ORIGINAL ARTICLE

Plant Breeding   WILEY

Hybrid breeding for biomass yield in winter triticale: II. Combining ability and hybrid prediction

Johannes Trini[1] | Hans Peter Maurer[1] | Sigrid Weissmann[2] | Tobias Würschum[1]

[1]State Plant Breeding Institute, University of Hohenheim, Stuttgart, Germany

[2]HegeSaat GmbH & Co.KG, Singen-Bohlingen, Germany

Correspondence
Hans Peter Maurer, State Plant Breeding Institute, University of Hohenheim, 70599 Stuttgart, Germany.
Email: h-p.maurer@uni-hohenheim.de

Funding information
Bundesministerium für Ernährung und Landwirtschaft, Grant/Award Number: FKZ 22015109 and FZK 22024515

Communicated by: Hermann Buerstmayr

Abstract

Accurate hybrid prediction and knowledge about the relative contribution of general (GCA) and specific combining ability (SCA) are of utmost importance for efficient hybrid breeding. We therefore evaluated 91 triticale single-cross hybrids in field trials at seven environments for plant height, heading time, fresh biomass, dry matter content and dry biomass. Fresh and dry biomass showed the highest proportion (23%) of variance due to SCA. Prediction accuracies based on GCA were slightly higher than based on mid-parent values. Utilizing parental kinship information yielded the highest prediction accuracies when both parental lines have been tested in other hybrid combinations, but still moderate-to-low prediction accuracies for two untested parents. Thus, hybrid prediction for biomass traits in triticale is currently promising based on mid-parent values as emphasized by our simulation study, but can be expected to shift to GCA-based prediction with an increasing importance of GCA due to selection in hybrid breeding. Moreover, the performance of potential hybrids between newly developed lines can be predicted with moderate accuracy using genomic relationship information.

KEYWORDS
biomass, genomic prediction, heterosis, hybrid prediction, prediction accuracy, simulation study, triticale

1 | INTRODUCTION

Owing to the increasing demand for renewable energy sources and a policy to promote them, the increase of plant biomass yield has become a breeding goal in crops suited for biomass production (Monforti, Bódis, Scarlat, & Dallemand, 2013; Wit & Faaij, 2010). A promising bioenergy crop is triticale (× *Triticosecale* Wittmack), as it possesses a high biomass and biogas yield potential compared to other small grain cereals (LTZ, 2013; Mergoum et al., 2009; Weiland, 2010). It could therefore become an attractive alternative or complement to the predominant, highly productive silage maize production

in Central Europe (Weiland, 2010). To date, triticale breeding programmes mainly focus on increasing grain yield, but also have a high potential to increase biomass yield due to the large genetic variation observed in European breeding material (Gowda et al., 2011).

Triticale is primarily improved by line breeding, but hybrid breeding has recently received increased attention to improve the trait performance of autogamous crops (Fischer et al., 2010; Gowda et al., 2013; Longin et al., 2012). Interestingly, Losert, Maurer, Weissmann, and Würschum (2016) reported a considerable commercial heterosis for biomass yield in triticale, indicating the potential economic value of hybrid breeding in this crop. Consequently, switching the variety type in triticale from line to hybrid cultivars appears worthwhile to investigate

further. A major bottleneck for hybrid breeding, however, is the fact that the number of possible hybrid combinations increases quadratically with the number of parental lines, thus resulting in huge numbers of possible hybrid combinations (Bernardo, 2010). Producing and testing all these combinations in the field is not feasible, and hence, approaches to preselect the most promising hybrid combinations are required.

Predicting the performance of single-cross hybrids based on mid-parent values is supposed to provide good results for qualitatively inherited and highly heritable traits as shown in previous studies (Boeven, Würschum, Weissmann, Miedaner, & Maurer, 2016; Gowda et al., 2013; Oettler, Tams, Utz, Bauer, & Melchinger, 2005). However, biomass yield is a highly quantitative trait (Alheit et al., 2014), and therefore, predictions based on mid-parent values are expected to be less powerful due to masking non-additive effects (Bernardo, 2010; Smith, 1986). Predictions of single-cross performance based on general combining ability (GCA) effects have shown promising results in triticale for complex traits such as grain yield or Fusarium head blight (Boeven et al., 2016; Fischer et al., 2010; Gowda et al., 2013). The ability to predict hybrid performance precisely by GCA effects strongly depends on the variance ratio between specific combining ability (SCA) effects and the total genotypic variance (Hallauer, Carena, & Miranda Filho, 2010; Melchinger, 1999). Previous results in triticale showed that this ratio strongly depended on the investigated trait (Boeven et al., 2016; Gowda et al., 2013; Oettler et al., 2005).

A limitation of the GCA-based hybrid prediction is that it is only possible when GCA estimates are available, which requires the production of hybrids and subsequent time- and resource-intensive field trials. Predicting hybrid performance from parental lines without their prior evaluation is possible, when pedigree or genomic relationship information is exploited. This can be used to preselect single-crosses to be tested in the field (Bernardo, 2010) and has previously been shown to result in promising prediction accuracies in different crops, including maize (Bernardo, 1994, 1995, 1996a, 1996b; Charcosset et al., 1998), sunflower (Reif et al., 2013) and triticale (Boeven et al., 2016; Gowda et al., 2013).

The aim of this study was to assess the potential of hybrid prediction for biomass yield in triticale. To this end, a set of 91 single-cross hybrids, derived from 33 parental inbred lines crossed in an incomplete factorial mating design, were evaluated in multi-location field trials. In particular, the goals of our study were to (a) evaluate the variances due to GCA and SCA effects, (b) compare different approaches to predict single-cross hybrid performance and (c) draw conclusions for triticale hybrid breeding.

2 | MATERIALS AND METHODS

2.1 | Phenotypic data

This study was based on a total of 178 winter triticale genotypes, comprising 23 female parental lines, 10 male parental lines, 91 single-cross hybrids and 54 check genotypes as described previously (Losert et al., 2016). The single-cross hybrids were derived by crosses between female and male parental lines in an unbalanced and incomplete factorial mating design (Table S1) using a cytoplasmic male sterility (CMS) inducing cytoplasm based on *Triticum timopheevii* Zhuk. The check genotypes comprised 8 triticale cultivars officially registered in Central Europe ('Agostino', 'Amarillo 105', 'Balu PZO', 'Cosinus', 'Grenado', 'SW Talentro', 'Tarzan' and 'Tulus') as well as 46 advanced breeding lines.

The field experiment was carried out in the growing seasons 2011/12 and 2012/13 at five locations: Bohlingen (BOH, 47°43′12″N, 8°53′46″E, 420 metres above sea level, masl), Eckartsweier (EWE, 48°31′18″N, 7°52′18″E, 140 masl), Hohebuch (HET, 49°12′29″N, 9°39′29″E, 360 masl), Hohenheim (HOH, 48°28′49″N, 9°11′16″E, 400 masl) and Oberer Lindenhof (OLI, 48°28′49″N, 9°18′56″E, 700 masl). The locations Hohebuch and Eckartsweier were only used for field trials during the growing season 2011/12.

Due to limited seed availability, not all single-cross hybrids were cultivated in both growing seasons (Table S1), and in addition, a partially replicated field design was chosen (Williams, Piepho, & Whitaker, 2011) with an average replication number across locations of 1.5 for females, hybrids and checks, and 1.6 for males. Plot sizes ranged from 6.0 to 9.0 m^2, and sowing density was 280 viable seeds/m^2. The following traits were evaluated in these trials: heading time (HT, EC stage according to Zadoks, Chang, & Konzak, 1974), plant height (PH, cm), fresh biomass (FBM, Mg/ha), dry matter content (DMC, %) and dry biomass (DBM, Mg/ha).

2.2 | Molecular data analysis

DNA of all parental lines was extracted from the leaves of young plants using a modified CTAB protocol (Doyle & Doyle, 1990) and genotyped by genotyping-by-sequencing by Diversity Arrays Technology, Canberra, Australia (www.diversityarrays.com). Markers showing a minor allele frequency of 5% or less, or more than 20% missing values were removed, resulting in a set of 56,722 dominant DArT markers used for further analyses.

A principal coordinate analysis (PCoA) based on Rogers' distances between pairs of inbred lines (Rogers, 1972) was performed according to Gower (1966). Elements of the additive relationship matrices of the female (A_F) and male (A_M) parental lines as well as the dominance relationship matrix D were calculated based on (a) the suggestions of Bernardo (1993) and (b) VanRaden (2008) method two. However, estimates exploiting relationship information according to VanRaden (2008) showed only minor differences compared to Bernardo (1993) and are therefore not reported further. Bernardo (1993) calculates the coefficients of coancestry θ_{ij} used for A_F, A_M and D between inbred lines i and j based on molecular marker data as:

$$\theta_{ij} = 1 + \frac{(S_{ij} - 1)}{1 - T}$$

where S_{ij} is the proportion of marker loci with shared variants between inbred lines i and j. The measure of T denotes the average probability that a variant from a parent of inbred line i and a variant

from a different inbred line *j* are alike in state, given that they are not identical by descent and was set to min $(1–S_{ij})$.

2.3 | Statistical analysis

The models used in this study follow the syntax outlined by Piepho, Büchse, and Emrich (2003). Briefly, crossed effects are denoted with a dot operator and fixed effects are separated from random effects by a colon, with fixed effects first. The model to obtain global variance components and best linear unbiased estimates (BLUEs) was:

$$G + E + E \cdot R{:}E \cdot G + E \cdot R \cdot B$$

with *G*, *E*, *R* and *B* denoting genotypes, environments, replications and incomplete blocks, respectively. Outliers were determined following method four (Bonferroni–Holm with rescaled MAD standardized residuals) presented by Bernal-Vasquez, Utz, and Piepho (2016). Best linear unbiased predictors (BLUPs) and variance components were calculated with a fully random model for each trait and BLUEs according to the model described above.

To estimate group-specific BLUPs and variance components, dummy variables were introduced for the genotypic groups females, males, hybrids and checks. This allowed to calculate GCA and SCA effects of the hybrids independently from the other genotypic groups. For simplification, dummy variables were suppressed in the following model used to estimate group-specific variance components:

$$E + GR7OUP + E \cdot R{:}FEMALE + MALE + CHECK +$$
$$FEMALE \cdot E + MALE \cdot E + CHECK \cdot E +$$
$$GCA_F + GCA_M + SCA +$$
$$GCA_F \cdot E + GCA_M \cdot E + SCA \cdot E +$$
$$GROUP \cdot E + E \cdot R \cdot B$$

with *E*, *R* and *B* as environments, replications and incomplete blocks, respectively, as well as *GROUP*, *FEMALE*, *MALE*, *CHECK*, *GCA* and *SCA* as the genotypic groups, females, males, checks, GCA effects of the female (GCA_F) or male (GCA_M) parental lines and the SCA effects.

Exploiting the relatedness between the parental genotypes, we specified a variance–covariance structure for the random effects of GCA_F, GCA_M and SCA (Piepho, Möhring, Melchinger, & Büchse, 2008). We assumed $\text{var}(GCA_F) = \mathbf{A}_F \sigma^2_{GCA-F}$, $\text{var}(GCA_M) = \mathbf{A}_M \sigma^2_{GCA-M}$ and $\text{var}(SCA) = \mathbf{D} \sigma^2_{SCA}$, where σ^2_{GCA-F}, σ^2_{GCA-M} and σ^2_{SCA} refer to the variance of GCA_F, GCA_M and SCA effects estimated by REML for the female or male as well as the hybrid combinations derived by crosses between the female and male parental lines, respectively. $\mathbf{A}_F$ was a 23 × 23 matrix and $\mathbf{A}_M$ a 10 × 10 matrix containing genomic relationship coefficient estimates defining the degree of genetic covariance between all pairs of female and male parental components, respectively. $\mathbf{D}$ was a 91 × 91 matrix containing the dominance relationships for the 91 single-cross hybrids derived from the probability that both alleles at a locus are identical by descent. Elements of $\mathbf{D}$ correspond to the

dominance relationship between two hybrids *a × b* and *c × d* with their parents *a* and *b* or *c* and *d*, respectively, and were calculated according to the suggestions of Reif et al. (2013) as $\delta_{ab,cd} = \left\{ \left(\theta_{ac}\theta_{bd} + \theta_{ad}\theta_{bd} \right) /2 \right\}$, where θ_{ac}, θ_{bd}, θ_{ad} and θ_{bd} refer to the coefficient of coancestry between corresponding parents estimated based on marker data. The matrices $\mathbf{A}_F$, $\mathbf{A}_M$ and $\mathbf{D}$ were calculated as outlined above.

Global and group-specific variance components were calculated with a fully random model using the restricted maximum likelihood method (REML) implemented by the software package ASReml-R 3.0 (Butler, Cullis, Gilmour, & Gogel, 2009). Residual error variances were assumed to be heterogeneous among environments (Kelly, Smith, Eccleston, & Cullis, 2007) for both, the global and the group-specific variance component estimation. Test of significance (*p* < .05, .01, and .001) of the variance components was performed with a likelihood ratio test according to Stram and Lee (1994). Global and group-specific heritabilities (H^2) were calculated as $H^2 = 1 - \bar{v}_{BLUP}/\left(2 * \sigma^2_G \right)$ (Cullis, Smith, & Coombes, 2006), with $\bar{v}_{BLUP}$ being the mean variance of a difference of two BLUPs and σ^2_G the genotypic variance estimate across all groups or the nested genotypic variance estimate of the corresponding genotypic group.

Hybrid performance was estimated based on GCA and SCA effects of the parental inbred lines *i* and *j* as:

$$\hat{y}_{hybrid(ij)} = \mu + \widehat{GCA}_{F(i)} + \widehat{GCA}_{M(j)} + \widehat{SCA}_{F(i) \times M(j)}$$

with $\hat{y}_{hybrid(ij)}$, μ, $\widehat{GCA}_{F(i)}$, $\widehat{GCA}_{M(j)}$ and $\widehat{SCA}_{F(i) \times M(j)}$ being the estimated phenotypic performance of the hybrid derived from the parental lines *i* and *j*, the overall mean, the GCA effect of the *i*-th female line, the GCA effect of the *j*-th male line and the SCA effect of the cross between the *i*-th female parental line with the *j*-th male parental line, respectively.

Pearson's product–moment correlations (*r*) were estimated between (a) Rogers' distance and the mid-parent heterosis of the corresponding hybrid, as well as (b) per se performance of the parental lines and the corresponding GCA effects of the female and male parental inbred lines. Mid-parent heterosis of each single-cross hybrid was estimated as the difference between the mean per se performance of the two parental inbred lines *i* and *j* of the corresponding single-cross and the hybrid's per se performance.

2.4 | Hybrid prediction

Prediction accuracies were calculated by dividing Pearson's product–moment correlation between predicted and observed values by the square root of the global heritability estimate (Table S2) of the corresponding trait (Legarra, Robert-Granié, Manfredi, & Elsen, 2008). Hybrids were predicted based on (a) mid-parent (MP) performance, (b) the sum of parental GCA effects, as well as the sum of parental (c) GCA (GCA_{kin}), or (d) GCA and SCA effects (GCA-SCA$_{kin}$) exploiting relationships between parental inbred lines as outlined above.

We evaluated the prediction accuracies by cross-validation approaches to correct for autocorrelation (Schrag, Melchinger,

Sørensen, & Frisch, 2006). First, we used a leave-one-hybrid-out cross-validation to predict each untested single-cross based on parental GCA effects, not exploiting genetic relationship (GCA$_{loho}$). For this scenario, all parental lines were included and only the hybrid to be predicted was not considered for the estimation of the parental GCA effects. Second, we applied four scenarios, where (a) both (T2), (b) only the male (T1$_{female}$), (c) only the female (T1$_{male}$), or (d) no (T0) parental component(s) contributed to the prediction of the untested single-cross, using relationship information based on (a) parental GCA effects only (GCA$_{kin}$) or (b) GCA and SCA effects (GCA-SCA$_{kin}$). For these three scenarios, one (T1$_{female}$ and T1$_{male}$) or two (T0) of the parental lines were removed from the training set, which was achieved with the dummy variables mentioned for the model above.

In addition, we expanded the dominance relationship matrix $\mathbf{D}$ to all 230 possible hybrid combinations, predicting the 139 untested hybrids by using the sum of GCA or GCA and SCA effects. For a better comparison between the five traits, we normalized the predicted phenotypic values for each trait, including tested and untested hybrids, as:

$$z_i = \frac{x_i - \bar{x}}{\sigma_x}$$

where z_i is the normalized phenotypic value, x_i is the predicted phenotypic value of the i-th hybrid, and $\bar{x}$ and σ_x are the mean and the standard deviation across the predicted values of tested and untested hybrids, respectively.

2.5 | Simulation study

We performed a simulation study to demonstrate what different prediction accuracies actually mean in a breeding programme, regarding the identification and thus the production of superior hybrids based on predicted hybrid values. We assessed the rate of success of having produced and tested (a) the best or (b) one of the five best hybrids from a 10 × 10 or a 100 × 100 factorial, when the hybrids are predicted with varying prediction accuracies, and only a subset of all possible hybrids from the factorial is actually produced. The simulation was based on 1,000,000 runs.

3 | RESULTS

Rogers' distance estimates based on genome-wide molecular marker data revealed a high variation in relatedness within as well as among the female and male parental lines (Figure 1a). Rogers' distances of the produced hybrids provided a good coverage of all possible hybrid combinations. Principal coordinate axes one and two explained 14.5% and 11.7% of total genotypic variation, respectively. A distinct clustering of the female and male parental lines was not observable (Figure 1b).
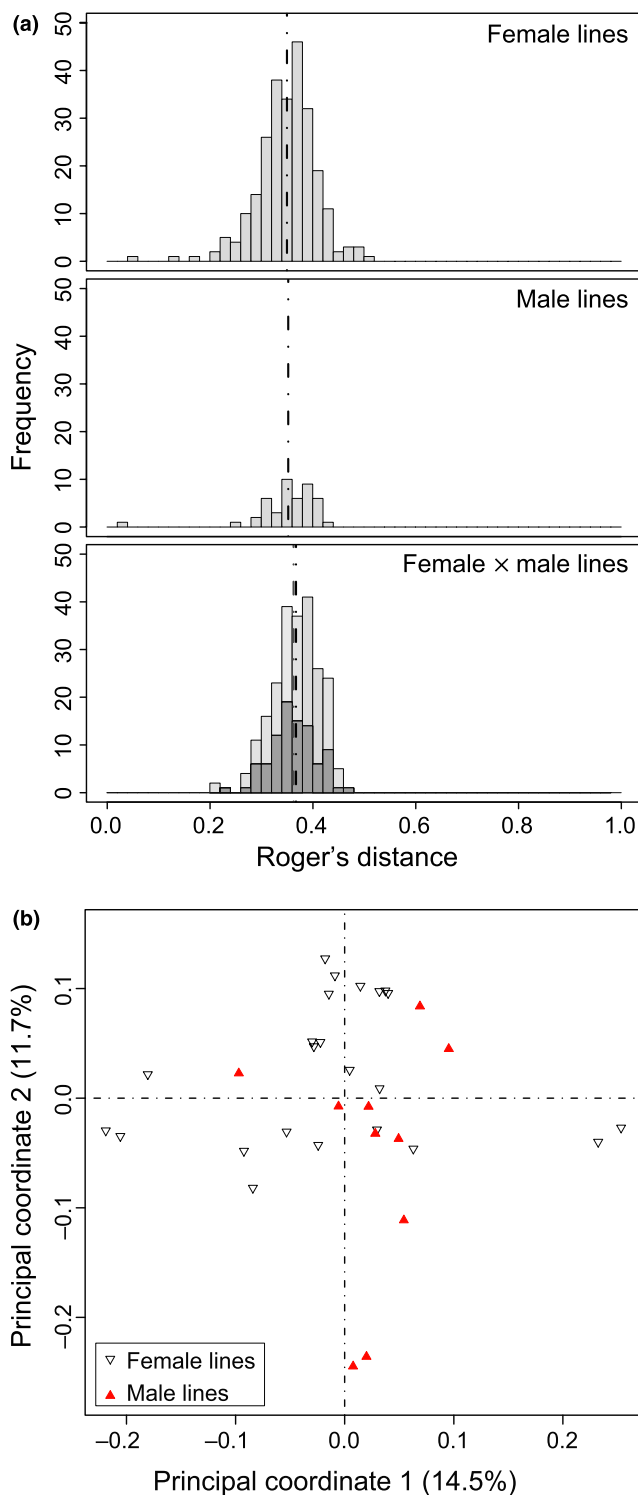

FIGURE 1 (a) Histogram of Rogers' distances based on genome-wide marker data among the 23 female and the 10 male parental lines, as well as between these parental lines. The dark grey bars indicate the distribution of Rogers' distances between the female and male parental lines of the triticale hybrids evaluated in this study. (b) Principal coordinate analysis of the 23 female and 10 male parental lines based on Rogers' distances. Values in parentheses refer to the proportion of variance explained by the principal coordinates [Colour figure can be viewed at wileyonlinelibrary.com]

Phenotypic assessment of the single-cross hybrids in multi-location trials yielded high heritability estimates for all five evaluated traits, ranging from 0.83 for heading time to 0.95 for plant height (Table 1). Also for the highly quantitative traits fresh biomass and dry biomass, high heritabilities of 0.91 and 0.92, respectively, could be obtained. Most variance components of the single-cross hybrids were significantly different from zero. The ratio of the SCA variance to the total genetic variance was low for heading time, plant height and dry matter content, ranging between 0.01 and 0.04. By contrast, for fresh biomass and for dry biomass this ratio was higher, amounting to 0.23 for both traits.

The correlation between the per se performance of the parental lines and their GCA effects was high for heading time, plant height and dry matter content, and moderate for fresh and dry biomass (Table 2). The correlation between Roger's distance between the two parental lines of a hybrid and the hybrid's mid-parent heterosis, by contrast, was low and non-significant. Hybrid prediction based on the mean of the parental lines yielded a high prediction accuracy between 0.83 and 0.89 for heading time, plant height and dry matter

content, and a moderate prediction accuracy of 0.67 and 0.63 for fresh and dry biomass, respectively (Table 3). Using GCA estimates to predict single-cross hybrids increased the prediction accuracy for all traits, but particularly for fresh and dry biomass with a relative increase of around 30% compared to prediction based on mid-parent values. Exploiting genetic relationships among the parental lines did not increase prediction accuracy based on GCA effects (GCA_{kin}), but for the prediction based on GCA and SCA effects ($GCA\text{-}SCA_{kin}$) the accuracy increased by around 0.15 for fresh and dry biomass (Table 3). The use of relationship information of all 230 potential hybrid combinations resulted in higher median estimates for the predicted phenotypic values of the 91 tested hybrids compared to the 139 untested hybrids for all traits (Figure S1). The inclusion of SCA effects in the prediction showed only minor changes.

To account for autocorrelation in the GCA-based prediction approaches, we applied different cross-validation schemes. In brief, we used a leave-one-hybrid-out cross-validation not using kinship information (GCA_{loho}) and two approaches exploiting genetic relationship information based on either GCA (GCA_{kin}) or GCA and SCA effects ($GCA\text{-}SCA_{kin}$). For the latter two, three scenarios were evaluated that represent the situations where both parental components (T2), only the female ($T1_{male}$) or the male ($T1_{female}$) parent, or none of the two parental components (T0) have been tested in other hybrid combinations before. Cross-validated prediction accuracies of GCA_{loho} were considerably lower compared to the GCA-based predictions (Table 4). For dry biomass yield, this cross-validated prediction accuracy dropped to nearly the level of the mid-parent-based prediction (Table 3). Estimates of the GCA_{kin} T2 scenario yielded similar prediction accuracies compared to GCA_{loho} for all traits, except fresh biomass yield, for which it dropped even below the corresponding value of mid-parent prediction (Table 3). As a general trend, we observed a decrease of the prediction accuracy from the T2 to the T1 and the T0 scenario. Including dominance relationships between parental inbred lines ($GCA\text{-}SCA_{kin}$) did not yield substantially higher estimates compared to using only GCA effect estimates (GCA_{kin}) (Table 4).

We performed a simulation study to demonstrate what the different prediction accuracies mean for hybrid prediction applied in a breeding programme (Figure 2). We assumed a 10 × 10 or a 100 × 100 factorial, for which the hybrids were predicted with a prediction accuracy ranging from 0 to 1. Out of these possible hybrids, only a certain fraction, that is selected based on the predicted hybrid values, is produced and tested in field trials. The question is then, what is the probability of having included the best or one of the five best hybrids in this selected fraction. Obviously, if the selected

TABLE 1 Means and ranges of best linear unbiased estimators, heritability estimates across environments and variance components for 91 single-cross hybrids for the traits heading time (HT), plant height (PH), fresh biomass (FBM), dry matter content (DMC) and dry biomass (DBM)

	HT	PH	FBM	DMC	DBM
Min	52.7	92.5	30.7	34.7	12.1
Max	60.5	128.7	44.5	39.9	16.7
Mean	56.8	113.7	39.5	37.2	14.7
σ^2_{GCA-F}	0.44***	44.39***	1.26**	0.21***	0.15*
$\sigma^2_{GCA-F \times E}$	0.11**	1.64***	0.00	0.14***	0.00
σ^2_{GCA-M}	1.43***	92.26***	5.36***	0.78***	0.65***
$\sigma^2_{GCA-M \times E}$	0.46***	6.24***	1.18***	0.29***	0.15***
σ^2_{SCA}	0.08**	4.12***	2.02***	0.01	0.24***
$\sigma^2_{SCA \times E}$	0.06*	1.48*	0.86+	0.00	0.10+
$\sigma^2_{SCA}/\sigma^2_{sum}$ [a]	0.04	0.03	0.23	0.01	0.23
H^2	0.83	0.95	0.91	0.89	0.92
σ^{2b}_e	0.97	9.76	6.89	1.27	0.89

[a]Sum of $\sigma^2_{GCA-F} + \sigma^2_{GCA-M} + \sigma^2_{SCA}$.
[b]Pooled residual error variance across all genotypic groups.
+, *, **, ***Significantly different from zero at the 0.1, 0.05, 0.01 and 0.001 probability level, respectively.

TABLE 2 Pearson's product–moment correlations (r) between parental per se performance (per se) and general combining ability (GCA) effects as well as between Rogers' distance (RD) and mid-parent heterosis (MPH), shown for heading time (HT), plant height (PH), fresh biomass (FBM), dry matter content (DMC) and dry biomass (DBM)

	HT	PH	FBM	DMC	DBM
r(per se, GCA)	0.81***	0.89***	0.57***	0.73***	0.54***
r(RD, MPH)	0.08	−0.07	−0.02	−0.00	−0.01

***Significantly different from zero at the 0.001 probability level.

fraction is 1.0, that is if all possible hybrids from the factorial are produced, these will include the best hybrid and the rate of success is 1. Likewise, if the prediction accuracy is 1.0, the prediction is perfect, and irrespective of the number of produced hybrids, these will include the best one, even if just one hybrid is produced. Now if, for example, we assume a moderate prediction accuracy of 0.6 and we expect a success rate of at least 0.95 of having included the best hybrid in the selected fraction, then we need to produce approximately 46% of all possible hybrids in the smaller factorial and 17% in the larger factorial. If, instead, we aim to identify at least one of the five best hybrids, which can be expected to not be significantly different anyhow, then the fraction of hybrids that needs to be produced decreases substantially to around 10% in the smaller and 2% in the larger factorial. As a general trend, the fraction of hybrids that has to be produced to include the best or at least one of the five best hybrids, increased with decreasing prediction accuracy.

TABLE 3 Prediction accuracy for heading time (HT), plant height (PH), fresh biomass (FBM), dry matter content (DMC) and dry biomass (DBM), for prediction based on mid-parent value (MP), the sum of general combining ability (GCA) effects, the sum of GCA effects (GCA$_{kin}$) and the sum of GCA and specific combining ability (SCA) effects (GCA-SCA$_{kin}$) using kinship estimates

	HT	PH	FBM	DMC	DBM
MP	0.89	0.86	0.67	0.83	0.63
GCA	0.97	0.99	0.87	0.99	0.85
GCA$_{kin}$	0.97	0.99	0.85	0.99	0.84
GCA-SCA$_{kin}$	1.00[a]	1.00[a]	1.00[a]	0.99	1.00[a]

[a]Prediction accuracy values ≥ 1 were set to 1.

TABLE 4 Cross-validated prediction accuracy for heading time (HT), plant height (PH), fresh biomass (FBM), dry matter content (DMC) and dry biomass (DBM), shown for GCA-based leave-one-hybrid-out cross-validation (GCA$_{loho}$), the sum of GCA effects (GCA$_{kin}$) and the sum of GCA and specific combining ability (SCA) effects (GCA-SCA$_{kin}$) using kinship estimates. For the cross-validation of the GCA$_{kin}$ and GCA-SCA$_{kin}$ approaches, three scenarios were simulated, where both (T2), only the male (T1$_{female}$) or only the female (T1$_{male}$), or no (T0) parental component(s) of the predicted hybrid has yet been tested in other hybrid combinations

	HT	PH	FBM	DMC	DBM
GCA$_{loho}$	0.89	0.95	0.74	0.93	0.67
GCA$_{kin}$					
T2	0.88	0.95	0.60	0.92	0.66
T1$_{female}$	0.72	0.72	0.64	0.74	0.60
T1$_{male}$	0.63	0.52	0.44	0.78	0.22
T0	0.08	0.31	0.39	0.52	0.27
GCA-SCA$_{kin}$					
T2	0.88	0.95	0.70	0.92	0.66
T1$_{female}$	0.72	0.77	0.64	0.74	0.60
T1$_{male}$	0.63	0.52	0.46	0.78	0.22
T0	0.09	0.31	0.41	0.52	0.28

4 | DISCUSSION

The first commercial triticale hybrids have already been registered (Longin et al., 2012), and the study of Losert et al. (2016) demonstrated that hybrids possess competitive biomass yield characteristics compared to line cultivars. However, the number of single-cross hybrids that could potentially be generated increases quadratically with the number of parental lines (Bernardo, 2010). Identifying the most promising single-cross hybrids out of the huge number of potential hybrid combinations is one of the biggest challenges for a successful hybrid breeding programme. The aim of this study was, therefore, to compare different approaches to predict single-cross hybrid performance based on mid-parent value, GCA and SCA effects, as well as by exploiting relationship information between the parental lines.

4.1 | Relationship between parental line per se performance and GCA effects

No significant correlation between mid-parent heterosis and Rogers' distance estimates was found for any of the traits evaluated in this study (Table 2), which is consistent with previous results (Gowda et al., 2013; Tams, Bauer, Oettler, Melchinger, & Schön, 2006). The correlation between parental line per se performance and their GCA effects, by contrast, was moderate to high, which corroborates earlier results from triticale and wheat (Gowda et al., 2013; Longin et al., 2013; Miedaner, Schulthess, Gowda, Reif, & Longin, 2017). This can be explained by the fact that the GCA reflects the additive and the additive × additive effects of a line, which are also selected during the improvement of the line per se performance in line breeding. As GCA estimates require the time- and resource-intensive establishment of hybrids and their evaluation in field trials, the parental line per se performance can be used as an efficient first predictor to select lines to be tested in hybrid combinations in triticale hybrid breeding programmes.

4.2 | Hybrid prediction based on mid-parent performance and general combining ability estimates

Estimates of the prediction accuracy based on mid-parent value were moderate for the genetically more complex traits fresh and dry biomass and high for the presumably less complex traits heading time, plant height and dry matter content. The use of GCA estimates instead of mid-parent values to predict hybrid performance resulted in an increased prediction accuracy for all traits, especially for fresh and dry biomass (Table 3). For the latter two traits, the prediction accuracy was still lower compared to the other three traits, which can be explained by a higher contribution of SCA as illustrated by the substantially higher ratio of the variance component of SCA to the total genetic variance. In contrast to prediction based on mid-parent value, GCA-based prediction
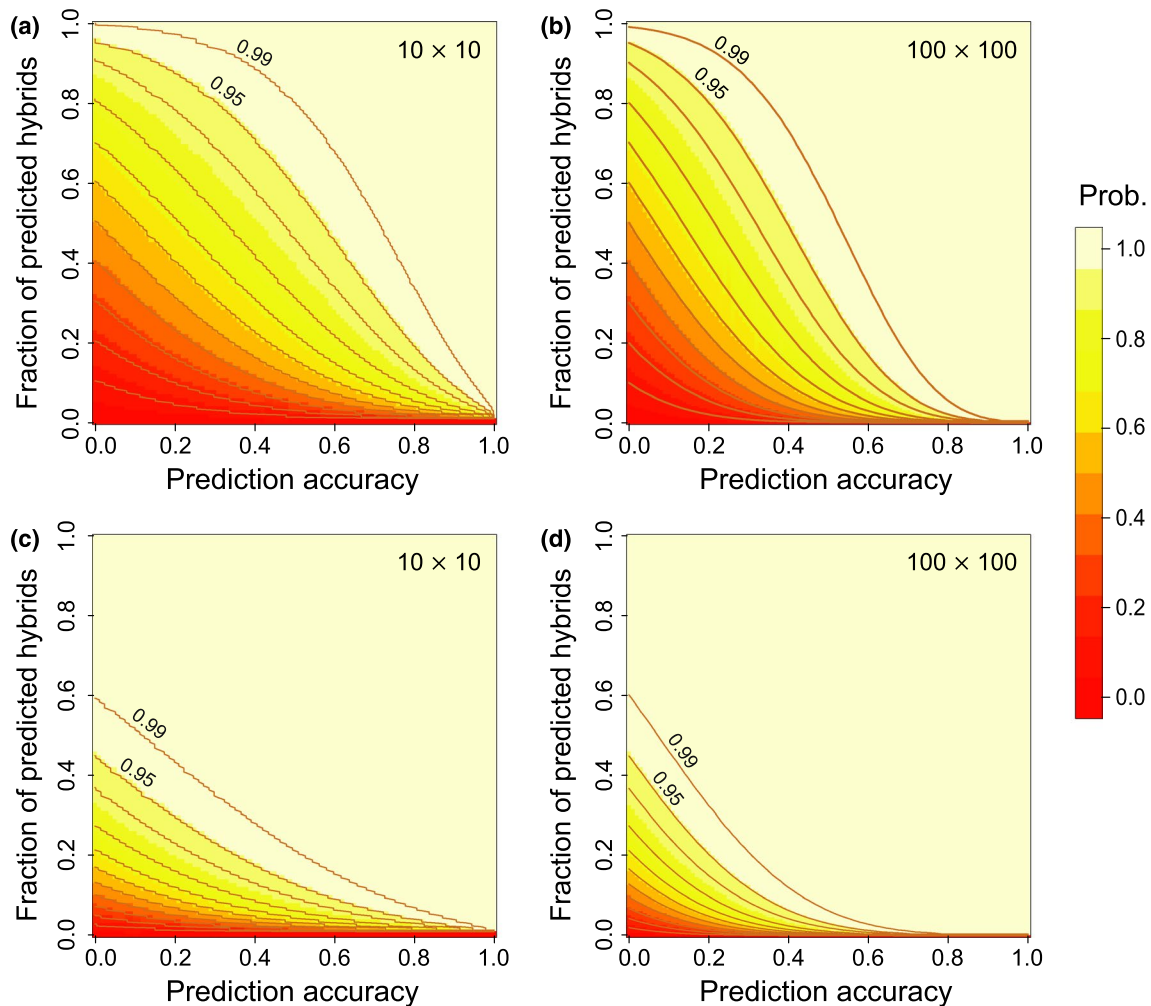
FIGURE 2 Probability of having selected and produced the best (a, b) or at least one of the five best hybrids (c, d), depending on the fraction of predicted hybrids that are actually produced and the prediction accuracy, shown for a 10 × 10 (a, c) or a 100 × 100 (b, d) factorial. Dark colour indicates low, bright colour a high rate of success, with the lines indicating different levels of success in steps of 0.1 and additionally of 0.95 and 0.99 [Colour figure can be viewed at wileyonlinelibrary.com]

also exploits non-additive effects, which was particularly evident for fresh and dry biomass (Bernardo, 2010; Hallauer et al., 2010; Reif et al., 2013; Smith, 1986; Wricke & Weber, 1986). However, dominance effects can only be exploited using SCA effects and seem to have a considerable influence on fresh and dry biomass (Reif, Gumpert, Fischer, & Melchinger, 2007). The influence of dominance effects on single-cross prediction was also reported in previous studies (Technow, Riedelsheimer, Schrag, & Melchinger, 2012; Zhao, Mette, Gowda, Longin, & Reif, 2014).

As a main goal of a hybrid breeding programme is to predict untested hybrids, we applied different cross-validation schemes to mimic this situation. The leave-one-hybrid-out approach (GCA$_{loho}$) predicts the hybrid of interest without including it in the estimation of the parental GCA effects. As expected, cross-validation reduced the GCA-based prediction accuracy (GCA$_{loho}$), but it was still higher than the prediction accuracy obtained by mid-parent values.

Mid-parent heterosis for fresh and dry biomass averaged around 5% in this panel, illustrating that hybrid performance for the

biomass yield traits in triticale is to a large extent based on effects that are captured by the GCA (Losert et al., 2016). In line with the high correlation between the line per se performance and the GCA estimates, hybrid prediction based on mid-parent value provided somewhat lower, but still promising prediction accuracies compared to GCA-based prediction. This, along with the results from our simulation, again underscores the potential to use the per se performance as a first criterion to select not only promising lines, but also hybrid combinations in current triticale hybrid breeding programmes as proposed for maize (Longin, Mi, Melchinger, Reif, & Würschum, 2014) and wheat (Longin, Utz, Melchinger, & Reif, 2007).

4.3 | Hybrid prediction exploiting genetic relationship

Genome-wide markers nowadays offer the possibility to estimate genetic kinship, which can be exploited for hybrid prediction.

The use of relationship information did not improve the prediction accuracy based on GCA effects alone (GCA$_{kin}$), but did so when also SCA effects were included for prediction (GCA-SCA$_{kin}$), particularly for the traits fresh and dry biomass with their higher contribution of SCA (Table 3). However, this difference was not discernible any more when cross-validation was applied, which is in line with previous studies in triticale (Boeven et al., 2016; Gowda et al., 2013), rapeseed (Werner et al., 2018) and sunflower (Reif et al., 2013). Furthermore, no differences between the GCA$_{kin}$ or GCA-SCA$_{kin}$ and the GCA$_{loho}$ estimates were observed, which is probably due to the already high prediction accuracies of the latter (Table 4).

Establishing new hybrid combinations is especially time- and labour-consuming in triticale, as the female component needs to be introgressed into a CMS cytoplasm and tested for sterility, whereas the male component requires a sufficient restoration ability. Predicting hybrids of yet untested parental lines is therefore of high interest in triticale hybrid breeding. We, therefore, considered three scenarios to predict untested single-cross combinations, where either the female (T1$_{female}$), the male (T1$_{male}$), or both (T0) parental lines have not been tested in hybrid combinations before. The reference is the T2 scenario, where both parents have already been used as hybrid components. In general, we observed a decrease in prediction accuracy from T2 to T0 and only slight differences between models considering GCA alone or GCA and SCA estimates (Table 4). This decrease in prediction accuracy following the decreasing number of highly related genotypes in the training set is well known in hybrid prediction and has, for example, been reported in triticale (Boeven et al., 2016; Gowda et al., 2013), wheat (Gowda et al., 2014; Liu et al., 2016; Mirdita et al., 2015; Zhao, Zeng, Fernando, & Reif, 2013), maize (Technow et al., 2012, 2014) and sunflower (Reif et al., 2013). The in part considerably higher prediction accuracies for the T1$_{female}$ compared to the T1$_{male}$ scenario are probably caused by a higher σ^2_{GCA} of the male parental lines (Table 1) (Boeven et al., 2016; Gowda et al., 2013; Reif et al., 2013). Interestingly, except for heading time, the prediction accuracies for the T0 scenario were still promising, ranging between ~0.3 and 0.5. As demonstrated by our simulation study, even such seemingly moderate prediction accuracies allow to substantially reduce the number of hybrids to be established while maintaining a high probability of including one of the best hybrids. Notably, this prediction accuracy can be expected to be improved further if larger training data sets for effect estimation become available.

The medians of predicted phenotypic values of tested compared to untested hybrids differed substantially, and only a small fraction of untested hybrids appear to have promising performance for future testing in the field (Figure S1). This is likely due to the high correlation between the per se performance of the parental lines and their GCA, so that a focus on better performing lines as parental components resulted in the establishment of better performing hybrids. Nevertheless, prediction of hybrids between as yet untested lines is promising and has the potential to become a valuable tool for triticale hybrid breeding.

5 | CONCLUSIONS

We did not observe any population structure separating the female and male lines used in this study (Figure 1), which corroborates previous results in triticale (Fischer et al., 2010; Gowda et al., 2013; Losert, Maurer, Marulanda, & Würschum, 2017; Tams, Bauer, Oettler, & Melchinger, 2004; Tams et al., 2006). This was suspected, as all parental lines were taken from the same triticale breeding germplasm and grouped into the female or male parents based on their maintainer and restorer characteristics, respectively. The establishment of heterotic pools is always mentioned as one of the major goals in hybrid breeding, as they allow to increase the contribution of GCA relative to SCA (Fischer et al., 2010; Reif et al., 2007). This in turn would increase the accuracy of GCA-based predictions. Thus, the establishment of heterotic groups remains a long-term goal in triticale hybrid breeding, but given the difficulties in identifying female and male lines with the required characteristics, the goal for the near future is to identify additional lines for each parental group as well as to improve their performance through intra-group recurrent as well as inter-group reciprocal recurrent selection.

Regarding the considerably high prediction accuracies obtained by mid-parent value raises the question whether the laborious and costly estimation of GCA effects and marker data for genetic relationship information are worthwhile. Indeed, at present the high prediction accuracies make mid-parent value a suitable instrument to predict single-cross hybrids, under the prerequisite of available field data for per se performance of both parental lines. This is in accordance with the results from our simulation, as well as findings from triticale (Boeven et al., 2016; Gowda et al., 2013), durum wheat (Gowda et al., 2010) and wheat (Liu et al., 2016; Longin et al., 2013; Miedaner et al., 2017). However, if selection in hybrid triticale breeding programmes is shifted to GCA instead of per se performance and potentially accompanied by the establishment of heterotic groups, GCA-based prediction can be expected to increasingly outperform prediction based on mid-parent value. Then, the further improvement of the promising prediction accuracies for the T1 and T0 scenarios through targeted compilation of larger training sets with higher predictive power will become a task for hybrid triticale breeding, as this circumvents the necessity to develop parental lines to a level, where testing of hybrid performance in multi-environment trials is possible.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ORCID

Johannes Trini [iD] https://orcid.org/0000-0002-1637-4248
Tobias Würschum [iD] https://orcid.org/0000-0002-7397-7731

REFERENCES

Alheit, K. V., Busemeyer, L., Liu, W., Maurer, H. P., Gowda, M., Hahn, V., ... Würschum, T. (2014). Multiple-line cross QTL mapping for biomass yield and plant height in triticale (× *Triticosecale* Wittmack). *Theoretical and Applied Genetics*, *127*(1), 251–260. https://doi.org/10.1007/s00122-013-2214-6

Bernal-Vasquez, A. M., Utz, H. F., & Piepho, H. P. (2016). Outlier detection methods for generalized lattices: A case study on the transition from ANOVA to REML. *Theoretical and Applied Genetics*, *129*(4), 787–804. https://doi.org/10.1007/s00122-016-2666-6

Bernardo, R. (1993). Estimation of coefficient of coancestry using molecular markers in maize. *Theoretical and Applied Genetics*, *85*(8), 1055–1062. https://doi.org/10.1007/BF00215047

Bernardo, R. (1994). Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Science*, *34*(1), 20–25. https://doi.org/10.2135/cropsci1994.0011183X003400010003x

Bernardo, R. (1995). Genetic models for predicting maize single-cross performance in unbalanced yield trial data. *Crop Science*, *35*(1), 141–147. https://doi.org/10.2135/cropsci1995.0011183X003500010026x

Bernardo, R. (1996a). Best linear unbiased prediction of maize single-cross performance. *Crop Science*, *36*(1), 50–56. https://doi.org/10.2135/cropsci1996.0011183X003600010009x

Bernardo, R. (1996b). Best linear unbiased prediction of the performance of crosses between untested maize inbreds. *Crop Science*, *36*(4), 872–876. https://doi.org/10.2135/cropsci1996.0011183X003600040009x

Bernardo, R. (2010). *Breeding for quantitative traits in plants*, 2nd ed. Woodbury, MN: Stemma Press.

Boeven, P. H. G., Würschum, T., Weissmann, S., Miedaner, T., & Maurer, H. P. (2016). Prediction of hybrid performance for Fusarium head blight resistance in triticale (× *Triticosecale* Wittmack). *Euphytica*, *207*(3), 475–490. https://doi.org/10.1007/s10681-015-1498-9

Butler, D. G., Cullis, B. R., Gilmour, A. R., & Gogel, B. J. (2009). *Asreml-R reference manual*.

Charcosset, A., Bonnisseau, B., Touchebeuf, O., Burstin, J., Dubreuil, P., Barrière, Y., ... Denis, J.-B. (1998). Prediction of maize hybrid silage performance using marker data: Comparison of several models for specific combining ability. *Crop Science*, *38*(1), 38–44. https://doi.org/10.2135/cropsci1998.0011183X003800010007x

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*, *11*(4), 381–393. https://doi.org/10.1198/108571106X154443

de Wit, M., & Faaij, A. (2010). European biomass resource potential and costs. *Biomass and Bioenergy*, *34*(2), 188–202. https://doi.org/10.1016/j.biombioe.2009.07.011

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*(13), 13–15.

Fischer, S., Maurer, H. P., Würschum, T., Möhring, J., Piepho, H.-P., Schön, C. C., ... Reif, J. C. (2010). Development of heterotic groups in triticale. *Crop Science*, *50*(2), 584–590. https://doi.org/10.2135/cropsci2009.04.0225

Gowda, M., Hahn, V., Reif, J. C., Longin, C. F. H., Alheit, K., & Maurer, H. P. (2011). Potential for simultaneous improvement of grain and biomass yield in Central European winter triticale germplasm. *Field Crops Research*, *121*(1), 153–157. https://doi.org/10.1016/j.fcr.2010.12.003

Gowda, M., Kling, C., Würschum, T., Liu, W., Maurer, H. P., Hahn, V., & Reif, J. C. (2010). Hybrid breeding in durum wheat: Heterosis and combining ability. *Crop Science*, *50*(6), 2224–2230. https://doi.org/10.2135/cropsci2009.10.0637

Gowda, M., Zhao, Y., Maurer, H. P., Weissmann, E. A., Würschum, T., & Reif, J. C. (2013). Best linear unbiased prediction of triticale hybrid performance. *Euphytica*, *191*(2), 223–230. https://doi.org/10.1007/s10681-012-0784-z

Gowda, M., Zhao, Y., Würschum, T., Longin, C. F., Miedaner, T., Ebmeyer, E., ... Reif, J. C. (2014). Relatedness severely impacts accuracy of marker-assisted selection for disease resistance in hybrid wheat. *Heredity*, *112*(5), 552–561. https://doi.org/10.1038/hdy.2013.139

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, *53*(3–4), 325–338. https://doi.org/10.1093/biomet/53.3-4.325

Hallauer, A. R., Carena, M. J., & Miranda Filho, J. B. (2010). *Quantitative genetics in maize breeding*, Vol. 6. New York, NY, London, UK: Springer.

Kelly, A. M., Smith, A. B., Eccleston, J. A., & Cullis, B. R. (2007). The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science*, *47*(3), 1063–1070. https://doi.org/10.2135/cropsci2006.08.0540

Legarra, A., Robert-Granié, C., Manfredi, E., & Elsen, J. M. (2008). Performance of genomic selection in mice. *Genetics*, *180*(1), 611–618. https://doi.org/10.1534/genetics.108.088575

Liu, G., Zhao, Y., Gowda, M., Longin, C. F. H., Reif, J. C., & Mette, M. F. (2016). Predicting hybrid performances for quality traits through genomic-assisted approaches in Central European Wheat. *PLoS ONE*, *11*(7), https://doi.org/10.1371/journal.pone.0158635

Longin, C. F. H., Gowda, M., Mühleisen, J., Ebmeyer, E., Kazman, E., Schachschneider, R., ... Reif, J. C. (2013). Hybrid wheat: Quantitative genetic parameters and consequences for the design of breeding programs. *Theoretical and Applied Genetics*, *126*(11), 2791–2801. https://doi.org/10.1007/s00122-013-2172-z

Longin, C. F. H., Mi, X., Melchinger, A. E., Reif, J. C., & Würschum, T. (2014). Optimum allocation of test resources and comparison of breeding strategies for hybrid wheat. *Theoretical and Applied Genetics*, *127*(10), 2117–2126. https://doi.org/10.1007/s00122-014-2365-0

Longin, C. F. H., Mühleisen, J., Maurer, H. P., Zhang, H., Gowda, M., & Reif, J. C. (2012). Hybrid breeding in autogamous cereals. *Theoretical and Applied Genetics*, *125*(6), 1087–1096. https://doi.org/10.1007/s00122-012-1967-7

Longin, C. F. H., Utz, H. F., Melchinger, A. E., & Reif, J. C. (2007). Hybrid maize breeding with doubled haploids: II. Optimum type and number of testers in two-stage selection for general combining ability. *Theoretical and Applied Genetics*, *114*(3), 393–402. https://doi.org/10.1007/s00122-006-0422-z

Losert, D., Maurer, H. P., Marulanda, J. J., & Würschum, T. (2017). Phenotypic and genotypic analyses of diversity and breeding progress in European triticale (× *Triticosecale* Wittmack). *Plant Breeding*, *136*(1), 18–27. https://doi.org/10.1111/pbr.12433

Losert, D., Maurer, H. P., Weissmann, S., & Würschum, T. (2016). Hybrid breeding for biomass yield in winter triticale: I. Hybrid performance, trait correlations and heterosis. *Plant Breeding*, *135*(5), 560–566. https://doi.org/10.1111/pbr.12402

LTZ (2013). *Energiepflanzen für Biogasanlagen*. Retrieved from https://pflanzen.fnr.de/

Melchinger, A. E. (1999). Genetic diversity and heterosis. In J. G. Coors, & S. Pandey (Eds.), *The genetics and exploitation of heterosis in crops* (pp. 99–118). Madison, WI: American Society of Agronomy. https://doi.org/10.2134/1999.geneticsandexploitation.c10

Mergoum, M., Singh, P. K., Peña, R. J., Lozano-del Río, A. J., Cooper, K. V., Salmon, D. F., & Gómez Macpherson, H. (2009). Triticale: A "New" crop with old challenges. In M. J. Carena (Ed.), *Cereals* (pp. 267–287). New York, NY: Springer US.

Miedaner, T., Schulthess, A. W., Gowda, M., Reif, J. C., & Longin, C. F. H. (2017). High accuracy of predicting hybrid performance of Fusarium head blight resistance by mid-parent values in wheat. *Theoretical and Applied Genetics*, *130*(2), 461–470. https://doi.org/10.1007/s00122-016-2826-8

Mirdita, V., Liu, G., Zhao, Y., Miedaner, T., Longin, C. F. H., Gowda, M., … Reif, J. C. (2015). Genetic architecture is more complex for resistance to Septoria tritici blotch than to Fusarium head blight in Central European winter wheat. *BMC Genomics*, *16*, 430. https://doi.org/10.1186/s12864-015-1628-8

Monforti, F., Bódis, K., Scarlat, N., & Dallemand, J.-F. (2013). The possible contribution of agricultural crop residues to renewable energy targets in Europe: A spatially explicit study. *Renewable and Sustainable Energy Reviews*, *19*, 666–677. https://doi.org/10.1016/j.rser.2012.11.060

Oettler, G., Tams, S. H., Utz, H. F., Bauer, E., & Melchinger, A. E. (2005). Prospects for hybrid breeding in winter triticale: I. Heterosis and combining ability for agronomic traits in European elite germplasm. *Crop Science*, *45*(4), 1476–1482. https://doi.org/10.2135/cropsci2004.0462

Piepho, H. P., Büchse, A., & Emrich, K. (2003). A hitchhiker's guide to mixed models for randomized experiments. *Journal of Agronomy and Crop Science*, *189*(5), 310–322. https://doi.org/10.1046/j.1439-037X.2003.00049.x

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, *161*(1–2), 209–228. https://doi.org/10.1007/s10681-007-9449-8

Reif, J. C., Gumpert, F. M., Fischer, S., & Melchinger, A. E. (2007). Impact of interpopulation divergence on additive and dominance variance in hybrid populations. *Genetics*, *176*(3), 1931–1934. https://doi.org/10.1534/genetics.107.074146

Reif, J. C., Zhao, Y., Würschum, T., Gowda, M., Hahn, V., & Léon, J. (2013). Genomic prediction of sunflower hybrid performance. *Plant Breeding*, *132*(1), 107–114. https://doi.org/10.1111/pbr.12007

Rogers, J. S. (1972). Measures of genetic similarity and genetic distance. In M. R. Wheeler (Ed.), *Studies in genetics VII* (pp. 145–154). Austin, TX: University of Texas Publications.

Schrag, T. A., Melchinger, A. E., Sørensen, A. P., & Frisch, M. (2006). Prediction of single-cross hybrid performance for grain yield and grain dry matter content in maize using AFLP markers associated with QTL. *Theoretical and Applied Genetics*, *113*(6), 1037–1047. https://doi.org/10.1007/s00122-006-0363-6

Smith, O. S. (1986). Covariance between line *per se* and testcross performance. *Crop Science*, *26*(3), 540–543. https://doi.org/10.2135/cropsci1986.0011183X002600030023x

Stram, D. O., & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics*, *50*(4), 1171. https://doi.org/10.2307/2533455

Tams, S. H., Bauer, E., Oettler, G., & Melchinger, A. (2004). Genetic diversity in European winter triticale determined with SSR markers and coancestry coefficient. *Theoretical and Applied Genetics*, *108*(7), 1385–1391. https://doi.org/10.1007/s00122-003-1552-1

Tams, S. H., Bauer, E., Oettler, G., Melchinger, A. E., & Schön, C. C. (2006). Prospects for hybrid breeding in winter triticale: II. Relationship between parental genetic distance and specific combining ability. *Plant Breeding*, *125*(4), 331–336. https://doi.org/10.1111/j.1439-0523.2006.01218.x

Technow, F., Riedelsheimer, C., Schrag, T. A., & Melchinger, A. E. (2012). Genomic prediction of hybrid performance in maize with models incorporating dominance and population specific marker effects. *Theoretical and Applied Genetics*, *125*(6), 1181–1194. https://doi.org/10.1007/s00122-012-1905-8

Technow, F., Schrag, T. A., Schipprack, W., Bauer, E., Simianer, H., & Melchinger, A. E. (2014). Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics*, *197*(4), 1343–1355. https://doi.org/10.1534/genetics.114.165860

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Weiland, P. (2010). Biogas production: Current state and perspectives. *Applied Microbiology and Biotechnology*, *85*(4), 849–860. https://doi.org/10.1007/s00253-009-2246-7

Werner, C. R., Qian, L., Voss-Fels, K. P., Abbadi, A., Leckband, G., Frisch, M., & Snowdon, R. J. (2018). Genome-wide regression models considering general and specific combining ability predict hybrid performance in oilseed rape with similar accuracy regardless of trait architecture. *Theoretical and Applied Genetics*, *131*(2), 299–317. https://doi.org/10.1007/s00122-017-3002-5

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

Wricke, G., & Weber, E. (1986). *Quantitative genetics and selection in plant breeding*. Berlin, New York: Walter de Gruyter. https://doi.org/10.1515/9783110837520

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*(6), 415–421. https://doi.org/10.1111/j.1365-3180.1974.tb01084.x

Zhao, Y., Mette, M. F., Gowda, M., Longin, C. F. H., & Reif, J. C. (2014). Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity*, *112*(6), 638–645. https://doi.org/10.1038/hdy.2014.1

Zhao, Y., Zeng, J., Fernando, R., & Reif, J. C. (2013). Genomic prediction of hybrid wheat performance. *Crop Science*, *53*(3), 802–810. https://doi.org/10.2135/cropsci2012.08.0463

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.