iheshiulor, Oscar O. M., John A. Woolliams, Xijiang Yu, Robin Wellmann, and Theo H. E. Meuwissen. 2016. "Within - and across - Breed Genomic Prediction Using Whole - Genome Sequence and Single Nucleotide Polymorphism Panels." *Genetics Selection Evolution* 48(15):1–15.

Jannink, Jean Luc, Aaron J. Lorenz, and Hiroyoshi Iwata. 2010. "Genomic Selection in Plant Breeding: From Theory to Practice." *Briefings in Functional Genomics and Proteomics* 9(2):166–77.

Jia, Zhongtao, Ying Liu, Benjamin D. Gruber, Kerstin Neumann, Benjamin Kilian, Andreas Graner, and Nicolaus von Wirén. 2019. "Genetic Dissection of Root System Architectural Traits in Spring Barley." *Frontiers in Plant Science* 10(April):1–14.

Kholová, Jana, T. Nepolean, C. Tom Hash, A. Supriya, V. Rajaram, S. Senthilvel, Aparna Kakkera, Rattan Yadav, and Vincent Vadez. 2012. "Water Saving Traits Co-Map with a Major Terminal Drought Tolerance Quantitative Trait Locus in Pearl Millet [Pennisetum Glaucum (L.) R. Br.]." *Molecular Breeding* 30(3):1337–53.

Klápště, Jaroslav, Heidi S. Dungey, Emily J. Telfer, Mari Suontama, Natalie J. Graham, Yongjun Li, and Russell McKinley. 2020. "Marker Selection in Multivariate Genomic Prediction Improves Accuracy of Low Heritability Traits." *Frontiers in Genetics* 11(October):1–15.

Liu, Zhigang, Kun Gao, Shengchen Shan, Riling Gu, Zhangkui Wang, Eric J. Craft, Guohua Mi, Lixing Yuan, and Fanjun Chen. 2017. "Comparative Analysis of Root Traits and the Associated QTLs for Maize Seedlings Grown in Paper Roll, Hydroponics and Vermiculite Culture System." *Frontiers in Plant Science* 8(March):1–13.

Lobell, D. B., and S. M. Gourdji. 2012. "The Influence of Climate Change on Global Crop Productivity." *Plant Physiology* 160(4):1686–97.

Manschadi, Ahmad M., Graeme L. Hammer, John T. Christopher, and Peter DeVoil. 2008. "Genotypic Variation in Seedling Root Architectural Traits and Implications for Drought Adaptation in Wheat (Triticum Aestivum L.)." *Plant and Soil* 303(1–2):115–29.

Mathew, Boby, Jens Léon, Wiebke Sannemann, and Mikko J. Sillanpää. 2018. "Detection of Epistasis for Flowering Time Using Bayesian Multilocus Estimation in a Barley MAGIC Population." *Genetics* 208(2):525–36.

Medina, Susan, Rubén Vicente, Maria Teresa Nieto-Taladriz, Nieves Aparicio, Fadia Chairi, Omar Vergara-Diaz, and José Luis Araus. 2019. "The Plant-Transpiration Response to Vapor Pressure Deficit (VPD) in Durum Wheat Is Associated with Differential Yield Performance and Specific Expression of Genes Involved in Primary Metabolism and Water Transport." *Frontiers in Plant Science* 9(January):1–19.

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):1819–29.

Montesinos-López, Abelardo, Humberto Gutierrez-Pulido, Osval Antonio Montesinos-López, and José Crossa. 2020. "Maximum a Posteriori Threshold Genomic Prediction Model for Ordinal Traits." *G3: Genes, Genomes, Genetics* 10(11):4083–4102.

Montesinos-López, Abelardo, Osval A. Montesinos-López, José Crossa, Juan Burgueño, Kent M. Eskridge, Esteban Falconi-Castillo, Xinyao He, Pawan Singh, and Karen Cichy. 2016. "Genomic Bayesian Prediction Model for Count Data with Genotype × Environment Interaction." *G3: Genes, Genomes, Genetics* 6(5):1165–77.

Montesinos-López, Osval A., Abelardo Montesinos-López, Paulino Pérez-Rodríguez, Kent Eskridge, Xinyao He, Philomin Juliana, Pawan Singh, and José Crossa. 2015. "Genomic Prediction Models for Count Data." *Journal of Agricultural, Biological, and Environmental Statistics* 20(4):533–54.

Montesinos-López, Osval A., Abelardo Montesinos-López, Paulino Pérez-Rodríguez, Gustavo de los Campos, Kent Eskridge, and José Crossa. 2015. "Threshold Models for Genome-Enabled

Prediction of Ordinal Categorical Traits in Plant Breeding." *G3: Genes, Genomes, Genetics* 5(2):291–300.

Mu, Xiaohuan, Fanjun Chen, Qiuping Wu, Qinwu Chen, Jingfeng Wang, Lixing Yuan, and Guohua Mi. 2015. "Genetic Improvement of Root Growth Increases Maize Yield via Enhanced Post-Silking Nitrogen Uptake." *European Journal of Agronomy* 63:55–61.

Muggeo, V. M. R. 2008. "Segmented: An R Package to Fit Regression Models with Broken-Line Relationships." *R News* 8:20–25.

Muggeo, V. M. R. 2017. "Interval Estimation for the Breakpoint in Segmented Regression: A Smoothed Score-Based Approach." *Australian and New Zealand Journal of Statistics* 59(3):311–22.

Nielsen, Nanna Hellum, Ahmed Jahoor, Jens Due Jensen, Jihad Orabi, Fabio Cericola, Vahid Edriss, and Just Jensen. 2016. "Genomic Prediction of Seed Quality Traits Using Advanced Barley Breeding Lines." *PLoS ONE* 11(10):1–18.

Olivoto, Tiago, and Alessandro Dal Col Lúcio. 2020. "Metan: An R Package for Multi-Environment Trial Analysis." *Methods in Ecology and Evolution* 11(6):783–89.

Pérez, Paulino, and G. de los Campos. 2014. "BGLR : A Statistical Package for Whole Genome Regression and Prediction." *Genetics* 198(2):483–95.

Pérez, Paulino, and Gustavo De Los Campos. 2014. "Genome-Wide Regression and Prediction with the BGLR Statistical Package." *Genetics* 198(2):483–95.

Puglisi, Damiano, Stefano Delbono, Andrea Visioni, Hakan Ozkan, İbrahim Kara, Ana M. Casas, Ernesto Igartua, Giampiero Valè, Angela Roberta Lo Piero, Luigi Cattivelli, Alessandro Tondelli, and Agostino Fricano. 2021. "Genomic Prediction of Grain Yield in a Barley MAGIC Population Modeling Genotype per Environment Interaction." *Frontiers in Plant Science* 12(May):1–18.

R Core Team. 2020. "R: A Language and Environment for Statistical Computing." *R Foundation for Statistical Computing* Vienna(Austria):https://www.r-project.org/.

Richard, Cecile Al, Lee T. Hickey, Susan Fletcher, Raeleen Jennings, Karine Chenu, and Jack T. Christopher. 2015. "High-Throughput Phenotyping of Seminal Root Traits in Wheat." *Plant Methods* 11(1):1–11.

Robinson, Hannah, Lee Hickey, Cecile Richard, Emma Mace, Alison Kelly, Andrew Borrell, Jerome Franckowiak, and Glen Fox. 2016. "Genomic Regions Influencing Seminal Root Traits in Barley." *The Plant Genome* 9(1):1–13.

Robinson, Hannah, Alison Kelly, Glen Fox, Jerome Franckowiak, Andrew Borrell, and Lee Hickey. 2018. "Root Architectural Traits and Yield: Exploring the Relationship in Barley Breeding Trials." *Euphytica* 214(151):1–14.

Rötter, R. P., F. Tao, J. G. Höhn, and T. Palosuo. 2015. "Use of Crop Simulation Modelling to Aid Ideotype Design of Future Cereal Cultivars." *Journal of Experimental Botany* 66(12):3463–76.

Sadok, Walid, and Thomas R. Sinclair. 2009a. "Genetic Variability of Transpiration Response to Vapor Pressure Defi Cit among Soybean Cultivars." *Crop Science* 49(3):955–60.

Sadok, Walid, and Thomas R. Sinclair. 2009b. "Genetic Variability of Transpiration Response to Vapor Pressure Deficit among Soybean (Glycine Max [L.] Merr.) Genotypes Selected from a Recombinant Inbred Line Population." *Field Crops Research* 113(2):156–60.

Sadok, Walid, and Bishal G. Tamang. 2019. "Diversity in Daytime and Night-Time Transpiration Dynamics in Barley Indicates Adaptation to Drought Regimes across the Middle-East." *Journal of Agronomy and Crop Science* 205(4):372–84.

Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. 2012. "NIH Image to ImageJ: 25 Years of Image Analysis." *Nature Methods* 9(7):671–75.

Schoppach, R., D. Fleury, T. R. Sinclair, and W. Sadok. 2017. "Transpiration Sensitivity to Evaporative Demand Across 120 Years of Breeding of Australian Wheat Cultivars." *Journal of Agronomy and Crop Science* 203(3):219–26.

Schoppach, Rémy, and Walid Sadok. 2012. "Differential Sensitivities of Transpiration to Evaporative

Demand and Soil Water Deficit among Wheat Elite Cultivars Indicate Different Strategies for Drought Tolerance." *Environmental and Experimental Botany* 84:1–10.

Schoppach, Rémy, and Walid Sadok. 2013. "Transpiration Sensitivities to Evaporative Demand and Leaf Areas Vary with Night and Day Warming Regimes among Wheat Genotypes." *Functional Plant Biology* 40(7):708–18.

Schoppach, Remy, Julian D. Taylor, Elisabeth Majerus, Elodie Claverie, Ute Baumann, Radoslaw Suchecki, Delphine Fleury, and Walid Sadok. 2016. "High Resolution Mapping of Traits Related to Whole-Plant Transpiration under Increasing Evaporative Demand in Wheat." *Journal of Experimental Botany* 67(9):2847–60.

Sinclair, Thomas R., Jyostna Devi, Avat Shekoofa, Sunita Choudhary, Walid Sadok, Vincent Vadez, Mandeep Riar, and Thomas Rufty. 2017. "Limited-Transpiration Response to High Vapor Pressure Deficit in Crop Species." *Plant Science* 260(April):109–18.

Sinclair, Thomas R., Graeme L. Hammer, and Erik J. Van Oosterom. 2005. "Potential Yield and Water-Use Efficiency Benefits in Sorghum from Limited Maximum Transpiration Rate." *Functional Plant Biology* 32(10):945–52.

Sivasakthi, Kaliamoorthy, Murugesan Tharanya, Jana Kholová, Ruth Wangari Muriuki, Thiyagarajan Thirunalasundari, and Vincent Vadez. 2017. "Chickpea Genotypes Contrasting for Vigor and Canopy Conductance Also Differ in Their Dependence on Different Water Transport Pathways." *Frontiers in Plant Science* 8(September):1–16.

Sorensen, DA A., S. Andersen, D. Gianola, and I. Korsgaard. 1995. "Bayesian Inference in Threshold Models Using Gibbs Sampling." *Genetics, Selection, Evolution* 27(3):229–49.

Tamang, Bishal G., and Walid Sadok. 2018. "Nightly Business: Links between Daytime Canopy Conductance, Nocturnal Transpiration and Its Circadian Control Illuminate Physiological Trade-Offs in Maize." *Environmental and Experimental Botany* 148(November 2017):192–202.

Tao, Fulu, Reimund P. Rötter, Taru Palosuo, C. G. H. Díaz-Ambrona, M. Inés Mínguez, Mikhail A. Semenov, Kurt Christian Kersebaum, Claas Nendel, Davide Cammarano, Holger Hoffmann, Frank Ewert, Anaelle Dambreville, Pierre Martre, Lucía Rodríguez, Margarita Ruiz-Ramos, Thomas Gaiser, Jukka G. Höhn, Tapio Salo, Roberto Ferrise, Marco Bindi, and Alan H. Schulman. 2017. "Designing Future Barley Ideotypes Using a Crop Model Ensemble." *European Journal of Agronomy* 82:144–62.

Tao, Fulu, Masayuki Yokozawa, and Zhao Zhang. 2009. "Modelling the Impacts of Weather and Climate Variability on Crop Productivity over a Large Area: A New Process-Based Model Development, Optimization, and Uncertainties Analysis." *Agricultural and Forest Meteorology* 149(5):831–50.

Tsai, Hsin Yuan, Luc L. Janss, Jeppe R. Andersen, Jihad Orabi, Jens D. Jensen, Ahmed Jahoor, and Just Jensen. 2020. "Genomic Prediction and GWAS of Yield, Quality and Disease-Related Traits in Spring Barley and Winter Wheat." *Scientific Reports* 10(1):1–15.

VanRaden, P. M. 2008. "Efficient Methods to Compute Genomic Predictions." *Journal of Dairy Science* 91(11):4414–23.

Villanueva, B., J. Fernández, L. A. García-Cortés, L. Varona, H. D. Daetwyler, and M. A. Toro. 2011. "Accuracy of Genome-Wide Evaluation for Disease Resistance in Aquaculture Breeding Programs." *Journal of Animal Science* 89(11):3433–42.

Wahbi, A., and P. J. Gregory. 1995. "Growth and Development of Young Roots of Barley (Hordeum Vulgare L.) Genotypes." *Annals of Botany* 75(5):533–39.

Wang, C. L., X. D. Ding, J. Y. Wang, J. F. Liu, W. X. Fu, Z. Zhang, Z. J. Yin, and Q. Zhang. 2013. "Bayesian Methods for Estimating GEBVs of Threshold Traits." *Heredity* 110(3):213–19.

Wang, Xin, Yang Xu, Zhongli Hu, and Chenwu Xu. 2018. "Genomic Selection Methods for Crop Improvement: Current Status and Prospects." *Crop Journal* 6(4):330–40.

Yan, Weikai, and James B. Holland. 2010. "A Heritability-Adjusted GGE Biplot for Test Environment Evaluation." *Euphytica* 171(3):355–69.

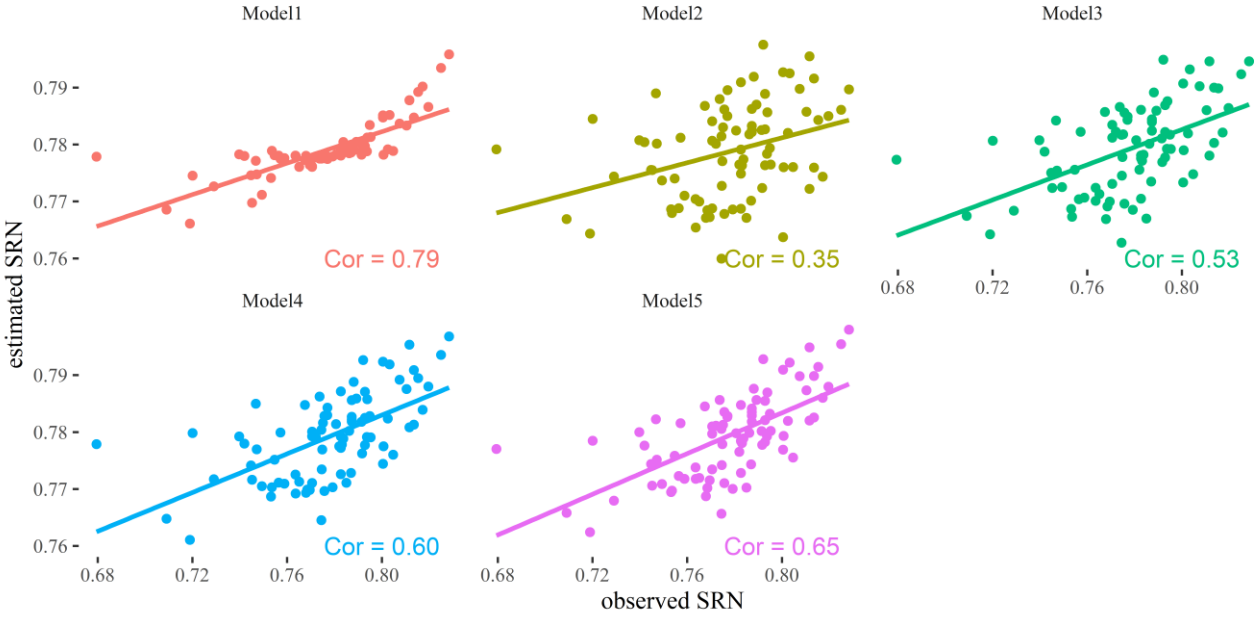
Yan, Weikai, L. A. Hunt, Qinglai Sheng, and Zorka Szlavnics. 2000. "Cultivar Evaluation and Mega-

Environment Investigation Based on the GGE Biplot." *Crop Science* 40(3):597–605.
Zadoks, J. C., T. T. Chang, and C. F. Konzak. 1974. "A Decimal Code for the Growth Stages of Cereals." *World Pumps* 14:415–21.
Zhang, Ao, Hongwu Wang, Yoseph Beyene, Kassa Semagn, Yubo Liu, Shiliang Cao, Zhenhai Cui, Yanye Ruan, Juan Burgueño, Felix San Vicente, Michael Olsen, Boddupalli M. Prasanna, José Crossa, Haiqiu Yu, and Xuecai Zhang. 2017. "Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 Bi-Parental Tropical Maize Populations." *Frontiers in Plant Science* 8(November):1–12.

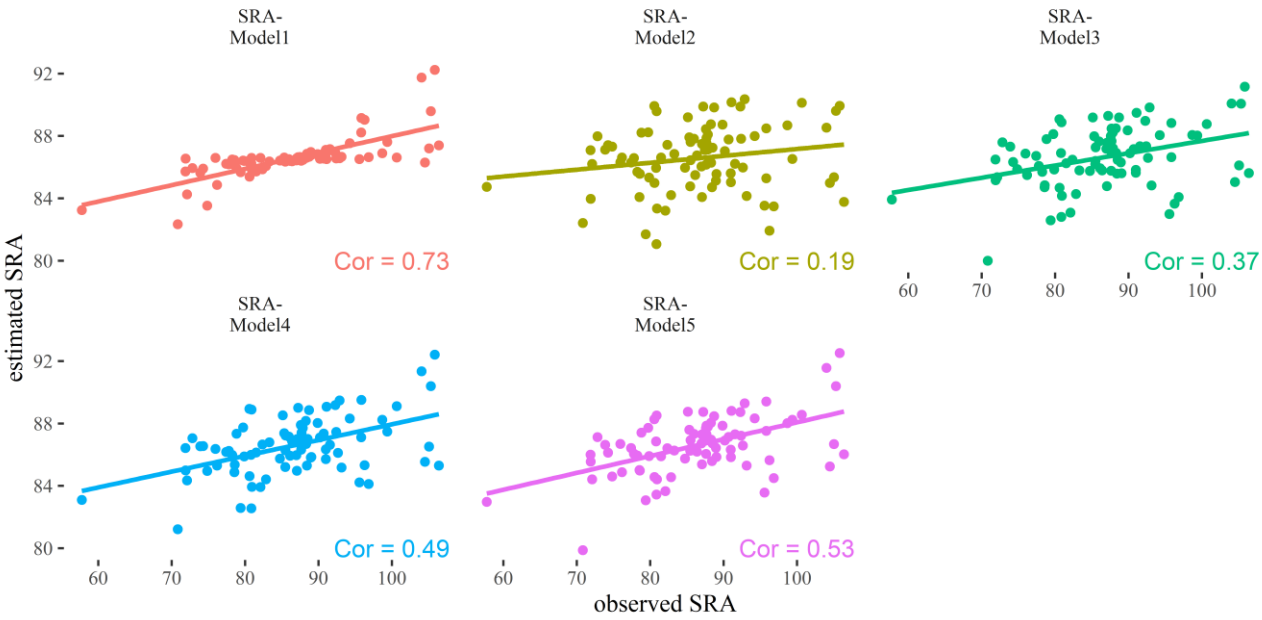
# 7 Supplementary Material

## 7.1 Supplementary Figures

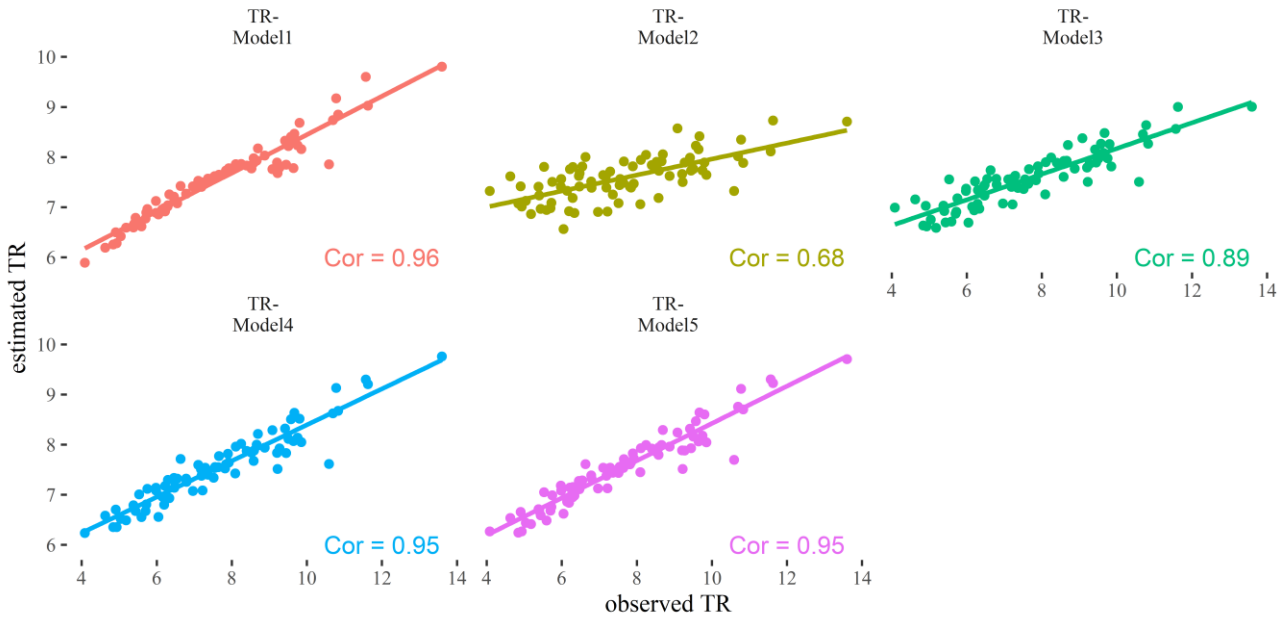
**Supplementary Figure 1: Scatter plot of prediction ability of the log-normal Models 1-5 used to predict SRN.** Pearson's correlation between estimated and observed SRN were reported for each model.




**Supplementary Figure 2: Scatter plot of prediction ability of the SRA-Models 1-5 used to predict SRA.** Pearson's correlation between estimated and observed SRA were reported for each model.



**Supplementary Figure 3: Scatter plot of prediction ability of the TR- Models 1-5 used to predict TR at high evaporative demand.** Pearson's correlation between estimated and observed TR were reported for each model.




## 7.2 Supplementary Tables

**Supplementary Table 1: Phenotypic data obtained for the panel of MAGIC lines.** For each line, the raw data of SRN was reported. Maximum (Max), minimum (Min) average (Mean) values were calculated at the end of the table.

| Genotype | Rep 1 | Rep 2 | Rep 3 | Rep 4 | Rep 5 | Rep 6 | Rep 7 | Rep 8 | Rep 9 | Rep 10 | Rep 11 | Rep 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Aldebaran** | 5 | 5 | 5 | 5 | 4 | 5 | 3 | 4 | 4 | 5 | 3 | 5 |
| **Athene** | 6 | 5 | 6 | 4 | NA | 4 | 6 | 4 | NA | 5 | 5 | 5 |
| **Dea** | 5 | 5 | 5 | 6 | 5 | 5 | 4 | 5 | 5 | 6 | 5 | 5 |
| **Fridericus** | 4 | 4 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 2 | 4 |
| **Hatif de Grignon** | 5 | 5 | 6 | NA | 5 | 5 | NA | 6 | NA | 6 | 5 | NA |
| **Ketos** | 5 | 5 | 6 | 5 | 5 | 4 | 6 | 5 | 3 | 3 | 4 | 5 |
| **M101** | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 4 | 6 | NA | 6 | 5 |
| **M102** | 4 | 6 | 5 | NA | NA | 5 | 5 | 5 | 4 | NA | 4 | 5 |
| **M106** | 6 | 6 | 4 | 5 | 5 | 5 | 4 | 5 | 2 | 6 | 5 | 3 |
| **M11** | 5 | 4 | 5 | 4 | 4 | 5 | NA | NA | NA | 6 | 4 | NA |
| **M118** | 5 | 6 | 6 | 6 | 6 | 4 | 6 | 5 | 5 | NA | 6 | 6 |
| **M121** | 6 | 6 | 6 | 5 | 5 | 5 | 6 | 4 | 4 | 5 | 5 | 5 |
| **M131** | 5 | 4 | 6 | 6 | NA | 4 | 6 | 6 | 5 | 4 | 7 | NA |
| **M132** | 5 | 4 | 5 | 4 | 5 | 6 | NA | 5 | 6 | 5 | NA | 5 |
| **M136** | 6 | NA | 5 | 5 | 5 | 5 | 6 | NA | 5 | 4 | NA | NA |
| **M140** | 4 | 5 | 6 | 4 | 6 | 5 | 6 | 6 | 5 | 5 | 5 | 6 |
| **M147** | 6 | 6 | 6 | 5 | 4 | 5 | 3 | 4 | NA | 5 | 5 | 4 |
| **M148** | 4 | 6 | 5 | 5 | 5 | 5 | NA | 6 | 6 | 5 | 6 | 4 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **M149** | 5 | 5 | 6 | 5 | 5 | 6 | 6 | NA | 6 | 4 | 5 | 2 |
| **M150** | 5 | 5 | 5 | 6 | 6 | 6 | 6 | NA | 5 | NA | 5 | 6 |
| **M151** | 6 | 6 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 6 | 4 | 4 |
| **M154** | 5 | 5 | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| **M156** | 6 | 5 | 6 | 5 | 5 | 5 | 5 | 6 | 5 | 2 | 6 | NA |
| **M158** | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 6 | 5 | 5 | 5 | 4 |
| **M162** | 5 | 4 | 6 | 4 | NA | 4 | 5 | 4 | 6 | 4 | 5 | 4 |
| **M163** | 6 | 5 | 6 | 5 | 4 | 6 | NA | 6 | 6 | 5 | 6 | 5 |
| **M169** | NA | 5 | 6 | 6 | 5 | 4 | 6 | 5 | 6 | 4 | 5 | 5 |
| **M177** | 5 | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 4 | 5 |
| **M18** | 6 | 5 | 5 | 5 | 4 | 5 | 5 | NA | 5 | 5 | 5 | NA |
| **M180** | 6 | 4 | 5 | 5 | 4 | 6 | 6 | 4 | 6 | 5 | 4 | 4 |
| **M181** | 5 | 6 | 5 | 5 | 3 | 6 | 3 | 5 | 6 | 5 | 3 | 6 |
| **M190** | 5 | 6 | 5 | 5 | NA | 6 | 5 | 5 | 4 | 5 | 5 | 6 |
| **M194** | 6 | 5 | 6 | 6 | 5 | 6 | 6 | 6 | 5 | 6 | 5 | 3 |
| **M200** | 5 | 5 | 4 | 5 | 4 | 5 | 3 | 4 | 4 | 4 | NA | 4 |
| **M201** | 5 | 6 | 6 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| **M215** | 6 | 5 | 5 | 5 | 6 | 5 | 5 | 5 | 5 | 6 | 6 | 5 |
| **M216** | 6 | 4 | 5 | 5 | 5 | 6 | 6 | 5 | 4 | 5 | 5 | 6 |
| **M22** | 4 | 6 | 4 | 4 | 5 | 3 | 3 | 5 | 4 | NA | 4 | 4 |
| **M224** | 6 | 4 | 5 | 5 | 6 | 6 | 5 | 5 | 5 | 6 | 5 | 5 |
| **M233** | 6 | 5 | NA | 6 | NA | 6 | 6 | 5 | 6 | 4 | 5 | 6 |
| **M237** | 4 | NA | 4 | 5 | 5 | 6 | 2 | 6 | 5 | 6 | 6 | 4 |
| **M242** | 5 | 5 | 4 | 5 | 6 | 5 | 5 | 5 | 4 | 5 | 5 | 5 |
| **M246** | 5 | 6 | 6 | 4 | 5 | 5 | 6 | 5 | 4 | NA | 4 | NA |
| **M253** | 5 | 6 | 6 | 6 | 6 | 6 | 6 | NA | 5 | 5 | 4 | 6 |
| **M254** | 6 | 6 | 5 | 6 | 5 | 5 | 3 | 3 | 4 | 7 | 3 | 5 |
| **M258** | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | NA |
| **M259** | 6 | 5 | 5 | 5 | 6 | 5 | 3 | 6 | 5 | 6 | 5 | 4 |
| **M262** | 6 | 5 | 4 | NA | 6 | 4 | 6 | 5 | 4 | 4 | 5 | 6 |
| **M264** | 5 | 5 | 5 | 5 | NA | 5 | 6 | 6 | NA | 5 | NA | NA |
| **M278** | 5 | 6 | 5 | 5 | 3 | 6 | 6 | 4 | NA | 5 | 3 | 5 |
| **M281** | 6 | 4 | 5 | 5 | 4 | 6 | 5 | 5 | NA | 6 | 4 | 5 |
| **M285** | 6 | 6 | 5 | 6 | 6 | 6 | 6 | NA | 6 | 4 | 6 | 5 |
| **M287** | 5 | 6 | 6 | 6 | 5 | 5 | 4 | 5 | 6 | 3 | NA | 4 |
| **M288** | 6 | 6 | 6 | 5 | NA | 6 | 6 | 5 | 5 | NA | NA | 5 |
| **M289** | 5 | 8 | 5 | 5 | 4 | 3 | 5 | 5 | 6 | 5 | 5 | 6 |
| **M298** | 6 | 6 | 5 | 6 | 5 | 6 | 5 | 6 | 6 | 6 | 6 | 6 |
| **M299** | 6 | 4 | 6 | NA | 4 | 5 | 5 | 5 | 6 | 5 | 5 | 5 |
| **M30** | 5 | 5 | 4 | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 |
| **M304** | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | NA | 5 |
| **M313** | 7 | 6 | 5 | 6 | 6 | 4 | 5 | 5 | 5 | 5 | 5 | 4 |
| **M314** | 6 | 6 | 6 | NA | 6 | 6 | 5 | 6 | 6 | 5 | NA | 5 |
| **M321** | 5 | 5 | 5 | 5 | 6 | 5 | 6 | 6 | 6 | 5 | NA | 4 |
| **M324** | 6 | 5 | 5 | 5 | 4 | 5 | 6 | 5 | 5 | 5 | 5 | 5 |
| **M328** | 5 | 6 | 6 | 5 | 6 | 5 | 6 | 5 | 6 | 5 | 4 | 4 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **M332** | 6 | 5 | 5 | 6 | NA | 6 | NA | 6 | 6 | 5 | 5 | 5 |
| **M338** | 5 | 6 | 5 | 5 | 5 | 5 | 5 | 4 | NA | 5 | 5 | 5 |
| **M339** | 5 | 5 | 6 | 4 | 6 | 4 | 5 | 6 | 5 | 5 | 6 | 5 |
| **M345** | 5 | NA | 6 | 6 | NA | 6 | 3 | 6 | 6 | 5 | 5 | 6 |
| **M349** | 6 | 4 | 5 | NA | 3 | 5 | 6 | 4 | NA | 6 | 6 | 4 |
| **M350** | 6 | 6 | 5 | 4 | 5 | 4 | NA | NA | 5 | 5 | 5 | 5 |
| **M355** | 6 | 6 | 6 | 6 | NA | 6 | 5 | 4 | 5 | 5 | 4 | 4 |
| **M357** | 4 | 6 | 4 | 5 | 5 | NA | 4 | 6 | 2 | 4 | 2 | 2 |
| **M36** | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 5 | 6 | 6 |
| **M365** | 6 | 5 | 6 | 5 | 6 | 5 | 6 | 5 | 4 | 5 | 4 | 6 |
| **M373** | 5 | 4 | 6 | 5 | 5 | 5 | 4 | 4 | NA | 5 | NA | 4 |
| **M374** | 5 | 5 | 5 | 5 | 3 | 6 | 5 | 6 | 4 | NA | 2 | 5 |
| **M376** | 5 | 5 | 6 | 5 | 5 | 6 | 4 | 4 | 6 | 6 | 5 | 5 |
| **M382** | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 6 | 5 |
| **M383** | 5 | 4 | 6 | 5 | 6 | NA | NA | 6 | 5 | NA | 5 | 5 |
| **M42** | 5 | 6 | 5 | 6 | 6 | 5 | 4 | 5 | 6 | 6 | 4 | 4 |
| **M52** | 4 | 3 | 5 | 5 | NA | 5 | NA | 5 | 5 | 5 | 5 | 4 |
| **M53** | 5 | 4 | NA | 5 | NA | 4 | NA | NA | 6 | NA | 5 | 5 |
| **M68** | 6 | 6 | 5 | 6 | 5 | 4 | 4 | 6 | 6 | 3 | 5 | 3 |
| **M78** | 5 | 6 | 5 | 5 | 6 | 6 | 6 | 5 | 5 | 6 | 6 | 6 |
| **M84** | 5 | 5 | 6 | 3 | 4 | 5 | 6 | 6 | NA | 6 | 6 | 6 |
| **M92** | 4 | 4 | 5 | 4 | 6 | 5 | 5 | 5 | 5 | 5 | 6 | 5 |
| **M95** | 5 | 5 | 5 | 5 | 6 | 6 | 5 | 4 | 5 | 5 | 5 | NA |
| **Ponente** | 6 | 5 | 5 | 5 | 6 | 6 | 6 | 4 | 4 | 5 | 5 | 4 |
| **Rihane03** | 5 | 4 | 5 | 4 | 5 | 5 | 4 | 5 | 3 | 6 | 5 | 3 |
| **Robor** | 5 | 6 | 6 | 6 | 6 | NA | 6 | 3 | 5 | 6 | 6 | 5 |
| **Statistical parameters** | **Mean (SD)** | | | | | 5 (0.84) | | | | | | |
| | **Max** | | | | | 8 | | | | | | |
| | **Min** | | | | | 2 | | | | | | |

**Supplementary Table 2: Phenotypic data obtained for the panel of MAGIC lines.** For each line, transpiration rate measured at a vapor pressure deficit of 2.70 KPa (TR), linear or segmented trend transpiration rate (TR trend), the adjusted means of seminal root angle (SRA) measured in sexagesimal degrees were reported. Numbers between bracktes point out the standard deviation of each measurement. For each trait, maximum (Max), minimum (Min) average (Mean) values were calculated at the end of the table.

| Genotype | TR ($mg_{H2O}m^{-2}s^{-1}$) | TR trend | SRA (Sexagesimal degrees) |
|---|---|---|---|
| **Aldebaran** | 6.76 (1.61) | Linear | 78.25° (31.86°) |
| **Athene** | 11.63 (2.29) | Linear | 71.88° (26.37°) |
| **Dea** | 8.69 (0.22) | Segmented | 87.63° (27.87°) |
| **Fridericus** | 7.51 (0.93) | Linear | 100.66° (32.46°) |
| **Hatif de Grignon** | 7.36 (4.11) | Segmented | 70.83° (13.91°) |
| **Ketos** | 5.04 (0.99) | Segmented | 82.35° (16.45°) |

| | | | |
|---|---|---|---|
| **M101** | NA | NA | NA |
| **M102** | NA | NA | NA |
| **M106** | 9.67 (1.6) | Segmented | 92.88° (29.84°) |
| **M11** | NA | NA | NA |
| **M118** | 5.98 (2.48) | Linear | 80.59° (15.79°) |
| **M121** | 8.59 (2.83) | Linear | 96.26° (13.66°) |
| **M131** | 8.27 (6.78) | Segmented | 84.74° (26.34°) |
| **M132** | 4.08 (0.51) | Segmented | 105.83° (13.71°) |
| **M136** | 6.13 (0.66) | Segmented | 104.03° (12.87°) |
| **M140** | 11.57 (0.26) | Linear | 98.69° (27.41°) |
| **M147** | 5.75 (2.07) | Segmented | 85.14° (28.89°) |
| **M148** | NA | NA | NA |
| **M149** | 4.63 (0.67) | Segmented | 86.39° (26.38°) |
| **M150** | 13.6 (2.36) | Segmented | 73.88° (24.52°) |
| **M151** | 7.82 (3.14) | Segmented | 105.06° (30.47°) |
| **M154** | 6.2 (1.12) | Segmented | 81.5° (18.1°) |
| **M156** | 9.8 (0.56) | Segmented | 76.17° (18.02°) |
| **M158** | 5.18 (1.43) | Linear | 80.93° (17.78°) |
| **M162** | 9.27 (4.37) | Segmented | 87.87° (30.9°) |
| **M163** | 5.98 (1.12) | Segmented | 80.86° (22.91°) |
| **M169** | 9.42 (1.8) | Linear | 96.84° (35.14°) |
| **M177** | 6.78 (1.32) | Segmented | 87.86° (28.1°) |
| **M18** | 6.54 (0.28) | Segmented | 57.75° (28.91°) |
| **M180** | 9.57 (1.7) | Segmented | 82.85° (18.33°) |
| **M181** | 7.66 (0.61) | Segmented | 75.95° (37.33°) |
| **M190** | 8.25 (1.82) | Segmented | 89.04° (19.21°) |
| **M194** | 5.53 (0.96) | Segmented | 88.43° (22.19°) |
| **M200** | 10.83 (1.98) | Segmented | 80.87° (32.99°) |
| **M201** | 7.85 (1.08) | Segmented | 88.72° (20.67°) |
| **M215** | NA | NA | 95.58° (31.11°) |
| **M216** | 6.79 (1.16) | Linear | 82.09° (41.83°) |
| **M22** | 6.3 (1.22) | Segmented | 80.81° (38.32°) |
| **M224** | 9.85 (3.14) | Linear | 79.39° (19.19°) |
| **M233** | 7.1 (3.5) | Segmented | 72.1° (19.02°) |
| **M237** | 5.38 (1.11) | Linear | 93.14° (25.78°) |
| **M242** | 6.22 (0.33) | Segmented | 86.99° (23.44°) |
| **M246** | 9.08 (4.49) | Linear | 78.82° (22.03°) |
| **M253** | 6.47 (1.35) | Segmented | 92.4° (28.79°) |
| **M254** | 7.64 (0.53) | Segmented | 86.13° (17.98°) |
| **M258** | 10.78 (0.54) | Segmented | 91.09° (26.07°) |
| **M259** | 9.2 (3.75) | Segmented | 83.3° (24.15°) |
| **M262** | 6.29 (1.05) | Segmented | 78.59° (24.21°) |
| **M264** | NA | NA | 88.27° (17.23°) |
| **M278** | 9.22 (6.03) | Segmented | 74.83° (13.68°) |
| **M281** | 5.72 (1.81) | Segmented | 72.82° (27.13°) |
| **M285** | 6.05 (0.35) | Linear | 87.22° (23.97°) |

| | | | | |
|---|---|---|---|---|
| **M287** | | 7.55 (2.29) | Segmented | 87.6° (22.25°) |
| **M288** | | 7.22 (2.42) | Segmented | 104.5° (44.59°) |
| **M289** | | 7.92 (0.25) | Segmented | 95.83° (12.55°) |
| **M298** | | 6.48 (1.55) | Linear | 87.07° (25.29°) |
| **M299** | | 8.4 (2.63) | Segmented | 88.31° (23.65°) |
| **M30** | | 7.19 (0.44) | Segmented | 87.68° (30.76°) |
| **M304** | | 5.44 (1.86) | Segmented | 90.19° (29.96°) |
| **M313** | | 5.7 (1.27) | Segmented | 87.68° (22.78°) |
| **M314** | | 9.51 (2.46) | Segmented | 105.32° (20.61°) |
| **M321** | | 8.87 (2.29) | Linear | 91.55° (20.34°) |
| **M324** | | 7.54 (2.53) | Segmented | 106.4° (31.21°) |
| **M328** | | 7.29 (0.89) | Segmented | 77.79° (25.49°) |
| **M332** | | 6.96 (2.12) | Segmented | 88.96° (15.21°) |
| **M338** | | 8.52 (3.83) | Segmented | 80.58° (27.34°) |
| **M339** | | 6.34 (2.66) | Segmented | 85.47° (18.24°) |
| **M345** | | 6.14 (1.15) | Linear | 99.37° (22.49°) |
| **M349** | | 8.09 (2.1) | Segmented | 77.41° (24.34°) |
| **M350** | | 4.84 (0.23) | Linear | 92.29° (19.98°) |
| **M355** | | 8.58 (1.75) | Segmented | 94.25° (17.99°) |
| **M357** | | 6.67 (0.72) | Linear | 71.9° (32.73°) |
| **M36** | | 7.17 (2.21) | Linear | 91° (20.5°) |
| **M365** | | 8.66 (2.58) | Segmented | 92.65° (28.82°) |
| **M373** | | 6.63 (3.37) | Segmented | 90.7° (29.11°) |
| **M374** | | 4.91 (1.62) | Linear | 79.74° (29.34°) |
| **M376** | | 9.75 (2.7) | Segmented | 85.51° (23.17°) |
| **M382** | | 4.05 (0.96) | Linear | 89.87° (18.59°) |
| **M383** | | 10.59 (5.17) | Linear | 87.06° (21.57°) |
| **M42** | | 10.7 (2.1) | Linear | 85.5° (17.51°) |
| **M52** | | 6.46 (2.04) | Linear | 78.55° (27.04°) |
| **M53** | | NA | NA | 95.83° (16.36°) |
| **M68** | | 7.89 (1.39) | Linear | 85.97° (25.48°) |
| **M78** | | 8.11 (0.68) | Segmented | 90.99° (18.97°) |
| **M84** | | 4.93 (0.2) | Linear | 74.27° (25.71°) |
| **M92** | | 5.4 (1.4) | Linear | 85.33° (30.05°) |
| **M95** | | 9.64 (4.71) | Segmented | 79.87° (27.96°) |
| **Ponente** | | 9.44 (4.2) | Segmented | 87.45° (21.54°) |
| **Rihane03** | | NA | NA | NA |
| **Robor** | | 5.59 (0.27) | Linear | 88.5° (19.38°) |
| **Statistical parameters** | **Mean (SD)** | 7.63 (1.99) | / | 86.51° (9.06°) |
| | **Max** | 13.6 | / | 106.40° |
| | **Min** | 4.08 | / | 57.75° |

**Supplementary Table 3: Estimated parameters of the SRN-Models 1-5 used to predict SRN.** Posterior average (Mean) values along with the standard deviation (SD) between brackets of fixed effects ($\beta_1, \beta_2, \ldots \beta_{12}$) were estimated.

| Model | β1 | β2 | β3 | β4 | β5 | β6 | β7 | β8 | β9 | β10 | β11 | β12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SRN-Model 1** | -0.90 (0.41) | -1.15 (0.42) | -0.96 (0.41) | -1.25 (0.43) | -1.30 (0.43) | -1.18 (0.43) | -1.30 (0.43) | -1.33 (0.43) | -1.32 (0.43) | -1.41 (0.43) | -1.53 (0.44) | -1.62 (0.44) |
| **SRN-Model 2** | -0.68 (0.35) | -0.94 (0.36) | -0.75 (0.36) | -1.03 (0.37) | -1.08 (0.37) | -0.96 (0.37) | -1.08 (0.37) | -1.12 (0.38) | -1.10 (0.37) | -1.19 (0.37) | -1.32 (0.38) | -1.40 (0.38) |
| **SRN-Model 3** | -0.88 (0.39) | -1.14 (0.40) | -0.95 (0.40) | -1.23 (0.41) | -1.28 (0.41) | -1.17 (0.41) | -1.28 (0.41) | -1.33 (0.41) | -1.31 (0.41) | -1.40 (0.41) | -1.52 (0.42) | -1.61 (0.42) |
| **SRN-Model 4** | -0.93 (0.38) | -1.19 (0.39) | -1.00 (0.39) | -1.28 (0.40) | -1.33 (0.40) | -1.21 (0.40) | -1.33 (0.40) | -1.37 (0.41) | -1.36 (0.40) | -1.45 (0.40) | -1.57 (0.41) | -1.66 (0.41) |
| **SRN-Model 5** | -0.81 (0.33) | -1.07 (0.34) | -0.88 (0.34) | -1.16 (0.35) | -1.21 (0.35) | -1.10 (0.35) | -1.22 (0.35) | -1.25 (0.35) | -1.24 (0.35) | -1.33 (0.35) | -1.45 (0.35) | -1.55 (0.36) |

**Supplementary Table 4: Estimated thresholds of the SRN-Models 1-5 used to predict SRN.** Posterior average (Mean) values along with the standard deviation (SD) between brackets of threshold parameters ($\gamma_1, \gamma_2, .. \gamma_6$) were calculated.

| Model | $\gamma_1$ | $\gamma_2$ | $\gamma_3$ | $\gamma_4$ | $\gamma_5$ | $\gamma_6$ |
|---|---|---|---|---|---|---|
| **SRN-Model 1** | -3.77 (0.47) | -3.09 (0.46) | -2.11 (0.45) | -0.72 (0.41) | 1.13 (0.17) | 1.30 (0.10) |
| **SRN-Model 2** | -3.52 (0.40) | -2.85 (0.39) | -1.88 (0.38) | -0.50 (0.36) | 1.21 (0.14) | 1.34 (0.07) |
| **SRN-Model 3** | -3.76 (0.44) | -3.07 (0.43) | -2.10 (0.42) | -0.71 (0.39) | 1.13 (0.17) | 1.31 (0.10) |
| **SRN-Model 4** | -3.81 (0.43) | -3.13 (0.42) | -2.15 (0.41) | -0.75 (0.39) | 1.11 (0.16) | 1.30 (0.10) |
| **SRN-Model 5** | -3.70 (0.37) | -3.01 (0.35) | -2.03 (0.35) | -0.63 (0.33) | 1.17 (0.13) | 1.32 (0.08) |

**Supplementary Table 5: Prediction ability of the SRN-Models 1-5 used to predict SRN:** Proportion of cases correctly classified (PCCC) using five TGBLUP models; Brier scores obtained from LOO cross-validation of the five TGBLUP models. For each model, maximum (Max), minimum (Min) average (Mean) values along with the standard deviation (SD) between brackets were calculated.

| Model | Prediction ability by Brier Score | | | Proportion of cases correctly classified (PCCC) | |
|---|---|---|---|---|---|
| | Max | Mean | Min | Correctly predicted | Incorrectly predicted |
| **SRN-Model 1** | 0.65 | 0.36 (0.08) | 0.21 | 451 (47.93%) | 490 (52.07%) |
| **SRN-Model 2** | 0.65 | 0.36 (0.08) | 0.21 | 494 (47.50%) | 447 (52.50%) |
| **SRN-Model 3** | 0.65 | 0.36 (0.08) | 0.21 | 491 (47.82%) | 450 (52.18%) |

| | | | | | |
|---|---|---|---|---|---|
| **SRN-Model 4** | 0.65 | 0.36 (0.08) | 0.21 | 491 (47.82%) | 450 (52.18%) |
| **SRN-Model 5** | 0.65 | 0.36 (0.08) | 0.21 | 490 (47.93%) | 451 (52.07%) |

**Supplementary Table 6: Prediction ability by Pearson correlation and Estimated variance component of the log-normal Models 1-5 used to predict SRN:** Genomic heritability; Pearson correlation between estimated and observed SRN using LOO cross-validation of the five extended GBLUP models; L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG points out the variance of additive x additive epistatic effects. Numbers between brackets point out the standard deviation (SD) of the estimated variances.

| Model | Genomic heritability $H^2$ ($h^2$) | Prediction ability by Pearson correlation | Estimated variance component | | | | |
|---|---|---|---|---|---|---|---|
| | | | L | G | GxG | Error variance | Total Variance |
| **Model 1** | 0.50 | 0.79 | 0.38 (0.12) | | | 0.37 (0.12) | 0.75 |
| **Model 2** | 0.36 | 0.35 | | 0.29 (0.09) | | 0.51 (0.10) | 0.80 |
| **Model 3** | 0.51 (0.23) | 0.53 | | 0.19 (0.08) | 0.22 (0.09) | 0.39 (0.11) | 0.80 |
| **Model 4** | 0.52 | 0.60 | 0.22 (0.09) | 0.19 (0.08) | | 0.37 (0.11) | 0.78 |
| **Model 5** | 0.58 (0.38) | 0.65 | 0.15 (0.07) | 0.15 (0.09) | 0.16 (0.07) | 0.33 (0.11) | 0.79 |

**Supplementary Table 7: Prediction ability by Pearson correlation and Estimated variance component of the SRA-Models 1-5 used to predict SRA:** Genomic heritability; Pearson correlation between estimated and observed SRA using LOO cross validation strategy of the five extended GBLUP models; L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG points out the variance of additive x additive epistatic effects. Numbers between brackets point out the standard deviation (SD) of the estimated variances.

| Model | Genomic heritability $H^2$ ($h^2$) | Prediction ability by Pearson correlation | Estimated variance component | | | | |
|---|---|---|---|---|---|---|---|
| | | | L | G | GxG | Error variance | Total Variance |
| **SRA-Model 1** | 0.5 | 0.73 | 44.38 (14.74) | | | 44.34 (14.68) | 88.72 |
| **SRA-Model 2** | 0.32 | 0.19 | | 31.93 (11.80) | | 66.76 (13.24) | 98.70 |
| **SRA-Model 3** | 0.45 (0.2) | 0.37 | | 20.27 (8.94) | 24.60 (11.33) | 54.35 (14.08) | 99.23 |
| **SRA-Model 4** | 0.5 | 0.49 | 28.09 (12.07) | 20.13 (8.53) | | 46.98 (14.12) | 95.22 |
| **SRA-Model 5** | 0.46 (0.37) | 0.53 | 21.01 (10.47) | 14.87 (6.98) | 17.34 (8.81) | 44.14 (13.92) | 97.37 |

**Supplementary Table 8: Prediction ability by Pearson correlation and Estimated variance component of the TR-Models 1-5 used to predict TR at high evaporative demand:** Genomic heritability; Pearson correlation between estimated and observed TR using LOO cross-validation of the five extended GBLUP models; L indicates the estimated variance of line effects; G is the estimated variance of marker effects while GxG points out the variance of additive x additive epistatic effects. Numbers between brackets point out the standard deviation (SD) of the estimated variances.

| Model | Genomic heritability $H^2$ ($h^2$) | Prediction ability by Pearson correlation | Estimated variance component | | | | |
|---|---|---|---|---|---|---|---|
| | | | L | G | GxG | Error variance | Total Variance |
| **TR-Model 1** | 0.49 | 0.96 | 0.31 (0.10) | | | 0.32 (0.11) | 0.63 |
| **TR-Model 2** | 0.3 | 0.68 | | 0.22 (0.08) | | 0.49 (0.09) | 0.72 |
| **TR-Model 3** | 0.41 (0.19) | 0.89 | | 0.14 (0.06) | 0.16 (0.07) | 0.42 (0.10) | 0.73 |
| **TR-Model 4** | 0.49 | 0.95 | 0.21 (0.09) | 0.13 (0.05) | | 0.34 (0.10) | 0.69 |
| **TR-Model 5** | 0.5 (0.35) | 0.95 | 0.15 (0.08) | 0.10 (0.05) | 0.11 (0.05) | 0.34 (0.10) | 0.71 |

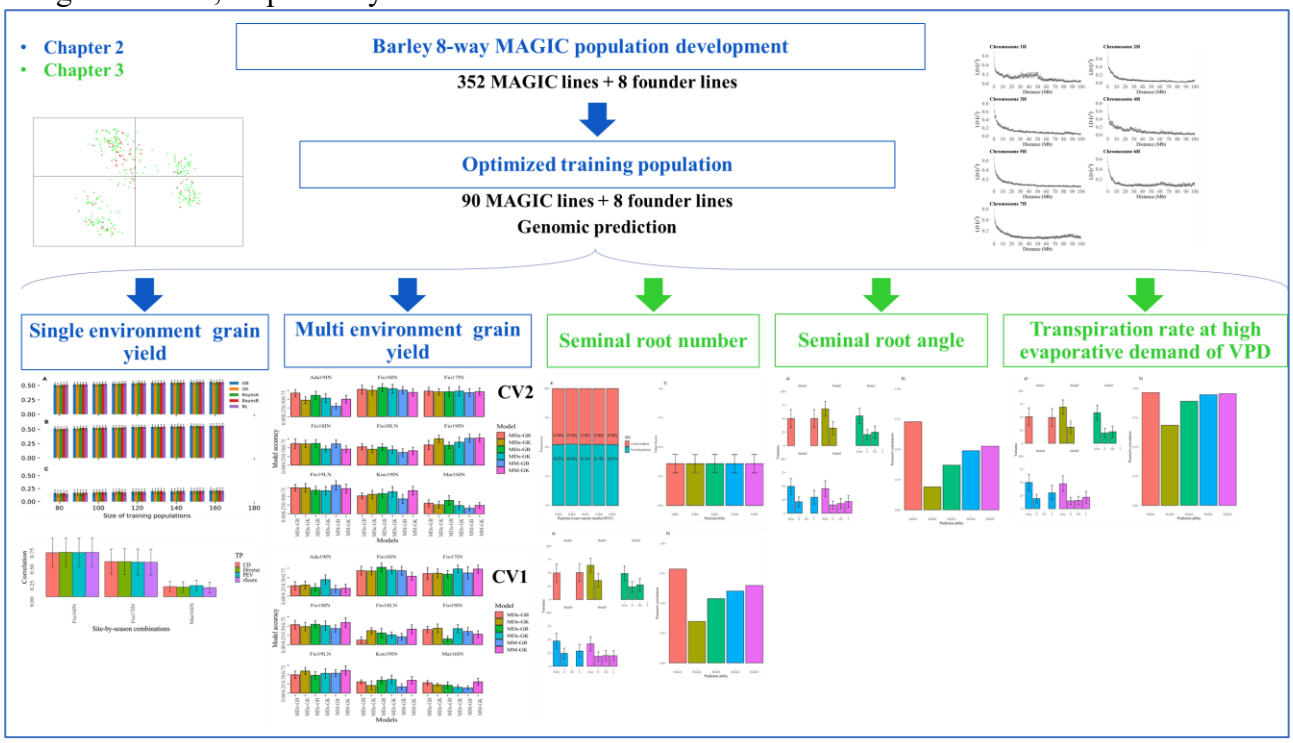
## 8 Data Availability Statement

The data underpinning the results presented in this manuscript are available from the authors upon request.

## General Conclusions

The results presented in this research work have contributed to acquiring new knowledge in the field of genomic prediction. Particularly, the present work has demonstrated that genomic prediction can be successfully applied to MAGIC populations as, in general, models fitted to these populations show high predictive ability and explain a large fraction of the observed phenotypic variability (**Figure 1**; **Chapter 2**). To date, actual breeding programs that implement genomic prediction are focused on grain yield. In **Chapter 3**, we have shown that genome-enabled prediction can be successfully applied to belowground (seminal root number and seminal root angle) and physiological (Transpiration rate at high evaporative demand of VPD) traits paving the way for using this methodology to other traits of agronomic interest. Interestingly, our empirical analysis (**Chapter 2**) shows that a limited panel of 90 lines is large enough to fit genomic prediction models with high predictive ability (**Figure 1**). This latter finding might potentially facilitate the use of MAGIC and genomic prediction in real breeding programs, where optimal trade-offs between costs and advantages are desirable.

**Figure 1: Flowchart of genomic prediction for grain yield, seminal root number, seminal root angle and transpiration rate at high evaporative demand of VPD develop for 8-way MAGIC population lines investigated in this Ph.D. thesis.** Chapter 2 and Chapter 3 are highlighted by blue and green colors, respectively.




To date, MAGIC population have been used to map QTL and have been limited impact on actual plant breeding programs. Moreover, MAGIC lines have never been used along with genomic prediction methodology: as shown in the present work, this approach might allow recycling data for several breeding populations derived from the same parents used for constructing MAGIC lines.

In cereal crops, direct selection for grain yield and above-ground traits has been used as the method of choice to select and develop superior genotypes[1]. Our empirical analyses showed that GP models can be used to predict grain yield, across an ample range of site-by-season combinations and site-by-season-by-management combinations characterized by different temperature and precipitation patterns, with values of predictive ability that range between 0.25 and 0.6 for both single and multi-environments (**Figure 1**).

In drought-prone or stressful environments direct selection based on grain yield is largely ineffective because this trait exhibits, in general, low levels of heritability, which in turn does not allow to realize the desired rate of genetic gain[2]. Multi-environment genomic prediction models allow the evaluation of GxE interaction and contribute selecting high-yielding genotypes. Alternatively, for harsh environments, indirect selection, that is the selection of grain yield by scoring one or more directly correlated secondary traits, is an alternative approach. This methodology could be used for breeding ideotypes designed for specific target environments and cropping systems and has been proposed as an alternative of empirical breeding[3]. Different ideotypes of future cereal cultivars have been proposed to cope with drought and climate change[4]. Designing and breeding for crop ideotypes require a priori knowledge of sets of belowground and physiological traits that can cope with the peculiar environmental conditions of target environments to increase yield. For instance, crop simulation models have pointed out that the belowground and physiological traits investigated in this research study can both increase grain yield and cope with drought tolerance, water-saving strategy, and thermal stress in stressful environments. Phenotyping these traits is time consuming, and no tools have been used to predict these targets quickly and for a large panel of lines. For this reason, despite they are key element to support varietal innovation and indirect selection, they are rarely used in breeding programs. Our empirical analysis showed that genomic prediction models can be used to predict seminal root number (0.36 and 0.58 of brier score and of Pearson correlation, respectively), seminal root angle (0.46 of Pearson correlation) and transpiration rate (0.94 of Pearson correlation) (**Figure 1**).

Recently, the evaluation of multiple traits in multiple environments is prompting breeders to implement the use of multi-traits genomic prediction models. These models have been used to predict agronomic and malting quality in barley[5] since, in some cases, the selection is based on trade-offs among the agronomic traits of interest. Therefore, it might be interesting to implement multi-trait genomic prediction models using the traits investigated in this research work to benchmark single and multi-trait models. The GP models implemented in **Chapter 3**, probably could be able to predict the grain yield of the breeding population of **Chapter 2**. Conversely, these models would allow to predict the grain yield starting from belowground and physiological traits.

[1] Lalić A, Novoselović D, Kovačević J, Drezner G, Babić D, Abičić I, Dvojković K. 2010. "Genetic gain and selection criteria effects on yield and yield components in barley (Hordeum vulgare L.)" Period Biol 112:311–316.

[2] Samarah. NH, 2005. "Effects of drought stress on growth and yield of barley". Agron Sustain Dev Springer Verlag/EDP Sci 25:145–149. https://doi.org/10.1051/agro.

[3] Donald CM. 1968. "The breeding of crop ideotypes". Euphytica 17:385–403. https://doi.org/10.1007/BF00056241

[4] Rötter RP, Tao F, Höhn JG, Palosuo T. 2015. "Use of crop simulation modelling to aid ideotype design of future cereal cultivars". J Exp Bot 66:3463–3476. https://doi.org/10.1093/jxb/erv098.

[5] Bhatta, M, Gutierrez L, Cammarota C, Cardozo F, Germán S, Gómez-Guerrero B, Pardo M F, Lanaro V, Sayas M, and Castro A J. 2020. "Multi-Trait Genomic Prediction Model Increased the Predictive Ability for Agronomic and Malting Quality Traits in Barley (Hordeum Vulgare L.)". G3: Genes, Genomes, Genetics 10(3):1113–24.

## Annex 1: R script implemented

**1 GBLUP for GY - Single Environment**

```
rm(list = ls())
install.packages(pkg= "BGLR", repos= "https://cran.r-project.org/")
library(BGLR)

####loading genotypic SNP dataset
SNP_data <- read.csv("/MyFiltred_SNP.csv")
####loading phenotypic GY dataset BLUPs
data <- read.table ("/MyBLUPs.csv", sep=",",header=TRUE,
na.strings="NA")
####deleting monomorphic SNP
SNP_data [,-1] <- SNP_data[,1][vapply(SNP_data[,-1], function(x)
length(unique(x))>1, logical(1L))]
data$Samples <- toupper(data$Samples)
SNP_data$Samples <- toupper (SNP_data$Samples)
data$Samples <- gsub ("^M[0]{1}", "M", data$Samples)
data$Samples <- gsub ("M0", "M", data$Samples)
data <- data[complete.cases(data),]
data$Samples <- as.character(data$Samples)
SNP2 <- merge (data, SNP_data, by.y = "Samples")
SNP3 <-SNP2 [,-c(1,2)]
y <- SNP2$BLUPs

####Example of Random Cross Validation strategy
yNA <- SNP2$BLUPs
set.seed(129)
tst <- sample(1:n, size=round(n/3,0), replace=FALSE)
yNA[tst] <- NA
ite<-52000
bur<-6000
t<-1

####Parametric Random Regression
#Bayesian ridge regression (BRR)
ETA <- list(list(X=SNP3, model="BRR"))
fm <- BGLR(y=yNA, ETA=ETA,nIter=ite, burnIn=bur,thin= t,saveAt= 'BRR_')
cor_pred <- cor(fm$yHat[tst], y[tst],use= "pairwise.complete.obs")
cor_unpred <- cor(fm$yHat[-tst], y[-tst])

####Semi-Parametric Regression
#Reproducing kernel Hilbert space (RKHS)
X <- scale (SNP3, center=TRUE, scale=FALSE)
D <- (as.matrix(dist(X, method='euclidean'))^2)/p
```

```
h <- 0.5
K <- exp (-h*D)
ETA<-list(list(K= K, model= 'RKHS'))
fm1<-BGLR(y= yNA, ETA= ETA, nIter= ite, burnIn= bur, thin= t, saveAt=
'RKHS_h=0.5_')
cor_predicted <- cor(fm1$yHat[tst], y[tst])
cor_unpredicted <- cor(fm1$yHat[-tst], y[-tst])

#Genomic Best Linear Unbiased Prediction (GBLUP)
#Van Raden Matrix
G<-tcrossprod(X)/n
ETA<-list(list(K= G, model= 'RKHS'))
fm2<-BGLR(y= yNA, ETA= ETA, nIter= ite, burnIn= bur, thin= t, saveAt=
'RKHS_')
cor_predicted<-cor(fm2$yHat[tst], y[tst])
cor_unpredicted<-cor(fm2$yHat[-tst], y[-tst])
```

## 2 GBLUP for GY - Multi Environment

```
rm(list = ls())
install.packages(pkg = "BGLR", repos = "https://cran.r-project.org/")
library (BGLR)

####loading genotypic dataset
SNP_data <- read.csv("/MyFiltred_SNP.csv")
####loading phenotypic GY dataset BLUPs
training_pop <- read.table ("/MyBLUPs.csv", sep = ",", header = TRUE,
na.strings = "NA")
SNP_data <- data.frame(SNP_data, stringsAsFactors = FALSE)
Samples <- rownames(SNP_data)
SNP_data <- data.frame(Samples, SNP_data, stringsAsFactors = FALSE)
SNP_data <- SNP_data[SNP_data$Samples %in% training_pop,]
####deleting monomorphic SNP
SNP_data[,-1] <- SNP_data[,1][vapply(SNP_data[,-1], function(x)
length(unique(x))>1, logical(1L))]
X <- scale (SNP_data[,-1],)
X <- scale (SNP_data[,-1],center=TRUE, scale=TRUE) #G-BLUP kernel (GB)
GB <- tcrossprod(X)/ncol(X)
Zg <- model.matrix(~factor(pheno_geno1$Entry)-1)
Ze <- model.matrix(~factor(pheno_geno1$Location)-1)
K1 <- Zg %*% GB %*% t(Zg)
ZEZE <- tcrossprod(Ze)
K2 <- K1*ZEZE
ite <- 80000
bur <- 10000
t <- 1
```

```
##Fitting “Multi-Environment, main genotypic effect” (MM)
ETA <- list(ENV=list(X=Ze, model="FIXED"), Grm=list(K=K1, model="RKHS",
df0 = 0.00001, S0 = 0.0001))
y <- scale(pheno_geno1$BLUP, center = TRUE, scale = TRUE)
fm_MM_FIXED <- BGLR(y=y, ETA=ETA, nIter=ite, burnIn=bur, thin=t, saveAt
= "MM__", df0 = 0.00001, S0 = 0.0001)

##Fitting “Multi-Environment, single variance GxE deviation model”(MDs)
ETA <- list(ENV=list(X=Ze, model="FIXED"), Grm = list(K=K1, model =
"RKHS", df0 = 0.00001, S0 = 0.0001), EGrm = list(K=K2, model="RKHS",
df0 = 0.00001, S0 = 0.0001))
y <- scale(pheno_geno1$BLUP,center = TRUE, scale = TRUE)
fm_MDs_FIXED <- BGLR(y=y, ETA=ETA, nIter=ite, burnIn=bur, thin=t,
saveAt = "MDs", df0 = 0.00001, S0 = 0.0001)

##Fitting “Multi-Environment, environment-specific variance GxE
#deviation model” (MDe)
ETA<-list(ENV=list(X=Ze, model = "FIXED", df0 = 0.00001, S0 = 0.0001),
Grm=list(K=K1, model = "RKHS", df0 = 0.00001, S0 = 0.0001))
for (k in 1:nEnv) {
    ZEE <- matrix(0, nrow = nrow(Ze), ncol = ncol(Ze))
    ZEE[,k] <- Ze[,k]
    ZEEZ <- (ZEE %*% t(Ze))
    K3 <- K1*ZEEZ
    ETA[[k+2]] <- list (K=K3, model="RKHS", df0 = 0.00001, S0 = 0.0001)
}
y <- scale(pheno_geno1$BLUP,center = TRUE, scale = TRUE)
fm_MDe_FIXED <- BGLR(y = y ,ETA = ETA, nIter = ite, burnIn = bur, thin=
t, df0 = 0.00001, S0 = 0.0001, saveAt = "MDe_")
```

## 3 TGBLUP for BGTs

```
rm(list = ls())
install.packages(pkg = "BGLR", repos = "https://cran.r-project.org/")
install.packages(pkg = "measures",repos ="https://cran.r-project.org/")
install.packages(pkg ="matrixcalc",repos="https://cran.r-project.org/")
library (BGLR)
library (measures)
library (matrixcalc)
####loading genotypic dataset
SNP_data <- read.csv("/MyFiltred_SNP.csv")
####loading phenotypic categorical dataset from raw data
data <- read.table("/MyCategorical_Variable_SRN.csv", sep = ",", header
= TRUE, na.strings = "NA")
data$Rep <- as.factor(data$Rep)
data$SRN <- as.factor(data$SRN)
```

```
####subset SNP data to select only genotypes with phenotypic records
SNP_data_filtered <- SNP_data[SNP_data$Samples %in% data$Geno,]
data_filtered <- data[data$Geno %in% SNP_data_filtered$Samples,]
length(unique(data_filtered$Geno))
data_filtered$Geno <- factor(data_filtered$Geno, ordered = T)

####deleting monomorphic SNP
SNP_data_filtered[,-1]<-SNP_data_filtered[,1][vapply(SNP_data_filtered
[,-1], function(x) length(unique(x))>1, logical(1L))]
SNPscaled <- scale (SNP_data_filtered[,-1], scale = TRUE,center = TRUE)

#KL line matrix
ZL<-model.matrix(~factor(data_filtered$Geno)-1)
KL<-tcrossprod(ZL)
#KG marker covariate matrix
GB<-tcrossprod(SNPscaled[,-1])/ncol(SNPscaled[,-1])
KG<- ZL %*% GB %*% t(ZL)
#KGG additive x additive epistasis matrix
GA<-hadamard.prod(GB,GB)
KGG<-ZL %*% GA %*% t(ZL)

ite <- 52000
bur <- 6000
t <- 1
BS <- ""

####Model used to fit the dataset based on KL
####Example of Leave-one-out (LOO) Cross Validation approach
ETA1 <- list(list(~factor(Rep), data= data_filtered, model="FIXED"),
list(K=KL, model="RKHS"))
for (i in 1:length(SNP_data_filtered$Samples)){
yNA <- data_filtered$SRN
 tst<- which(data_filtered$Geno == SNP_data_filtered$Samples[i])
 yNA[tst] <- NA
 print(SNP_data_filtered$Samples[i])
 fm1 <- BGLR(y=yNA, response_type = "ordinal", ETA = ETA1, nIter = ite,
burnIn= bur, thin= t, saveAt = "Model_1")
 probs <- fm1$probs[tst,]
 BS2 <- multiclass.Brier(probs, data_filtered$SRN[tst])/2
 BS <- c(BS, BS2)
 write.table(BS, "BrierScore_SRN_LOO_Model1.txt")
}

####Model used to fit the dataset based on KG
####Example of Leave-one-out (LOO) Cross Validation approach
ETA2 <- list(list(~factor(Rep), data=data_filtered, model="FIXED"),
list(K=KG, model="RKHS"))
```

```
for (i in 1:length(SNP_data_filtered$Samples)){
yNA <- data_filtered$SRN
 tst<- which(data_filtered$Geno == SNP_data_filtered$Samples[i])
 yNA[tst] <- NA
 print(SNP_data_filtered$Samples[i])
 fm1 <- BGLR(y=yNA, response_type = "ordinal", ETA = ETA2, nIter = ite,
burnIn= bur, thin= t, saveAt = "Model_2")
 probs <- fm1$probs[tst,]
 BS2 <- multiclass.Brier(probs, data_filtered$SRN[tst])/2
 BS <- c(BS, BS2)
 write.table(BS, "BrierScore_SRN_LOO_Model2.txt")
}

####Model used to fit the dataset based on KG and KGG
####Example of Leave-one-out (LOO) Cross Validation approach
ETA3 <- list(list(~factor(Rep), data=data_filtered, model="FIXED"),
list(K=KG, model="RKHS"), list(K=KGG, model="RKHS"))
for (i in 1:length(SNP_data_filtered$Samples)){
yNA <- data_filtered$SRN
 tst<- which(data_filtered$Geno == SNP_data_filtered$Samples[i])
 yNA[tst] <- NA
 print(SNP_data_filtered$Samples[i])
 fm1 <- BGLR(y=yNA, response_type = "ordinal", ETA = ETA3, nIter = ite,
burnIn= bur, thin= t, saveAt = "Model_3")
 probs <- fm1$probs[tst,]
 BS2 <- multiclass.Brier(probs, data_filtered$SRN[tst])/2
 BS <- c(BS, BS2)
 write.table(BS, "BrierScore_SRN_LOO_Model3.txt")
}

####Model used to fit the dataset based on KL and KG
####Example of Leave-one-out (LOO) Cross Validation approach
ETA4 <- list(list(~factor(Rep), data=data_filtered, model="FIXED"),
list(K=KL, model="RKHS"), list(K=KG, model="RKHS"))
for (i in 1:length(SNP_data_filtered$Samples)){
yNA <- data_filtered$SRN
 tst<- which(data_filtered$Geno == SNP_data_filtered$Samples[i])
 yNA[tst] <- NA
 print(SNP_data_filtered$Samples[i])
 fm1 <- BGLR(y=yNA, response_type = "ordinal", ETA = ETA4, nIter = ite,
burnIn= bur, thin= t, saveAt = "Model_4")
 probs <- fm1$probs[tst,]
 BS2 <- multiclass.Brier(probs, data_filtered$SRN[tst])/2
 BS <- c(BS, BS2)
 write.table(BS, "BrierScore_SRN_LOO_Model4.txt")
}
```

```
####Model used to fit the dataset based on KL and KG and KGG
####Example of Leave-one-out (LOO) Cross Validation approach
ETA5 <- list(list(~factor(Rep), data=data_filtered, model="FIXED"),
list(K=KL, model="RKHS"), list(K=KG, model="RKHS"), list(K=KGG,
model="RKHS"))
for (i in 1:length(SNP_data_filtered$Samples)){
yNA <- data_filtered$SRN
 tst<- which(data_filtered$Geno == SNP_data_filtered$Samples[i])
 yNA[tst] <- NA
 print(SNP_data_filtered$Samples[i])
 fm1 <- BGLR(y=yNA, response_type = "ordinal", ETA = ETA5, nIter = ite,
burnIn= bur, thin= t, saveAt = "Model_5")
 probs <- fm1$probs[tst,]
 BS2 <- multiclass.Brier(probs, data_filtered$SRN[tst])/2
 BS <- c(BS, BS2)
 write.table(BS, "BrierScore_SRN_LOO_Model5.txt")
}
```

## Annex 2: Published article

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 May 2021
doi: 10.3389/fpls.2021.664148

# Genomic Prediction of Grain Yield in a Barley MAGIC Population Modeling Genotype per Environment Interaction

*Damiano Puglisi[1†], Stefano Delbono[2†], Andrea Visioni[3], Hakan Ozkan[4], İbrahim Kara[5], Ana M. Casas[6], Ernesto Igartua[6], Giampiero Valè[7], Angela Roberta Lo Piero[1], Luigi Cattivelli[2], Alessandro Tondelli[2] and Agostino Fricano[2*]*

[1] Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A), Università di Catania, Catania, Italy, [2] Council for Agricultural Research and Economics–Research Centre for Genomics and Bioinformatics, Fiorenzuola d'Arda, Italy, [3] Biodiversity and Crop Improvement Program, International Center for Agricultural Research in the Dry Areas, Avenue Hafiane Cherkaoui, Rabat, Morocco, [4] Department of Field Crops, Faculty of Agriculture, University of Cukurova, Adana, Turkey, [5] Bahri Dagdas International Agricultural Research Institute, Konya, Turkey, [6] Aula Dei Experimental Station (EEAD-CSIC), Spanish Research Council, Zaragoza, Spain, [7] DiSIT, Dipartimento di Scienze e Innovazione Tecnologica, Università del Piemonte Orientale, Vercelli, Italy

OPEN ACCESS

**Edited by:**
*Sukhjiwan Kaur,*
*Agriculture Victoria, Australia*

**Reviewed by:**
*Ezio Portis,*
*University of Turin, Italy*
*Sivakumar Sukumaran,*
*International Maize and Wheat Improvement Center (Mexico), Mexico*

***Correspondence:**
*Agostino Fricano*
*agostino.fricano@crea.gov.it*

*†These authors have contributed equally to this work*

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

**Received:** *04 February 2021*
**Accepted:** *26 April 2021*
**Published:** *24 May 2021*

**Citation:**
*Puglisi D, Delbono S, Visioni A, Ozkan H, Kara İ, Casas AM, Igartua E, Valè G, Piero ARL, Cattivelli L, Tondelli A and Fricano A (2021) Genomic Prediction of Grain Yield in a Barley MAGIC Population Modeling Genotype per Environment Interaction. Front. Plant Sci. 12:664148. doi: 10.3389/fpls.2021.664148*

Multi-parent Advanced Generation Inter-crosses (MAGIC) lines have mosaic genomes that are generated shuffling the genetic material of the founder parents following pre-defined crossing schemes. In cereal crops, these experimental populations have been extensively used to investigate the genetic bases of several traits and dissect the genetic bases of epistasis. In plants, genomic prediction models are usually fitted using either diverse panels of mostly unrelated accessions or individuals of biparental families and several empirical analyses have been conducted to evaluate the predictive ability of models fitted to these populations using different traits. In this paper, we constructed, genotyped and evaluated a barley MAGIC population of 352 individuals developed with a diverse set of eight founder parents showing contrasting phenotypes for grain yield. We combined phenotypic and genotypic information of this MAGIC population to fit several genomic prediction models which were cross-validated to conduct empirical analyses aimed at examining the predictive ability of these models varying the sizes of training populations. Moreover, several methods to optimize the composition of the training population were also applied to this MAGIC population and cross-validated to estimate the resulting predictive ability. Finally, extensive phenotypic data generated in field trials organized across an ample range of water regimes and climatic conditions in the Mediterranean were used to fit and cross-validate multi-environment genomic prediction models including G×E interaction, using both genomic best linear unbiased prediction and reproducing kernel Hilbert space along with a non-linear Gaussian Kernel. Overall, our empirical analyses showed that genomic prediction models trained with a limited number of MAGIC lines can be used to predict grain yield with values of predictive ability that vary from 0.25 to 0.60 and that beyond QTL mapping and analysis of epistatic effects, MAGIC population might be used to successfully fit genomic prediction models.

We concluded that for grain yield, the single-environment genomic prediction models examined in this study are equivalent in terms of predictive ability while, in general, multi-environment models that explicitly split marker effects in main and environmental-specific effects outperform simpler multi-environment models.

**Keywords: genomic prediction, MAGIC, barley, GBLUP, genotype x environment interaction**

## INTRODUCTION

The experimental design that underlies Multi-parent Advanced Generation Intercrosses (MAGIC) populations traces its origins to the advanced inter-cross lines, which were originally developed in animal model species (Yalcin et al., 2005). MAGIC populations are developed crossing multiple inbred parents or founders, which are subsequently inter-mated several times following pre-defined crossing schemes to shuffle founder genomes in each single line (Huang et al., 2015). In plants, MAGIC populations have been explicitly developed for genetic research purposes as they allow to increase power and precision for detecting and mapping quantitative trait loci (QTLs) (Cavanagh et al., 2008; Huang et al., 2015; Scott et al., 2020). Theoretically, MAGIC populations have the potential to dissect the genetic bases of complex traits at sub-centimorgan scale, allowing to overcome common issues related to the use of biparental families for QTL mapping and detection such as low-resolution power, low genetic diversity of parents and limited number of recombination events (Valdar et al., 2006). In cereal crops, MAGIC populations have been developed and established for rice (Bandillo et al., 2013; Ponce et al., 2018), bread wheat (Mackay et al., 2014; Sannemann et al., 2018; Stadlmeier et al., 2018), maize (Dell'Acqua et al., 2015; Jiménez-Galindo et al., 2019) and barley (Mathew et al., 2018) and to date they have been deployed for unraveling the genetic bases of biotic and abiotic stresses, grain yield (GY) and seed quality traits. Beyond the aforementioned applications, barley MAGIC populations have been recently exploited to disentangle the effect of epistasis on flowering time (Mathew et al., 2018; Sannemann et al., 2018; Afsharyan et al., 2020).

Similarly to MAGIC, the theory underlying genomic prediction (GP) was originally developed and deployed in animal species. The pivotal component of GP is a population of individuals having phenotypic and genotypic information, which is known as training population (TP) and is used to regress genome-wide single nucleotide polymorphisms (SNPs) or other types of DNA markers on phenotypes to simultaneously predict their effects (Meuwissen et al., 2001), that is for training GP models. Trained GP models are subsequently used in combination with the genotypic information of candidate individuals that must be selected for computing their genomic estimated breeding values (GEBVs) and ranking them to apply truncation selection (Meuwissen et al., 2001; Heffner et al., 2009). This latter population of candidate individuals having only genotypic information is known as breeding population (BP) (Meuwissen et al., 2001; Heffner et al., 2009). To date, GP has been largely applied for crop improvement fitting GP models trained with individuals from either biparental families or diversity panels of mostly unrelated accessions. As the genetic relatedness of TP and BP affects the prediction ability of GP models (Ben Hassen et al., 2018; Norman et al., 2018), these two approaches have profound differences in terms of versatility as DNA marker effects estimated on diversity panels have the potential of a broader applicability and might be used in different breeding programs (Bassi et al., 2015), while GP models trained with individuals of biparental families can allow to accurately predict the performance of offspring produced within the same cross.

Typically, GP models require to regress a number of predictors (DNA markers) that greatly exceeds the number of observations or phenotypes and several parametric and non-parametric models have been proposed to deal with overfitting and the "large $p$, small $n$" problem (Meuwissen et al., 2001; Jannink et al., 2010; Pérez and de los Campos, 2014) as in these conditions the estimation of marker effects using ordinary least squares method is not practicable. A commonly used solution is to estimate marker effects jointly using the Least Absolute Shrinkage and Selection Operator (LASSO) method (Tishbirani, 1996) and its Bayesian counterpart (Bayesian Lasso or BL), which uses a penalizing or regularization parameter ($\lambda$) that denotes the amount of shrinkage for regressing markers (De Los Campos et al., 2009). Other popular whole genome regression methods based on Bayesian theory are BayesA and BayesB (Meuwissen et al., 2001), which relax the assumption of common variance across marker effects adopted in other models (e.g., ridge regression) and allow each marker to have its own variance. Differently to BayesA, BayesB allows having markers with no effects in the model and theoretically assumes more realistic conditions as it is plausible that a large fraction of genome-wide markers does not contribute to explaining the observed phenotypic variance. Beyond these methods, whole genome regression based on reproducing kernel Hilbert space (RKHS) has been proposed and applied to implement GP models (Gianola and Van Kaam, 2008; Gota and Gianola, 2014). In the RKHS regression, a reproducing kernel, that is any positive definite function for mapping from pairs of points in input space to other pairs of points, is used to transform DNA markers of individuals in square distance matrix that are used in a linear model (Gota and Gianola, 2014). The Gaussian Kernel (GK) is one of the most common function used as reproducing kernel and depends

**Abbreviations:** DH, Days-to Heading; PH, Plant Height; GY, Grain Yield; MAGIC, Multi-parent Advanced Generation Inter-crosses; TP, Training Population; BP, Breeding Population; GP, Genomic Prediction; GBLUP, Genomic Best Linear Unbiased Prediction; RKHS, Reproducing Kernel Hilbert Space; GK, Gaussian Kernel; LD, Linkage Disequilibrium; PCA, Principal Component Analysis; QTL, Quantitative Trait Locus; GEBV, Genomic Estimated Breeding Value; SNP, Single Nucleotide Polymorphism; SE-GP, Single Environment Genomic Prediction; ME-GP, Multi Environment Genomic Prediction.

on the bandwidth (or smoothing) parameter *h* that controls the decay rate of the kernel as two points step away. Several studies have shown that the use of GK in combination with RKHS improves the prediction of genetic values if the bandwidth parameter *h* is correctly chosen (Pérez-Elizalde et al., 2015). Moreover as RKHS regression does not assume linearity, this model might allow to better capture non-additive effects without explicitly including epistatic interactions and dominance in GP models (Gianola and Van Kaam, 2008). Differently from methods based on whole genome regression of markers, the genomic best linear unbiased prediction (GBLUP) method treats genomic values of individuals as random effects in a linear mixed model and uses a genomic relationship matrix based on DNA marker data to compute GEBVs (VanRaden, 2008; Wang et al., 2018). Notably, the use of RKHS along with the genomic relationship matrix is equivalent to the mixed linear model of GBLUP, that is GBLUP method represents a special case of RKHS regression (Gota and Gianola, 2014).

The effectiveness of GP depends, among other factors, on the degree of correlation between GEBVs and true genetic values that is the predictive ability of the model. In practice, the predictive ability is evaluated using the Pearson's correlation coefficient between GEBVs and the realized phenotypes or other estimators (e.g., adjusted means). To date several empirical studies have been conducted for fitting GP models on biparental populations and panels of mostly unrelated accessions across different species and traits, which point out that, depending on the genetic architecture of the trait, each statistical model has its own advantages and disadvantages in term of predictive ability and estimation of marker effects (Heslot et al., 2012; Ben Hassen et al., 2018). Other factors that strongly influence the predictive ability are the size of the TP, its structure, and its relatedness with the BP (Desta and Ortiz, 2014). Several targeted and untargeted methods have been developed to optimize the composition of TP for maximizing the predictive ability for a given set of individuals (Rincent et al., 2012; Akdemir et al., 2015). Nevertheless, these methods generally generate trait-dependent TPs which might hamper the implementation of these procedures in real breeding programs.

The first objective of the present study was to create a new barley MAGIC population using a diverse founder set of old and new 6-rowed, winter cultivars showing contrasting GY, which was examined across an ample range of site-by-season combinations characterized by different temperature and precipitation patterns. The second objective of this study was to combine data collected across these field trials with genotypic information to fit different single-environment genomic prediction (SE-GP) and multi environment genomic prediction (ME-GP) models for empirically assessing the predictive ability in multi-parent populations. Moreover, we applied different untargeted optimization methods to this MAGIC population for assembling and benchmarking the performance of optimized TPs. Fitting SE-GP and ME-GP models to MAGIC lines, we aimed at broadening the use of these experimental populations beyond classical QTL mapping and analysis of epistatic effects for sustaining and accelerating barley breeding.

# MATERIALS AND METHODS

## Development of the Barley MAGIC Population

The MAGIC population used in this study was developed using a founder set of eight 6-rowed barley genotypes with a winter growth habit, which were selected on the basis of their pedigrees and similarity in days-to-heading (DH) (**Table 1**). At the first stage of MAGIC development, four $F_1$ populations were created crossing one of the four old 6-rowed barley varieties (Hatif de Grignon, Dea, Robur and Athene) with one of the four 6-rowed modern barley varieties (Ponente, Ketos, Aldebaran and Fridericus). At the second stage of MAGIC development, half-diallel crosses of these four $F_1$ individuals were carried out to generate six sets of plants. Finally, these six sets of genotypes, each of which contained the alleles of four out eight founder parents, were appropriately crossed in predefined funnel schemes to combine the genome of the eight founders in single lines. Differently from the original crossing schemes developed for constructing MAGIC populations (Cavanagh et al., 2008), instead of recursively self-fertilizing these plants for several generations, seeds of the eight-way inter-crosses were sent to an external lab (SAATEN-UNION GmbH, Germany) to generate 352 inbred MAGIC lines using doubled haploid technology.

## Field Trials and Plant Phenotyping

The MAGIC population of 352 inbred individuals and the eight founder parents (**Table 1**) were sown during the fall of two consecutive growing seasons (2015–2016 and 2016–2017) in Fiorenzuola d'Arda (Italy) at CREA-Centro di Genomica e Bioinformatica (44°55′39.0"N 9°53′40.6"E, 78 m above sea level), using an alpha-lattice design with two-replicates. The whole set of MAGIC and the founder parents were also sown during the fall of 2015–2016 growing season in Marchouch (Morocco) at the Experimental station (33°36′43.5" N 6°42′53.0"W, 390 m above sea level) of the "International Center for Agricultural Research in the Dry Areas" using the same experimental design. Similarly, the subset of 82 MAGIC lines included in the optimized TP (TP-Diverse) and the eight founder parents were sown during the fall in 2017–2018 and 2018–2019 growing seasons in Fiorenzuola d'Arda under two different levels of nitrogen fertilization using alpha lattice experimental designs with two replicates. Trials conducted under ideal nitrogen conditions were fertilized with 100 kg/ha of nitrogen applied in two doses: 50 kg/ha were used at the sowing and 50 kg/ha were applied at the stem elongation stage. Field trials conducted under low nitrogen conditions received 50 kg/ha of nitrogen, 25 of which were applied at sowing while the remaining amount was applied at the stem elongation stage. In the growing season 2018–2019, other two field trials were conducted in Konya (Turkey) (37°53′37.9"N 32°37′26.0"E, 1,005 m above sea level) and in Adana (Turkey) (36°59′52.9"N 35°20′28.0"E, 24 m above sea level) to phenotype the optimized TP (TP-Diverse) using the same experimental design. For each trial considered in this study, plots of three square meters and a sowing density of 350 seeds per square meter were adopted, respectively. Local check cultivars were included

**TABLE 1 |** Founder set of barley varieties that were intermated for creating the barley MAGIC population.

| Genotype | Year of release | Country of release | Pedigree | DH (days) | PH (cm) | GY (t/ha) |
|---|---|---|---|---|---|---|
| Hatif de Grignon | 1937 | France | Selection from French landraces | 208.3 | 95.9 | 4.1 |
| Dea | 1953 | Germany | [(Ragusa x Peragis12) × (Heils Franken × Frw.Berg)] x [(Ragusa × Mahnd.Viktoria) (Ragusa × Bolivia)] | 212.1 | 95.3 | 6.0 |
| Robur | 1973 | France | Ager × (Hatif de Grignon × Ares) | 208.3 | 78.8 | 6.3 |
| Athene | 1977 | Germany | (Herfodia × Hord.sp.nigrum H204) × (Madru x Weissenhaus-Stamm) | 211.5 | 94.0 | 6.0 |
| Ponente | 2001 | Italy | (Vetulkio × Arma) × Express | 209.7 | 85.0 | 6.3 |
| Ketos | 2002 | France | (Gotic x Orblonde) × (12813 × 91H595) | 208.6 | 81.9 | 6.8 |
| Aldebaran | 2003 | Italy | Rebelle × Jaidor | 208.5 | 83.0 | 7.2 |
| Fridericus | 2006 | Germany | Carola × LP 6–564 | 211.6 | 89.3 | 7.3 |

*For each genotype of the founder set, the adjusted means of days to heading (DH), plant height (PH) and grain yield (GY) scored in eight different trials were reported along with available pedigree information.*

as internal checks in all experiments to compare phenotypes with trait observations collected in past seasons. Common protocols were adopted for each trial to phenotype plant genotypes for GY and DH. Phenotyping of MAGIC lines for GY was conducted as follows: from each plot grains were collected using a combine harvester and the total grain weight recorded in each plot was converted in tons per hectare. DH was measured as the number of days between sowing date and the date of heading stage, which was defined when 50% of the plants in a plot were at Zadoks' 55 growth stage (Zadoks et al., 1974). For each trial, phenotypic data of GY used in GP models were centered by subtracting the overall mean and standardized dividing by the sample standard deviation.

## Statistical Models for Computing the Adjusted Means of GY

The adjusted means of GY were computed in each site-by-season combination and across environments including DH as fixed covariate using the approach described in Emrich et al., 2008. The resulting model for computing the adjusted means of GY collected in field trials organized according to alpha-lattice design was:

$$y_{ijk} = 1\mu + Rep_i + Block_j(Rep_i) + Gen_k + DH_k + e_{ijk} \quad (1)$$

where $y_{ijk}$ is the response variable, that is the raw GY, $\mu$ is the general mean, $Rep_i$ is the effect of the $i^{th}$ replicate, $Block_j(Rep_i)$ is the effect of the $j^{th}$ incomplete block within the $i^{th}$ replicate, $Gen_k$ is the random effect of the $k^{th}$ genotype and DH is the effect of "Days-to-heading" covariate measured in each plot. In this model it is supposed that the random effects of $Gen_k$ follow a normal distribution with mean 0 and variance $\sigma_g^2$, that is $Gen_k \sim NIID\left(0, \sigma_g^2\right)$, and similarly, the residual terms $e_{ijk}$ are normally distributed with mean 0 and variance equals to $\sigma^2$, that is $e_{ijk} \sim NIID(0, \sigma^2)$. The adjusted GY values obtained predicting the random terms $Gen_k$ from the aforementioned model were used as phenotypes for training GP models. The linear mixed model reported in Equation 1 was fitted for each site-by-season combination using R 3.6.2 statistical environment and lme4 package (Bates et al., 2015) and variance components of fitted models were used to compute broad sense heritability ($H^2$) of GY.

## Genotyping of Genetic Materials

DNA was extracted from plant leaves using the Macherey Nagel Plant II extraction kit (Macherey Nagel, Dueren, Germany) and analyzed using gel electrophoresis and Quant-iT$^{TM}$ PicoGreen$^{TM}$ dsDNA Assay Kit (ThermoFisher, Grand Island, NY, United States) following manufacturer's instructions to assess quality and concentration, respectively. DNA samples were shipped to a propel-certified service provider (Trait Genetics GmbH, Gatersleben, Germany) and fingerprinted using the Illumina Infinium technology along with the Barley 50 k iSelect SNP Array (Bayer et al., 2017). To update the physical positions of SNP markers interrogated with the Barley 50 k iSelect SNP Array, probe sets used to design this array were mapped against the new reference sequence of barley (Monat et al., 2019). The raw genotyping table was imported in R software using "synbreed" package (Wimmer et al., 2012) to filter out markers with more than 10% of missing data and impute remaining missing data using Beagle 4.1 (Browning and Browning, 2016). 20 random leaf samples from field trials organized in Adana and Marchouch were genotyped using Illumina Infinium technology and Barley 50 k iSelect SNP Array to assess whether mislabelling of genotypes occurred during phenotyping operations and data collection.

## Clustering and Linkage Disequilibrium Analyses of the MAGIC Population

Principal component analysis was used to assess the diversity of the whole MAGIC population and was carried on imputed SNP data of the 352 MAGIC lines and the eight founders using ade4

package along with R version 3.6.2 (Thioulouse et al., 2018; R Core Team, 2019)., 2018). The first two principal components were used to visualize the dispersion of MAGIC lines in a graph. Linkage disequilibrium between pairs of markers was measured using $r^2$ (Hill and Robertson, 2008) in the subset of MAGIC genotypes included in the optimized TP and computed using Plink 1.9 software (Purcell et al., 2007; Chang et al., 2015).$r^2$ values showing $p$-values above 0.001 were filtered out, while the remaining pairwise $r^2$ values were imported and examined with a custom script developed for R 3.6.2 (R Core Team, 2019) to compute the mean $r^2$ in 100 kb windows, which was plotted in R 3.6.2 using ggplot2 package (Wickham, 2016).

## Statistical Models Used for Fitting SE-GP

SE-GP models were fitted using BayesA, BayesB and BL models (Tishbirani, 1996; Meuwissen et al., 2001; Park and Casella, 2008). Moreover, RKHS regression models were fitted using a linear GBLUP kernel (GB) and a non-linear GK (Gianola and Van Kaam, 2008; Gota and Gianola, 2014). For the GK, that is $K\left(x_i, x_i^{'}\right) = e^{-\left(h*d_{ii}^{2\prime}\right)}$, where $d_{ii}^{2\prime}$ points out the squared Euclidean distance between individuals $i$ and $i^{'}$, the rate of decay imposed by the bandwidth parameter $h$, was estimated using an empirical Bayesian methodology (Pérez-Elizalde et al., 2015) modifying published R codes (Cuevas et al., 2016).

## Statistical Models Used for Fitting ME-GP

Beyond SE-GP models, the adjusted means of GY computed across different site-by-season combinations were fitted to three previously described ME-GP models. Following the model nomenclature reported in Bandeira e Sousa et al. (2017), these three models were indicated in this study as "multi-environment, main genotypic effect" (MM) model (Jarquín et al., 2014; López-Cruz et al., 2015; Bandeira e Sousa et al., 2017), "multi-environment, single variance G×E deviation model" (MDs) (Jarquín et al., 2014; Bandeira e Sousa et al., 2017) and the "multi-environment, environment-specific variance G×E deviation model" (MDe) (López-Cruz et al., 2015; Bandeira e Sousa et al., 2017). Site-by-season combinations were considered as environments in MM, MDs and MDe regression models, which are briefly defined and summarized as follows. In the MM model, environments were considered as fixed effects while the random genetic effects were considered constant across all environments without modeling marker x environment interactions. Following matrix notation, the MM regression model is defined as follows:

$$y = 1\mu + Z_e\beta_e + Z_u u + \varepsilon \quad (2)$$

where y is the vector of observations collected in all environments, is the overall mean, $Z_e$ is the incidence matrix that connects observed phenotypes to the environments in which they were measured, $\beta_e$ is the vector of environmental fixed effects that must be estimated, $Z_u$ is an incidence matrix connecting genotypes with phenotypes for each environment, u is the vector of random genetic effects that must be predicted while ε is a vector of model residuals. In this model, marker genetic effects are assumed as $u \sim N\left(0, \sigma_{\mu 0}^2 K\right)$, that is, they follow a multivariate normal distribution with mean and variance-covariance matrix equal to zero and $\sigma_{\mu 0}^2 K$, respectively. The term $\sigma_{\mu 0}^2$ of the variance-covariance matrix is the variance of additive genetic effects across environments, while $K$ can be either a genomic relationship matrix (VanRaden, 2008) or a kernel function as discussed below. Model residuals of the vector are assumed to be independent and normally distributed with null mean and variance equal to$\sigma_e^2$, that is$\varepsilon \sim N\left(0, I\sigma_e^2\right)$, where $I$ points out the identity matrix. Overall, the MM regression model estimates marker effects across all environments and does not split them in main marker effects and in environmental-specific effects as in MDs and MDe models. As already substantiated in López-Cruz et al. (2015), for balanced field trial designs, MM is equivalent to fitting a genomic regression model using the average performance of each line across environments as phenotype.

Differently from the MM model, the MDe model allows markers to assume different effects in each $j^{th}$ environment (López-Cruz et al., 2015; Bandeira e Sousa et al., 2017), and consequently allows to account for marker x environment interactions. This model assumes that the effects of the $j^{th}$ environments, and the effects of markers are separated into two components, which are the main effect of markers for all environments, names as $b_{0k}$, and the peculiar random effect $b_{ik}$, of the markers in each $j^{th}$ environment, that is the effects of marker x environment interactions (López-Cruz et al., 2015). Consequently, in MDe models, the effect of the $k^{th}$ marker on the $j^{th}$ environment ($\beta_{jk}$) is described as the sum of an effect common to all environments ($b_{0k}$), plus a random deviation ($b_{ik}$) peculiar to the $j^{th}$ environment, that is $\beta_{jk} = b_{0k} + b_{ik}$.

Following matrix notation, the MDe regression model is defined as follows:

$$y = 1\mu + Z_e\beta_e + Z_u u_o + u_E + \varepsilon \quad (3)$$

where, $Z_e$, $_e$ have the same meaning of the MM regression model, $u_o$ represents the main effect of markers across all environments with a variance–covariance structure similar to MM model, that is, $uN(0,\sigma_{\mu 0}^2 K)$. As pointed out by López-Cruz et al. (2015) $\sigma_{\mu 0}^2$ is common to all environments, and the borrowing of information among environments is generated through the kernel matrix $K$. $u_E$ points out the specific effects of marker x environment interactions, which follow a multi-variate normal distribution with null mean and a variance–covariance matrix $K_E$, that is, $u_E N(0, K_E)$. For $j$ environments, the variance-covariance matrix $K_E$ is defined as follows:

$$K_e = \begin{bmatrix} \sigma_{\mu E1}^2 K_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\mu Em}^2 K_m & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \sigma_{\mu Ej}^2 K_j \end{bmatrix}$$

As explained in Bandeira e Sousa et al. (2017), $K_E$ can be discomposed as a sum of $j$ matrices, one for each $j$ environment.

Consequently, the interaction term $u_E$ can be decomposed in $j$ environmental specific effects to transform equation 3 as follows:

$$y = 1\mu + Z_e\beta_e + Z_u u_0 + u_{E1} + u_{E2} + u_{E3} + \ldots + u_{Ej} + \varepsilon \quad (4)$$

where each interaction effect $u_{Ej}$ has a normal distribution with null mean and a variance-covariance structure $\sigma^2_{\mu\ Ej}K_j$.

Starting from the MM regression model, the MDs model adds the random interaction effect of the environments with the genetic information of the lines pointed out with ue. Following matrix notation, the MDs modes is described as follows:

$$y = 1\mu + Z_e\beta_e + Z_u u + ue + \varepsilon \quad (5)$$

where, $Z_e$, $\beta_e$, $Z_u$, u and $\varepsilon$ have the same meaning of the MM regression model. As substantiated in Jarquín et al. (2014) the interaction term ue has a multi-variate normal distribution with null mean and variance-covariance matrix equal to $\left[Z_u K Z_u^{'}\right]^{\circ}\left[Z_E Z_e^{'}\right]$, where the Haddamar product operator denotes the element to element product between the two matrices in the same order.

In the present study, MM, MDs and MDe regression models were fitted using either the linear GB kernel method (VanRaden, 2008) or the non-linear GK method (Bandeira e Sousa et al., 2017). For the linear GB kernel method, the matrix $K$ of the aforementioned models was the genomic relationship matrix and was computed as $K = \left(\frac{XX^{'}}{p}\right)$ (VanRaden, 2008), where X is the standardized matrix of molecular markers for the individuals, of order $n$ by $p$; where $n$ and $p$ are the number of observations and the number of markers, respectively. For GK method, the matrix $K$ of MM, MDs and MDe regression models was computed as $K_j\left(x_{ij}, x_{ij}^{'}\right) = e^{-\left(h_j * d_{ii}^{2'}\right)}$ where $d_{ii}^{2'}$ is the squared Euclidean distance of the markers genotypes in individuals $i$ and $i$'' for the $j^{th}$ environment. Similarly to SE-GP models, the bandwidth parameter $h$ was computed using an empirical Bayes method (Pérez-Elizalde et al., 2015; Cuevas et al., 2016).

MM, MDs and MDe regression models used in this study were fitted using BGLR package 1.08 (Pérez and de los Campos, 2014) in R 3.6.2 statistical environment, adapting scripts provided in the framework of other studies (Bandeira e Sousa et al., 2017). For each model implemented in this study, predictions were based on 500,000 iterations collected after discarding 10,000 iterations for burn-in period-and using a thinning interval of five iterations. Trace plots for each of the variance parameters were created to assess whether the number of burn-in iterations was sufficient.

## Optimization of the TPs

In this study three different untargeted optimization criteria based on coefficient of determination (Laloe, 1993), predictive error variance (Rincent et al., 2012) and rScore (Ou and Liao, 2019) were used to assemble three corresponding TPs, each of which groups a set of 90 MAGIC individuals. The R package TSDFGS (Ou and Liao, 2019) was used to assemble these three optimized TPs using the aforementioned criteria. A fourth empirical untargeted optimization criterion was adopted for assembling another TP from the whole MAGIC population and aimed at maximizing the average distance between each selected accession and the closest other line using the modified Roger's distance (Thachuk et al., 2009). This criterion was implemented in R 3.6.2 using the heuristic algorithm implemented in the package Core Hunter3 (De Beukelaer et al., 2018) and was used to select a subset of 82 out 352 MAGIC individuals along with the eight MAGIC founder parents.

## Cross Validation Schemes

In this study several cross-validation (CV) schemes were adopted for estimating the predictive ability of GP models along with their standard errors (Burgueño et al., 2012; Gianola and Schon, 2016). For estimating the predictive ability of SE-GP models implemented with BayesA, BayesB, Bayesian Lasso, GB and RKHS with GK, cross validation was carried out using 100 repeated random partitioning of MAGIC population into training and validation sets. Using increasingly larger TPs of 80, 90, 100, 110, 120, 130, 140, 150, and 160 individuals, CV schemes were applied to compute mean and standard deviation of predictive ability for each TP size. Totally 4,500 models were fitted to carry out this CV experiment, combining the five statistical models with the aforementioned dimensions of the TP and 100 repeated random partitioning of MAGIC in training and validation sets.

Cross-validation of SE-GP models fitted using optimized TPs was carried out using the standard leave-one-out (LOO) strategy to estimate their predictive ability (Gianola and Schon, 2016). Basically, using LOO strategy, N GP models are fitted using N-1 individuals excluding recursively one individual from the TP and the GEBV of the excluded line is predicted from a model trained using all other lines. In our LOO experiment, this was carried out separately for each group of 90 lines included in the optimized TPs, and the accuracy of these predictions was calculated as the Pearson's correlation coefficient between GEBVs and the corresponding adjusted means of GY.

The predictive ability of ME-GP models was assessed using cross-validation 1 (CV1) and cross-validation 2 (CV2) schemes (Burgueño et al., 2012), assigning 90% of lines to the training set and the remaining 10% to the validation set. In both CV schemes, all the parameters of the MM, MDs and MDe regression models were recursively re-estimated in each of 100 random partitions. For each random partitioning, models were fitted using genotypes included in the training sets and the predictive ability was computed as the Pearson's correlation coefficient between GEBVs and the corresponding adjusted means of GY. Overall, 100 Pearson's correlations were computed for each model and the mean and standard deviation of these values were computed to estimate the predictive ability of GP models.

# RESULTS

## Development of the Barley MAGIC Population

The barley genotypes included in the founder set of MAGIC were examined in field trials organized in height site-by-season combinations in Italy, Germany and Scotland (Xu et al., 2018) for assessing the diversity of European cultivars for GY, plant height and DH. These field trials showed that the founder set, which includes four elite and four old barley varieties with different genetic background, exhibits limited variation of DH values (**Table 1**). Following a modified version of the standard crossing design (Huang et al., 2015), this founder set was intermated to create an eight-way MAGIC population of 352 individuals, which were subsequently genotyped to assess the contribution of each founder parent to the mosaic genome of each line.

## Estimating the Predictive Ability of GP Models as a Function of TP Size

In GP models, the variation of predictive ability as a function of the TP size has been empirically investigated on segregating families and in collections of mostly unrelated accessions (Norman et al., 2018). Here, we investigated the relationship between TP size and the predictive ability of different GP statistical models fitted to the barley MAGIC population. To carry out this analysis, the whole panel of 352 MAGIC lines and the founder parents were genotyped using the Barley 50 k iSelect SNP Array (Bayer et al., 2017). SNPs with more than 10% of missing data were discarded, while the remaining missing genotypes were imputed using the algorithm implemented in BEAGLE (Browning and Browning, 2016). This procedure allowed to identify 19,723 polymorphic SNPs, which were combined to the adjusted means (BLUPs) of GY computed in three site-by-season combinations (**Table 2**) to fit and cross-validate SE-GP models. Overall, five different whole genome regression methods based on BayesA, BayesB, BL, GB and RKHS fitted with the non-linear GK (Gianola and Van Kaam, 2008; Gota and Gianola, 2014; Cuevas et al., 2016; Crossa et al., 2017) were compared.

These aforementioned SE-GP models were fitted to the MAGIC population and cross-validated for estimating the trend of predictive ability as a function of TP size (**Figure 1**). Specifically, CV was implemented randomly partitioning 100 times the whole panel of MAGIC lines in a TP and in a validating population (VP). Overall, nine different CV experiments were carried out, using TP sizes of 80, 90, 100, 110, 120, 130, 140, 150, and 160 MAGIC lines and the remaining genotypes as VPs (**Figure 1**). The CV of these GP models points out that in the three site-by-season combinations (**Table 2**), GB, GK, BayesA, BayesB and BL show comparable predictive abilities across the entire range of TP sizes considered (**Figure 1**). Moreover, these CV experiments point out that in temperate locations (Fio16IN, Fio17IN, **Table 2**), the predictive ability of SE-GP models exceeds 0.50 even using TPs of 80 or 90 individuals (**Figure 1**), while in the harsh and pre-desertic environment of Mar16IN (**Table 2**), it does not exceed 0.25 and shows larger standard deviation. Varying the size of TPs from 80 to 160 individuals slightly increases the values of predictive ability for GY in the remaining individuals of the MAGIC population (**Figure 1** and **Supplementary Table 1**) as already substantiated in other GP models fitted using collection of mostly unrelated genotypes (Norman et al., 2018). Overall, this empirical analysis shows that 80 or 90 MAGIC individuals are sufficient to fit SE-GP models yielding high values of predictive ability and that larger TPs do not significantly improve the predictive ability of GP models either in temperate or stressful environments (**Figure 1** and **Supplementary Table 1**).

## Designing Optimized TPs of MAGIC

The predictive ability of GP models fitted in collection of mostly unrelated accessions and in biparental populations depends on the size of TP, the genome distribution and number of molecular markers used for whole genome regression, the genetic composition of TP and its genetic relationship with the BP (Heffner et al., 2009; Jannink et al., 2010; Desta and Ortiz, 2014; Berro et al., 2019). Particularly, it was assessed that using a large reference panel of accessions, the predictive ability of GP models can be improved increasing the diversity of the TPs (Norman et al., 2018). Along with these empirical findings, several statistical criteria and algorithms have been proposed to optimize TPs for maximizing predictive ability using reference panels of accessions or sets of advanced lines (Akdemir et al., 2015; Berro et al., 2019; Ou and Liao, 2019).

Here, we examined three different untargeted optimization criteria based on the coefficient of determination (CD_mean) (Laloe, 1993), prediction error variance (PEV) (Rincent et al., 2012) and rScore (Ou and Liao, 2019) and benchmarked them against a method that samples a diverse TP from the whole MAGIC population using SNP markers (**Figure 2**). The rationale of this latter method is to maximize the average distance, computed using the modified Roger's method, between each selected accession and the closest other genotype (Thachuk et al., 2009). This criterion, named entry-to-nearest entry was maximized with a heuristic algorithm to construct a highly diverse TP in which all MAGIC lines are maximally different (De Beukelaer et al., 2018). The TP assembled with this latter untargeted optimization criterion, named "TP-Diverse" (**Figure 2**), was constructed using the panel of 19,723 polymorphic SNPs detected in the whole MAGIC population, and was subsequently used as optimized TP and benchmarked to TPs assembled using CD_mean, PEV and rScore optimization methods (**Figure 2**).

Following this "TP-Diverse" optimization, our procedure led to identify a set of 82 MAGIC lines as the smallest population subset fulfilling the aforementioned criterion, which was used as TP along with the eight founder parents. Overall, when applied to MAGIC populations, the four optimized TPs spawned similar predictive abilities across the three site-by-season combinations (**Figure 2**) and consequently the genetic makeup of this TP was further investigated. The genetic relationships between TP-Diverse and the remaining MAGIC lines was assessed conducting a principal component analysis (PCA) on genetic data, which pointed out that the first two principal components explain 22.3 and 5.5 percent of the total genetic variability of the

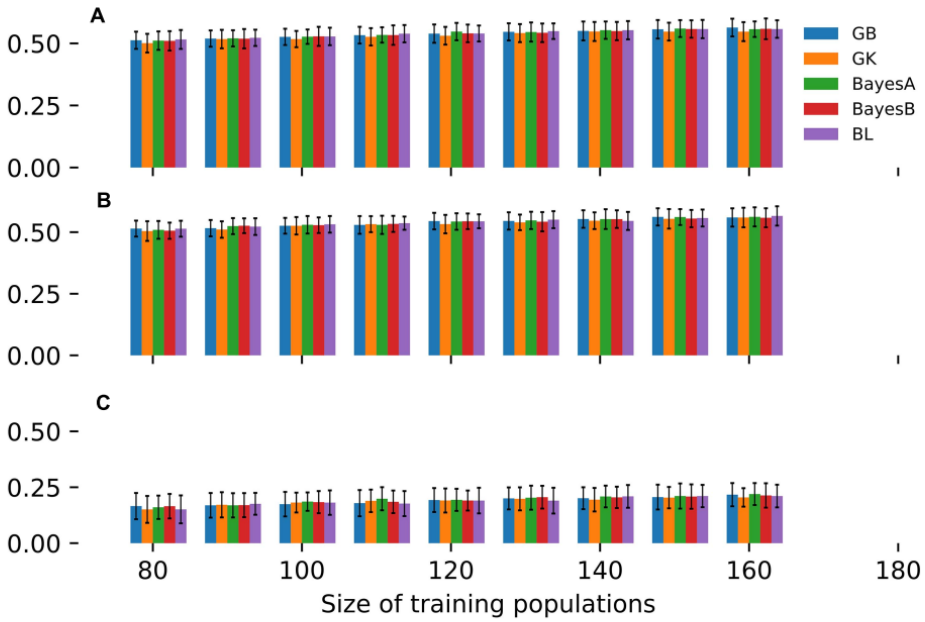



**FIGURE 1 |** CV of different SE-GP models fitted to GY measured in the MAGIC population. Bars report the values of predictive ability for GY computed in **(A)** Fio16IN, **(B)** Fio17IN, and **(C)** Mar16IN. Bars of different colors point out values of predictive ability computed using GB, GK, BayesA, BayesB and BL models as a function of TP sizes, while the error bars point out the standard deviation of predictive ability values.

**TABLE 2 |** Field trials carried out for phenotyping the whole MAGIC population and the founder set for GY.

| Acronym | Site | Country | Growing season | Populations | Traits |
|---|---|---|---|---|---|
| Fio16IN | Fiorenzuola d'Arda | Italy | 2015–2016 | 352 MAGIC and the founder set | DH, GY |
| Fio17IN | Fiorenzuola d'Arda | Italy | 2016–2017 | 352 MAGIC and the founder set | DH, GY |
| Mar16IN | Marchouch | Morocco | 2015–2016 | 352 MAGIC and founder set | DH, GY |

MAGIC population, respectively (**Figure 3**). PCA shows three main clusters of MAGIC lines and corroborates that individuals included in the TP-Diverse are representative of the whole diversity of MAGIC lines (red points).

In segregating families and collections of mostly unrelated accessions, a large number of molecular markers is often needed to capture the effects of all QTLs or alternatively, strong linkage disequilibrium (LD) between markers and causative variants that control the traits of interest is desirable to achieve high values of predictive ability in GP (Lorenzana and Bernardo, 2009; Heffner et al., 2011; Norman et al., 2018). Consequently, the extent of LD was investigated in TP-Diverse to assess its correlation with the predictive ability values of GP models. Firstly, SNP markers of the barley 50 K SNP chip used to fingerprint the whole MAGIC population were lifted over to the new barley reference sequence (Monat et al., 2019) and secondly, the average extent of $r^2$ was computed for each barley chromosome. Overall, a large fraction of the 44,040 SNPs of the barley 50 k SNP chip were lifted over and 18,248 out 19,723 polymorphic SNPs unambiguously mapped to the reference sequence of barley (**Supplementary Table 2**) were used to estimate the decay of average LD computed in bins of 100 kb (**Figure 4**). This analysis indicated that across the seven barley chromosomes $r^2$ decays relatively slowly as SNPs mapped more than 10 Mbp apart show $r^2$ values of circa 0.2, while the average $r^2$ values of markers within 1 MB or less exceed 0.4 (**Figure 4**). Considering the average number of markers per chromosome (**Supplementary Table 2**), the levels of LD measured in TP-Diverse are sufficiently high and higher marker densities might not significantly increase the predictive ability of GP models fitted in our MAGIC population of barley as empirically observed in other crops (Norman et al., 2018). Overall, the predictive ability values obtained with GP models fitted with the three optimization methods are substantially equivalent to the prediction accuracy obtained with TP-Diverse (**Figure 2**) and consequently this latter TP was chosen for fitting further single- and multi-environment GP models.

## Using the Optimized TP for Fitting SE-GP and ME-GP Models

Field trials of TP-Diverse were organized in nine site-by-season combinations and phenotypic data for GY and DH were collected using common phenotyping protocols, while the remaining set of MAGIC lines were used in Fio16IN, Fio17IN

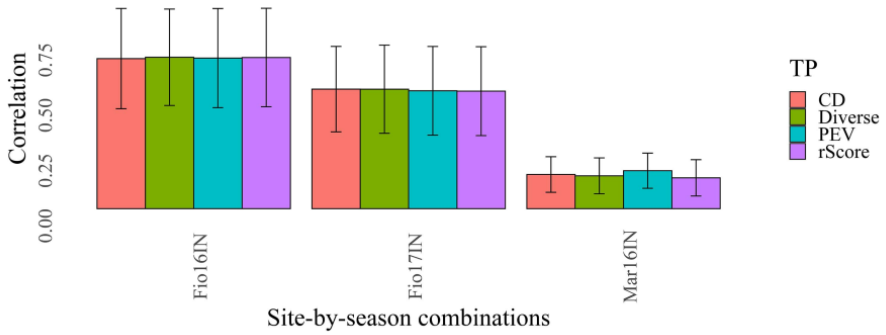



**FIGURE 2 |** Benchmarking of different methods for optimizing TPs of MAGIC. Bars of different colors report the values of predictive ability obtained with GP models fitted using CD_mean (CD), prediction error variance (PEV), rScore and Diverse optimization criteria. The error bars of each plot point out the standard deviation of the predictive ability values.

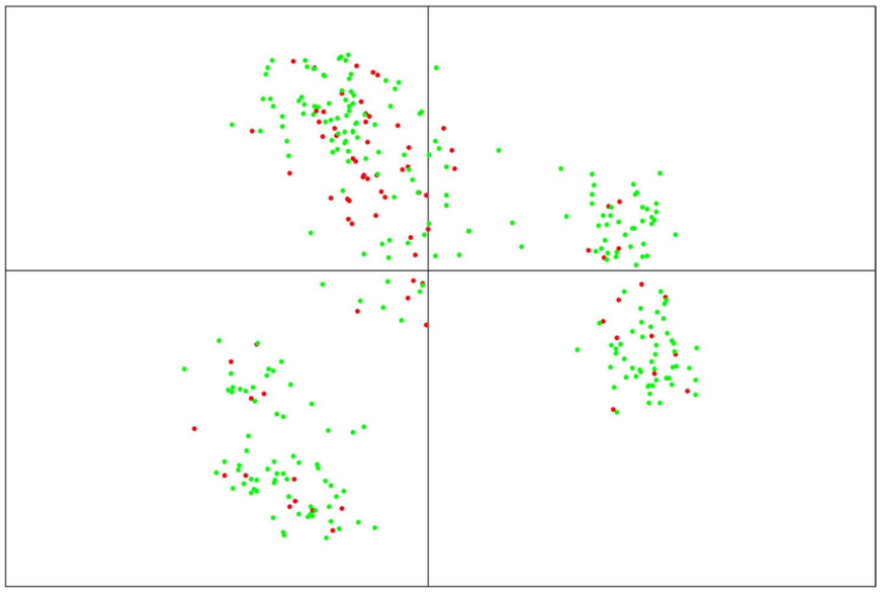


**FIGURE 3 |** Principal component analysis (PCA) of the MAGIC population based on 19,723 SNPs. The first two axes of PCA explain 22.3 and 5.5% of the total variability, respectively. Red points represent the subset of MAGIC lines included in TP-Diverse, while green points represent the remaining MAGIC lines.

and Mar16IN as VP (**Table 3**). Alpha-lattice experimental designs were adopted for all field trials and mixed linear models were used to compute adjusted means of GY and broad sense heritability ($H^2$) for each site-by-season combination considering genotypes as random variables (BLUPs) (**Table 3**). This analysis indicated that $H^2$ varies significantly across the nine field trials and spans from 0.805 in Kon19IN to 0.122 in Mar16IN (**Table 3**). The adjusted means of GY were subsequently used as phenotypes for fitting GP models along with genotypic information.

To assess the performance of MAGIC lines included in TP-Diverse, across different locations and years, a pairwise correlation analysis of the adjusted means of GY computed in the nine site-by-season combinations considered in this study was carried out (**Figure 5**). The correlations of GY across environments spanned from −0.030 to 0.553 and, as expected, values were higher between field trials carried out in the same environments but in different years, while lower values were observed among Mar16IN and other site-by-season combinations, corroborating the hypothesis that the climatic peculiarity of this environment imposes higher levels of stress to MAGIC lines (**Figure 5**). Similarly, the adjusted means of GY computed in Fio18LN exhibited lower correlation values with other site-by-season combinations (**Figure 5**). These adjusted means of GY were used to train SE-GP and ME-GP models using "TP-Diverse." For each site-by-season combination, phenotypic and genotypic data were standardized, and nine different SE-GP models were fitted using GB and GK statistical models (**Table 4**).

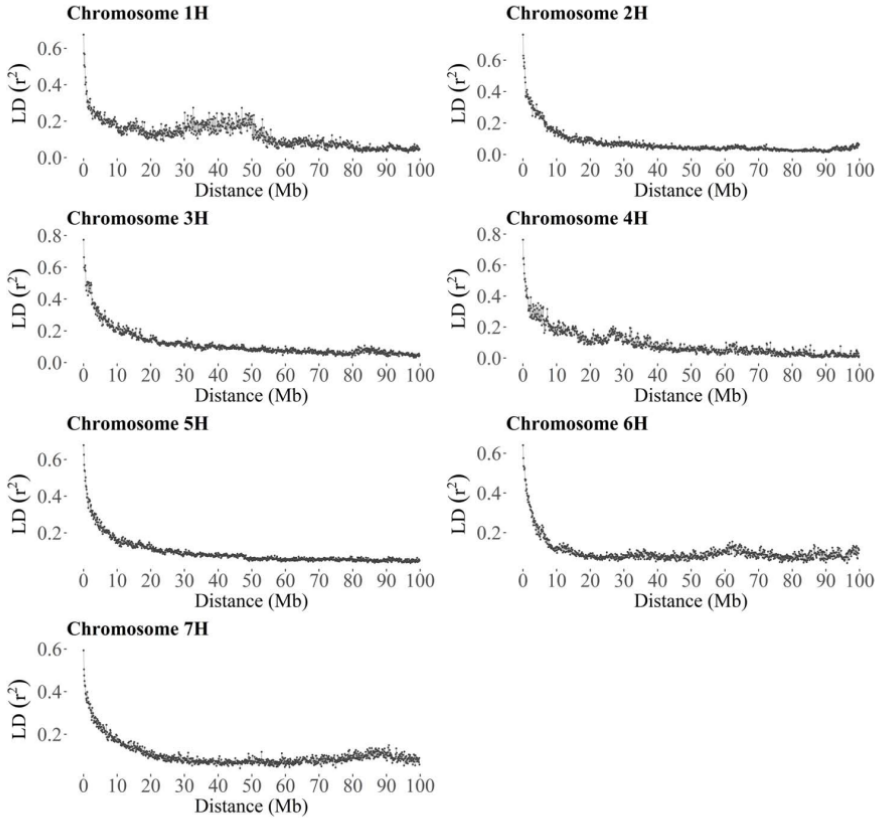



**FIGURE 4 |** Extent of the average linkage disequilibrium in TP-Diverse. For each barley chromosome, each point shows the average $r^2$ computed in 100 kb windows as a function of marker distance.

**TABLE 3 |** Summary of field trials carried out for phenotyping TP and VP for GY.

| Acronym | Site | Country | Growing season | Populations | $H^2$ |
|---|---|---|---|---|---|
| Fio16IN | Fiorenzuola d'Arda | Italy | 2015–2016 | TP and VP | 0.660 |
| Fio17IN | Fiorenzuola d'Arda | Italy | 2016–2017 | TP and VP | 0.472 |
| Fio18IN | Fiorenzuola d'Arda | Italy | 2017–2018 | TP | 0.532 |
| Fio18LN | Fiorenzuola d'Arda–Low Nitrogen | Italy | 2017–2018 | TP | 0.395 |
| Fio19IN | Fiorenzuola d'Arda | Italy | 2018–2019 | TP | 0.652 |
| Fio19LN | Fiorenzuola d'Arda–Low Nitrogen | Italy | 2018–2019 | TP | 0.663 |
| Mar16IN | Marchouch | Morocco | 2015–2016 | TP and VP | 0.122 |
| Ada19IN | Adana | Turkey | 2018–2019 | TP | 0.737 |
| Kon19IN | Konya | Turkey | 2018–2019 | TP | 0.805 |

*For each site-by-season combination, the estimates of broad sense heritability ($H^2$) of GY were reported. $H^2$ was computed for the whole panel of MAGIC lines for Fio16IN, Fio17IN and Mar16IN.*

As expected after standardization, for models fitted using GB, the summation of variance components was circa 1 (**Table 4**), while the distribution of the residuals after fitting all GP models to the nine site-by-season combinations was approximately normal. The analysis of variance components of SE-GP models showed that the values of error variance in GK models are lower than those obtained for the corresponding GB models (**Table 4**), and similarly in GK models the values of genetic component variance are always higher than the corresponding quantities computed for GB models (**Table 4**).

The adjusted means of GY computed at the nine site-by-season combinations were used to fit ME-GP, particularly three models were fitted, which were named "Multi-environment, main genotypic effect" (MM), "Multi-environment, single variance GxE deviation" (MDs) (Jarquín et al., 2014) and "Multi-environment, environment specific variance GxE deviation"


**FIGURE 5 |** Pairwise correlations of GY obtained in the nine site-by-season combinations for TP and VP. Numbers reported in black, red, and blue on the upper graph show pairwise Pearson correlations computed between adjusted means of GY for the whole set of lines tested, TP and VP, respectively. The lower graph shows scatter plots of GY adjusted means computed in pairs of site-by-season combinations.

**TABLE 4 |** Variance components of SE-GP models fitted using GBLUP (GB) and GK statistical model.

| Site-by-season combination | GB | | GK | |
|---|---|---|---|---|
| | Genetic effect variance | Residual variance | Genetic effect variance | Residual variance |
| Kon19IN | 0.586 (0.010) | 0.557 (0.045) | 0.660 (0.016) | 0.489 (0.068) |
| Mar16IN | 0.467 (0.089) | 0.719 (0.067) | 0.590 (0.013) | 0.588 (0.078) |
| Fio18IN | 0.491 (0.059) | 0.560 (0.029) | 0.632 (0.086) | 0.455 (0.043) |
| Fio18LN | 0.412 (0.048) | 0.752 (0.050) | 0.544 (0.000) | 0.611 (0.069) |
| Fio17IN | 0.537 (0.072) | 0.480 (0.016) | 0.655 (0.084) | 0.417 (0.041) |
| Fio16IN | 0.618 (0.066) | 0.336 (0.094) | 0.680 (0.065) | 0.348 (0.011) |
| Ada19IN | 0.561 (0.086) | 0.543 (0.036) | 0.654 (0.019) | 0.498 (0.070) |
| Fio19IN | 0.480 (0.079) | 0.651 (0.049) | 0.659 (0.005) | 0.469 (0.054) |
| Fio19LN | 0.479 (0.058) | 0.566 (0.024) | 0.632 (0.083) | 0.446 (0.041) |

*For each site-by-season combination, the estimated variance components of genetic effects and residuals fitted with GB and GK models are reported, while bracketed numbers point out the corresponding standard deviation.*

(MDe) (López-Cruz et al., 2015) following recent model nomenclature (Bandeira e Sousa et al., 2017). Similarly to SE-GP models, MM, MDs, and MDe models were fitted using GB and GK methods and totally six model method combinations were used to fit multi-environment predictions. The analysis of variance components showed that for all three models (MM, MDs, and MDe), GK methods exhibit lower values of the estimated residual variances pointing out a better model fitting (**Table 5**). Moreover, model comparisons showed that the inclusion of the interaction term (GxE) in MDe model induces a reduction in the estimated residual variance for GY compared to MM models either using GB or GK methods, but MDs models fitted better the data compared to MDe. For the MDe models, the residual variance components of MDe-GK were smaller than those of the MDe-GB, whereas the estimated variance components for the genetic main effect and genetic environment specific effect variances were higher for the GK than for the GB (**Table 5**).

## Predictive Ability of ME-GP Models With GB and GK Methods

The predictive ability of MM, MDs, and MDe models implemented using GB and GK methods was estimated

with cross-validation 1 (CV1) and cross-validation 2 (CV2) schemes using 100 random partitions. For each of the six multi-environment model-method combinations, the values of predictive ability for CV1 and CV2 schemes were obtained for the set of 100 random partitions, which were used to compute the average predictive ability and the associated standard deviation. Overall, CV2 showed that in four site-by-season combinations (Fio16IN, Fio17IN, Fio19IN, and Fio19LN) the predictive ability is generally higher and exceed 0.70 for certain ME-GP models, while for Mar16IN the six model-method combinations exhibit, on average, the lowest values of predictive ability as for this site-by-season combination the lowest values of 0.161 and 0.236 were observed for MM-GB and MDs-GK models, respectively (**Figure 6** and **Supplementary Table 3**).

As in most of the case, the standard deviations associated to the values of predictive ability were overlapping (**Figures 6**, **7**), Welch's *t*-tests were applied to determine whether pairwise comparisons of predictive ability values obtained with ME-GP models were statistically different (**Supplementary Figures 1**, **2**). CV2 experiments showed that in Fio17IN the values of predictive ability computed with the six-model method combinations were comparable except for MM-GB, which was significantly lower than the predictive ability of MDs-GK, while in Fio16IN the predictive ability of MM-GK was significantly lower than the predictive ability obtained with the remaining model-method combinations (**Figure 6** and **Supplementary Figure 2**). In Fio16IN, CV2 showed that MDe-GB and MDe-GK have similar performance and significantly higher values of predictive ability compared to MM models, either implemented with GB or GK statistical methods (**Figure 6**, **Supplementary Table 3**, and **Supplementary Figure 2**). In Ada19IN the best model predictive ability using CV2 scheme was obtained with MDe-GB, while for Fio18LN the best values of predictive ability were obtained with MDe-GB and MDs-GB models. Overall, CV2 experiments indicated that in four out nine site-by-season combinations (Fio16IN, Fio17IN, Fio18IN, and Mar16IN) MDe-GB and MDe-GK models have higher values of predictive ability compared to MM models, either implemented with GB or GK statistical methods (**Figure 6**, **Supplementary Table 3**, and **Supplementary Figure 2**). Differently, Fio19IN, Fio19LN, and Kon19IN deviate from this trend as for these site-by-season combinations the values of predictive ability for MM models were higher (**Supplementary Table 3**). In Fio19IN, MM-GB and MM-GK had the higher predictive ability values along with MDe-GK, while for Fio19LN the higher value of predictive ability was found for MM-GB.

The values of predictive ability obtained for random CV1 decreased (**Figure 7** and **Supplementary Table 4**) as compared with those computed for CV2 for all models. Similarly to the results obtained for CV2, CV1 experiments indicated that in four site-by-season combinations (Fio16IN, Fio17IN, Fio18IN, and Fio19LN) the predictive ability of GP-ME models is generally higher than the values of predictive ability observed in other site-by-season combinations for all models. MDs-GB and MD-GK yielded the higher values of predictive ability in Ada19IN, Fio16IN, and Fio17IN, respectively. In Fio18IN, Fio18LN, Mar16IN, and Fio19LN, the higher predictive ability values were found for MM-GK, although in this latter site-by-season combination the accuracy of MDe-GK does not differ significantly (**Supplementary Figure 1**). In Fio19IN, the highest values of predictive ability were obtained for MDe-GB and MD-GK models (**Figure 7** and **Supplementary Figure 1**).

**TABLE 5 |** Variance components of ME-GP models fitted using GBLUP (GB) and RKHS along with the Gaussian Kernel (GK) methods.

| Component | Environment | GB | GK |
|---|---|---|---|
| **Multi-environment, main genotypic effect (MM) model** | | | |
| Residual ($\sigma^2_e$) | – | 0.758 (0.047) | 0.746 (0.045) |
| Genetic main effect ($\sigma^2_{\mu 0}$) | – | 0.249 (0.069) | 0.373 (0.088) |
| **Multi-environment, single variance GxE deviation (MDs) model** | | | |
| Residual ($\sigma^2_e$) | – | 0.516 (0.056) | 0.389 (0.071) |
| Genetic main effect ($\sigma^2_{u0}$) | – | 0.281 (0.077) | 0.374 (0.089) |
| Genetic interaction effect ($\sigma^2_{ue}$) | – | 0.247 (0.066) | 0.589 (0.140) |
| **Multi-environment, environment specific variance GxE deviation (MDe) model** | | | |
| Residual ($\sigma^2_e$) | – | 0.602 (0.016) | 0.592 (0.018) |
| Genetic main effect ($\sigma^2_{u0}$) | – | 0.292 (0.026) | 0.402 (0.031) |
| Genetic environment specific effect ($\sigma^2_{uEj}$) | Ada19IN | 0.251 (0.054) | 0.353 (0.083) |
| | Fio16IN | 0.035 (0.027) | 0.054 (0.046) |
| | Fio17IN | 0.010 (0.066) | 0.024 (0.023) |
| | Fio18LN | 0.062 (0.054) | 0.116 (0.085) |
| | Fio18IN | 0.007 (0.006) | 0.018 (0.015) |
| | Mar16IN | 0.549 (0.085) | 0.873 (0.122) |
| | Kon19IN | 0.217 (0.050) | 0.312 (0.079) |
| | Fio19LN | 0.008 (0.007) | 0.053 (0.018) |
| | Fio19IN | 0.004 (0.003) | 0.055 (0.011) |

*For each of the three regression models (MM, MDs and MDe), the estimated variance components fitted with GB and GK methods are reported, while bracketed numbers point out the corresponding standard deviation of variance component estimates.*

## DISCUSSION

### Broadening the Use of MAGIC Populations for Plant Breeding

Multi-parent Advanced Generation Intercrosses populations were conceived to improve precision and efficiency of QTL mapping in plants and animals as they allow overcoming limitations of biparental populations and association mapping panels (Huang et al., 2015). In cereal crops, these experimental populations have been extensively used for research purpose and contributed to dissecting the genetic bases of several traits among which biotic stress resistance (Stadlmeier et al., 2018; Jiménez-Galindo et al., 2019; Riaz et al., 2020), GY, grain quality (Zaw et al., 2019) and DH (Afsharyan et al., 2020). Recently, these genomic resources have been established in barley to investigate the effects of epistasis and environmental interactions on flowering time

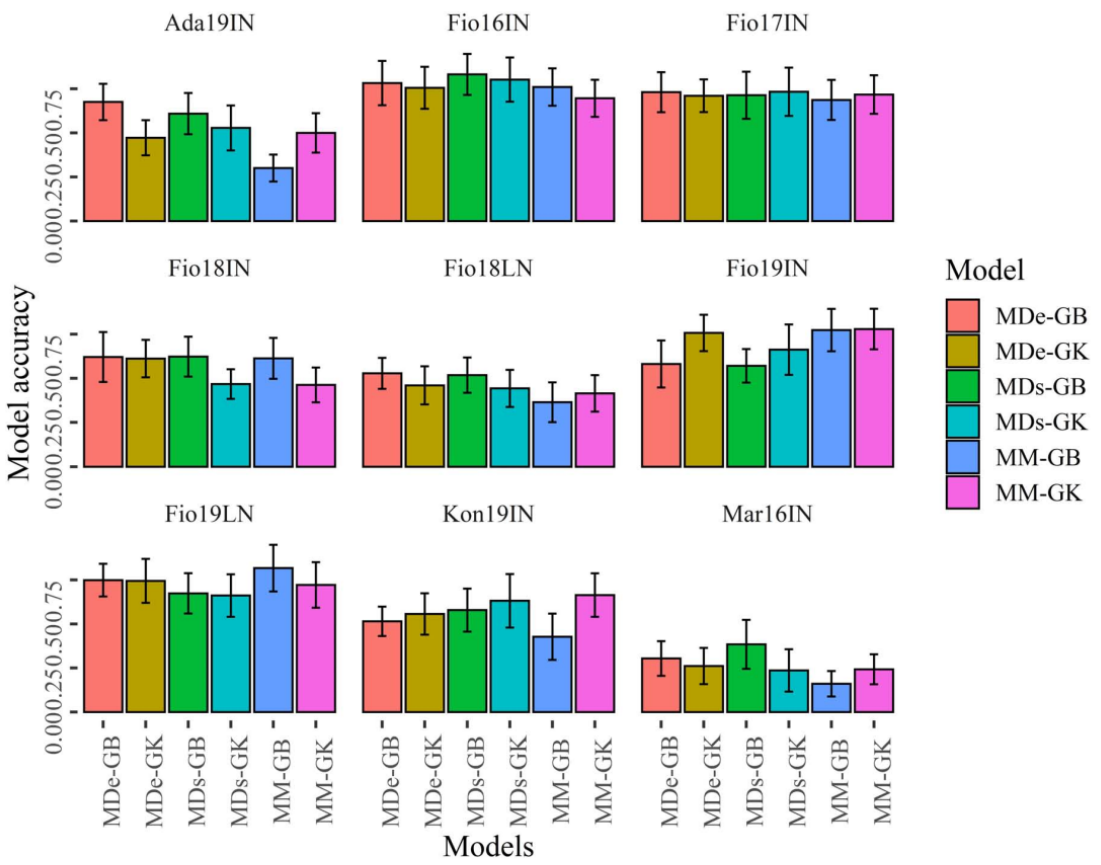



**FIGURE 6 |** Bar plots of the predictive ability values obtained with CV2. Bar plots show the mean correlation between observed and predicted values of GY obtained with 100 random CV2 partitions for MM, MDs and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Error bars point out the standard deviation of predictive ability values.

(Mathew et al., 2018; Afsharyan et al., 2020), further broadening the original scope for which they were devised.

In the present study, we constructed a new MAGIC population shuffling alleles of winter 6-rowed barley varieties, and demonstrated that, along with biparental populations and collections of mostly unrelated accessions, these genomic resources might be used to train GP models with high predictive ability and might speed up barley breeding. Under this point of view, the large number of MAGIC populations developed in the last years in several crops (Kover et al., 2009; Rebetzke et al., 2014; Mathew et al., 2018; Stadlmeier et al., 2018) can be considered as untapped resources that would contribute to further strengthening and stimulating the application of GP in plant breeding. On the other side, *de novo* creation of MAGIC populations to train GP models for actual breeding purposes is hampered because of their time consuming and costly development, which requires to intermate and self-fertilize the founder parents for several cycles. The results presented in this study show that these limitations might be softened using doubled haploid technology, which allows to short self-fertilization stages to obtain fully homozygous lines. Similarly, speed breeding might contribute to accelerating the development of new MAGIC populations (Watson et al., 2018).

To examine the genetic relationship between the whole set of MAGIC and the subset of lines included in the "TP-Diverse," a PCA was carried out using 19,723 SNPs, which detected genetic structure in the MAGIC population and three main clusters of individuals. The nature of these clusters is unclear, but it is plausible that they might reflect subgroups of individuals showing segregation distortion for one or more founders. In our eight-way MAGIC population, the expected segregation rate of the eight founder haplotypes is 1:1:1:1:1:1:1:1, but the haplotypes of some founders (e.g., Dea) deviate from the expected ratio (**Data not shown**). Segregation distortion is a common phenomenon that occurs in MAGIC populations as pointed out in other studies (Sannemann et al., 2018). Although this did not hamper our ability to train GP models with this population, this phenomenon might explain the genetic structure pointed out with PCA.

Overall, the use of SE-GP and ME-GP models trained with MAGIC populations might find effective applications when the diversity of BPs originates from the same parents included in

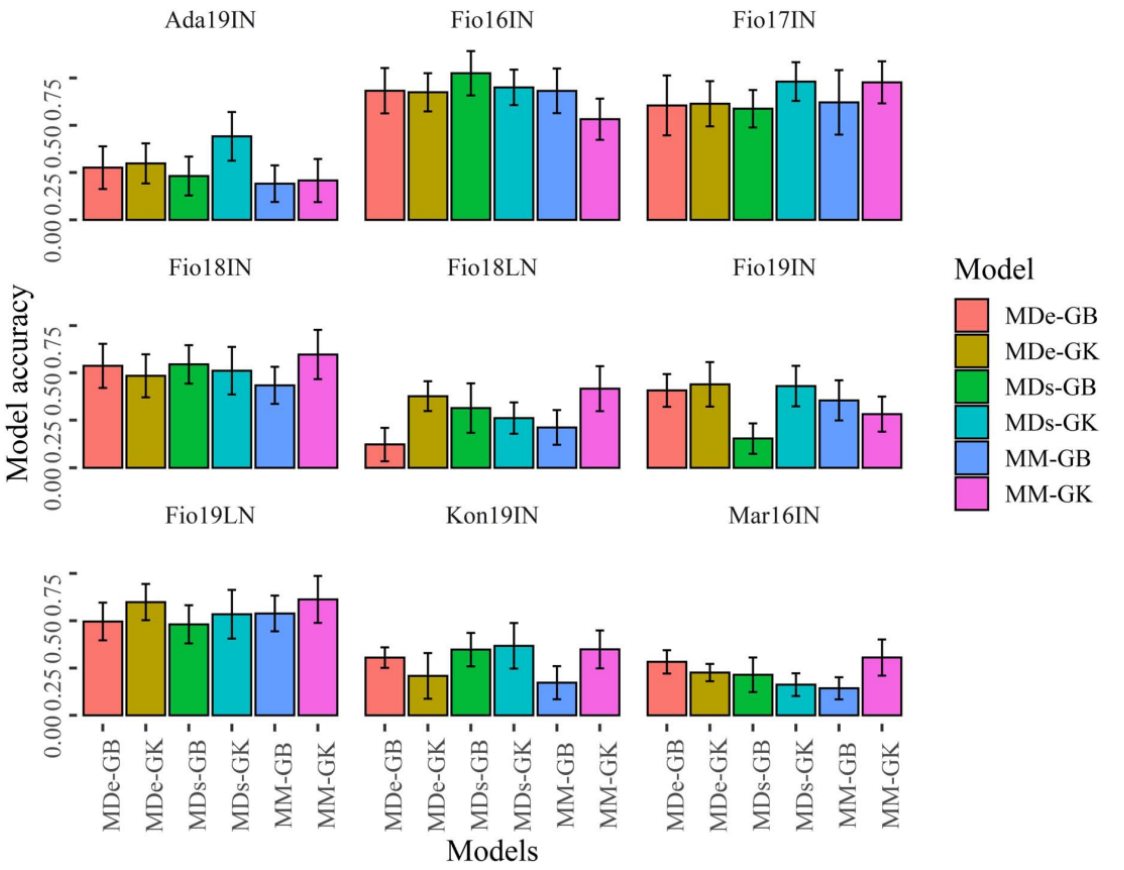



**FIGURE 7 |** Bar plots of the predictive ability values obtained with CV1. Bar plots show the mean correlation between observed and predicted values of GY obtained with 100 random CV1 partitions for MM, MDs, and MDe models implemented with GBLUP (GB) and Gaussian Kernel (GK) methods. Error bars point out the standard deviation of predictive ability values.

the founder set. In this case, GP models based on MAGIC populations might be applied to select the best offspring from crosses obtained with the MAGIC founders.

## Benchmarking of Different TPs to Improve the Predictive Ability of GP Models

The composition of TPs and their genetic relationship with BPs affect the predictive ability of GP models as pointed out in several studies (Desta and Ortiz, 2014; Norman et al., 2018; Edwards et al., 2019) and to date several algorithms for optimizing TPs have been developed to increase the predictive ability of GP models (Akdemir et al., 2015). Untargeted and targeted optimization criteria based on GBLUP have been so far developed and tested in biparental populations and panel of mostly unrelated accessions. Nevertheless, the use of these optimization methods in actual breeding programs is hampered as the optimization process can lead to different optimized TP per each trait of interest. These optimization algorithms require *a priori* information (knowledge of the BP genotypes and traits for which GP models must be developed) and output trait-dependent TPs (Akdemir et al., 2015). Moreover, in real breeding programs, BPs change over time and it might be difficult to implement these optimization procedures. Previous studies have shown that the relatedness between TPs and BPs has a large impact on the predictive ability of GP models, which can be improved increasing the genetic diversity of TPs (Norman et al., 2018). In fact, when the TP exhibits a narrow genetic diversity, low values of the predictive ability are often obtained in GP as it becomes impossible to predict all the marker effects that contribute to determining the phenotypic variations (Norman et al., 2018). Following these empirical findings, in this study we assembled a TP of 90 barley genotypes, which was named "TP-Diverse," maximizing the genetic diversity among MAGIC lines and assessing its predictive ability using random CV schemes. Surprisingly, the predictive ability obtained with TP-Diverse was comparable with the predictive ability of GP models trained with the other three optimized TPs used in this study (**Figure 2**). One of the main advantages of using this approach is that the criterion adopted to assemble "TP-Diverse" depends only on genetic data and does not generate trait-dependent TPs. On the

other side, in this study we have not developed mathematical models to demonstrate or justify the rationale of this empirical criterion and consequently its validity should be further validated in other studies.

## Fitting SE-GP and ME-GP Models Using the MAGIC Population of Barley

Several empirical analyses have been conducted to benchmark the predictive ability of different GP models in barley, maize and wheat panels of mostly unrelated accessions, biparental populations of *A. thaliana* and diallel crosses of maize and wheat to predict GY and other traits (Heslot et al., 2012). In this study, we presented another empirical analysis to assess the most promising GP models for MAGIC populations, implementing CV schemes for estimating the standard deviation of predictive ability values.

Three out five models fitted in this study (BayesA, BayesB, and BL) belong to the group of so called "Bayesian alphabet," which denotes Bayesian linear regressions that differ in their prior density distribution (Gianola, 2013). In these Bayesian regression models, the prior density distribution assigned to marker effects controls the shrinkage of estimates and then different priors induce different types of shrinkage of marker effects. In the original description both BayesA and BayesB were introduced as hierarchical structures (Meuwissen et al., 2001) and it was later demonstrated that BayesA adopts a scaled t-distribution prior, while BayesB adopts priors that are mixtures of a peak in the vicinity of zero and of a continuous density priors (e.g., t, or normal density distribution) (Gianola et al., 2009). BL adopts a double exponential prior density distribution, which behaves similar to that of BayesA as both priors used in these models do not allow marker effects to be equal to zero and shrink estimates of the remaining marker effects. While the priors adopted in BL and BayesA prevent to have marker effects equal to zero, the prior used in BayesB allows to have null marker effects. The rationale of this prior is that in GP many markers might have a null contribution to the observed phenotypic variation. Although marker effects might be estimated differently, the predictive ability of the Bayesian models fitted in this study does not differ significantly (**Figure 1**). Moreover, our empirical analysis shows that the predictive ability of Bayesian models fitted to MAGIC populations is comparable with that of GB and GK models (**Figure 1**). Several empirical analyses have been carried out in cereal crops to highlight advantages and limits of different whole genome regression methods. In rice, SE-GP models fitted with BayesA, GB, and GK for three traits were compared using a reference panel of 284 accessions under different linkage disequilibrium scenarios (Ben Hassen et al., 2018). These results showed that under high linkage disequilibrium scenarios GK models slightly outperform GB in terms of prediction ability. Differently, when a subset of rice reference panel was used to predict the performance of 97 advanced lined derived from biparental crosses, GK and GB prediction ability showed comparable results for the three traits considered (Ben Hassen et al., 2018). Anyway, the results obtained in this study are limited to one (complex) trait and it might plausible that for simpler traits GP models fitted in MAGIC might have different trend of the predictive ability.

Beyond SE-GP models, in this study we used the MAGIC population of barley to fit three different ME-GP models, two of which (MDs and MDe models) include terms for incorporating GxE interaction. In plant breeding, multi-environment field trials are routinely carried out to evaluate and exploit GxE interaction as it contributes to creating high-yielding genotypes. Consequently, modeling GxE interaction in GP has the potential to differentiate marker effects. MDe models used in this study (López-Cruz et al., 2015; Bandeira e Sousa et al., 2017) partition marker effects in main effects, that is effects that are stable across environments and environment-specific effects, that is interaction effects between markers and specific genotypes. As pointed out in other studies, MDe models are known to be more efficient when used along with sets of environments that have positive correlations. This limit arises as the pairwise correlation between environments is represented by the variance of the main marker effects, which in turn forces the co-variance between a pair of environments to be positive (López-Cruz et al., 2015; Bandeira e Sousa et al., 2017). This requirement is not trivial and might not allow to fit correctly MDe models. In our study, the adjusted means of GY in Mar16IN showed low or negative correlation with the other site-by-season combinations tested in this study and this might be the reason for which we have found that MDs models fit better the data, particularly when used in combination with the non-linear GK.

GP models based on reproducing kernel Hilbert Space along with the non-linear GK have the potential to capture non-additive genetic effects and potentially might outperform GB in terms of model fitting and predictive ability. In maize and wheat, comparison between the same GP models fitted with GB and the nonlinear GK for GY, unveiled that the latter method outperforms GB in terms of predictive ability in both single environment and multi-environment models (Cuevas et al., 2016; Bandeira e Sousa et al., 2017). In cereal crops, GY is a complex trait controlled by nonlinearity effects between genotypes and phenotypes owing to epistasis, environmental interactions (Bandeira e Sousa et al., 2017; Cuevas et al., 2018) and other interactions that are not considered in standard quantitative genetic models (Gianola et al., 2006). GK models have the potential to capture small and complex interactions, which are more evident in quantitative traits and this can explain the higher prediction ability of GK for GY. The empirical analysis presented in this study using barley MAGIC population corroborates that, for complex traits like GY, the predictive ability of GK outperforms that of GB. Overall, considering the number of models and methods fitted and the extensive field trials carried out across the Mediterranean, this study has delivered the most comprehensive empirical analysis of GP models fitted with MAGIC populations.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories

and accession number(s) can be found in the article/ **Supplementary Material**.

## AUTHOR CONTRIBUTIONS

AF designed and supervised the research along with the help of EI and LC. AF wrote the manuscript along with DP and significant contributions from AV, EI, HO, AC, AT, GV, LC, and AP. DP performed the research, while GV, AT, and SD developed the MAGIC population. AV, HO, EI, AC, SD, AT, İK, AP, and DP carried out field trials and plant phenotyping. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was carried out in the framework of the iBarMed project, which has been funded through the ARIMNet2 initiative and the Italian "Ministry of Agricultural, Food and Forestry Policies" under grant agreement "DM n. 20120." ARIMNet2 has received funding from the EU 7th Framework Programme for research, technological development and demonstration under grant agreement no. 618127. The work was also supported by SYSTEMIC_1063 (An integrated approach to the challenge of sustainable food systems: adaptive and mitigatory strategies to address climate change and malnutrition: From cereal diversity to plant breeding), a research project funded by Italian "Ministry of Agricultural, Food and Forestry Policies" in the frame of the Knowledge Hub on Food and Nutrition Security.

## ACKNOWLEDGMENTS

This article is dedicated to the memory of Antonio Michele Stanca, who passed away on March 19th, 2020. Beyond supporting this study, Stanca's research encompassed a wide range of plant science topics and contributed to exploiting morphological barley mutants to unveil the mechanisms that underlie plant developmental biology.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.664148/full#supplementary-material

## REFERENCES

Afsharyan, N. P., Sannemann, W., Léon, J., and Ballvora, A. (2020). Effect of epistasis and environment on flowering time in barley reveals a novel flowering-delaying QTL allele. *J. Exp. Bot.* 71, 893–906. doi: 10.1093/jxb/erz477

Akdemir, D., Sanchez, J. I., and Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *G Select. Evol.* 47, 1–10. doi: 10.1186/s12711-015-0116-6

Bandeira e Sousa, M., Cuevas, J., Couto, E. G., de, O., Pérez-Rodríguez, P., Jarquín, D., et al. (2017). Genomic-enabled prediction in maize using kernel models with genotype × environment interaction. *G3 Genes Genom. Genet.* 7, 1995–2014. doi: 10.1534/g3.117.042341

Bandillo, N., Raghavan, C., Muyco, P. A., Sevilla, M. A. L., Lobina, I. T., Dilla-Ermita, C. J., et al. (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice* 6:11. doi: 10.1186/1939-8433-6-11

Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R., and Crossa, J. (2015). Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). *Plant Sci.* 242, 23–36. doi: 10.1016/j.plantsci.2015.08.021

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Statist. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* 8:1792. doi: 10.3389/fpls.2017.01792

Ben Hassen, M., Cao, T. V., Bartholomé, J., Orasen, G., Colombi, C., Rakotomalala, J., et al. (2018). Rice diversity panel provides accurate genomic predictions for complex traits in the progenies of biparental crosses involving members of the panel. *Theoret. Appl. Genet.* 131, 417–435. doi: 10.1007/s00122-017-3011-4

Berro, I., Lado, B., Nalin, R. S., Quincke, M., and Gutiérrez, L. (2019). Training population optimization for genomic selection. *Plant Genome* 12:190028. doi: 10.3835/plantgenome2019.04.0028

Browning, B. L., and Browning, S. R. (2016). Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* 98, 116–126. doi: 10.1016/j.ajhg.2015.11.020

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

Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* 11, 215–221. doi: 10.1016/j.pbi.2008.01.002

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7. doi: 10.1186/s13742-015-0047-8

Crossa, J., Pérez-rodríguez, P., Cuevas, J., Montesinos-lópez, O., Jarquín, D., Campos, G. D. L., et al. (2017). Genomic selection in plant breeding: methods. *Models Perspect.* 22, 961–975. doi: 10.1016/j.tplants.2017.08.011

Cuevas, J., Crossa, J., Soberanis, V., Pérez-Elizalde, S., Pérez-Rodríguez, P., Campos, G., et al. (2016). Genomic prediction of genotype × environment interaction kernel regression models. *Plant Genome* 9:lantgenome2016.03.0024. doi: 10.3835/plantgenome2016.03.0024

Cuevas, J., Granato, I., Fritsche-Neto, R., Montesinos-Lopez, O. A., Burgueño, J., Sousa, M. B., et al. (2018). Genomic-enabled prediction Kernel models with random intercepts for multi-environment trials. *G3 Genes Genomes Genet.* 8, 1347–1365. doi: 10.1534/g3.117.300454

De Beukelaer, H., Davenport, G. F., and Fack, V. (2018). Core hunter 3: flexible core subset selection. *BMC Bioinform.* 19:203. doi: 10.1186/s12859-018-2209-z

De Los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., et al. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* 182, 375–385. doi: 10.1534/genetics.109.101501

Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., et al. (2015). Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* 16, 1–23. doi: 10.1186/s13059-015-0716-z

Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006

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

Emrich, K., Wilde, F., Miedaner, T., and Piepho, H. P. (2008). REML approach for adjusting the Fusarium head blight rating to a phenological date in inoculated

selection experiments of wheat. *Theoret. Appl. Genet.* 117, 65–73. doi: 10.1007/s00122-008-0753-z

Gianola, D. (2013). Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* 194, 573–596. doi: 10.1534/genetics.113.151753

Gianola, D., and Schon, C. C. (2016). Cross-validation without doing cross-validation in genome-enabled prediction. *G3 Genes Genom Genet.* 6, 3107–3128. doi: 10.1534/g3.116.033381

Gianola, D., and Van Kaam, J. B. C. H. M. (2008). Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178, 2289–2303. doi: 10.1534/genetics.107.084285

Gianola, D., De Los Campos, G., Hill, W. G., Manfredi, E., and Fernando, R. (2009). Additive genetic variability and the Bayesian alphabet. *Genetics* 183, 347–363. doi: 10.1534/genetics.109.103952

Gianola, D., Fernando, R. L., and Stella, A. (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics* 173, 1761–1776. doi: 10.1534/genetics.105.049510

Gota, M., and Gianola, D. (2014). Kernel-based whole-genome prediction of complex traits: a review. *Front. Genet.* 5:363. doi: 10.3389/fgene.2014.00363

Heffner, E. L., Jannink, J.-L., and Sorrells, M. E. (2011). Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *Plant Genome* 4, 65–75. doi: 10.3835/plantgenome2010.12.0029

Heffner, E. L., Sorrells, M. E., and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512

Heslot, N., Yang, H. P., Sorrells, M. E., and Jannink, J. L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Sci.* 52, 146–160. doi: 10.2135/cropsci2011.06.0297

Hill, W. G., and Robertson, A. (2008). The effect of linkage on limits to artificial selection. *Genet. Res.* 89, 311–336. doi: 10.1017/S001667230800949X

Huang, B. E., Verbyla, K. L., Verbyla, A. P., Raghavan, C., Singh, V. K., Gaur, P., et al. (2015). MAGIC populations in crops: current status and future prospects. *Theoret. Appl. Genet.* 128, 999–1017. doi: 10.1007/s00122-015-2506-0

Jannink, J. L., Lorenz, A. J., and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genom. Proteom.* 9, 166–177. doi: 10.1093/bfgp/elq001

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

Jiménez-Galindo, J. C., Malvar, R. A., Butrón, A., Santiago, R., Samayoa, L. F., Caicedo, M., et al. (2019). Mapping of resistance to corn borers in a MAGIC population of maize. *BMC Plant Biol.* 19:431. doi: 10.1186/s12870-019-2052-z

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., et al. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5:e1000551 doi: 10.1371/journal.pgen.1000551

Laloe, D. (1993). Precision and information in linear models of genetic evaluation. *Genet. Select. Evol.* 25, 557–576. doi: 10.1051/gse:19930604

López-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J.-L., et al. (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3&#58 Genes Genomes Genet.* 5, 569–582. doi: 10.1534/g3.114.016097

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

Mackay, I. J., Bansept-Basler, P., Bentley, A. R., Cockram, J., Gosman, N., Greenland, A. J., et al. (2014). An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. *G3 Genes Genomes Genet.* 9, 1603–1610. doi: 10.1534/g3.114.012963

Mathew, B., Léon, J., Sannemann, W., and Sillanpää, M. J. (2018). Detection of epistasis for flowering time using bayesian multilocus estimation in a barley MAGIC population. *Genetics* 208, 525–536. doi: 10.1534/genetics.117.300546

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.

Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., et al. (2019). TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open- source tools. *Genome Biol.* 20:284.

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

Ou, J. H., and Liao, C. T. (2019). Training set determination for genomic selection. *Theoret. Appl. Genet.* 132, 2781–2792. doi: 10.1007/s00122-019-03387-0

Park, T., and Casella, G. (2008). The bayesian lasso. *J. Am. Statist. Associat.* 103, 681–686. doi: 10.1198/016214508000000337

Pérez, P., and de los Campos, G. (2014). BGLR: a statistical package for whole genome regression and prediction. *Genetics* 198, 483–495. doi: 10.1534/genetics.114.164442

Pérez-Elizalde, S., Cuevas, J., Pérez-Rodríguez, P., and Crossa, J. (2015). Selection of the bandwidth parameter in a bayesian kernel regression model for genomic-enabled prediction. *J. Agricult. Biol. Environ. Statist.* 20, 512–532. doi: 10.1007/s13253-015-0229-y

Ponce, K. S., Ye, G., and Zhao, X. (2018). QTL identification for cooking and eating quality in indica rice using multi-parent advanced generation intercross (MAGIC) population. *Front. Plant Sci.* 9:868. doi: 10.3389/fpls.2018.00868

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

R Core Team (2019). *R: A Language and Environment for Statistical Computing. Industrial and Commercial Training*. Vienna: R Core Team. doi: 10.1108/eb003648

Rebetzke, G. J., Verbyla, A. P., Verbyla, K. L., Morell, M. K., and Cavanagh, C. R. (2014). Use of a large multiparent wheat mapping population in genomic dissection of coleoptile and seedling growth. *Plant Biotechnol. J.* 12, 219–230. doi: 10.1111/pbi.12130

Riaz, A., KockAppelgren, P., Hehir, J. G., Kang, J., Meade, F., Cockram, J., et al. (2020). Genetic analysis using a multi-parent wheat population identifies novel sources of septoria tritici blotch resistance. *Genes* 11, 1–26. doi: 10.3390/genes11080887

Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., et al. (2012). Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics* 192, 715–728. doi: 10.1534/genetics.112.141473

Sannemann, W., Lisker, A., Maurer, A., Léon, J., Kazman, E., Cöster, H., et al. (2018). Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800. *BMC Genomics* 19:3. doi: 10.1186/s12864-018-4915-3

Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., et al. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity* 125, 396–416. doi: 10.1038/s41437-020-0336-6

Stadlmeier, M., Hartl, L., and Mohler, V. (2018). Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. *Front. Plant Sci.* 871:e01825. doi: 10.3389/fpls.2018.01825

Thachuk, C., Crossa, J., Franco, J., Dreisigacker, S., Warburton, M., and Davenport, G. F. (2009). Core hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinformatics* 10:243. doi: 10.1186/1471-2105-10-243

Thioulouse, J., Dufour, A. B., Jombart, T., Dray, S., Siberchicot, A., and Pavoine, S. (2018). Multivariate analysis of ecological data with ade4. *Multivar. Anal. Ecol. Data ade4* 18:58850. doi: 10.1007/978-1-4939-8850-1

Tishbirani, R. (1996). Regression shrinkage and selection via the Lasso. *J. R. Statist. Soc. Ser. B (Methodol.)* 58, 267–288.

Valdar, W., Flint, J., and Mott, R. (2006). Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* 172, 1783–1797. doi: 10.1534/genetics.104.039313

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Wang, X., Xu, Y., Hu, Z., and Xu, C. (2018). Genomic selection methods for crop improvement: current status and prospects. *Crop J.* 6, 330–340. doi: 10.1016/j.cj.2018.03.001

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M. D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8

Wickham, H. (2016). ggplot 2: Elagant graphics for data analysis. *Media* 35:98141. doi: 10.1007/978-0-387-98141-3

Wimmer, V., Albrecht, T., Auinger, H. J., and Schön, C. C. (2012). Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* 28, 2086–2087. doi: 10.1093/bioinformatics/bts335

Xu, X., Sharma, R., Tondelli, A., Russell, J., Comadran, J., Schnaithmann, F., et al. (2018). Genome-wide association analysis of grain yield-associated traits in a pan-european barley cultivar collection. *Plant Genome* 11:170073. doi: 10.3835/plantgenome2017.08.0073

Yalcin, B., Flint, J., and Mott, R. (2005). Using progenitor strain information to identify quantitative trait nucleotides in outbred mice. *Genetics* 171:673–81. doi: 10.1534/genetics.104.028902

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zaw, H., Raghavan, C., Pocsedio, A., Swamy, B. P. M., Jubay, M. L., Singh, R. K., et al. (2019). Exploring genetic architecture of grain yield and quality traits in a 16-way indica by japonica rice MAGIC global population. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-55357-7

**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.

## Annex 3: List of articles according to PubMed format

Influence of the genetic background on the performance of molecular markers linked to seedlessness in table grapes. Bennici S, Di Guardo M, Distefano G, La Malfa S, **Puglisi D**, Arcidiacono F, Ferlito F, Deng Z, Gentile A, Nicolosi E. Scientia Horticulturae (2019) 252, 316–323 | https://doi.org/10.1016/j.scienta.2019.03.060.

Genomic prediction of grain yield in a barley MAGIC population modelling genotype per environment interaction. **Puglisi D**, Delbono S, Visioni A, Ozkan H, Kara İ, Casas A M, Igartua E, Valè G, Lo Piero A R, Cattivelli L, Tondelli A, Fricano A. Frontiers in Plant Science (2021) | https://doi.org/10.3389/fpls.2021.664148.

Malting Quality of ICARDA Elite Winter Barley (*Hordeum vulgare L.*) Germplasm Grown in Moroccan Middle Atlas. Bouhlal O, Affricot J R, **Puglisi D**, El-Baouchi A, El-Otmani F, Kandil M, Hafidi A, Keser M, Sanchez-Garcia M, Visioni A. Journal of the American Society of Brewing Chemists (2021) |

## Annex 4: Ph.D. Training

During the industrial Ph.D. course, Dr. Damiano Puglisi participated in several training courses, summer school and national/international congresses in order to deep all topics covered by the Ph.D. research program.

The impact of climate change on agriculture was deepened by attending summer school "*Climate change and crop productivity: The role of plant physiology, breeding and biotechnology*", which took place at Polvese's Island in Trasimeno's Lake, Perugia (Italy), from 12th to 15th June 2018.

Theoretical-practical training on statistical models was deepened through several national and international statistical courses including "*Metodologia statistica per le Scienze Agrarie. I modelli lineari generali e generalizzati*", promoted by the Società Italiana di Agronomia (SIA), held in Rome (Italy) from 21th to 25th January 2019; "*Gestione ed interpretazione di dati biologici complessi: basi teoriche ed utilizzo di software di analisi. Excel (avanzato) e R*", organized by the University of Catania (Italy) from 11th to 22th February 2019; "*Genome Wide Association Mapping with R*", promoted by INRA and ICARDA, held in Rabat (Morocco) 11th April 2019; "*SPSS and R*" organized by ENSET Ph.D. students, held in Rabat (Université Mohammed V, Morocco), 15rh February 2020; "*The Data Scientist's Toolbox*" and "*R programming*" online courses authorized by Johns Hopkins University and offered through Coursera 12th May and 10th June 2020, respectively; "*Enhancing discovery and creativity with AI*" online course promoted by Swiss Data Science Center 1th October 2020; "*Samsung Innovation Camp*", promoted by University of Catania (Italy), 25 hours online course and two days of classroom training for the realization of a Project Work, selected among the 60 best students of the University of Catania who obtained the higher score, 30th November 2020.

The quantitative genetics theory behind crop breeding was deepened by attending international specialized training courses including "*Classical and molecular approaches in wheat breeding*", promoted by ICARDA and held in Rabat (Morocco) from 15th April to 3th May 2019. During this event, Dr. Damiano Puglisi participated to both theoretical and practical sessions covering wheat breeding approaches and strategies, marker assisted selection, doubled haploid, and speed breeding techniques; and "*First International Experts Workshop on Pre-breeding utilizing Crop Wild Relatives (1st PBCWR)*" which took place at INRA-Institute National de la Recherche Agronomique, in Rabat (Morocco), from 24th April to 26th April 2019.

Finally, abstracts and posters submitted for participation in a summer school and scientific national/international congresses in which Dr. Damiano Puglisi took part during his industrial Ph.D. course:

The development and validation of methodologies for the genetic improvement of barley based on genomic selection to support varietal innovation and agriculture of the future. **Puglisi D.** Summer school Polvese's Island on Trasimeno's Lake, Perugia (Italy), Abstract 12th-15th June 2018.

Il programma di miglioramento genetico per l'uva da tavola dell'Università di Catania. Nicolosi E, Ferlito F, Domina F, **Puglisi D**, Salonia F, Zingale N, La Malfa S, Gentile A. XII Giornate Scientifiche SOI, Bologna (Italy), 19th-22th June 2018. Book of abstract and poster: 75. Acta Italus Hortus ISBN: 978-88-940276-8-6.

Multi-parental genomic prediction for improving barley yield in harsh Mediterranean environments. Delbono S, Tondelli A, Visioni A, Jilal A, Iguarta E, Ozkan H, **Puglisi D**, Cattivelli L, Fricano A. LXII Congress of the Italian Society of Agricultural Genetics (SIGA), Verona (Italy), 25th-28th September 2018. Book of abstract and poster: 7.07. ISBN 978-88-904570-8-1.

iBarMed: Innovative barley breeding approaches to tackle the impact of climatic change in the Mediterranean region. **Puglisi D**, Delbono S, Tondelli A, Visioni A, Jilal A, Iguarta E, Ozkan H, Cattivelli L, Fricano A. International conference: ARIMNet2 ends, but our heritage serves the future. INRA, 2 Place Pierre Viala, 34060 Montpellier (France), 18th-19th June 2019.

Prediction of grain yield and complex traits to assess the potential of a MAGIC population for genomic selection in barley. **Puglisi D**, Tondelli A, Visioni A, Ozkan H, Lo Piero A R, Cattivelli L, Fricano A. SIGA (Italian Society of Agricultural Genetics) Young Web Meeting, 7th July 2020. Book of abstract: SY32. ISBN 978-88-944843-0-4.

Genome-enabled prediction models for grain yield, transpiration rate and below-ground traits using a barley MAGIC population. **Puglisi D**, Delbono S, Visioni A, Ozkan H, Kara Ä, Casas A M, Igartua E, Valè G, Lo Piero A R, Cattivelli L, Tondelli A, Fricano A. LXIV Congress of the Italian Society of Agricultural Genetics (SIGA), *Plant genetic innovation for food security in a climate change scenario*. 14th-16th September 2021. Book of abstract and oral communication: 4.07.

## Ringraziamenti

La borsa di dottorato è stata cofinanziata con risorse del Programma Operativo Nazionale Ricerca e Innovazione 2014 – 2020 (CCI 2014IT16M2OP005), Fondo Sociale Europeo, Azione I.1 “Dottorati innovativi con caratterizzazione industriale”.

UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell' Istruzione, dell'Università e della Ricerca

PON RICERCA E INNOVAZIONE 2014 - 2020

L’attività di ricerca è stata condotta nell'ambito dei progetti iBarMed (https://www.ibarmed.com/) e Whealbi (https://www.whealbi.eu/).

Ringrazio con affetto il mio tutor Dott. Agostino Fricano per essere stato sempre presente, giorno dopo giorno, in questi tre bellissimi anni di dottorato, per avermi supportato e avermi dato la possibilità di approfondire le mie conoscenze con grande professionalità. A lui devo il raggiungimento di questo traguardo.

Ringrazio il Dott. Luigi Cattivelli, tutor e referente, per avermi guidato e inspirato lungo tutto il mio percorso, con fermezza e affetto.

Ringrazio di cuore la mia co-tutor Prof.ssa Angela Roberta Lo Piero per la sua costante presenza e disponibilità.

Ringrazio il Dott. Andrea Visioni, amico e punto di riferimento, che ha reso unica e indimenticabile la mia esperienza in Marocco. Non posso non ringraziare le splendide persone che ho conosciuto ad ICARDA, le quali sono diventate per me amici e compagni di avventura: in particolare, Outmane, Youness e Fatima.

Ringrazio il Dott. Nelson Nazzicari e il Prof. Matteo dell'Acqua, i revisori della tesi di dottorato, per l'entusiasmo mostratomi e gli utili consigli. Sono felice e onorato di aver ricevuto i loro apprezzamenti.

Voglio esprimere la mia gratitudine a tutta la mia sezione di Arboricoltura e Genetica Agraria con una menzione speciale ai professori e ai colleghi tra cui: Prof. Distefano, Prof.ssa Gentile, Prof.ssa La Malfa, Prof.ssa Nicolosi, Prof. Continella, Prof.ssa Domina, Giuseppina, Giulia, Chiara, Mario, Lara, Stefania, Giuseppe, Fabrizio, Fabio, Riccardo, Angelo, Danilo.

Un ringraziamento speciale va alla mia ragazza Alessandra per avermi supportato e consigliato con amore, condividendo insieme a me tutti i momenti chiave che mi hanno permesso di raggiungere questo traguardo. Sono legato a lei con il mio cuore e la mia anima.

Ringrazio, infine, la mia splendida famiglia: i miei genitori per essermi stati sempre accanto e per non aver mai smesso di credere in me. Tutto quello che sono, lo devo a loro; i miei fratelli Giovanni e Francesco e le mie cognate Doriana e Martina.

## Data Availability Statement

The data underpinning the results (**Chapter 2** and **Chapter 3**) and R script (**Annex 1**) presented in this Ph.D. thesis are available upon request (damianopuglisi@gmail.com; damiano.puglisi@unict.it).