

Accounting for Genotype-by-Environment Interactions and Residual Genetic Variation in Genomic Selection for Water-Soluble Carbohydrate Concentration in Wheat

Ben Ovenden,^{*,1} Andrew Milgate,[†] Len J. Wade,^{*,2} Greg J. Rebetzke,[§] and James B. Holland^{**}

^{*}NSW Department of Primary Industries, Yanco Agricultural Institute, Yanco NSW 2703, Australia, [†]NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga NSW 2650, Australia, [‡]Charles Sturt University, Graham Centre, Wagga Wagga NSW 2678, Australia, [§]CSIRO Agriculture and Food, Canberra, ACT 2601, Australia, and ^{**}USDA-ARS Plant Science Research Unit and North Carolina State University Department of Crop and Soil Sciences, Raleigh, NC 27695-7620

ORCID ID: 0000-0003-2015-1650 (B.O.)

ABSTRACT Abiotic stress tolerance traits are often complex and recalcitrant targets for conventional breeding improvement in many crop species. This study evaluated the potential of genomic selection to predict water-soluble carbohydrate concentration (WSCC), an important drought tolerance trait, in wheat under field conditions. A panel of 358 varieties and breeding lines constrained for maturity was evaluated under rainfed and irrigated treatments across two locations and two years. Whole-genome marker profiles and factor analytic mixed models were used to generate genomic estimated breeding values (GEBVs) for specific environments and environment groups. Additive genetic variance was smaller than residual genetic variance for WSCC, such that genotypic values were dominated by residual genetic effects rather than additive breeding values. As a result, GEBVs were not accurate predictors of genotypic values of the extant lines, but GEBVs should be reliable selection criteria to choose parents for intermating to produce new populations. The accuracy of GEBVs for untested lines was sufficient to increase predicted genetic gain from genomic selection per unit time compared to phenotypic selection if the breeding cycle is reduced by half by the use of GEBVs in off-season generations. Further, genomic prediction accuracy depended on having phenotypic data from environments with strong correlations with target production environments to build prediction models. By combining high-density marker genotypes, stress-managed field evaluations, and mixed models that model simultaneously covariances among genotypes and covariances of complex trait performance between pairs of environments, we were able to train models with good accuracy to facilitate genetic gain from genomic selection.

KEYWORDS

Genomic Selection
residual genetic variation
genotype-by-environment interaction
factor analytic model
relative accuracy
GenPred
Shared Data
Resources

Abiotic stresses such as water deficit during the growing season are a major limitation to crop production worldwide (Fischer *et al.* 2014; Ray *et al.* 2015; Foulkes and Reynolds 2015). However, the incidence and severity, as well as the timing of water deficit can differ markedly between sites and years, contributing to the historically low rates of genetic gain for yield in water deficit environments compared to well-watered environments (Richards *et al.* 2010; Araus *et al.* 2002).

An alternative to selection for grain yield directly is to identify useful traits that confer physiological adaptation to water deficit conditions (Rebetzke *et al.* 2009; Bernier *et al.* 2008; Reynolds *et al.* 2016; Lopes *et al.* 2011). For example, selection for major genes for reduced height

and photoperiod insensitivity has been used to alter plant architecture and provide drought escape potential in wheat (Passioura 1996; Bennett *et al.* 2012; Kamran *et al.* 2014). Water soluble carbohydrate (WSC) accumulation and remobilization can contribute to performance under water deficit as a source of assimilate for grain filling in many crop species (Slewinski 2012). Carbohydrate accumulation occurs when the crop synthesizes assimilate at a rate greater than needed by the various sinks (for example, developing florets, elongating shoots and roots). For wheat, the excess carbohydrate is stored mainly in the lower parts of stems and culms (Gebbing 2003), where the quantity of WSC can reach as much as 40% of total stem weight (Schnyder 1993).

The main sink for remobilization of WSC is the developing grain (Schnyder 1993; Takahashi *et al.* 2001; van Herwaarden *et al.* 1998). Remobilized WSC can contribute as much as 30–50% of grain yield under terminal drought conditions, and around 10–20% under well-watered conditions (Bidinger *et al.* 1977; Schnyder 1993; Pheloung and Siddique 1991; Gebbing and Schnyder 1999).

Selection for increased WSC in cereal breeding programs has been advocated for some time (Blum 1998), and may be an alternative to direct selection for grain yield under water deficit conditions (Asseng and van Herwaarden 2003). Previous studies indicate higher broad-sense heritability for WSC concentration (WSCC) compared to grain yield, and more stable genotype rankings within a target population of environments (Ruuska *et al.* 2006; Dreccer *et al.* 2013; Piaskowski *et al.* 2016). Indirect selection for WSCC by breeders has already occurred in some environments. For example, Rebetzke *et al.* (2009) reviewed the trend for increasing WSCC with year of variety release observed for Western Australian and the International Maize and Wheat Improvement Center (CIMMYT) wheat breeding programs. Shearman *et al.* (2005) also reported that UK wheat cultivars showed increased WSCC with progressive year of release. However, WSCC is not predicted to be a useful indirect selection criterion for yield improvement in some cases (Ovenden *et al.* 2017); furthermore, phenotyping for abiotic stress characteristics such as WSCC is resource intensive and phenotypic selection may not always be feasible in a breeding program. The practical utility of WSCC for breeding appears to be context-dependent and requires empirical investigation for specific populations and environments.

New genomic approaches may be more efficient to select for abiotic stress traits than conventional phenotypic selection, as genomic selection can leverage the information obtained from difficult and expensive phenotyping (Lin *et al.* 2014; Cooper *et al.* 2014; Leplat *et al.* 2016), to enable more rapid and inexpensive selection for many loci that may be involved in the inheritance of these complex traits (Vivek *et al.* 2016; de los Campos *et al.* 2013). Comprehensive phenotyping is still an important part of the breeding process when genomic selection is applied, as it is essential to build accurate prediction models. However, early generation nursery screens can be replaced with genomic screens (Sallam *et al.* 2015), and expensive field trials can be augmented by genetic material ‘enriched’ through genomic selection (Heffner *et al.* 2009).

Genomic selection methods use marker-based measures of realized relatedness from whole-genome marker profiles to predict genomic estimated breeding values (GEBVs) of progeny (Habier *et al.* 2013; Meuwissen *et al.* 2001; de los Campos *et al.* 2013). For lines that have both genotype and phenotype data, GEBVs can be combined with the residual line effect estimates to produce Genomic Best Linear Unbiased Predictors (G-BLUPs) for the total genotypic value, hereafter referred to as the genotypic value (GV) as described in Oakey *et al.* (2016). As reviewed by Hill (2012); Nakaya and Isobe (2012) and de los Campos

et al. (2013), a key difference between genomic and phenotypic selection is the potential reduction in breeding cycle time. Although genomic predictions of untested individuals are typically less accurate than well-replicated phenotypic evaluations, genomic selection may be more effective over time because of the opportunity to implement selection in additional generations per unit of time. If breeding cycles can be shortened, then relative accuracy can also decrease while still achieving greater genetic gain than phenotypic selection (Desta and Ortiz 2014).

Developing suitable genomic selection models for abiotic stress tolerance characteristics requires the consideration of complex genotype \times environment interactions (hereafter $G \times E$ interactions) within and across a target population of environments because expression of these traits is often environmentally-dependent. Complex patterns of $G \times E$ interactions can be incorporated into genomic prediction models, and although $G \times E$ interactions will necessarily limit gains from selection for wide adaptation, models that incorporate $G \times E$ effects can help breeders select sets of lines optimally adapted to different subsets of environments. One such approach is to use parsimonious mixed models, such as factor analytic (FA) models, that attempt to capture most of the $G \times E$ signal with a reduced number of parameters compared to a full unstructured covariance model (Guo *et al.* 2013; Burgueño *et al.* 2012; Oakey *et al.* 2016).

The objective of this study was to evaluate the ability of genomic selection models to predict a complex physiological trait (WSCC) in both untested lines and in new environments using a mixed-model that accounts for variation in the pairwise correlations of performance in different pairs of environments.

MATERIALS AND METHODS

Genotypes

The set of 358 lines used in this study was selected from a multi-site, multi-year irrigated winter cereals evaluation trial with a total of 1,314 genotypes. The genetic entries included both elite breeding lines and contemporary commercial varieties from Australian wheat breeding companies and CIMMYT representing a range of maturity types.

As WSC accumulation varies according to development stage (Ehdaie *et al.* 2008), this study aimed to assess genotypes as close as practicable to a common anthesis date. At the Yanco irrigated experiment in 2009, a total of 358 breeding lines and varieties out of the 990 grown in the experiment were selected based on common Zadoks’ development score (Zadoks *et al.* 1974) taken at approximately mid-anthesis. Lines selected were between Z49 (early head emergence) and Z56 (60% heading) which corresponds to a range of approximately 3–5 days difference in anthesis date in south-eastern Australia. For the second year of this study in 2010, the same breeding lines were selected for WSCC measurement except for 11 breeding lines that were excluded from the overall experiment in that year.

Lines were genotyped using the Illumina 9k Infinium iSelect bead-chip array (Cavanagh *et al.* 2013) resulting in 4,883 polymorphic SNPs across the population (File S1). Missing values were imputed using Beagle (Browning and Browning 2009) implemented in the R package Synbreed (Wimmer *et al.* 2012). The resulting 4,162 SNP markers (excluding markers that were duplicated, monomorphic, and those with minor allele frequency of less than 5%) were used to compute a scaled identity by descent relationship matrix (K) after Endelman and Jannink (2012) (File S2). There was little evidence of population structure in the set of lines used in this study, with the first two eigenvectors of the K matrix accounting for approximately 15% of the observed variation in genomic relationships. A principal components plot of these eigenvectors showed no obvious clustering of lines (File S3).

Copyright © 2018 Ovenden *et al.*

doi: <https://doi.org/10.1534/g3.118.200038>

Manuscript received January 11, 2018; accepted for publication March 29, 2018; published Early Online April 16, 2018.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at Figshare: <https://doi.org/10.25387/g3.6143243>.

¹Corresponding author: NSW Department of Primary Industries, Yanco Agricultural Institute, Private Mail Bag, Yanco NSW 2703, Australia, E-mail: ben.ovenden@dpi.nsw.gov.au

²Present address: The University of Queensland, School of Agriculture and Food Sciences, Brisbane QLD 4072, Australia

Experimental design

Experiments in this study were grown in south-eastern Australia at Coleambally and Yanco in 2009 and 2010. A split-plot design was used, in which the main-plot factor was irrigation treatment (irrigated or rainfed), and the 990 genotype entries (including the subset of genotypes for WSCC phenotyping) were the sub-plot factor. There were two replicates of each treatment at each site. The placement of genotypes within field experiment layouts was optimized with the spatial design package DiGGer (Coombes 2002). For the laboratory phase measuring WSCC using near-infrared spectroscopy (NIRS), an experimental design structured by day of measurement and NIRS instrument carousel and well was implemented to account for laboratory as well as field sources of experimental error. Samples from both field sites were pooled into one experimental design for each year, and the placement of genotypes within the laboratory experimental phase was also optimized with DiGGer (Coombes 2002), with partial replication of 20% of experiment field plots sampled (*i.e.*, a replication level of 1.20), following the methods in Cullis *et al.* (2006) and Smith *et al.* (2006).

Experiments were sown on a full profile of moisture, achieved by flood irrigating each site four to six weeks before sowing, so that the focus on water deficit conditions would be in the later stages of crop growth. Sowing dates were targeted for the first two weeks of May. Pre-sowing nitrogen was targeted to be approximately 120 kg N ha⁻¹. Irrigated experiments were fertilized further to a total of approximately 300 kg N ha⁻¹, consistent with the estimated nitrogen demand by the crops. Experiments were subject to a strict weed, pest and disease control regime to maximize yield potential. Irrigation scheduling for the irrigated treatments was intended to maintain soil water potentials above -100 kPa during the growing season, with irrigations commencing as soil water potential fell below -75 kPa.

Water-soluble carbohydrate measurement

Tissue for WSCC analysis was sampled from a 50 cm long section of row (0.09 m²) in each plot when the irrigated treatments at each site were approximately 180° d post-anthesis, following the methods of Rebetzke *et al.* (2008). Approximately 5-10 stalks (including leaves, leaf sheaths and heads, but not senesced plant material) from each sample were ground to pass through a 2 mm sieve. Ground biomass samples were homogenized and subsampled for scanning by NIRS with a Bruker Multi-purpose Analyzer (Bruker Optik GmbH, Ettlingen, Germany) and OPUS software (version 5.1), and WSCC for the NIRS calibration samples (10% of the full set) was determined using the alkaline ferricyanide method (Piltz and Law 2007). For the 2009 experiment, the coefficient of determination for the calibration linear model was $r^2 = 0.92$ and the root-mean-square error of cross-validation (RMSECV) was 15.4. For the 2010 experiment, the calibration linear model $r^2 = 0.92$ and the RMSECV = 16.0. The WSCC phenotype data are provided in File S4.

Statistical methods

A single-stage, multiplicative linear mixed model was used to analyze the multi-experiment data with the molecular marker data following the approaches of Beec *et al.* (2010) and Oakey *et al.* (2016). The linear mixed model was:

$$y = X\tau + Z_g u_a + Z_g u_{\bar{a}} + Z_u u + \eta$$

where y is the $n \times 1$ data vector of the response variable across p experiments with N_j plots per experiment j . Each combination of year (2009, 2010), site (YANA, COLE) and irrigation treatment (IRR, RFD) was treated as a separate experiment so that $p = 8$. τ is a $t \times 1$ vector of fixed effects for the corresponding $n \times t$ design matrix (X), including

experiment main and design-based effects. The term u is a random component with associated design matrix Z_u and contains experiment-specific terms used to capture extraneous variation (after Gilmour *et al.* 1997), including the blocking structure of the field (row, range, replicate and irrigation bay), and laboratory (day of measurement, instrument carousel and carousel well) phases of the design. The $n \times 1$ residual vector η was modeled within each year of the laboratory phase design. The $m \times 1$ vector of genetic line within environment effects g , with corresponding design matrix Z_g , is partitioned into a vector of additive line within environment effects u_a and residual line within environment effects $u_{\bar{a}}$ such that $g = u_a + u_{\bar{a}}$, following the approach of Oakey *et al.* (2006). The difference between the 'additive' and 'residual' genetic effects is that the additive effects have a covariance structure proportional to the realized additive genetic relationship matrix, whereas the residual genetic effects are independent among lines. When we arrange the vectors of additive and residual genetic-within-environment effects ordered as m genotypes within each of p environments matrices, their covariance structures are:

$$\text{var}(u_a) = G_{ea} \otimes K$$

$$\text{var}(u_{\bar{a}}) = G_{\bar{e}\bar{a}} \otimes I_m$$

where K is the $m \times m$ realized genomic relationship matrix estimated from the marker data described above and I_m is an $m \times m$ identity matrix. For each variance model above, G_{ea} and $G_{\bar{e}\bar{a}}$ are the $p \times p$ matrices of variances and covariances of additive and residual genetic effects across environments, respectively. Factor analytic models of different orders (different k) can be used to model the two genetic components (Smith *et al.* 2001). For a factor analytic model, these matrices are decomposed as

$$G_{ea} = \left(\Lambda_{ea}^{p \times k} \Lambda_{ea}' + \psi_{ea} \right)$$

$$G_{\bar{e}\bar{a}} = \left(\Lambda_{\bar{e}\bar{a}}^{p \times k} \Lambda_{\bar{e}\bar{a}}' + \psi_{\bar{e}\bar{a}} \right).$$

Here, Λ is a $p \times k$ matrix of p environment loadings for k factors retained in the factor analytic model, and ψ is a diagonal matrix of the p environment specific variances.

A series of mixed models of increasing complexity of the G_{ea} and $G_{\bar{e}\bar{a}}$ terms were fitted to the data. First, we fitted diagonal covariance structures, in which the genetic variance was allowed to vary among experiments, but genetic effects were uncorrelated between experiments. Then a sequence of factor analytic models in which the covariances of genotype effects were allowed to vary within and among experiments were fitted to the data. Factor analytic models for each combination of $k = 1$ or 2 factors for each of the two genetic effects were used (Table 1). Selection of the final model was performed on the basis of Akaike's Information Criterion (AIC); (Akaike 1974), and log likelihood ratio tests comparing the nested FA models (Stram and Lee 1994). All data were analyzed using the software package ASReml-R (Butler *et al.* 2009), in the R environment (R Development Core Team 2012).

Experiment-specific GVs incorporating both additive and residual genetic effects were obtained from the final model for each line following Beec *et al.* (2010). The GV for line i at environment j was estimated from the random effect solutions in the final model as:

$$GV_{ij} = \hat{u}_{ajj} + \hat{u}_{\bar{a}ij} \quad (1)$$

These GVs were used later in cross-validation analyses as the best estimates of 'true' genotypic values at each experiment.

■ **Table 1** Factor analytic models fitted to the dataset for genomic prediction. Increasing order factor models were assessed using AIC and log likelihood ratio tests comparing nested models. The model with additive: FA2 and residual genetic: FA2 covariance structure shows a significant improvement in fit from both additive: FA1 / residual genetic: FA1 and additive: FA1 / residual genetic: FA2 models and is referred to as the final model. Higher order models were not possible to fit with the computing resources available

Covariance structure - Additive	Covariance structure - Residual genetic	REML Log Likelihood	AIC	Parameters	Full / reduced model parameters difference	Log likelihood ratio test model comparison:	Critical value	P value
DIAG	DIAG	4037.480	-8042.956	16	—	—	—	—
FA1	FA1	4405.638	-8747.276	32	16	to DIAG/DIAG	736.32	2.389×10 ⁻¹⁴⁶
FA1	FA2	4465.921	-8853.842	39	7	to FA1/FA1	120.57	5.840×10 ⁻²³
FA2	FA1	4453.568	-8829.136	39	7	to FA1/FA1	95.86	7.708×10 ⁻¹⁸
FA2	FA2	4473.418	-8854.836	46	7	to FA1/FA2	14.99	0.0361

Additive genomic estimated breeding values (GEBVs) were also obtained using the same methods but based only on the additive genetic component of the model. The GEBV for line i at environment j was estimated from the random effect solutions in the final model as:

$$GEBV_{ij} = \hat{u}_{aij} \quad (2)$$

Experiments were clustered based on the matrix of genetic correlations among experiments, using the agglomerative hierarchical clustering method given in Cullis *et al.* (2010).

Broad and narrow-sense heritability estimates were calculated from the final FA model (Table 1). Broad-sense heritability (H) for each trait at each experiment j was calculated following the generalized formula for unbalanced data in Cullis *et al.* (2006):

$$H_j = 1 - \frac{APPEV_{GVj}}{2\sigma_{aj}^2 + 2\sigma_{aj^2}}$$

where $APPEV_{GTj}$ is the average pairwise prediction error variance of GVs at experiment j (the variance of pairwise GV differences), and σ_{aj}^2 and σ_{aj^2} are the additive and residual genetic variance components for experiment j , respectively. Narrow-sense heritability was calculated for each trait at each experiment using:

$$h_j^2 = 1 - \frac{APPEV_{GEBVj}}{2\sigma_{aj}^2}$$

where $APPEV_{GEBVj}$ is the average variance of comparisons between GEBVs at experiment j , and σ_{aj}^2 is the additive genetic variance component at experiment j . The broad-sense heritabilities for total genotypic values of lines across all experiments and for genotypic values across experiments within each of the water deficit and well-watered environment clusters were also calculated using similar formulae. In these cases, however, we estimated the additive and residual genetic variances across experiments from the average of the additive and residual pairwise covariance estimates respectively between experiments for the experiments within each environment cluster (Zila *et al.* 2013; Isik *et al.* 2017).

Cross validation of genomic estimated breeding values

Fivefold cross-validation (Ogut *et al.* 2015; Crossa *et al.* 2014; Burgueño *et al.* 2012; Lorenz *et al.* 2011) was used to measure accuracy of breeding value predictions across environments. Lines were randomly assigned to five subsets for a 'fivefold' cross validation scheme across separate experiments in the model. The final FA model (Table 1) was fitted to four of the five ~80% subsets ('training set') to estimate model parameters and to predict GEBVs for the remaining subset ('validation set') (Equation 2). This procedure was repeated, holding out a different subset as the validation set each

time. The entire process of allocating lines to folds, estimating model parameters, and predicting GEBVs was replicated ten times.

The GVs from the full model including all the data (Equation 1) were considered the best estimates of the true values of total genotypic value for each line-experiment combination. The correlation between the GVs within an experiment and the GEBVs for each validation set was estimated. We refer to these as within-experiment prediction accuracies. In addition, the ability of experiment-specific GEBVs to predict genotypic values at other experiments was estimated for each validation set as the correlation between the GEBVs for experiment i and the GVs from the full model at experiment j . We refer to these as across-experiment prediction accuracies.

Relative accuracy of phenotypic value prediction (RAPV) at experiment j measures the relative accuracy of GEBVs for untested lines compared to the best estimates of their total genetic value within each experiment using both genomic and phenotypic data:

$$RAPV_j = \frac{\text{corr}(GEBV_{ij}, GV_{ij})}{\sqrt{H_j}}$$

In this case, the correlation value is divided by the square root of the broad-sense heritability estimate to estimate the accuracy of GEBVs relative to total genotypic value estimation with complete phenotypic data (Legarra *et al.* 2008). We also estimated relative accuracy of breeding value prediction at experiment j (RABV):

$$RABV_j = \frac{\text{corr}(GEBV_{ij}, GV_{ij})}{\sqrt{h_j^2}}$$

The RABV compares the accuracy of GEBVs for untested lines relative to GEBVs based on complete data. The GEBVs are more informative to predict gain from selection after intermating selected lines to generate a new population of breeding lines.

Data and reagent availability

The supplementary files contain the data used in this study. File S1 contains the SNP genotype information, File S2 contains the relationship matrix and File S3 contains a PCA plot of the first two eigenvectors of the relationship matrix. File S4 contains the phenotype dataset for WSCC. File S5 is script for the models used for genomic selection and cross-validation and File S6 contains the cross-validation sets. Supplemental material available at Figshare: <https://doi.org/10.25387/g3.6143243>.

RESULTS

Model selection and genotype × environment interactions

The 2-factor FA model for both additive and residual variance matrices was selected as the best model using the AIC (Table 1) and used for

■ **Table 2** Genetic variances, heritability, predictive ability and relative accuracy by experiment and environment cluster, with standard deviations given in parentheses. Experiment codes are given as year-site-irrigation treatment. The experiments 09YANA_RFD and 09COLE_RFD constitute the water deficit experiment cluster; all other experiments are included in the well-watered experiment cluster. The predictive ability of the GEBVs model at each experiment and environment cluster was divided by the broad-sense heritability to provide measures of accuracy relative to phenotypic selection (RAPV), and by the narrow-sense heritability to provide relative accuracy to total estimated breeding values (RABV)

Experiment or environment cluster	Additive genetic variance	Residual genetic variance	Proportion of genetic variance that is additive	Broad-sense heritability (H)	Narrow-sense heritability (h^2)	Predictive ability	Relative accuracy against H (RAPV)	Relative accuracy against h^2 (RABV)
All experiments	0.00824	0.01537	34.90%	0.778	0.363	0.480 (0.206)	0.544 (0.234)	0.797 (0.343)
Well-watered	0.01297	0.02299	36.07%	0.788	0.413	0.502 (0.192)	0.565 (0.220)	0.781 (0.304)
Water deficit	0.00914	0.03028	23.19%	0.810	0.181	0.455 (0.177)	0.506 (0.197)	1.070 (0.417)
09COLE_IRR	0.012713	0.043910	22.45%	0.853	0.345	0.503 (0.188)	0.545 (0.203)	0.857 (0.320)
09COLE_RFD	0.014623	0.031643	31.61%	0.760	0.242	0.471 (0.169)	0.540 (0.194)	0.958 (0.344)
09YANA_IRR	0.013688	0.023822	36.49%	0.791	0.427	0.535 (0.182)	0.602 (0.205)	0.819 (0.279)
09YANA_RFD	0.006774	0.039412	14.67%	0.891	0.260	0.445 (0.185)	0.471 (0.196)	0.873 (0.363)
10COLE_IRR	0.018323	0.021732	45.74%	0.722	0.392	0.474 (0.192)	0.558 (0.226)	0.757 (0.306)
10COLE_RFD	0.014472	0.014562	49.84%	0.777	0.552	0.466 (0.179)	0.529 (0.204)	0.627 (0.241)
10YANA_IRR	0.012508	0.042923	22.57%	0.862	0.388	0.520 (0.196)	0.560 (0.211)	0.835 (0.315)
10YANA_RFD	0.009967	0.020868	32.32%	0.806	0.400	0.481 (0.196)	0.536 (0.218)	0.760 (0.310)

genomic prediction. The proportion of within-experiment total genotypic variance that was due to additive polygenic variance in this model ranged from 15 to 50% across experiments (Table 2). In the selected model, environments clustered into two distinct environmental groups based on the correlations between total genetic effects (GV) across experiments (Figure 1). The groups consisted of a well-watered environment cluster, including the 2009 irrigated experiments and all 2010 experiments, and a water deficit environment represented by the two 2009 rainfed experiments. This corresponded with environmental conditions encountered in this study. In 2009, both sites experienced below average rainfall, and warmer air temperatures, particularly during the later stages of grain-filling. Weather conditions throughout 2010 were cooler than average, with substantially above average rainfall during late spring (coinciding with the crop grain-filling period). Thus the non-irrigated experiments in 2010 did not suffer water deficit.

Averaged across experiments in each environment cluster, the proportion of total genetic variance that was additive was 36% for the well-watered environment cluster, and 23% for the water deficit environment cluster (Table 2). Very similar patterns of relationships and clustering among environments were observed based on the correlations of additive-only genetic effects between environments (Figure 1), although the 09YANA_RFD experiment was less distinct from the well-watered cluster in this case than when the total genotypic correlations were considered. Correlations between the residual genetic effects of different environments also revealed a similar pattern, but with a slightly weaker correlation between 09COLE_IRR and the other well-watered environment experiments (Figure 1). The 09COLE_IRR experiment was the only experiment to be grown on raised beds with all other irrigated experiments utilizing a flat field layout.

Cross-validation of genomic estimated breeding values

The predictive ability (Sallam *et al.* 2015) of experiment-specific GEBVs was measured as the average correlation between experiment-specific GEBVs in the test sets and experiment-specific GVs estimated for the same lines when all trait data are used. Within-experiment prediction abilities ranged from $r = 0.474$ to 0.535 for the well-watered experiments and from $r = 0.445$ to $r = 0.481$ for the water deficit experiments (Table 2 and diagonal elements in Figure 2). The predictive ability across experiments within the well-watered environment cluster averaged $r = 0.502$, and across experiments within the water deficit cluster was $r = 0.455$ (Table 2 and diagonal elements of Figure 2). In contrast, GEBVs specific to an experiment within one cluster had much poorer predictive ability of GVs in the other cluster. We also estimated the ability of test set GEBVs specific to one experiment or cluster to predict GVs based on complete phenotype data in a different experiment or cluster (the off-diagonal elements of Figure 2). GEBVs for water-deficit experiments had only a weak correlation with the GVs in the well-watered environments (average $r = 0.196$). Similarly, GEBVs for the well-watered environments had an average correlation of $r = 0.211$ with GVs in the water deficit environments. Average GEBVs across all experiments also had low correlation with GVs within the water deficit environments ($r = 0.287$). Within each environment cluster, the GEBVs for individual experiments had moderate correlations with GVs at other experiments ($r = 0.442$ to $r = 0.536$ for the well-watered experiments, and $r = 0.400$ to $r = 0.447$ for the water deficit experiments).

The relative accuracy statistics were computed with both broad-sense and narrow-sense heritabilities as RAPV and RABV, respectively, so that a measure of accuracy relative to phenotypic selection can be compared to the accuracy relative to breeding value prediction (Table 2). As the broad-sense heritability estimates were much higher than narrow-sense, the

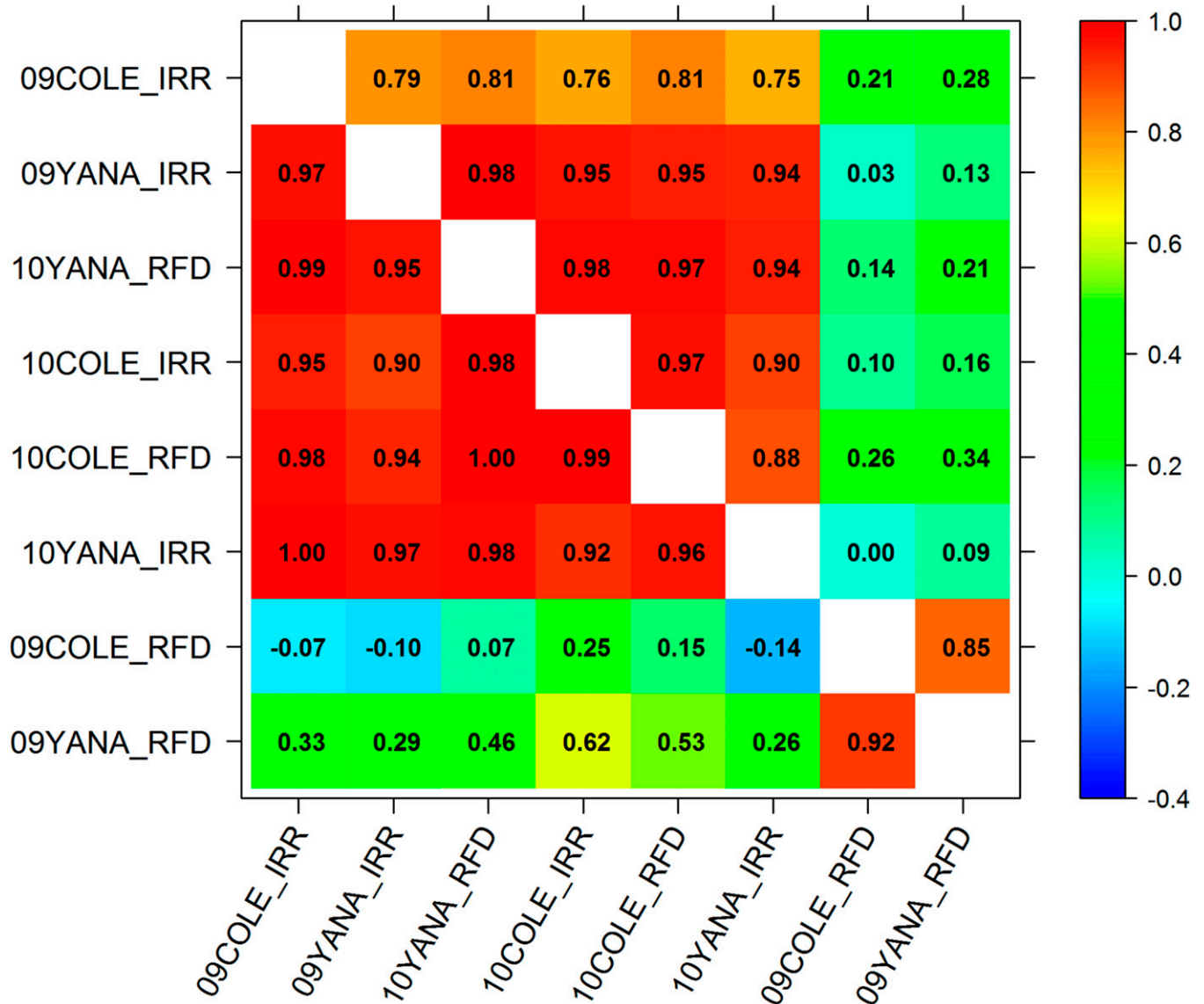


Figure 1 Correlations between total additive and residual genetic GV values in different experiments based on the full data set (above the diagonal) and correlations between additive GEBVs in different experiments, based on the full data set (below the diagonal). Experiment codes are given as year-site-irrigation treatment.

relative accuracy statistics indicate that the efficiency of genotype value prediction was poor (RAPV ranged from 0.471 to 0.602 at individual experiments, with a mean of 0.544, Table 2), whereas the relative accuracy of breeding value prediction with marker information only compared to complete phenotype data were much higher (RABV ranged from 0.627 to 0.958 at individual experiments, with a mean of 0.797, Table 2). The RABV for the water deficit experiment cluster was 1.07, a surprising but valid result analogous to relative efficiency of selection on a correlated trait, which can be >1 relative to direct selection.

DISCUSSION

Integration of genomic selection with genotype \times environment effects

Our approach in this study combines the concept of a Target Population of Environments (TPE) in regard to selection (Comstock 1977; Cooper *et al.* 1997; Basford and Cooper 1998) with genomic prediction methods, and

extends the work of Oakey *et al.* (2016) to a large field-based multi-environment trial. The TPE concept applies equally to genomic selection as phenotypic selection in the face of genotype-by-environment interactions. Our results show that predictive ability depends on the extent and nature of the genetic correlation between the training and the validation populations. Heslot *et al.* (2013) also demonstrated empirically in wheat that $G \times E$ patterns impact genomic selection in the same way that they impact phenotypic selection. Heslot *et al.* (2013) also observed that the main driver of prediction accuracy between environments were $G \times E$ effects and that genomic predictions are specific to the TPE they are predicted for – conclusions also supported by Lado *et al.* (2013) and Oakey *et al.* (2016), who also showed that multi-year models could give better prediction accuracy when environmental influence is large.

Genomic prediction models based on multi-environment trials may involve considerable complexity to allow heterogeneity of genetic correlations and genetic variances among environments, variable sources of extraneous non-genetic variation among environments, in

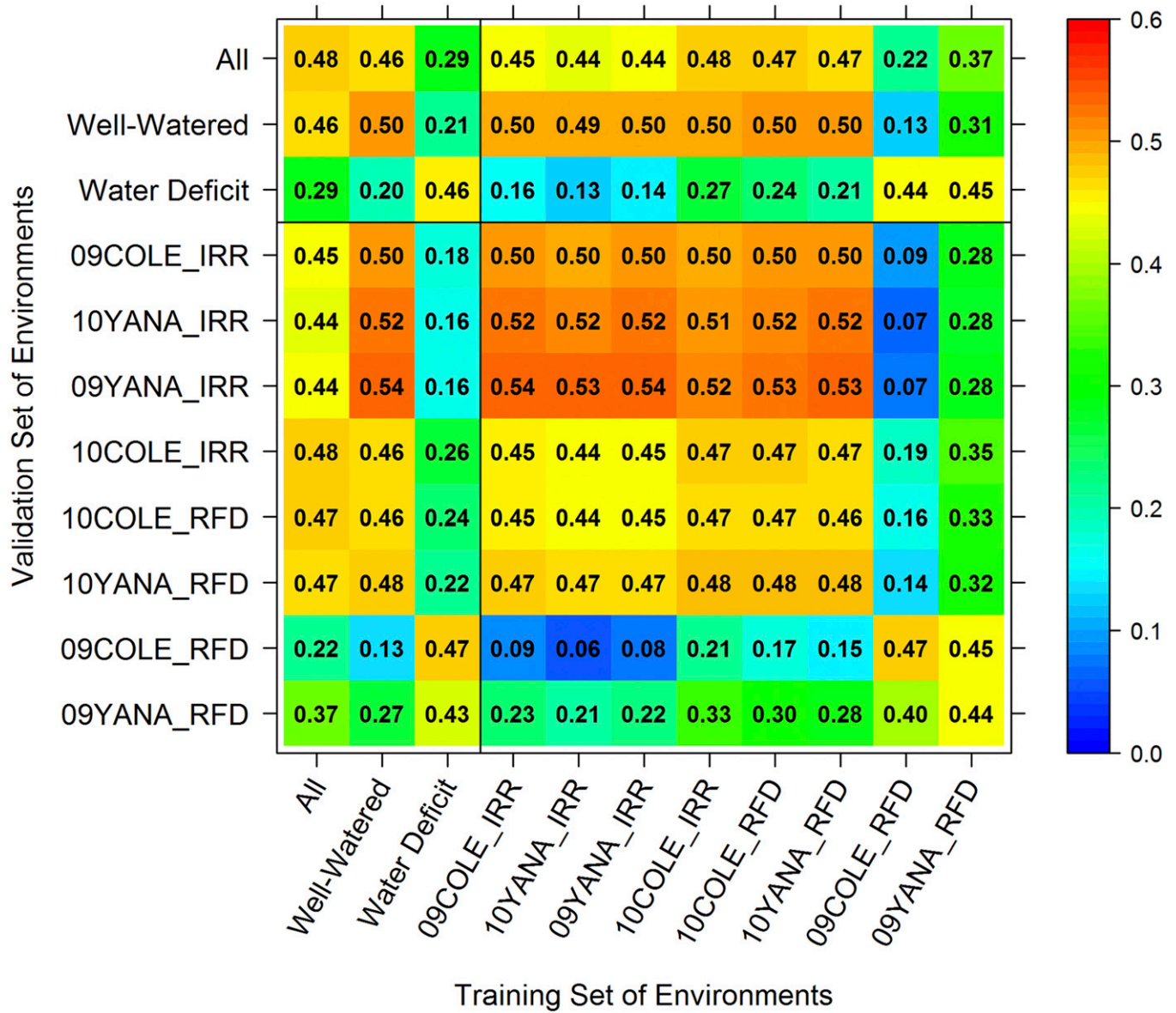


Figure 2 Predictive ability of GEBVs average across all experiments, averaged across experiments within each environment cluster (well-watered or water-deficit), or predicted for each specific experiment. The training set of environments is given by the X axis, and the validation set of environments is given by the Y axis. Experiment codes are given as year-site-irrigation treatment. Diagonal values represent ability of GEBVs within a given environment to predict GV in the same environment. Off-diagonal values represent the ability of GEBVs in a given environment to predict GV in a different environment.

addition to the high dimensionality of marker data. Researchers are faced with making choices about modeling greater complexity in the patterns of genetic correlations among environments vs. greater complexity in the genetic architecture modeled by marker data, based on tradeoffs between capturing more signal vs. overfitting and increasing computational demands as model complexity increases. A variety of modeling approaches have been proposed, reflecting different choices about which aspects of model complexity to emphasize. For example, Sallam *et al.* (2015) utilized an across-experiments model for TPE identified as having low $G \times E$ so the term in the model was minimized, whereas Crossa *et al.* (2010) modeled genetic effects within each environment separately. Lopez-Cruz *et al.* (2015) introduced a model with common $G \times E$ variance for all sets, but allowed variable marker effects through a Bayesian model. Heslot *et al.* (2014) and Jarquín *et al.* (2014)

introduced models that accounted for marker interactions with specific climate variables, adding another layer of complexity to the modeling of $G \times E$ patterns.

Our focus in this study was to emphasize the modeling of complex patterns of heterogeneity in the genetic variation expressed within environments and the pairwise genetic correlations between environments, along with extraneous non-genetic effects. We chose a FA model (Cullis *et al.* 2010; Smith *et al.* 2001) that can capture such heterogeneity more parsimoniously than fully unstructured models. The patterns of genetic correlations between environments observed in training data sets and patterns of the accuracy of GEBVs for one site to predict genetic values at other sites in test sets from the FA model provide a way of characterizing target environments, still following the ideas on exploiting $G \times E$ to make genetic gains outlined by Byth

(1981). The environment clusters in this study show that year effects are more important than location for WSCC (Figure 1), and this observation is borne out by other $G \times E$ studies with similar findings, particularly for traits where expression is significantly affected by seasonal conditions (Smith *et al.* 2015; Chenu *et al.* 2011; Milgate *et al.* 2015).

Using factor analytic models for genomic predictions

Heslot *et al.* (2014) demonstrate a method to predict genotype performance in untested environments based on climatic variables. In contrast, the FA model approach helps breeders to understand the groupings of environments within the TPE based on their genetic correlations, and breeders can predict average performance in subsets of these environments. Both Burgueño *et al.* (2012) and Rutkoski *et al.* (2015) utilized similar FA models for genomic prediction, however this study shows that the power of an FA model lies in the ability to identify environment subsets based on the loadings and predict genotype performance more accurately within those TPE rather than the across-experiments average. Since we predicted values for untested lines within tested environments, our estimated prediction abilities are biased upward compared to prediction of untested lines within untested environments. Our results, which reflect upper bounds on prediction abilities across environments, demonstrate that GEBVs have reasonable accuracy only within a clearly defined cluster of environments, and have substantially worse prediction accuracy of performance in environments outside of the group of related environments for which they were predicted. The FA model permits prediction of genetic values for untested lines within each site based on the genomic relationships between the untested and tested lines and also based on the genetic correlations observed among the tested set of environments. The FA model can improve prediction accuracy within a single site over what is possible from a single-environment analysis by information sharing among environments with high genetic correlations (Guo *et al.* 2013; Cullis *et al.* 2010; Kelly *et al.* 2007). In this study, GEBVs averaged across all environments were less accurate for environment-specific prediction than GEBVs averaged across subsets of sites within the same environmental cluster (Figure 2).

A drawback to the FA model is that the emphasis on including more complexity in the modeling of extraneous genetic variation and the heterogeneity of genetic covariation among environments may limit the complexity of genetic architecture models that can be tested. The total genotypic value of each line was modeled as the sum of a polygenic additive effect and a residual genetic effect (after Oakey *et al.* 2016). Our results demonstrated that the residual genetic effects were more important than the additive effects in this case, which greatly limits the effectiveness of prediction of untested varieties, as those predictions depend only on the additive effects. The residual genetic effects were modeled as independent among varieties, such that they cannot contribute to prediction of untested varieties. Alternative strategies include explicitly modeling epistatic genomic relationships; we attempted this but could not achieve model convergence. Oakey *et al.* (2016) discuss the implications of including additional relationship matrices to account for a proportion of non-additive genetic effects, however they also note the difficulty of fitting several relationship matrices to a MET. Bayesian models could be fitted to the genomic relationships, to capture residual genetic relationships and variation among marker effects, but this approach would be even more computationally difficult; future research could focus on integrating heterogeneity of genetic covariances among sites into such models.

Practical application of genomic selection for water soluble carbohydrate improvement

Breeders focusing on varietal development for target populations of environments that may experience abiotic stresses often seek to incorporate resistance to abiotic stresses, sometimes to contribute to grain yield *per se*,

but also to select for stable grain yield performance in the presence of variable levels of abiotic stress. Thus, abiotic stress resistance traits can be a selection target in their own right, especially in situations where $G \times E$ variation is driven by abiotic stresses and resistance to these stresses can ensure grain yield stability across years. Genomic selection may be a good way to select for abiotic stress traits such as WSCC, especially as a substitute for resource-intensive phenotypic selection. Levels of WSCC increase and decrease with crop development and growing conditions, meaning that full expression of the phenotype is not easily captured, and both wet chemistry and NIRS methods for measuring water soluble carbohydrates are time consuming and expensive (Gebbing and Schnyder 1999; Ruuska *et al.* 2006).

We estimated that most of the genotypic variance for WSCC did not fit a polygenic additive polygenic model, a surprising result considering numerous results showing that wheat grain yield is often adequately described with an additive genetic model (Heslot *et al.* 2012; Pérez-Rodríguez *et al.* 2012; Burgueño *et al.* 2012). However, other studies of wheat yield across diverse environments indicate that residual genetic variation may explain substantial fraction of the total genetic variation (Cuevas *et al.* 2017). Because of the prevalence of residual genetic variance, the GEBVs for WSCC were not very accurate estimates of total genotypic value, regardless of the information used to compute them (markers, phenotype records, or both). In this study, residual genetic effects were modeled by fitting a separate independent random term for residual genetic effects in addition to the additive effects whose covariance is proportional to the additive realized relationship matrix (after Oakey *et al.* 2016). Total genotypic prediction, which includes additive and residual genetic effects, is optimal for identifying the best available lines in a population, but this requires phenotypic records on each line to be predicted. Potentially, predictions can be improved by explicit modeling of non-additive as well as additive genetic relationships, with either parametric models including dominance and epistasis (Muñoz *et al.* 2014; Su *et al.* 2012; Da *et al.* 2014) or non-linear kernel methods (Gianola and van Kaam 2008). These models would allow total genotypic values to be predicted on lines in the absence of any phenotypic records.

If the total genotypic values are not closely correlated with the true breeding values, they will not be the best predictions to use for parental selection to generate a new population from intermating, as dominance, epistatic, and most other non-additive effects that may contribute to the residual genetic effects do not contribute to long-term genetic gain over meiotic outcrossing generations (Hill *et al.* 2008; Holland 2001; Oakey *et al.* 2016; Cockerham 1983; Cockerham and Matzinger 1985). GEBVs, on the other hand, while they may not be optimal for predicting the best existing lines, should be better estimators of the utility of lines as parents of new breeding populations. To distinguish the different uses of GEBVs in a breeding program, we computed two relative accuracy statistics that refer to the use of GEBVs to identify optimal extant lines or to identify lines with better breeding values for use as parents for intermating to create a new breeding population. Compared to phenotypic selection the RAPV for the well-watered environment cluster was 57%, and 51% for the water deficit environment cluster, indicating our GEBVs are marginal at predicting phenotypic performance in these environments for new genotypes (Table 2), and that genomic selection is unlikely to be a substitute for phenotypic selection of the best performing lines. In contrast, when compared with true breeding values for the purpose of selecting new parents and maximizing genetic gain over time, the relative accuracy (RABV) was higher. The RABV measures can be coupled with the assumptions of Desta and Ortiz (2014); Heffner *et al.* (2010); and Lorenz *et al.* (2011) suggesting genomic selection is predicted to facilitate a reduction in the breeding

cycle time of the average wheat breeding program of at least half due to the ability to select at earlier generations and also in off-season generations. Therefore, relative to selection using the true breeding values, and making the assumption of a decrease in the breeding cycle time of 50%, genetic gain per unit of time from genomic selection in this study is estimated to be $2 \times 78\% = 158\%$ for the well-watered environment cluster and $2 \times 107\% = 214\%$ for the water deficit environment cluster. This indicates that genomic selection methods can increase the rate of genetic gain for WSCC. The challenge in practical terms, as noted by Cooper *et al.* (2014), may be in implementing the enabling technologies needed to make genomic selection work and readily integrated into a commercial breeding program. Finally, if the ultimate goal of selection is to improve yield performance within or across abiotic stress environments, the genotypic and additive genetic correlations of WSCC with yield and other agronomic traits must be considered. Ovenden *et al.* (2017) previously demonstrated significant but relatively low genotypic correlations between WSCC and yield in this germplasm sample, suggesting that its utility as an indirect selection criterion needs to be assessed on a case by case basis.

Conclusions

This study provides empirical evidence that genomic selection methods could improve the rate of genetic gain for carbohydrate accumulation, provided that the TPE are carefully characterized and understood, and predictions are restricted to environment subsets of interest.

With additive variance being low in the models under study, the importance of relative accuracy for different genomic selection strategies becomes apparent. Compared to phenotypic selection, the relative accuracy of the GEBVs modeled here is low, however the GEBVs more accurately model the true breeding values. Therefore, genomic selection could well result in higher genetic gain per unit of time through the ability to better select parents for intermating, and genomic selection may be a useful tool for making genetic gains in complex abiotic stress characteristics.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Grains Research and Development Corporation (GRDC) of Australia funding for this project (ICF00007) and a Grains Industry Research Scholarship for Ben Ovenden. The American Australian Association is gratefully acknowledged for the ConocoPhillips Education Fellowship awarded to Ben Ovenden. The authors thank Kerry Schirmer and Aaron Hutchison for their expert technical contributions and data collection. The authors also thank the reviewers for their comments which improved the manuscript. The breeding lines used in this study were kindly provided by Australian Grains Technologies, InterGrain, HRZ Wheats (now Dow Seeds), Sydney University, LongReach Plant Breeders and CIMMYT.

LITERATURE CITED

- Akaike, H., 1974 A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19: 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Araus, J., G. Slafer, M. Reynolds, and C. Royo, 2002 Plant breeding and drought in C3 cereals: what should we breed for? *Ann. Bot. (Lond.)* 89: 925–940. <https://doi.org/10.1093/aob/mcf049>
- Asseng, S., and A. F. van Herwaarden, 2003 Analysis of the benefits to wheat yield from assimilates stored prior to grain filling in a range of environments. *Plant Soil* 256: 217–229. <https://doi.org/10.1023/A:1026231904221>
- Basford, K. E., and M. Cooper, 1998 Genotype×environment interactions and some considerations of their implications for wheat breeding in Australia. *Aust. J. Agric. Res.* 49: 153–174. <https://doi.org/10.1071/A97035>
- Beeck, C. P., W. A. Cowling, A. B. Smith, and B. R. Cullis, 2010 Analysis of yield and oil from a series of canola breeding trials. Part I. Fitting factor analytic mixed models with pedigree information. *Genome* 53: 992–1001. <https://doi.org/10.1139/G10-051>
- Bennett, D., A. IZANLOO, M. Reynolds, H. Kuchel, P. Langridge *et al.*, 2012 Genetic dissection of grain yield and physical grain quality in bread wheat (*Triticum aestivum* L.) under water-limited environments. *Theor. Appl. Genet.* 125: 255–271. <https://doi.org/10.1007/s00122-012-1831-9>
- Bernier, J., G. N. Atlin, R. Serraj, A. Kumar, and D. Spaner, 2008 Breeding upland rice for drought resistance. *J. Sci. Food Agric.* 88: 927–939. <https://doi.org/10.1002/jsfa.3153>
- Bidinger, F., R. B. Musgrave, and R. A. Fischer, 1977 Contribution of stored pre-anthesis assimilate to grain yield in wheat and barley. *Nature* 270: 431–433. <https://doi.org/10.1038/270431a0>
- Blum, A., 1998 Improving wheat grain filling under stress by stem reserve mobilisation. *Euphytica* 100: 77–83. <https://doi.org/10.1023/A:1018303922482>
- Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84: 210–223. <https://doi.org/10.1016/j.ajhg.2009.01.005>
- Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa, 2012 Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Sci.* 52: 707–719. <https://doi.org/10.2135/cropsci2011.06.0299>
- Butler, D., B. R. Cullis, A. R. Gilmour, and B. J. Gogel, 2009 ASReml-R reference manual, Queensland Department of Primary Industries and Fisheries, Brisbane.
- Byth, D. E., 1981 A conceptual basis of genotypic × environment interactions for plant improvement, *Refresher training course notes on interpretation of plant response and adaptation to agricultural environments*, edited by Byth, D. E., and V. E. Mungomery. University of Queensland, University of Queensland, Brisbane, Queensland.
- Cavanagh, C. R., S. Chao, S. Wang, B. E. Huang, S. Stephen *et al.*, 2013 Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. USA* 110: 8057–8062. <https://doi.org/10.1073/pnas.1217133110>
- Chenu, K., M. Cooper, G. L. Hammer, K. L. Mathews, M. F. Dreccer *et al.*, 2011 Environment characterization as an aid to wheat improvement: interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. *J. Exp. Bot.* 62: 1743–1755. <https://doi.org/10.1093/jxb/erq459>
- Cockerham, C. C., 1983 Covariances of Relatives from Self-Fertilization. *Crop Sci.* 23: 1177–1180. <https://doi.org/10.2135/cropsci1983.0011183X002300060035x>
- Cockerham, C. C., and D. F. Matzinger, 1985 Selection Response Based on Selfed Progenies. *Crop Sci.* 25: 483–488. <https://doi.org/10.2135/cropsci1985.0011183X002500030012x>
- Comstock, R. E., 1977 Quantitative genetics and the design of breeding programs in *Proceedings of the International Conference on Quantitative Genetics*, edited by Pollak, E., O. Kempthorne, and T.B. Bailey. Iowa State University Press, Ames, Iowa.
- Coombes, N. E., 2002 The reactive tabu search for efficient correlated experimental designs, Liverpool John Moores University, Liverpool, UK.
- Cooper, M., C. D. Messina, D. Podlich, L. R. Totir, A. Baumgarten *et al.*, 2014 Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop Pasture Sci.* 65: 311–336. <https://doi.org/10.1071/CP14007>
- Cooper, M., R. E. Stucker, I. H. DeLacy, and B. D. Harch, 1997 Wheat Breeding nurseries, target environments, and indirect selection for grain yield. *Crop Sci.* 37: 1168–1176. <https://doi.org/10.2135/cropsci1997.0011183X003700040024x>
- Crossa, J., G. de los Campos, P. Pérez, D. Gianola, J. Burgueño *et al.*, 2010 Prediction of genetic values of quantitative traits in plant breeding

- using pedigree and molecular markers. *Genetics* 186: 713–724. <https://doi.org/10.1534/genetics.110.118521>
- Crossa, J., P. Perez, J. Hickey, J. Burgueno, L. Ornella *et al.*, 2014 Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity* 112: 48–60. <https://doi.org/10.1038/hdy.2013.16>
- Cuevas, J., J. Crossa, O.A. Montesinos-López, J. Burgueño, P. Pérez-Rodríguez *et al.*, 2017 Bayesian Genomic Prediction with Genotype × Environment Interaction Kernel Models. *G3: Genes|Genomes|Genetics* 7(1): 41–53. <https://doi.org/10.1534/g3.116.035584>
- Cullis, B. R., A. B. Smith, C. P. Beeck, and W. A. Cowling, 2010 Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. *Genome* 53: 1002–1016. <https://doi.org/10.1139/G10-080>
- Cullis, B. R., A. B. Smith, and N. E. Coombes, 2006 On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11: 381–393. <https://doi.org/10.1198/108571106X154443>
- Da, Y., C. Wang, S. Wang, and G. Hu, 2014 Mixed Model Methods for Genomic Prediction and Variance Component Estimation of Additive and Dominance Effects Using SNP Markers. *PLoS One* 9: e87666. <https://doi.org/10.1371/journal.pone.0087666>
- de los Campos, G., J. M. Hickey, R. Pong-Wong, H. D. Daetwyler, and M. P. L. Calus, 2013 Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 193: 327–345. <https://doi.org/10.1534/genetics.112.143313>
- Desta, Z. A., and R. Ortiz, 2014 Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19: 592–601. <https://doi.org/10.1016/j.tplants.2014.05.006>
- Drecker, M. F., S. C. Chapman, A. R. Rattey, J. Neal, Y. Song *et al.*, 2013 Developmental and growth controls of tillering and water-soluble carbohydrate accumulation in contrasting wheat (*Triticum aestivum* L.) genotypes: can we dissect them? *J. Exp. Bot.* 64: 143–160. <https://doi.org/10.1093/jxb/ers317>
- Ehdaie, B., G. A. Alloush, and J. G. Waines, 2008 Genotypic variation in linear rate of grain growth and contribution of stem reserves to grain yield in wheat. *Field Crops Res.* 106: 34–43. <https://doi.org/10.1016/j.fcr.2007.10.012>
- Endelman, J.B., and J.-L. Jannink, 2012 Shrinkage estimation of the realized relationship matrix. *G3: Genes|Genomes|Genetics* 2(11): 1405–1413. <https://doi.org/10.1534/g3.112.004259>
- Fischer, R. A., D. R. Byerlee, and G. O. Edmeades, 2014 ACIAR Monograph No. 158, *Crop yields and global food security*, Australian Centre for International Agricultural Research, Canberra, ACT.
- Foulkes, M. J., and M. P. Reynolds, 2015 Breeding challenge: improving yield potential, pp. 397–421 in *Crop Physiology*, Ed. 2nd. chap 16., edited by V. O. Calderini, V. O. S. F. Academic Press, San Diego. <https://doi.org/10.1016/B978-0-12-417104-6.00016-9>
- Gebbing, T., 2003 The enclosed and exposed part of the peduncle of wheat (*Triticum aestivum*): spatial separation of fructan storage. *New Phytol.* 159: 245–252. <https://doi.org/10.1046/j.1469-8137.2003.00799.x>
- Gebbing, T., and H. Schnyder, 1999 Pre-anthesis reserve utilization for protein and carbohydrate synthesis in grains of wheat. *Plant Physiol.* 121: 871–878. <https://doi.org/10.1104/pp.121.3.871>
- Gianola, D., and J. B. C. H. M. van Kaam, 2008 Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178: 2289–2303. <https://doi.org/10.1534/genetics.107.084285>
- Gilmour, A. R., B. R. Cullis, and A. P. Verbyla, 1997 Accounting for natural and extraneous variation in the analysis of field experiments. *J. Agric. Biol. Environ. Stat.* 2: 269–293. <https://doi.org/10.2307/1400446>
- Guo, Z., D.M. Tucker, D. Wang, C.J. Basten, E. Ersoz *et al.*, 2013 Accuracy of across-environment genome-wide prediction in maize nested association mapping populations. *G3: Genes|Genomes|Genetics* 3(2): 263–272. <https://doi.org/10.1534/g3.112.005066>
- Habier, D., R. L. Fernando, and D. J. Garrick, 2013 Genomic BLUP decoded: A look into the black box of genomic prediction. *Genetics* 194: 597–607. <https://doi.org/10.1534/genetics.113.152207>
- Heffner, E. L., A. J. Lorenz, J.-L. Jannink, and M. E. Sorrells, 2010 Plant breeding with genomic selection: Gain per unit time and cost. *Crop Sci.* 50: 1681–1690. <https://doi.org/10.2135/cropsci2009.11.0662>
- Heffner, E. L., M. E. Sorrells, and J.-L. Jannink, 2009 Genomic selection for crop improvement. *Crop Sci.* 49: 1–12. <https://doi.org/10.2135/cropsci2008.08.0512>
- Heslot, N., D. Akdemir, M. Sorrells, and J.-L. Jannink, 2014 Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127: 463–480. <https://doi.org/10.1007/s00122-013-2231-5>
- Heslot, N., J.-L. Jannink, and M. E. Sorrells, 2013 Using genomic prediction to characterize environments and optimize prediction accuracy in applied breeding data. *Crop Sci.* 53: 921–933. <https://doi.org/10.2135/cropsci2012.07.0420>
- Heslot, N., H.-P. Yang, M. E. Sorrells, and J.-L. Jannink, 2012 Genomic selection in plant breeding: a comparison of models. *Crop Sci.* 52: 146–160. <https://doi.org/10.2135/cropsci2011.06.0297>
- Hill, W. G., 2012 Quantitative genetics in the genomics era. *Curr. Genomics* 13: 196–206. <https://doi.org/10.2174/138920212800543110>
- Hill, W. G., M. E. Goddard, and P. M. Visscher, 2008 Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 4: e1000008. <https://doi.org/10.1371/journal.pgen.1000008>
- Holland, J. B., 2001 Epistasis and plant breeding, pp. 27–92 in *Plant breeding reviews*, edited by J. Janick, John Wiley Son, New York.
- Isik, F., J. Holland, and C. Maltecca, 2017 *Genetic Data Analysis for Plant and Animal Breeding*, Springer International, Berlin. <https://doi.org/10.1007/978-3-319-55177-7>
- Jarquín, D., J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt *et al.*, 2014 A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127: 595–607. <https://doi.org/10.1007/s00122-013-2243-1>
- Kamran, A., M. Iqbal, and D. Spaner, 2014 Flowering time in wheat (*Triticum aestivum* L.): a key factor for global adaptability. *Euphytica* 197: 1–26. <https://doi.org/10.1007/s10681-014-1075-7>
- Kelly, A. M., A. B. Smith, J. A. Eccleston, and B. R. Cullis, 2007 The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Sci.* 47: 1063–1070. <https://doi.org/10.2135/cropsci2006.08.0540>
- Lado, B., I. Matus, A. Rodríguez, L. Inostroza, J. Poland *et al.*, 2013 Increased Genomic Prediction Accuracy in Wheat Breeding Through Spatial Adjustment of Field Trial Data. *G3: Genes|Genomes|Genetics* 3(12): 2105–2114. <https://doi.org/10.1534/g3.113.007807>
- Legarra, A., C. Robert-Granié, E. Manfredi, and J.-M. Elsen, 2008 Performance of genomic selection in mice. *Genetics* 180: 611–618. <https://doi.org/10.1534/genetics.108.088575>
- Leplat, F., J. Jensen, and P. Madsen, 2016 Genomic Prediction of Manganese Efficiency in Winter Barley. *Plant Genome* 9(1): 1–13. <https://doi.org/10.3835/plantgenome2015.09.0085>
- Lin, Z., B. J. Hayes, and H. D. Daetwyler, 2014 Genomic selection in crops, trees and forages: a review. *Crop Pasture Sci.* 65: 1177–1191. <https://doi.org/10.1071/CP13363>
- Lopes, M. S., J. L. Araus, P. D. R. van Heerden, and C. H. Foyer, 2011 Enhancing drought tolerance in C4 crops. *J. Exp. Bot.* 62: 3135–3153. <https://doi.org/10.1093/jxb/err105>
- Lopez-Cruz, M., J. Crossa, D. Bonnett, S. Dreisigacker, J. Poland *et al.*, 2015 Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3: Genes|Genomes|Genetics* 5(4): 569–582. <https://doi.org/10.1534/g3.114.016097>
- Lorenz, A. J., S. Chao, F. G. Asoro, E. L. Heffner, T. Hayashi *et al.*, 2011 Genomic selection in plant breeding: Knowledge and prospects, pp. 77–123 in *Advances in agronomy*, edited by Sparks, D. L., Academic Press, Waltham, Massachusetts. <https://doi.org/10.1016/B978-0-12-385531-2.00002-5>
- 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: 1819–1829.

- Milgate, A., B. Owenden, D. Adorada, C. Lisle, J. Lacy *et al.*, 2015 Genetic improvement of triticale for irrigated systems in south-eastern Australia: a study of genotype and genotype \times environment interactions. *Crop Pasture Sci.* 66: 782–792. <https://doi.org/10.1071/CP14357>
- Muñoz, P. R., M. F. R. Resende, S. A. Gezan, M. D. V. Resende, G. de los Campos *et al.*, 2014 Unraveling additive from nonadditive effects using genomic relationship matrices. *Genetics* 198: 1759–1768. <https://doi.org/10.1534/genetics.114.171322>
- Nakaya, A., and S. N. Isobe, 2012 Will genomic selection be a practical method for plant breeding? *Ann. Bot. (Lond.)* 110: 1303–1316. <https://doi.org/10.1093/aob/mcs109>
- Oakey, H., B. Cullis, R. Thompson, J. Comadran, C. Halpin *et al.*, 2016 Genomic Selection in Multi-environment Crop Trials. *G3: Genes|Genomes|Genetics* 6(5): 1313–1326. <https://doi.org/10.1534/g3.116.027524>
- Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel, 2006 Joint modeling of additive and non-additive genetic line effects in single field trials. *Theor. Appl. Genet.* 113: 809–819. <https://doi.org/10.1007/s00122-006-0333-z>
- Ogut, F., Y. Bian, P. J. Bradbury, and J. B. Holland, 2015 Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. *Heredity* 114: 552–563. <https://doi.org/10.1038/hdy.2014.123>
- Owenden, B., A. Milgate, C. Lisle, L. J. Wade, G. J. Rebetzke *et al.*, 2017 Selection for water-soluble carbohydrate accumulation and investigation of genetic \times environment interactions in an elite wheat breeding population. *Theor. Appl. Genet.* 130: 2445–2461. <https://doi.org/10.1007/s00122-017-2969-2>
- Passioura, J. B., 1996 Drought and drought tolerance. *Plant Growth Regul.* 20: 79–83. <https://doi.org/10.1007/BF00024003>
- Pérez-Rodríguez, P., D. Gianola, J.M. González-Camacho, J. Crossa, Y. Manes *et al.*, 2012 Comparison Between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. *G3: Genes|Genomes|Genetics* 2 (12):1595–1605. <https://doi.org/10.1534/g3.112.003665>
- Pheloung, P., and K. Siddique, 1991 Contribution of stem dry matter to grain yield in wheat cultivars. *Funct. Plant Biol.* 18: 53–64.
- Piaskowski, J. L., D. Brown, and K. Garland Campbell, 2016 Near-infrared calibration of soluble stem carbohydrates for predicting drought tolerance in spring wheat. *Agron. J.* 108: 285–293. <https://doi.org/10.2134/agronj2015.0173>
- Piltz, J., and D. Law, 2007 AFIA-Laboratory Methods Manual, Australian Fodder Industry Association Inc., Balwyn, Victoria.
- R Development Core Team, 2012 R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
- Ray, D. K., J. S. Gerber, G. K. MacDonald, and P. C. West, 2015 Climate variation explains a third of global crop yield variability. *Nat. Commun.* 6: 5989. <https://doi.org/10.1038/ncomms6989>
- Rebetzke, G. J., S. C. Chapman, C. L. McIntyre, R. A. Richards, A. G. Condon *et al.*, 2009 Grain Yield Improvement in Water-Limited Environments, pp. 215–249 in *Wheat: Science and Trade*, edited by Carver, B. F., Wiley-Blackwell, Ames, Iowa. <https://doi.org/10.1002/9780813818832.ch11>
- Rebetzke, G. J., A. F. van Herwaarden, C. Jenkins, M. Weiss, D. Lewis *et al.*, 2008 Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. *Aust. J. Agric. Res.* 59: 891–905. <https://doi.org/10.1071/AR08067>
- Reynolds, M. P., E. Quilligan, P. K. Aggarwal, K. C. Bansal, A. J. Cavalieri *et al.*, 2016 An integrated approach to maintaining cereal productivity under climate change. *Glob. Food Secur.* 8: 9–18. <https://doi.org/10.1016/j.gfs.2016.02.002>
- Richards, R. A., G. J. Rebetzke, M. Watt, A. G. Condon, W. Spielmeier *et al.*, 2010 Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Funct. Plant Biol.* 37: 85–97. <https://doi.org/10.1071/FP09219>
- Rutkoski, J., R. P. Singh, J. Huerta-Espino, S. Bhavani, J. Poland *et al.*, 2015 Efficient use of historical data for genomic selection: A case study of stem rust resistance in wheat. *Plant Genome* 8(1): 1–10. <https://doi.org/10.3835/plantgenome2014.09.0046>
- Ruuska, S. A., G. J. Rebetzke, A. F. van Herwaarden, R. A. Richards, N. A. Fettell *et al.*, 2006 Genotypic variation in water-soluble carbohydrate accumulation in wheat. *Funct. Plant Biol.* 33: 799–809. <https://doi.org/10.1071/FP06062>
- Sallam, A. H., J. B. Endelman, J.-L. Jannink, and K. P. Smith, 2015 Assessing genomic selection prediction accuracy in a dynamic barley breeding population. *Plant Genome* 8: 1–15. <https://doi.org/10.3835/plantgenome2014.05.0020>
- Schnyder, H., 1993 The role of carbohydrate storage and redistribution in the source-sink relations of wheat and barley during grain filling - a review. *New Phytol.* 123: 233–245. <https://doi.org/10.1111/j.1469-8137.1993.tb03731.x>
- Shearman, V. J., R. Sylvester-Bradley, R. K. Scott, and M. J. Foulkes, 2005 Physiological processes associated with wheat yield progress in the UK. *Crop Sci.* 45: 175–185.
- Slewinski, T. L., 2012 Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *J. Exp. Bot.* 63: 4647–4670. <https://doi.org/10.1093/jxb/ers124>
- Smith, A., B. Cullis, and R. Thompson, 2001 Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57: 1138–1147. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>
- Smith, A., A. Ganesalingam, H. Kuchel, and B. Cullis, 2015 Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theor. Appl. Genet.* 128: 55–72. <https://doi.org/10.1007/s00122-014-2412-x>
- Smith, A., P. Lim, and B. R. Cullis, 2006 The design and analysis of multi-phase plant breeding experiments. *J. Agric. Sci.* 144: 393. <https://doi.org/10.1017/S0021859606006319>
- Stram, D. O., and J. W. Lee, 1994 Variance components testing in the longitudinal mixed effects model. *Biometrics* 50: 1171–1177. <https://doi.org/10.2307/2533455>
- Su, G., O. F. Christensen, T. Ostensen, M. Henryon, and M. S. Lund, 2012 Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS One* 7: e45293. <https://doi.org/10.1371/journal.pone.0045293>
- Takahashi, T., P. M. Chevalier, and R. A. Rupp, 2001 Storage and remobilization of soluble carbohydrates after heading in different plant parts of a winter wheat cultivar. *Plant Prod. Sci.* 4: 160–165. <https://doi.org/10.1626/pps.4.160>
- van Herwaarden, A. F., J. F. Angus, R. A. Richards, and G. D. Farquhar, 1998 ‘Haying-off’, the negative grain yield response of dryland wheat to nitrogen fertiliser II. Carbohydrate and protein dynamics. *Aust. J. Agric. Res.* 49: 1083–1094. <https://doi.org/10.1071/A97040>
- Vivek, B. S., G. K. Krishna, V. Vengadessan, R. Babu, P. H. Zaidi *et al.*, 2016 Use of Genomic Estimated Breeding Values Results in Rapid Genetic Gains for Drought Tolerance in Maize. *Plant Genome* 10: 1–10. <https://doi.org/10.3835/plantgenome2016.07.0070>
- Wimmer, V., T. Albrecht, H.-J. Auinger, and C.-C. Schön, 2012 Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* 28: 2086–2087. <https://doi.org/10.1093/bioinformatics/bts335>
- Zadoks, J. C., T. T. Chang, and C. F. Konzak, 1974 A decimal code for the growth stages of cereals. *Weed Res.* 14: 415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x>
- Zila, C.T., L.F. Samayoa, R. Santiago, A. Butrón, and J.B. Holland, 2013 A genome-wide association study reveals genes associated with fusarium ear rot resistance in a maize core diversity panel. *G3: Genes|Genomes|Genetics* 3(11): 2095–2104. <https://doi.org/10.1534/g3.113.007328>

Communicating editor: E. Akhunov