High accuracy of genome-enabled prediction of belowground and physiological traits in barley seedlings

Damiano Puglisi [ID],[1] Andrea Visioni [ID],[2] Hakan Ozkan [ID],[3] İbrahim Kara,[4] Angela Roberta Lo Piero [ID],[1] Fatima Ezzahra Rachdad [ID],[2,5] Alessandro Tondelli [ID],[6] Giampiero Valè [ID],[7] Luigi Cattivelli [ID],[6] and Agostino Fricano [ID] [6,*]

[1]Dipartimento di Agricoltura, Alimentazione e Ambiente (Di3A), Università di Catania, 95123 Catania, Italy,
[2]Biodiversity and Crop Improvement Program, International Center for Agricultural Research in the Dry Areas, 6299 Rabat, Morocco,
[3]Faculty of Agriculture, Department of Field Crops, University of Cukurova, 01330 Adana, Turkey,
[4]Bahri Dagdas International Agricultural Research Institute, Km Karatay/Konya 42020, Turkey,
[5]Faculty of Sciences Ben M'sik, Department of Biology, Environment and Ecology Laboratory, Hassan II University of Casablanca, 7955 Casablanca, Morocco,
[6]Council for Agricultural Research and Economics—Research Centre for Genomics and Bioinformatics, 29017 Fiorenzuola d'Arda (PC), Italy,
[7]DiSIT, Dipartimento di Scienze e Innovazione Tecnologica, Università del Piemonte Orientale, 13100 Vercelli, Italy

*Corresponding author: Council for Agricultural Research and Economics—Research Centre for Genomics and Bioinformatics, Via San Protaso 302, 29017 Fiorenzuola d'Arda (PC), Italy. Email: agostino.fricano@crea.gov.it

Abstract

In plants, the study of belowground traits is gaining momentum due to their importance on yield formation and the uptake of water and nutrients. In several cereal crops, seminal root number and seminal root angle are proxy traits of the root system architecture at the mature stages, which in turn contributes to modulating the uptake of water and nutrients. Along with seminal root number and seminal root angle, experimental evidence indicates 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 at the seedling stage in a panel of 8-way Multiparent Advanced Generation Inter-Crosses lines of winter barley and correlated these traits with grain yield measured in different site-by-season combinations. Second, phenotypic and genotypic data of the Multiparent Advanced Generation Inter-Crosses population were combined to fit and cross-validate different genomic prediction models for these belowground and physiological traits. Genomic prediction models 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 was implemented using models based on extended genomic best linear unbiased predictors. The results presented in this study show that genome-enabled prediction models of seminal root number, seminal root angle, and transpiration rate data have high predictive ability and that the best models investigated in the present study include first-order additive × additive epistatic interaction effects. Our analyses indicate that beyond grain yield, genomic prediction models might be used to predict belowground and physiological traits and pave the way to practical applications for barley improvement.

Keywords: seminal root angle; seminal root number; transpiration rate; MAGIC; barley; genomic prediction; threshold GBLUP; MPP; Multiparental Populations; Multiparent Advanced Generation Inter-Cross (MAGIC)

Introduction

Despite their importance for the uptake of water and nutrients, 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 and the existence of a correlation between these traits (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 a deeper root system architecture can cope with drought and heat stresses, increasing 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). Consequently, the improvement of crops targeting the root system architecture and belowground traits is high desirable to enhance productivity and cope with climate change (Tracy et al. 2020).

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 2 outmost seminal roots at the seedling stage, are 2 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). Although phenotyping platforms and technologies to analyze the whole root system architecture at mature stage of development are progressively becoming widespread, the incorporation of belowground traits in actual breeding programs is still in its infancy and might benefit from using SRA and SRN, which can be easily scored.

The evaporative demand or vapor pressure deficit (VPD) points out 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 this phenomenon (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). In fact, 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 2 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 GY performances and biomass production in rainfed and irrigated field trials (Medina et al. 2019). 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 1 cultivar, which corroborate the existence of untapped diversity for TR in barley germplasm (Sadok and Tamang 2019) that might be exploited for barley improvement and development of more drought-tolerant genotypes.

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 known as training population (TP) and is used for constructing predictive models, which allow 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 nonparametric 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 2014b) as in these conditions, the estimation of marker effects using the 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). Interestingly, under certain assumptions, it has been demonstrated that GBLUP and ridge regression of markers are actually equivalent models (Habier et al. 2007). 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), root vigor (Biscarini et al. 2014), 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 the present study, we examined the diversity and distribution of belowground (SRA and SRN) and physiological (TR response to increasing VPD) traits at the seedling stage in an 8-way Multiparent Advanced Generation Inter-crosses (MAGIC) population of winter barley and in its founder parents (Puglisi et al. 2021). We correlated phenotypic data of SRA, SRN, and TR scored in controlled conditions with GY obtained in different site-by-season combinations to re-assess the relevance of belowground and physiological traits of seedlings for the uptake of nutrients and water. 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 incorporating these traits in actual breeding programs, underpinning ideotype breeding and characterizing large plant collections.

Materials and methods
Plant materials and genotyping
The MAGIC population examined in this study and used for genome-enabled predictions has been extensively described elsewhere (Puglisi *et al.* 2021). Particularly, the set of 89 MAGIC lines used in the present study (Supplementary Tables 1 and 2) corresponds to the TP previously assembled for fitting multienvironment GP models and genotyped using the barley 50K SNP chip (Puglisi *et al.* 2021). In the present study, 20,426 polymorphic SNPs were used for fitting GP models (Supplementary File 1).

Phenotyping of plant material
The set of MAGIC lines was phenotyped in the following site-by-season combinations to examine GY and heading date: 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, 1,005 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 2 different levels of nitrogen fertilization as previously described (Puglisi *et al.* 2021). The detailed procedure for analyzing field trial data and deriving the adjusted means of GY is described in the Supplementary text 1 "Procedure for deriving the adjusted means of grain yield."

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. The detailed procedure and the experimental design to phenotype SRA and SRN and for computing the adjusted means of SRA is described in the Supplementary text 2 "Procedure for phenotyping SRA and SRN and deriving the adjusted means of SRA."

The TR response under progressive VPD was examined at ICARDA's physiology laboratory under controlled conditions. This experiment was designed randomizing the 90 MAGIC lines using 3 biological replicates per genotype, 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. The day after, pots were subsequently closed with plastic bags and balls in order to limit evaporation. TR was measured under increasing VPD ranging from 0.4 to 5.4 kPa in a greenhouse under 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, 2009b; Schoppach and Sadok 2012; Schoppach *et al.* 2017; Tamang and Sadok 2018; Sadok and Tamang 2019) and described in detail in the Supplementary text 3 "Procedure for estimating the TR at different values of VPD."

Descriptive statistics and correlation analyses
Variation in GY and SRA in the panel MAGIC lines was investigated computing maximum and minimum values, mean, median, and standard deviation (SD). This descriptive analysis was computed using "metan" package (Olivoto and Lúcio 2020) implemented in R 4.0.3 statistical (R Core 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 latter set of correlation analyses was computed using "polycor" package (Fox 2010) implemented in R 4.0.3 statistical (R Core Team 2019).

GP models fitted to SRN
In the present study, 2 different GP models were fitted to SRN combining phenotypic data with genotypic information obtained with the Barley 50 k SNP chip (Puglisi *et al.* 2021). For this trait, GP models were fitted following 2 different hypotheses. First, 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 ith genotype in the kth 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 γ 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:

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

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 Φ^{-1} is the corresponding inverse function. Consequently, threshold models are specified with $C - 1$ linear predictors η_{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, β points out the vector of fixed effects, and b is the vector of random effects. The probabilities π_{ikc} are linked to the linear predictors η_{ikc} as follows:

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

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

$$l_{ik} = X_{ik}^T\beta + Z_{ik}^T u + e_{ik} \tag{1}$$

where the error terms e_{ik} are independent and identically distributed and follow a normal distribution with mean 0 and SD 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 5 extended threshold models (Table 1), which were already substantiated and described in other studies (Jarquín *et al.* 2014; Montesinos-López *et al.* 2015a).

The resulting 5 models include the following sets of linear predictors:

$$\mathbf{SRN - Model}\ 1: l_{ik} = R_k + L_i + \varepsilon_{ik} \tag{2}$$
$$\mathbf{SRN - Model}\ 2: l_{ik} = R_k + g_i + \varepsilon_{ik} \tag{3}$$
$$\mathbf{SRN - Model}\ 3: l_{ik} = R_k + g_i + g_{Ai} + \varepsilon_{ik} \tag{4}$$
$$\mathbf{SRN - Model}\ 4: l_{ik} = R_k + L_i + g_i + \varepsilon_{ik} \tag{5}$$
$$\mathbf{SRN - Model}\ 5: l_{ik} = R_k + L_i + g_i + g_{Ai} + \varepsilon_{ik} \tag{6}$$

where l_{ik} is the latent "liability" variable of kth replicates in the ith line. SRN-Model 1 includes R_k, which is the fixed effect of kth replicates and L_i that is the random effect of the ith line

Table 1. Summary of the linear predictors incorporated in the GBLUP and TGBLUP models used to analyze SRN, SRA, and TR.

Model	Main effects			Interaction
	R	L	G	G × G
SRN-Model 1	×	×		
SRN-log-Model 1				
SRA-Model 1				
TR-Model 1				
SRN-Model 2	×		×	
SRN-log-Model 2				
SRA-Model 2				
TR-Model 2				
SRN-Model 3	×		×	×
SRN-log-Model 3				
SRA-Model 3				
TR-Model 3				
SRN-Model 4	×	×	×	
SRN-log-Model 4				
SRA-Model 4				
TR-Model 4				
SRN-Model 5	×	×	×	×
SRN-log-Model 5				
SRA-Model 5				
TR-Model 5				

G, marker covariates; G × G, first-order additive × additive epistasis; L, line; R, replicate.

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 ith line, that is $g_i = \sum_{n=1}^{p} x_{in} b_n$, where x_{in} is the genotype of the ith line at marker n and b_n is the corresponding effect of marker n. The vector of additive genetic value $\mathbf{g} = (g_1, g_2, g_3 \dots g_i)$ is supposed to be normally distributed as $\mathbf{g} \sim N\left(0, G\sigma_g^2\right)$ with mean 0 and variance–covariance structure $G\sigma_g^2$, where σ_g^2 points out the additive genetic variance σ_g^2 and $\mathbf{G}$ is the genomic marker relationship matrix (VanRaden 2008). SRN-Model 3 extends SRN-Model 2 including first-order multiplicative epistatic effects $\mathbf{g_A} = (g_{A1}, g_{A2}, \dots g_{Ai})$, which are assumed to be distributed as $\mathbf{g_A} \sim N\left(0, G_A\sigma_{g_A}^2\right)$, that is the vector of epistatic effects follows a normal distribution with mean 0 and epistatic additive × additive genetic variance $\sigma_{g_A}^2$ (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 2014a) in R 4.0.3 statistical (R Core Team 2019) using default prior distributions and modifying codes published in other studies (Montesinos-López *et al.* 2015a).

Second, we handled SRN as count data for predicting this trait *per se*, fitting 5 log-normal GP models based on GBLUP and indicated as SRN-log-Model 1–5 (Table 1) (Montesinos-López *et al.* 2015b). In SRN-log-Model 1–5, the response variable is the logarithm of SRN, that is $log(y_{kj} + 1)$, and was fitted using the same sets of linear predictors (R_k, L_j, g_j, g_{Aj}) described for the 5 extended TGBLUP models [Table 1; Equations (2–6)]. 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 ε_{ik} of Kth replicates in i^{th} line, which in these models is distributed as $\varepsilon_{ik} \sim N(0, \sigma_e^2)$, that is the residuals are independent and normally distributed with mean 0 and variance σ_e^2. Like TGBLUP models, log-normal models were implemented using BGLR package (Pérez and De Los Campos 2014b) in R 4.0.3 statistical (R Core Team 2019).

GP models fitted for SRA and TR

The 5 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:

$$\textbf{SRA} - \textbf{Model 1}: \ y_i = L_i + \varepsilon_i \tag{7}$$
$$\textbf{SRA} - \textbf{Model 2}: \ y_i = g_i + \varepsilon_i \tag{8}$$
$$\textbf{SRA} - \textbf{Model 3}: \ y_i = g_i + g_{Ai} + \varepsilon_i \tag{9}$$
$$\textbf{SRA} - \textbf{Model 4}: \ y_i = L_i + g_i + \varepsilon_i \tag{10}$$
$$\textbf{SRA} - \textbf{Model 5}: \ y_i = L_i + g_i + g_{Ai} + \varepsilon_i \tag{11}$$
$$\textbf{TR} - \textbf{Model 1}: \ y_i = L_i + \varepsilon_i \tag{12}$$
$$\textbf{TR} - \textbf{Model 2}: \ y_i = g_i + \varepsilon_i \tag{13}$$
$$\textbf{TR} - \textbf{Model 3}: \ y_i = g_i + g_{Ai} + \varepsilon_i \tag{14}$$
$$\textbf{TR} - \textbf{Model 4}: \ y_i = L_i + g_i + \varepsilon_i \tag{15}$$
$$\textbf{TR} - \textbf{Model 5}: \ y_i = L_i + g_i + g_{Ai} + \varepsilon_i \tag{16}$$

where y_i is the adjusted mean of SRA [Equations (7–11)] or TR [Equations (12–16)], ϵ_i is the error term of the ith 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 σ_ε^2. In these models, the linear predictors L_j, g_j, and g_{Aj} follow the same distribution defined for TGBLUP models. These extended GBLUP models (SRA-Models 1–5, TR-Model 1–5) were implemented using BGLR package (Pérez and De Los Campos 2014a) in R 4.0.3 statistical (R Core 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 best linear unbiased estimator (BLUE), a_i is the lower bound estimate of y_i computed as the difference between y_i and 2 SD, and b_i is the upper bound estimate of y_i computed as the sum of y_i with 2 SD.

Cross-validation of GP models

For the extended TGBLUP models, leave-one-out (LOO) 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:

$$\textbf{BS} = \textbf{\textit{n}}^{-1} \sum\nolimits_{\textbf{i=1}}^{\textbf{\textit{n}}} \ \sum\nolimits_{\textbf{c=1}}^{\text{g}} \ (\hat{\boldsymbol{\pi}}_{\textbf{ic}} - \boldsymbol{d}_{\textbf{ic}})^2, \tag{17}$$

where $(\hat{\pi}_{ic} - d_{ic})^2$ is the average square difference between $\hat{\pi}_{ic}$ predictions and d_{ic} classes for observation i into category c. BS obtained with Equation (17) was divided by 2 in order to have a range that varies from 0 to 1 (Brier 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.

Results

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

To assess the variability of SRA and SRN, the panel of MAGIC lines was phenotyped at the seedling stage under controlled conditions. This analysis showed that SRN varies greatly in the MAGIC population as it ranges between 2 and 8 with a mean of 5 seminal roots and a SD of 0.84 (Fig. 1a; Supplementary Table 1). The adjusted means of SRA and the corresponding 95% confidence intervals of estimates were computed using BLUEs, analyzing the adopted experimental design with a linear mixed model, which was fitted to the raw measurements of SRA. 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 86.61° (Fig. 1b; Supplementary Table 2). Both belowground traits exhibit a bell-shaped distribution (Fig. 1, a and b) and particularly, Shapiro–Wilk normality test showed that, for SRA, the null hypothesis, that is that the adjusted means of SRA follow a normal distribution, cannot be rejected (P-value = 0.1844).

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 5.4 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 showed a segmented response in a large fraction of genotypes, while in 28 MAGIC lines model fitting and selection indicated a linear response of TR to increasing levels of VPD (Supplementary Table 2). The breakpoint values of genotypes showing a segmented trend of TR to increasing VPD ranged from 2.3 to circa 2.5 kPa, in agreement with the results presented in previous studies (Sadok and Tamang 2019). Consequently, we investigated the variability of whole-plant TR at a VPD of 2.7 kPa, which is a VPD value higher than the breakpoints of genotypes showing a segmented TR response. This analysis showed that the panel of MAGIC exhibits circa a 5-fold variation of TR as measured values range from 4.05 (genotype "M382") to 13.6 (genotype "M150") $mg_{H2O}m^{-2}s^{-1}$ (Fig. 1c; Supplementary Table 2).

Correlation of belowground and physiological traits with GY

A genotype × environment (GGE) biplot analysis was carried out using the adjusted means of GY to assess the level of correlation among site-by-season combinations and identify environments with peculiar bio-climatic parameters (Fig. 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) owing to the dryer and hotter conditions of this site. For assessing whether belowground and physiological traits might contribute to determining yield formation under limiting nitrogen (FioLN) and water conditions (MarIN), we correlated GY with SRN, SRA, and TR measured at a VPD of 2.7 kPa, 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 (Fig. 2b). This pairwise correlation
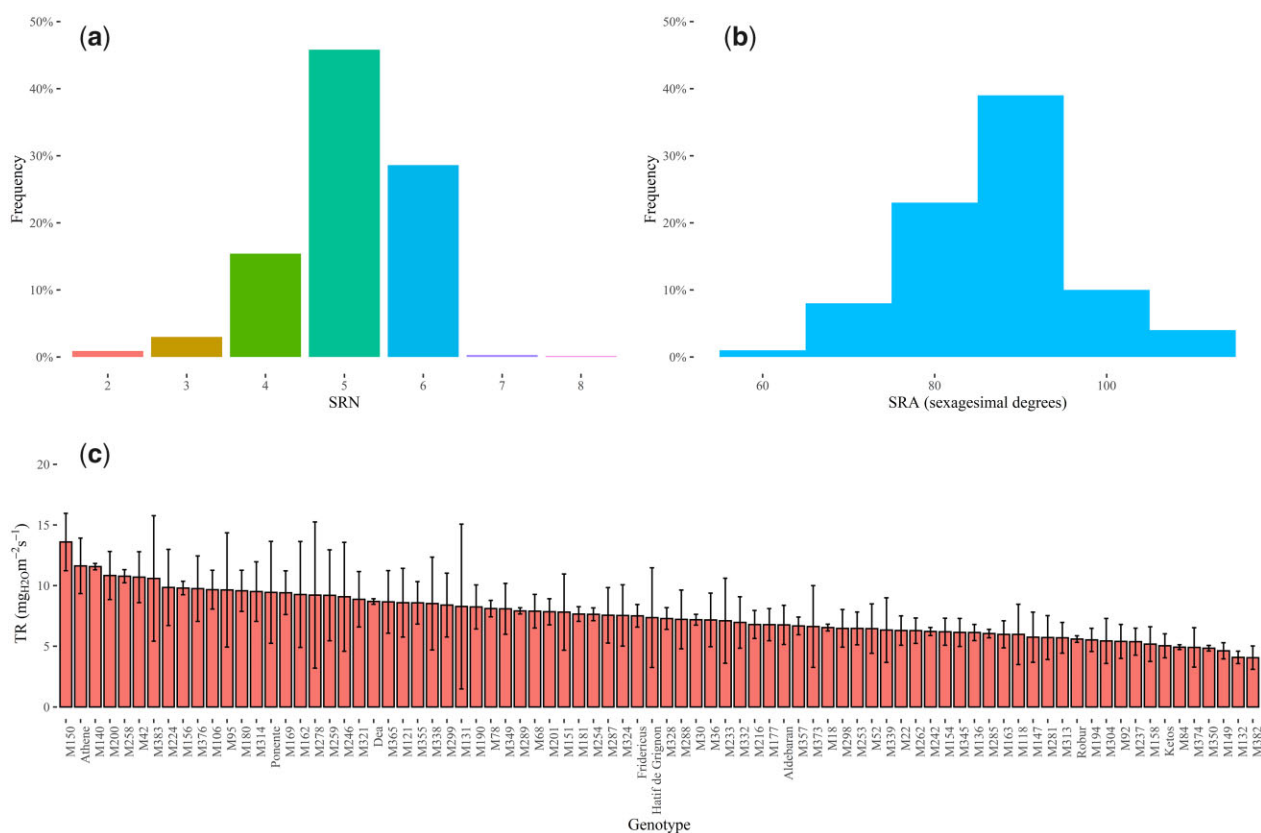
Fig. 1. Phenotypic distribution of SRA, SRN, and whole-plant TR measured at a VPD of 2.7 kPa in the panel of MAGIC lines: a) histogram of SRN counted in the MAGIC lines; b) histogram of the adjusted means of SRA measured in sexagesimal degrees; c) bar plot of TR measured at a VPD of 2.7 kPa. Error bars point out the 95% confidence interval of TR values.
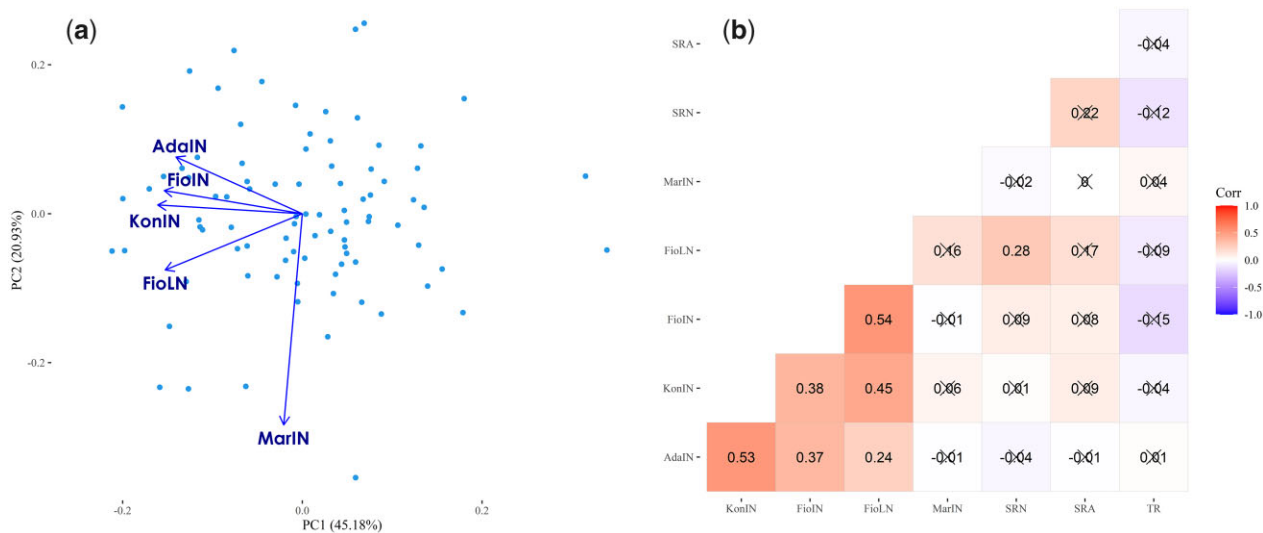

Fig. 2. GGE biplot of GY and pairwise correlations of SRA, SRN, GY, and TR at a VPD of 2.7 kPa. a) The environment-vector view of the GGE biplot indicates similarities among test environments in discriminating the genotypes. b) Depending on the trait distribution type (discrete or continuous), values indicate pairwise Pearson's correlation or polyserial correlation between GY, SRA, SRN, and TR measured under high evaporative demand of VPD (2.7 kPa). Correlation values showing P-values larger than 0.05 are marked by a cross.

analysis indicated that SRN shows a positive and moderate correlation with SRA ($r = 0.22$, P-value $= 0.07$) (Fig. 2b; Supplementary Fig. 1), while GY in FioLN exhibited positive correlations with SRN, showing a value of 0.28 (P-value $= 0.02$) (Fig. 2b; Supplementary Fig. 1). Unexpectedly, no significant

correlations were observed between GY measured in MarIN with belowground and physiological traits. Overall, GY showed positive correlations among KonIN, AdaIN, and FioIN, while no significant correlations were observed between MarIN and the remaining sites (KonIN, AdaIN, FioIN, and FioLN) (Fig. 2b).
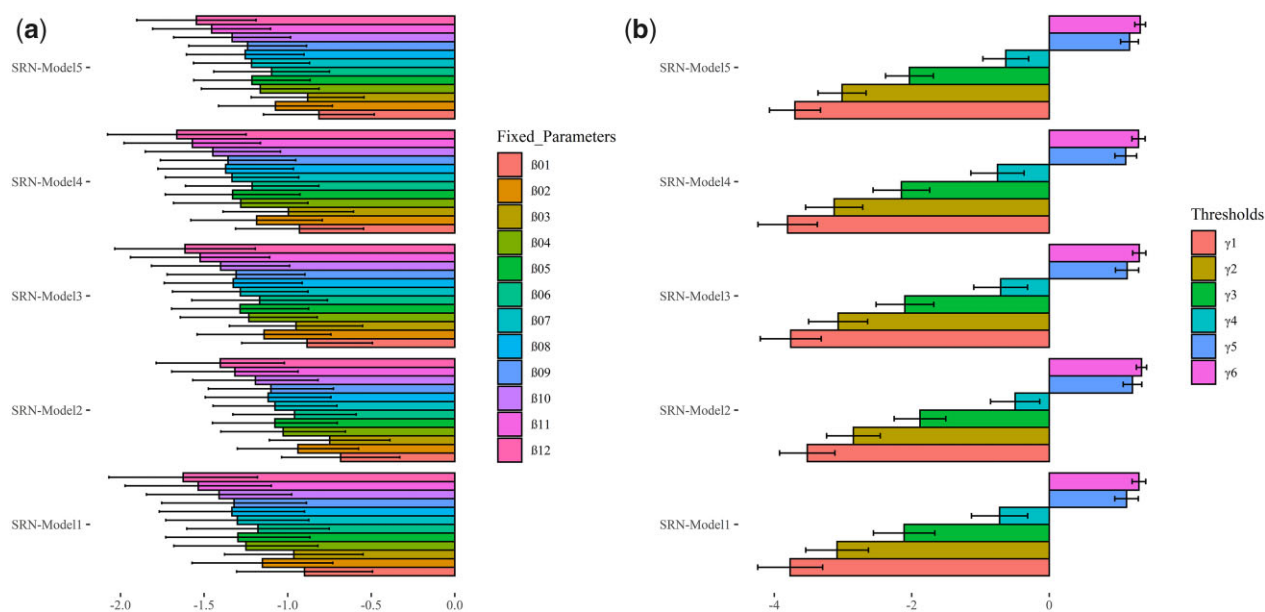
Fig. 3. Bar plots of the estimated parameters and thresholds in SRN-Models 1–5. a) Posterior mean and 95% confidence interval of the 12 fixed effects (β_1, β_2,…β_{12}) estimated in the 5 TGBLUP models. b) Posterior mean and 95% confidence interval of threshold parameters (γ_1, γ_2, .γ_6). In both graphs, the error bars point out the posterior 95% confidence interval of parameter values.

Using SRN as proxy trait to predict GY under limiting nitrogen conditions

Here, following the moderate correlation observed between GY obtained under limiting nitrogen conditions and the number of seminal roots in FioLN (Fig. 2b), we assumed that SRN might serve as a proxy trait for predicting GY of MAGIC lines under nitrogen or nutrient 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 for fitting 5 TGBLUP GP models fitted with different sets of linear predictors, 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; Equations (2–6)] (Montesinos-Lopes *et al.* 2015b). 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 first-order additive × additive epistasis [Table 1; Equations (2–6)]. The estimates of fixed effects and their 95% confidence intervals, that is the effects of the 12 replicates implemented for phenotyping SRN, showed similar values in the 5 threshold models (Fig. 3a; Supplementary Table 3). Similarly, the estimates of the 6 model thresholds (γ_1, γ_2,…γ_6) indicated similar values across the 5 threshold models (Fig. 3b; Supplementary Table 4). Overall, the fitting of SRN-Models 1–5 showed similar posterior means of fixed effects and thresholds (Fig. 3; Supplementary Tables 3 and 4).

The probabilities for each ordinal categorical phenotype estimated in the 5 TGBLUP models for the whole data set are shown in Fig. 4. These boxplots showed that the average probabilities for category 5 (5 seminal roots) are circa 0.50 in the whole data set for all 5 models followed by categories 6 and 4 (Fig. 4). Unlike the distribution estimated from raw data, these probability estimates

consider the effect of replicates but, overall, show similar trends from the distributions obtained based on raw frequencies (Fig. 1a).

The analysis of the estimated variance components in the 5 TGBLUP models fitted to SRN data shows that overall, the total variance explained in SRN-Models 1–5 varies from 1.05 to 1.07 (Table 2). Molecular markers explain circa 4.76% of the total variance in SRN-Model 2, and circa 1.8% in SRN-Model 3, 4, and 5 (Table 2). Similarly, the first-order additive × additive epistasis explains 4.76% of the total variance for SRN-Model 3 and 1.88% for SRN-Model 5.

To assess the predictive ability of SRN-Models 1–5, LOO cross-validation was implemented to compute BS between predicted and observed categorical values and the PCCC (Gianola and Schon 2016; Montesinos-López *et al.* 2020). Cross-validation analysis pointed out the PCCC is circa 50% for the 5 models considered in the present study (Fig. 5a; Supplementary Table 5). Similarly, BS points out a high predictive ability as a value of circa 0.3 was estimated for SRN-Models 1–5 (Fig. 5b).

GP of SRN using log-normal transformation of count data

Beyond using SRN as a proxy trait for predicting GY under nutrient scarcity, GP models for count data were fitted to predict this trait independently from its association with drought tolerance and GY under nitrogen deficiency at the mature stage, using log-normal models (SRN-log-Models 1–5) incorporating the same combinations of linear predictors included in SRN-Models 1–5 (Table 1). The analysis of variance components of these 5 log-normal models showed that "SRN-log-Model 5," which incorporates line (L), marker (G), and first-order additive × additive epistasis (G × G), has a lower error variance compared to the other models considered in the present study, and allows to better fit the data (Fig. 6a; Supplementary Table 6). The variance of G × G is 27.52% for SRN-log-Model 3 and 19.83% for SRN-log-Model 5 (Fig. 6a; Supplementary Table 6). LOO cross validation pointed out that the predictive ability
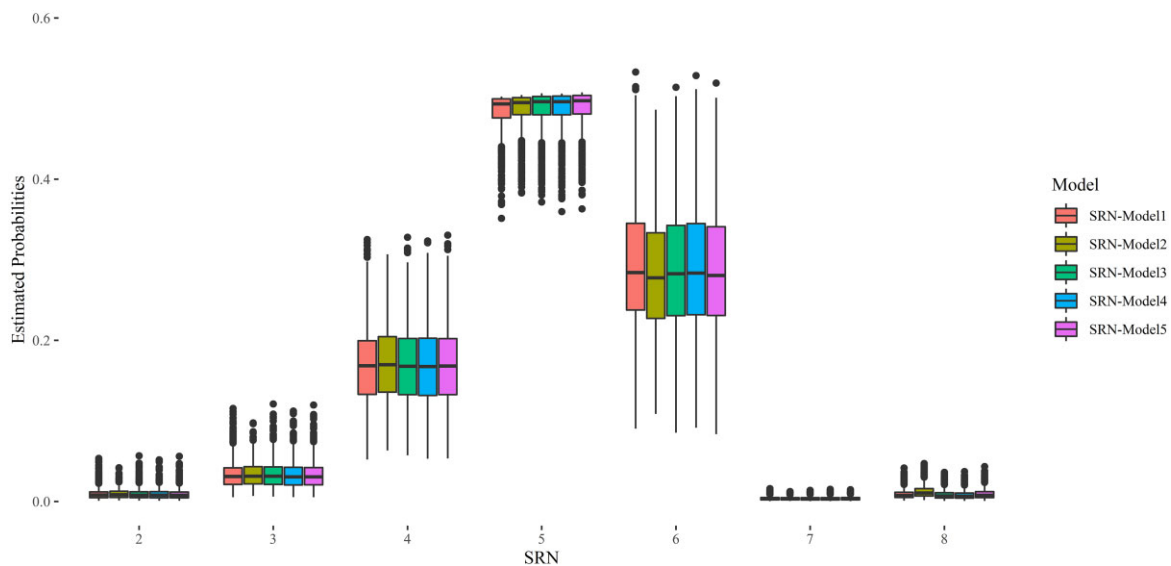
Fig. 4. Boxplots of the estimated probabilities of SRN in SRN-Models 1–5. Each boxplot summarizes the distribution of the estimated probability in each category in the 5 TGBLUP models.

Table 2. Estimated variance components of SRN-Models 1–5.

Model	L	G	G × G	Error variance	Total variance
SRN-Model 1	0.07 (6.54%)			1	1.07
SRN-Model 2		0.05 (4.76%)		1	1.05
SRN-Model 3		0.02 (1.86%)	0.04 (4.76%)	1	1.07
SRN-Model 4	0.04 (3.77%)	0.02 (1.88%)		1	1.06
SRN-Model 5	0.02 (1.88%)	0.02 (1.88%)	0.02 (1.88%)	1	1.06

L, estimated variance of line effects; G, estimated variance of marker effects; G × G points out the variance of first-order additive × additive epistasis. Numbers between brackets point out percentage of explained variance of each model predictor.
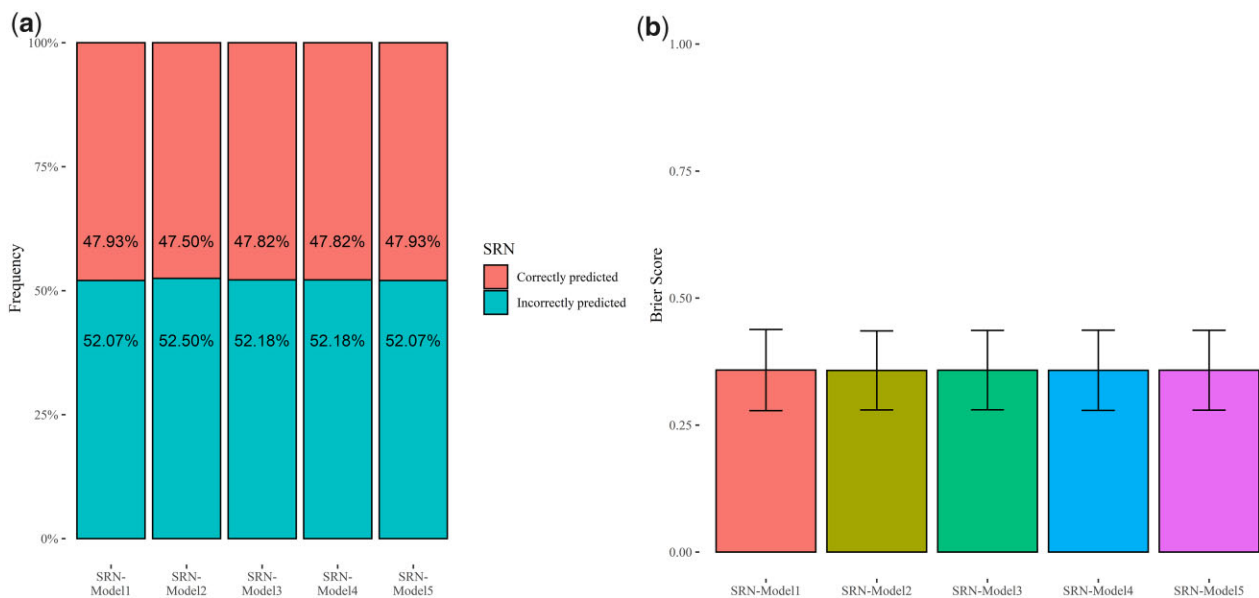

Fig. 5. Predictive ability of SRN-Models 1–5. a) Proportion of cases correctly classified using the LOO cross-validation of the 5 TGBLUP models. b) Brier scores obtained from LOO cross-validation of the 5 TGBLUP models.
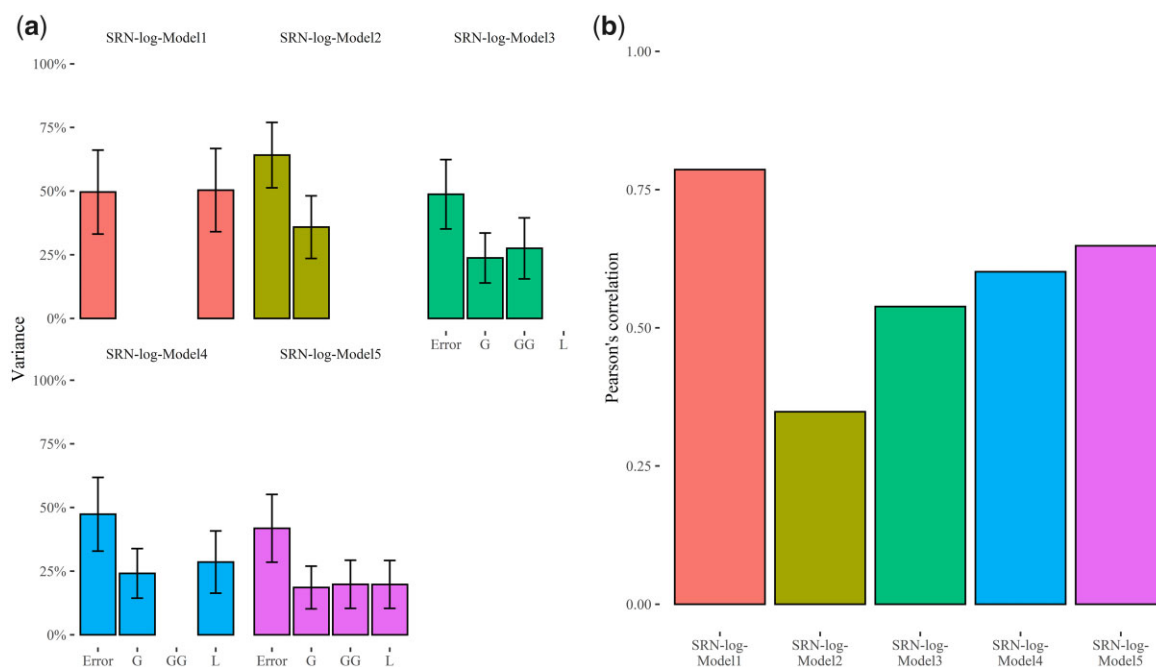
Fig. 6. Estimated variance components and predictive ability of SRN-log-Models 1–5. a) Bar plots indicate the explained variance of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects while G × G and "Error" point out the variance of first-order additive × 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.

values of these models, measured using Pearson's correlation coefficient between predicted and observed data, range from 0.35 (SRN-log-Model 2) to 0.79 (SRN-log-Model 1), while SRN-log-Model 3, SRN-log-Model 4, and SRN-log-Model 5 show predictive ability values of 0.54, 0.60, and 0.65, respectively (Fig. 6b; Supplementary Table 6). Overall, model comparison indicated that for SRN, log normal models that explicitly incorporate markers and first-order additive × additive epistatic interactions capture a larger fraction of the total phenotypic variability and have better predictive ability.

Prediction of SRA in barley seedlings

As with log-normal models, the GBLUP counterpart of the 5 TGBLUP models used for predicting SRN were fitted to SRA. These 5 models incorporate the main effects and interactions used for SRN [Table 1; Equations (7–11)] but assume that the response variable, that is SRA, is continuous and follows a normal distribution. The adjusted means of SRA were combined with 20,426 polymorphic SNPs detected in this panel of MAGIC population to fit these extended GBLUP models. The analysis of variance components of these 5 models showed that "SRA-Model 5," which incorporates line, marker, and first-order additive × additive epistatic interaction effects, has a lower error variance compared to the other models considered in the present study and allows to better fit data (Fig. 7a; Supplementary Table 7). The variance of first-order additive × additive epistatic interaction was 27.79% for SRA-Model 3 and 17.81% for SRA-Model 5 (Fig. 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 (Fig. 7b; Supplementary Table 7). As observed for log-normal models

(SRN-log-Models 1–5), models that predict SRA incorporating marker and first-order epistatic effects in the linear predictors show higher predictive ability and better model fitting.

Models for predicting TR under high evapotranspiration demand

The 5 combinations of linear predictors incorporated in GP models fitted to SRA and SRN data [Table 1; Equations (12–16)] were used to predict TR at a VPD of 2.7 kPa. The analysis of variance components of these 5 models showed that "TR-Model 5," which incorporates line, marker, and first-order additive × additive epistatic interaction effects as linear predictors, has a lower error variance compared to the other models considered in the present study and allows to better fit the data (Fig. 8a; Supplementary Table 8). The analysis of variance components showed that first-order additive × additive epistatic interaction effects explain 21.93% and 15.69% of the total variance for TR-Model 3 and TR-Model 5, respectively (Fig. 8a; Supplementary Table 8). 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 0.68 (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 (Fig. 8b; Supplementary Table 8). Like observed for GP models fitted to belowground traits, this analysis demonstrated that models that explicitly incorporate marker and interaction effects fit better the TR data and have better predictive ability.

Discussion

In the present study, the phenotypic variability of SRN, SRA, and TR under increasing evaporative demand was surveyed in a panel of MAGIC lines of barley (Supplementary Tables 1 and 2). The results showed that, in this genetic material, SRN can vary from 2
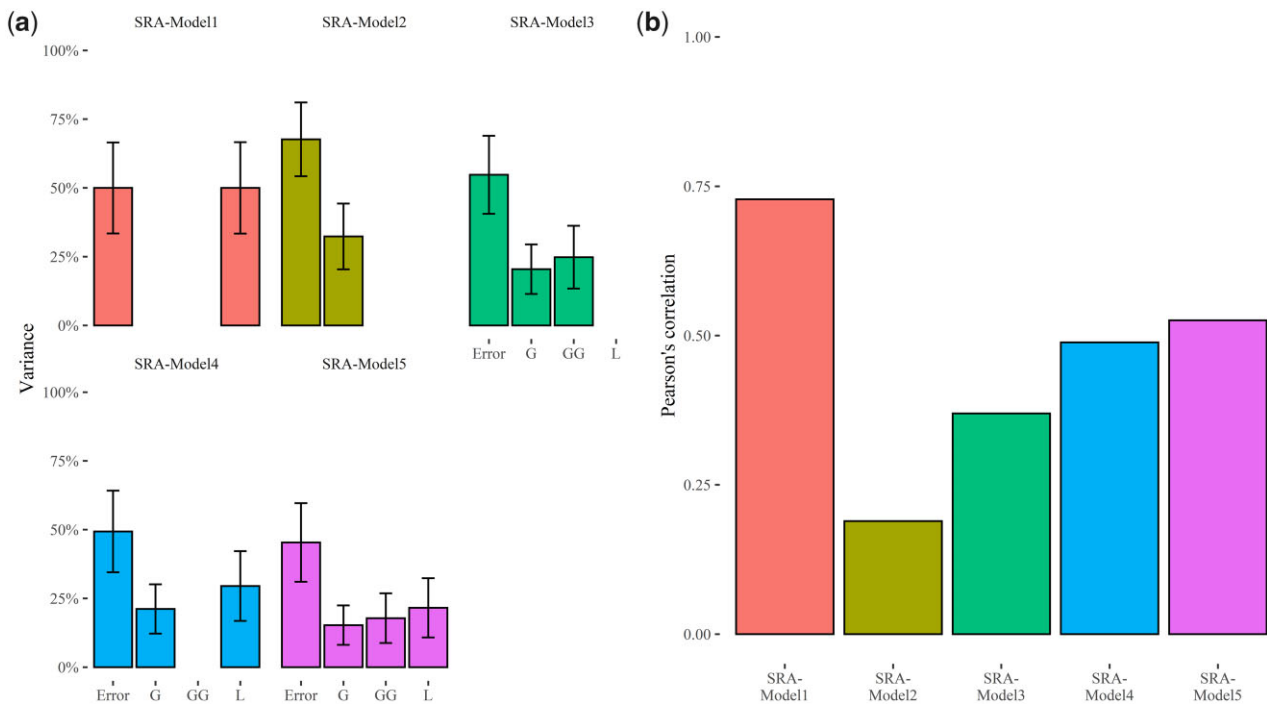
Fig. 7. Estimated variance components and predictive ability of the SRA-Models 1–5. a) Bar plots indicate the variance of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects while G × G and "Error" point out the variance of additive × 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.
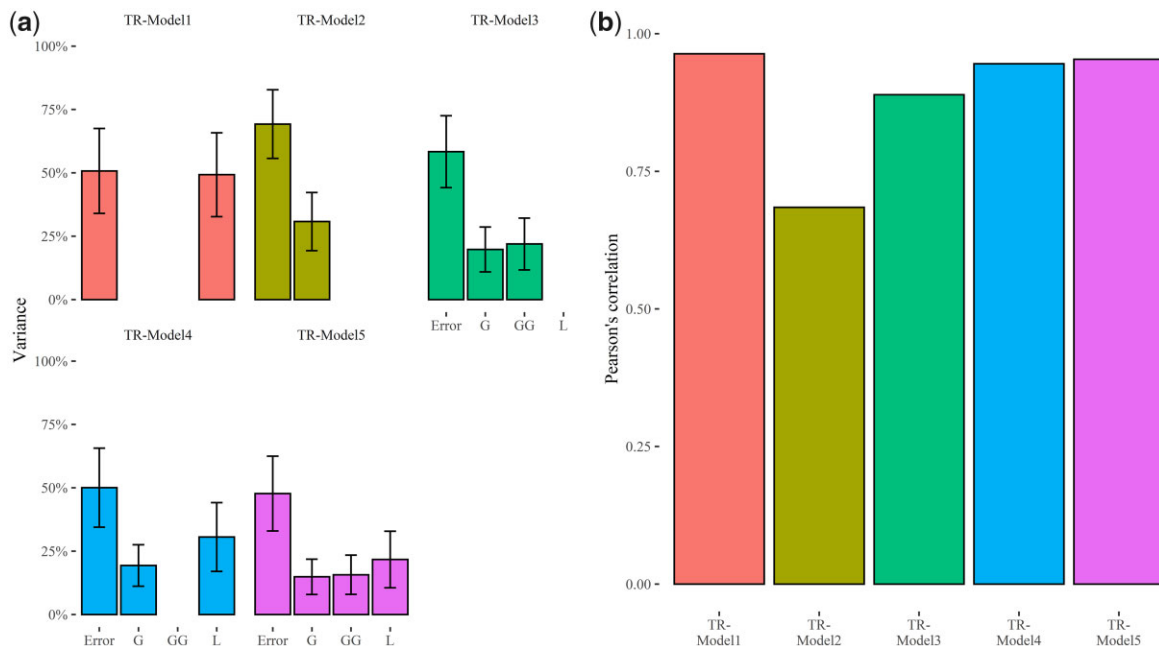

Fig. 8. Estimated variance components and predictive ability of the TR-Models 1–5: a) bar plots indicate the variance of each component, expressed as percentage of the total model variance. L indicates the estimated variance of line effects; G is the estimated variance of marker effects, while G × G and "Error" point out the variance of additive × 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).

to 8 (Fig. 1a), consistently with other studies carried out on populations of mostly unrelated accessions of barley (Robinson *et al.* 2016; Jia et al. 2019).

In the scientific literature, protocols for SRA phenotyping propose to examine the first pair of seminal roots (Robinson *et al.*

2016) or the 2 outmost seminal roots (Jia et al. 2019). Following the first protocol, in the present study, SRA between the first pair of seminal roots has been measured, and the comparison of our results with those obtained in other works carried out following the same methodology (Robinson *et al.* 2016, 2018) points out that

the SRA has a wider phenotypic distribution in MAGIC lines. On the other side, the lack of common genotypes between these experiments hampers our ability to compare results across different studies and does not allow to attribute the wider distribution of SRA detected in our genetic material to confounding or genetic effects. Unlike the findings reported in other studies (Robinson *et al.* 2016), SRA and SRN measured in barley seedlings showed a moderate correlation ($r = 0.22$; Fig. 2b), slightly above the threshold of significance (P-value = 0.07; Supplementary Fig. 1). Nevertheless, this comparison is not exhaustive as in other studies SRN was 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. The analysis of SRN as continuous or discrete phenotypic trait implies different assumptions, which in turn hamper the comparisons of trait correlations; as in our analysis, SRN was considered as a discrete trait and consequently the polyserial correlation between SRA and SRN was computed instead of using Pearson's correlation coefficient. Interestingly, the present study indicated a moderate and positive and significant correlation of SRN ($r = 0.28$; P-value = 0.02) with GY measured in FioLN, which was managed with reduced amount of nitrogen (Fig. 2b; Supplementary Fig. 1). As highlighted for other cereal crops, this correlation suggests a link between the root architecture system and the ability of barley to grow in soils with a reduced nitrogen fertilization without experiencing yield penalty. Anyway, the sample size of this analysis along with the inclusion of only 2 site-by-season combinations organized with this nitrogen management impose to carry out other studies to definitively underpin the tie between these belowground traits measured at the seedling stage and the ability to promote yield formation under limiting nitrogen conditions.

In recent years, TR has been widely targeted in different crops 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 1 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 substantiated with SRA and SRN analyses, technical causes and the lack of common genotypes hamper cross-study comparison of TR in different barley germplasm, 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).

Unexpectedly, our analyses did not detect significant correlation between TR and GY measured in MarIN, which is the hottest and driest environment investigated in this study pointing out that, in barley, this trait might exhibit correlation with GY under harsher conditions. In field trials, GY depends on the genotypic values of plants and environmental factors that exert their influence from sowing to harvest and consequently it is not surprising to detect inconsistent correlations with the belowground and physiological traits examined in the present study. The present study does not address the correlations between SRN, SRA, and

TR measured at the juvenile stage and the root architecture of mature barley, which remain still unclear and deserve additional analyses.

In the present study, genomic-enabled prediction of SRA, SRN, and TR measured at a VPD of 2.7 kPa was carried out, leveraging on prediction models for different types of data (continuous, count, and ordinal) and we showed that these traits can be modeled with high predictive ability. These findings are relevant for barley improvement and ideotype breeding as can pave the way to exploit untapped plant collection minimizing phenotyping costs. For standard, log normal, and TGBLUP GP models, we used 5 sets of linear predictors, which differ for the types of effects considered (Table 1). Using TGBLUP models, we observed that the inclusion of additive × additive epistatic interaction (G × G) in the set of linear predictors (SRN-Model 3 and SRN-Model 5) increases the total variance that models account for (Table 2). These results are corroborated in log-normal GP models, as also these models benefit of the inclusion of additive × additive epistatic interaction in the set of linear predictors (Figs. 6–8). Overall, for all traits examined in the present study, we found that the inclusion of interaction effects brings advantages both in model fitting and predictive ability, as already substantiated in several studies (Varona *et al.* 2018). Despite being called additive × additive epistatic interactions, the functional interpretation of these effects and of the variance component counterpart may be misleading as their existence in nonadditive GP models and genome-wide association studies does not prove the role of epistasis in the actual genetic architecture of these traits (de los Campos *et al.* 2019; Schrauf *et al.* 2020). Particularly, the low marker density used in the present study to fingerprint the panel of MAGIC and the incomplete linkage disequilibrium of markers (Puglisi *et al.* 2021) favor the detection of phantom epistasis in nonadditive GP models, that is the portion of additive effects that is not captured in the models due to incomplete linkage disequilibrium generates apparent epistasis, which is in turn detected in our models including additive × additive epistatic interactions (de los Campos *et al.* 2019; Schrauf *et al.* 2020). While the functional interpretation of first-order additive × additive effects might be questionable, the inclusion of epistatic effects in GP should be considered in conditions that foster the emergence of apparent epistasis to improve predictive ability of models (Schrauf *et al.* 2020).

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 used TGBLUP models assuming that SNR varies as an ordinal discrete variable indicating the ability of plants to grow under water or nutrient scarcity. (Montesinos-López *et al.* 2015a). Despite the low values of variance explained by molecular markers (Table 2), the TGBLUP models fitted to SRN showed high predictive ability values (Brier Score equals to 0.36) (Fig. 5), highlighting that GP can be successfully applied to traits showing low heritability 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).

Count data arise in plant breeding 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). Currently, for this type of data, GP models are fitted using standard GBLUP along with log transformed data or specific generalized mixed linear models that use a link function

based on Poisson or Negative Binomial distributions (Montesinos-López *et al.* 2016). This latter group of models have been recently formalized, implemented in a Bayesian framework, and compared with log-normal models for count data (Montesinos-López *et al.* 2016). In the present study, GP for SRN was modeled as count data fitting log-normal GP models as preliminary tests showed that the computational demand of these models does not counterbalance the higher predictive ability of these models. Multitrait GP might be a natural approach to exploit SRA, SRN, and TR response at high VPD as multivariate GP has been shown, using either real or simulated data sets, to have superior predictive ability compared to univariate GP. While the use of traits varying on a continuous scale is straightforward in multivariate GP, count data still pose several challenges that currently hamper our ability to exploit this trait in multivariate analyses.

Data availability

The phenotypic data underlying this article are available in the online Supplementary material (Supplementary Tables 1 and 2 File 2). The genotyping data underlying this article are available in the comma delimited file named "Supplementary File 1," which contains the genotypes detected at 20,426 loci.
Supplemental material available at G3 online.

Acknowledgments

Funding

Conflicts of interest

The authors declare that there is no conflict of interest.

Literature cited

Alahmad S, El Hassouni K, Bassi FM, Dinglasan E, Youssef C, Quarry G, Aksoy A, Mazzucotelli E, Juhász A, Able JA, *et al.* A major root architecture QTL responding to water limitation in durum wheat. Front Plant Sci. 2019;10:1–18. doi:10.3389/fpls.2019.00436.

Bhatta M, Gutierrez L, Cammarota L, Cardozo F, Germán S, Gómez-Guerrero B, Pardo MF, Lanaro V, Sayas M, Castro AJ, *et al.* Multi-trait genomic prediction model increased the predictive ability for agronomic and malting quality traits in barley (*Hordeum vulgare* L.). G3 (Bethesda). 2020;10(3):1113–1124. doi:10.1534/g3.119.400968.

Biscarini F, Stevanato P, Broccanello C, Stella A, Saccomani M. Genome-enabled predictions for binomial traits in sugar beet populations. BMC Genet. 2014;15:87–89. doi:10.1186/1471-2156-15-87.

Brier GW. Verification of forecasts expressed in terms of probability. Mon Wea Rev. 1950;78(1):1–3. doi:10.1175/1520-0493(1950)078<0001:vofeit>2.0.co;2.

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de Los Campos G, Burgueño J, González-Camacho JM, Pérez-Elizalde S, Beyene Y, *et al.* Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. 2017; 22(11):961–975. doi:10.1016/j.tplants.2017.08.011.

de Dorlodot S, Forster B, Pagès L, Price A, Tuberosa R, Draye X. Root system architecture: opportunities and constraints for genetic improvement of crops. Trends Plant Sci. 2007;12(10):474–481. doi:10.1016/j.tplants.2007.08.012.

de los Campos G, Sorensen DA, Toro MA. Imperfect linkage disequilibrium generates phantom epistasis (& perils of big data). G3(Bethesda). 2019;9:1429–1436. doi:10.1534/g3.119.400101.

Desta ZA, Ortiz R. Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci. 2014;19(9):592–601.

Devi MJ, Sinclair TR, Taliercio E. 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. 2015; 10(10):e0139134.doi:10.1371/journal.pone.0139134.

Drasgow F. Polychoric and Polyserial Correlations. 2nd edn. In: Kotz S, Read CB, Balakrishnan N, Vidakovic B, Johnson NL, editors. Encyclopedia of Statistical Sciences. Hoboken, NJ, USA: John Wiley and Sons, Inc.; 2006. p. 1–6.

El Hassouni K, Alahmad S, Belkadi B, Filali-Maltouf A, Hickey LT, Bassi FM. Root system architecture and its association with yield under different water regimes in durum wheat. Crop Sci. 2018; 58(6):2331–2346. doi:10.2135/cropsci2018.01.0076.

Fletcher AL, Sinclair TR, Allen LH. Transpiration responses to vapor pressure deficit in well watered "slow-wilting" and commercial soybean. Environ Exp Bot. 2007;61(2):145–151. doi:10.1016/j.envexpbot.2007.05.004.

Fox J. polycor: Polychoric and Polyserial Correlations. R package version 0.7-10. 2010. https://CRAN.R-project.org/package=polycor.

Gianola D, Schon CC. Cross-validation without doing cross-validation in genome-enabled prediction. G3 (Bethesda). 2016; 6(10):3107–3128. doi:10.1534/g3.116.033381.

Guo G, Zhao F, Wang Y, Zhang Y, Du L, Su G. Comparison of single-trait and multiple-trait genomic prediction models. BMC Genet. 2014;15:30–37. doi:10.1186/1471-2156-15-30.

Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship information on genome-assisted breeding values. Genetics. 2007;177(4):2389–2397. doi:10.1534/genetics.107.081190.

Hammer GL, Dong Z, McLean G, Doherty A, Messina C, Schussler J, Zinselmeier C, Paszkiewicz S, Cooper M. Can changes in canopy and/or root system architecture explain historical maize yield trends in the U.S. Corn Belt? Crop Sci. 2009;49(1):299–312. doi:10.2135/cropsci2008.03.0152.

Herder G, Den Van Isterdael G, Beeckman T, De Smet I. The roots of a new green revolution. Trends Plant Sci. 2010;15(11):600–607. doi: 10.1016/j.tplants.2010.08.009.

Iheshiulor OOM, Woolliams JA, Yu X, Wellmann R, Meuwissen THE. Within- and across-breed genomic prediction using whole-genome sequence and single nucleotide polymorphism panels. Genet Sel Evol. 2016;48:15. doi:10.1186/s12711-016-0193-1.

Jannink JL, Lorenz AJ, Iwata H. Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics. 2010;9(2):166–177. doi:10.1093/bfgp/elq001.

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

Jia Z, Liu Y, Gruber BD, Neumann K, Kilian B, Graner A, von Wirén N. Genetic dissection of root system architectural traits in spring barley. Front Plant Sci. 2019;10:400–414. doi: 10.3389/fpls.2019.00400.

Kholová J, Nepolean T, Tom Hash C, Supriya A, Rajaram V, Senthilvel S, Kakkera A, Yadav R, Vadez V. Water saving traits co-map with a major terminal drought tolerance quantitative trait locus in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Mol Breeding. 2012;30(3):1337–1353. doi:10.1007/s11032-012-9720-0.

Klápště J, Dungey HS, Telfer EJ, Suontama M, Graham NJ, Li Y, McKinley R. Marker selection in multivariate genomic prediction improves accuracy of low heritability traits. Front Genet. 2020;11: 499094–499015. doi:10.3389/fgene.2020.499094.

Liu Z, Gao K, Shan S, Gu R, Wang Z, Craft EJ, Mi G, Yuan L, Chen F. Comparative analysis of root traits and the associated QTLs for maize seedlings grown in paper roll, hydroponics and vermiculite culture system. Front Plant Sci. 2017;8:436–413. doi: 10.3389/fpls.2017.00436.

Lobell DB, Gourdji SM. The influence of climate change on global crop productivity. Plant Physiol. 2012;160(4):1686–1697. doi: 10.1104/pp.112.208298.

Manschadi AM, Hammer GL, Christopher JT, DeVoil P. Genotypic variation in seedling root architectural traits and implications for drought adaptation in wheat (*Triticum aestivum* L.). Plant Soil. 2008;303(1–2):115–129. doi:10.1007/s11104-007-9492-1.

Medina S, Vicente R, Nieto-Taladriz MT, Aparicio N, Chairi F, Vergara-Diaz O, Araus JL. 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. Front Plant Sci. 2019;9:1994–1919. doi:10.3389/fpls.2018.01994.

Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001; 157(4):1819–1829.

Montesinos-López A, Gutierrez-Pulido H, Montesinos-López OA, Crossa J. Maximum a posteriori threshold genomic prediction model for ordinal traits. G3 (Bethesda). 2020;10(11):4083–4102. doi:10.1534/g3.120.401733.

Montesinos-López A, Montesinos-López OA, Crossa J, Burgueño J, Eskridge KM, Falconi-Castillo E, He X, Singh P, Cichy K. Genomic Bayesian prediction model for count data with genotype × environment interaction. G3 (Bethesda). 2016;6(5):1165–1177. doi: 10.1534/g3.116.028118.

Montesinos-López OA, Montesinos-López A, Pérez-Rodríguez P, Eskridge K, He X, Juliana P, Singh P, Crossa J. Genomic prediction models for count data. J Agric Biol Environ Stat. 2015b;20(4): 533–554. doi:10.1007/s13253-015-0223-4.

Montesinos-López OA, Montesinos-López A, Pérez-Rodríguez P, de Los Campos G, Eskridge K, Crossa J. Threshold models for genome-enabled prediction of ordinal categorical traits in plant breeding. G3 (Bethesda). 2015a;5(2):291–300. doi:10.1534/g3. 114.016188.

Mu X, Chen F, Wu Q, Chen Q, Wang J, Yuan L, Mi G. Genetic improvement of root growth increases maize yield via enhanced post-silking nitrogen uptake. Eur J Agron. 2015;63:55–61. doi: 10.1016/j.eja.2014.11.009.

Nielsen NH, Jahoor A, Jensen JD, Orabi J, Cericola F, Edriss V, Jensen J. Genomic prediction of seed quality traits using advanced barley breeding lines. PLoS One. 2016;11(10):e0164494.doi:10.1371/journal.pone.0164494.

Olivoto T, Lúcio ADC. metan: an R package for multi-environment trial analysis. Methods Ecol Evol. 2020;11(6):783–789. doi: 10.1111/2041-210X.13384.

Pérez P, de los Campos G. BGLR: a statistical package for whole genome regression and prediction. Genetics. 2014a;198(2):483–495. doi:10.1534/genetics.114.164442.

Pérez P, De Los Campos G. Genome-wide regression and prediction with the BGLR statistical package. Genetics. 2014b;198(2): 483–495. doi:10.1534/genetics.114.164442.

Puglisi D, Delbono S, Visioni A, Ozkan H, Kara İ, Casas AM, Igartua E, Valè G, Lo Piero AR, Cattivelli L, *et al.* Genomic prediction of grain yield in a barley MAGIC population modeling genotype per environment interaction. Front Plant Sci. 2021;12:1–18. doi: 10.3389/fpls.2021.664148.

R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2019. [accessed 2021 Nov 11] https://www.R--project.org.

Richard CA, Hickey LT, Fletcher S, Jennings R, Chenu K, Christopher JT. High-throughput phenotyping of seminal root traits in wheat. Plant Methods. 2015;11:13–11. doi:10.1186/s13007-015-0055-9.

Robinson H, Hickey L, Richard C, Mace E, Kelly A, Borrell A, Franckowiak J, Fox G. Genomic regions influencing seminal root traits in barley. Plant Genome. 2016;9(1). doi:10.3835/plantgenome2015.03.0012.

Robinson H, Kelly A, Fox G, Franckowiak J, Borrell A, Hickey L. Root architectural traits and yield: exploring the relationship in barley breeding trials. Euphytica. 2018;214(9):1–16. doi: 10.1007/s10681-018-2219-y.

Sadok W, Sinclair TR. Genetic variability of transpiration response to vapor pressure deficit among soybean cultivars. Crop Sci. 2009a; 49(3):955–960. doi:10.2135/cropsci2008.09.0560.

Sadok W, Sinclair TR. Genetic variability of transpiration response to vapor pressure deficit among soybean (*Glycine max* [L.] Merr.) genotypes selected from a recombinant inbred line population. F Crop Res. 2009b;113(2):156–160. doi:10.1016/j.fcr.2009.05.002.

Sadok W, Tamang BG. Diversity in daytime and night-time transpiration dynamics in barley indicates adaptation to drought regimes across the Middle-East. J Agro Crop Sci. 2019;205(4): 372–384. doi:10.1111/jac.12331.

Schoppach R, Fleury D, Sinclair TR, Sadok W. Transpiration sensitivity to evaporative demand across 120 years of breeding of Australian wheat cultivars. J Agro Crop Sci. 2017;203(3):219–226. doi:10.1111/jac.12193.

Schoppach R, Sadok W. Differential sensitivities of transpiration to evaporative demand and soil water deficit among wheat elite cultivars indicate different strategies for drought tolerance. Environ Exp Bot. 2012;84:1–10. doi:10.1016/j.envexpbot.2012.04.016.

Schoppach R, Sadok W. Transpiration sensitivities to evaporative demand and leaf areas vary with night and day warming regimes

among wheat genotypes. Funct Plant Biol. 2013;40(7):708–718. doi:10.1071/FP13028.

Schoppach R, Taylor JD, Majerus E, Claverie E, Baumann U, Suchecki R, Fleury D, Sadok W. High resolution mapping of traits related to whole-plant transpiration under increasing evaporative demand in wheat. J Exp Bot. 2016;67(9):2847–2860.

Schrauf MF, Martini JWR, Simianer H, de Los Campos G, Cantet R, Freudenthal J, Korte A, Munilla S. Phantom epistasis in genomic selection: on the predictive ability of epistatic models. G3 (Bethesda). 2020;10(9):3137–3145. doi:10.1534/g3.120.401300.

Sinclair TR, Devi J, Shekoofa A, Choudhary S, Sadok W, Vadez V, Riar M, Rufty T. Limited-transpiration response to high vapor pressure deficit in crop species. Plant Sci. 2017;260:109–118.

Sinclair TR, Hammer GL, Van Oosterom EJ. Potential yield and water-use efficiency benefits in sorghum from limited maximum transpiration rate. Funct Plant Biol. 2005;32(10):945–952. doi:10.1071/FP05047.

Sivasakthi K, Tharanya M, Kholová J, Wangari Muriuki R, Thirunalasundari T, Vadez V. Chickpea genotypes contrasting for vigor and canopy conductance also differ in their dependence on different water transport pathways. Front Plant Sci. 2017;8:1663. doi:10.3389/fpls.2017.01663.

Sorensen DA, Andersen S, Gianola D, Korsgaard I. Bayesian inference in threshold models using Gibbs sampling. Genet Sel Evol. 1995;27(3):229–249. doi:10.1016/0999-193X(96)80477-0.

Tamang BG, Sadok W. Nightly business: links between daytime canopy conductance, nocturnal transpiration and its circadian control illuminate physiological trade-offs in maize. Environ Exp Bot. 2018;148:192–202. doi:10.1016/j.envexpbot.2017.11.016.

Tao F, Rötter RP, Palosuo T, Díaz-Ambrona CGH, Mínguez MI, Semenov MA, Kersebaum KC, Nendel C, Cammarano D, Hoffmann H, et al. Designing future barley ideotypes using a crop model ensemble. Eur J Agron. 2017;82:144–162. doi:10.1016/j.eja.2016.10.012.

Tao F, Yokozawa M, Zhang Z. 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. Agric for Meteorol. 2009;149(5):831–850. doi:10.1016/j.agrformet.2008.11.004.

Tracy SR, Nagel KA, Postma JA, Fassbender H, Wasson A, Watt M. Crop improvement from phenotyping roots: highlights reveal expanding opportunities. Trends Plant Sci. 2020;25(1):105–118. doi:10.1016/j.tplants.2019.10.015.

Tsai H-Y, Janss LL, Andersen JR, Orabi J, Jensen JD, Jahoor A, Jensen J. Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat. Sci Rep. 2020;10(1):1–15. doi:10.1038/s41598-020–60203-2.

VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91(11):4414–4423. doi:10.3168/jds.2007-0980.

Varona L, Legarra A, Toro MA, Vitezica ZG. Non-additive effects in genomic selection. Front Genet. 2018;9:78. doi:10.3389/fgene.2018.00078.

Wahbi A, Gregory PJ. Growth and development of young roots of barley (*Hordeum vulgare* L.) genotypes. Ann Bot. 1995;75(5):533–539. doi:10.1006/anbo.1995.1055.

Wang X, Xu Y, Hu Z, Xu C. Genomic selection methods for crop improvement: current status and prospects. Crop J. 2018;6(4):330–340. doi:10.1016/j.cj.2018.03.001.

Zadoks JC, Chang TT, Konzak CF. A decimal code for the growth stages of cereals. Weed Res. 1974;14(6):415–421. doi:10.1111/j.1365-3180.1974.tb01084.x.

Zhang A, Wang H, Beyene Y, Semagn K, Liu Y, Cao S, Cui Z, Ruan Y, Burgueño J, San Vicente F, et al. Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. Front Plant Sci. 2017;8:1916. doi:10.3389/fpls.2017.01916.