Theor Appl Genet (2015) 128:291–301
DOI 10.1007/s00122-014-2429-1

ORIGINAL PAPER

Yield performance and stability of CMS-based triticale hybrids

Jonathan Mühleisen · Hans-Peter Piepho · Hans Peter Maurer · Jochen Christoph Reif

Received: 26 September 2014 / Accepted: 13 November 2014 / Published online: 16 December 2014

Abstract

Key message **CMS-based triticale hybrids showed only marginal midparent heterosis for grain yield and lower dynamic yield stability compared to inbred lines.**

Abstract Hybrids of triticale (×*Triticosecale* Wittmack) are expected to possess outstanding yield performance and increased dynamic yield stability. The objectives of the present study were to (1) examine the optimum choice of the biometrical model to compare yield stability of hybrids versus lines, (2) investigate whether hybrids exhibit a more pronounced grain yield performance and yield stability, and (3) study optimal strategies to predict yield stability of hybrids. Thirteen female and seven male parental lines and their 91 factorial hybrids as well as 30 commercial lines were evaluated for grain yield in up to 20 environments. Hybrids were produced using a cytoplasmic male sterility (CMS)-inducing cytoplasm that originated from *Triticum-timopheevii* Zhuk. We found that the choice of the biometrical model can cause contrasting results and concluded that a group-by-environment interaction term should be added to the model when estimating stability variance of hybrids and lines. midparent heterosis for grain yield was on average 3 % with a range from −15.0 to 11.5 %. No hybrid outperformed the best inbred line. Hybrids had, on average, lower dynamic yield stability compared to the inbred lines. Grain yield performance of hybrids could be predicted based on midparent values and general combining ability (GCA)-predicted values. In contrast, stability variance of hybrids could be predicted only based on GCA-predicted values. We speculated that negative effects of the used CMS cytoplasm might be the reason for the low performance and yield stability of the hybrids. For this purpose a detailed study on the reasons for the drawback of the currently existing CMS system in triticale is urgently required comprising also the search of potentially alternative hybridization systems.

Introduction

The success of hybrid breeding in outcrossing crops such as maize and rye (Coors and Pandey 1999) coupled with an improved molecular breeding toolbox (Whitford et al. 2013; Kempe et al. 2014) has stimulated the interest in hybrid breeding in the selfing species wheat, barley, and triticale. Expected advantages of hybrid varieties versus the dominating line varieties are increased grain or biomass yield combined with higher yield stability (Longin et al. 2012). Hybrids of adapted European winter wheat lines, which were produced using chemical hybridizing agents (CHA), showed a grain yield advantage of around 10 % compared to their parental lines (Longin et al. 2013) as well as a significantly increased yield stability (Mühleisen et al. 2014a). The same amount of heterosis was also

---

Communicated by Ian Mackay.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2429-1) contains supplementary material, which is available to authorized users.

J. Mühleisen · H. P. Maurer
State Plant Breeding Institute, University of Hohenheim,
70593 Stuttgart, Germany

H.-P. Piepho
Biostatistics Unit, Institute of Crop Science, University of Hohenheim, 70593 Stuttgart, Germany

J. C. Reif (✉)
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, 06466 Stadt Seeland, Germany
e-mail: reif@ipk-gatersleben.de

observed for European durum wheat hybrids produced with CHA (Gowda et al. 2010). In accordance with this, hybrids of adapted European barley lines produced using a cytoplasmic male sterility system (CMS, Ahokas 1979) showed also a midparent heterosis of around 10 % (Mühleisen et al. 2013a). Moreover, re-analysis of registration trials revealed higher yield stability of barley hybrids in contrast to line varieties (Mühleisen et al. 2014b). For the average level of midparent heterosis in triticale, however, results varied depending on the underlying hybridization system: Hybrids based on CHA showed an average midparent heterosis of around 10 % (Oettler et al. 2005; Fischer et al. 2010). In contrast, CMS-based hybrids had an average midparent heterosis of only around 2 % (Gowda et al. 2013). The level of heterosis is a crucial parameter determining the long-term perspective of hybrid versus line breeding (Longin et al. 2014). Therefore, it is important to verify or falsify the differences in average hybrid performance for CHA- versus CMS-based hybrids in triticale.

First results on the yield stability of hybrids versus inbred lines based on a limited set of five environments suggested that triticale hybrids are more yield stable in contrast to the inbred lines (Mühleisen et al. 2014a). Estimating yield stability, however, is a challenging task and heritability or repeatability of stability parameters was often found to below and reliable estimation requires intensive phenotyping in a large number of diverse environments (Becker 1987; Léon and Becker 1988; Pham and Kang 1988; Jalaluddin and Harrison 1993; Sneller et al. 1997; Kumar et al. 1998; Robert 2002; Mühleisen et al. 2014b). Several stability measures have been suggested, which can be grouped in static and dynamic yield stability parameters (Becker and Leon 1988). While static yield stability requires a constant yield across environments, dynamic yield stability allows predictable fluctuations depending on the environmental yield levels. Since for grain yield, natural limitations such as water or nitrogen supply restrict the yield formation, high yield combined with pronounced static yield stability across varying environments is hardly possible (compare Lin et al. 1986; Becker and Leon 1988). In the present study, we focused therefore on the stability variance (Shukla 1972), which is a commonly used measure of dynamic yield stability.

The results of the contrast between dynamic yield stability of hybrid versus lines are strongly influenced by the implementation of the biometrical models. Mühleisen et al. (2014a), for instance, have shown that the magnitude of the deviation variances estimated with the regression approach suggested by Eberhart and Russell (1966) was impaired by the definition of the environmental index. Stability variance of a genotype depends on the reference population (Shukla 1972). Thus, if two groups of genotypes with different reactions on environmental growth factors are evaluated together, traditional stability analysis according to Shukla (1972) may become biased. Therefore, application of Shukla's stability variance model to comparisons among groups might require some modifications.

One major challenge in hybrid breeding is the efficient prediction of hybrid performance (Zhao et al. 2014). Previous studies have shown that prediction of grain yield performance of triticale hybrids based on midparent values was possible (Fischer et al. 2010). Prediction accuracy of hybrid grain yield performance increased using general combining ability (GCA) effects. Whether the same holds true for yield stability was not yet investigated.

Our study is based on an extensive phenotypic data set of triticale hybrids, their parental lines, and commercial lines, i.e., advanced breeding lines and registered varieties, which were evaluated for grain yield in up to 20 environments. The objectives of the present study were to (1) examine the optimum choice of the biometrical model to compare yield stability of hybrids versus lines, (2) investigate whether hybrids exhibit a more pronounced grain yield performance and yield stability, and (3) study optimal strategies to predict yield stability of hybrids.

Materials and methods

Plant materials and field experiments

The present study comprised 141 winter triticale (×*Triticosecale* Wittmack) genotypes belonging to four groups: 91 single-cross hybrids, 13 female parental lines, 7 male parental lines, and 30 commercial lines. The 91 hybrids were produced by factorial crosses between the 13 female and 7 male parental lines. Hybrid seeds were produced by pollination of the female lines converted into the cytoplasmic male sterility (CMS)-inducing cytoplasm with the male restorer lines in isolation plots. The CMS-inducing cytoplasm was provided by Saatzucht Dr. Hege GbR and originated from *Triticumtimopheevii* Zhuk. Isolation plots were spatially separated to avoid pollination of female lines with a pollen mixture of several males.

The genotypes were evaluated in multi-location field trials in the years 2011, 2012, and 2013. All field trials were laid out as plot trials with application of fertilizers, fungicides and growth regulators according to locally adapted practices. Plots were drilled with a plot size between 5 and 12 m^2, 6–10 rows per plot, 12–18 cm row spacing and sowing density between 250 and 280 seeds per m^{-2}. The experimental design of field trials and number of entries varied across years and locations (supplementary Table S1). In total, the trials were conducted at nine locations in Germany and one location in Central France. Yield data were obtained from 20 environments (year-by-location

Table 1 Description of the environments

Location	Year	Yield level (Mg ha^{-1})	Average midparent heterosis (%)	Soil texture	Soil quality index (Ackerzahl)	Precipitation	Growing degree days	Sunshine hours	Comments
Stuttgart	2011	9.85	2.68	Silty loam	60	251	1,811	1,065	
St, Johann	2011	7.57	−2.95	Sandy loam	54	283	1,532	984	
Willstätt	2011	7.01	3.11	Loamy sand	54	203	1,913	1,118	Spatial heterogeneous drought stress
Weimar	2011	8.25	0.28	Loam	58	245	1,719	1,047	
Issoudun	2011	5.12	2.63	Loamy sand	n/a	n/a	n/a	n/a	Dry conditions during heading caused lower yields
Ranzin	2011	8.63	7.22	Sandy loam	50	233	1,628	1,072	
Stuttgart	2012	8.35	5.08	Silty loam	60	252	1,764	900	Nitrogen fertilization was too low and caused reduced yields
St, Johann	2012	8.16	6.36	Sandy loam	55	334	1,488	850	
Weimar	2012	8.44	5.65	Sandy loam	66	227	1,663	824	
Ranzin	2012	10.43	6.73	Sandy loam	55	127	1,568	845	Several plots were affected by frost damages. Yields of plots with very strong frost damages were treated as missing values
Stuttgart	2013	8.25	1.44	Clay loam	n/a	285	1,516	751	
St, Johann	2013	7.93	−4.64	Sandy loam	n/a	398	1,234	644	
Willstätt	2013	9.85	4.30	Very loamy sand	59	417	1,666	784	
Weimar	2013	10.25	2.14	Sandy loam	58	302	1,380	753	
Issoudun	2013	10.07	4.48	Sandy loam	n/a	n/a	n/a	n/a	
Ranzin	2013	10.11	3.68	Sandy loam	63	202	1,445	944	
Schwäbisch Hall	2013	9.84	5.09	Loamy silt	40	306	1,465	n/a	
Langenstein	2013	9.34	2.25	Silty loam	60	223	1,445	754	
Vipperow	2013	8.52	0.76	n/a	n/a	220	1,444	928	
Celle	2013	9.63	4.83	Loamy sand	50	n/a	n/a	n/a	

combinations) (Table 1). Limited seed availability, frost damage and other reasons resulted in some missing plots with an average of 19.3 environments in which a genotype has been successfully evaluated for grain yield. Data from 2011 were already used in earlier publications (Mühleisen et al. 2013b, 2014a).

Weather data

Data for maximum and minimum temperature of the day in degree celsius (°C), daily precipitation in millimeters and daily sunshine hours were provided from the German Meteorological Service (DeutscherWetterdienst) from a nearby weather station. If the minimum or maximum temperature was negative, it was fixed to zero. If the maximum temperature was larger than 25 °C it was fixed to 25 °C. Subsequently, the growing degree days were estimated by summing up the mean of maximum and minimum temperatures of each day from 1st of March till 15th of June following McMaster and Wilhelm (1997) (Table 1). Sunshine hours and precipitation were also summed up for the period from 1st of March till 15th of June (Table 1).

Statistical analysis of the field trials

Each trial, except the trial Willstätt in the year 2011, was analyzed with the following mixed model:

$$y_{ikm} = \mu + g_i + r_k + b_{km} + \varepsilon_{ikm}, \tag{1}$$

where y_{ikm} is the grain yield of the ith genotype in the mth incomplete block within the kth replicate, μ the general mean, g_i the effect of the ith genotype, r_k the effect of the kth replicate, b_{km} the effect of the mth incomplete block within the kth replicate, and ε_{ikm} was the error of y_{ikm}. The effects μ and g_i were assumed to be fixed. The random effects r_k, b_{km} and ε_{ikm} were assumed to follow a normal distribution with zero mean and variances σ_r^2, σ_b^2, and $\sigma_{\varepsilon p}^2$, respectively. The variance $\sigma_{\varepsilon p}^2$ denoted the residual variance of plot yields. Empirical best linear unbiased estimators (BLUEs) of genotype means were calculated. The trial in Willstätt in the year 2011 was analyzed using a covariance model described in detail elsewhere (Mühleisen et al. 2013b).

In order to provide a general overview about the different sources of variation, we fitted the following model:

$$\begin{aligned} y_{tipqj.} = {} & \mu + a_t \\ & + d_{\text{hybrid}} \times f_p + d_{\text{hybrid}} \times m_q + d_{\text{hybrid}} \times g_i \\ & + d_{\text{female}} \times g_i + d_{\text{male}} \times g_i + d_{\text{commercial line}} \times g_i \\ & + e_j + (ae)_{tj} + d_{\text{hybrid}} \times (fe)_{pj} + d_{\text{hybrid}} \times (me)_{qj} \\ & + d_{\text{hybrid}} \times (ge)_{ij} + d_{\text{female}} \times (ge)_{ij} + d_{\text{male}} \times (ge)_{ij} \\ & + d_{\text{commercial line}} \times (ge)_{ij} + \varepsilon_{tipqj.} \end{aligned} \tag{2}$$

In this model, $y_{tipqj.}$ refers to the adjusted mean yield in the jth environment of the ith genotype belonging to the tth group with pth female parent and qth male parent. The effect μ denotes the intercept, a_t the main effect of the tth group, d_{hybrid} is a quantitative dummy variable that is equal to one for hybrids and zero for other genotypes. Note that the dummy variable acts as a switch that allows fitting genotypic effects only to specified groups (Piepho et al. 2012). The effect f_p is the GCA effect of pth female parent, m_q is the GCA effect of the qth male parent, g_i is the specific combining ability (SCA) effect of the ith genotype, provided that it is a hybrid, d_{female} is a dummy variable that is one for female parents and zero for other genotypes, d_{male} a dummy variable that is one for male parents and zero for other genotypes, $d_{\text{commercial line}}$ a dummy variable that is one for the commercial lines and zero for other genotypes, and e_j is the main effect of the jth environment. The following seven terms in model (2) denote interaction effects with the environment and $\varepsilon_{tipqj.}$ is the error of $y_{tipqj.}$. The general mean μ was assumed to be fixed. The other effects were assumed to be independent random normal deviates with zero mean and variances σ_a^2, $\sigma_{\text{GCA(female)}}^2$, $\sigma_{\text{GCA(male)}}^2$, σ_{SCA}^2, $\sigma_{g(\text{female})}^2$, $\sigma_{g(\text{male})}^2$, $\sigma_{g(\text{commercial})}^2$, σ_e^2, $\sigma_{a\times e}^2$, $\sigma_{\text{GCA(female)}\times e}^2\sigma_{a\times e}^2$, $\sigma_{\text{GCA(male)}\times e}^2$, $\sigma_{\text{SCA}\times e}^2$, $\sigma_{g(\text{female})\times e}^2$, $\sigma_{g(\text{male})\times e}^2$, $\sigma_{g(\text{commercial})\times e}^2$, $\sigma_{\varepsilon m(\text{pooled})}^2$, respectively. The variance $\sigma_{\varepsilon m(\text{pooled})}^2$ denoted the pooled residual variance of adjusted mean yields of genotypes at individual environments. The residual variance of adjusted means was fixed to an approximated value as follows: For each environment the residual variance of plot yields ($\sigma_{\varepsilon p}^2$) from model (1) was divided by the number of replicates. Subsequently the mean of these values was calculated and used as residual variance of adjusted means.

As measure of dynamic yield stability we used Shukla's stability variance (1972). The stability variance is the genotype-specific genotype-by-environment interaction variance. A genotype with a small stability variance is stable according to the dynamic concept of stability (Becker and Leon 1988). The complete stability analysis was based on the adjusted genotype means (BLUEs) estimated with model (1) and with the analysis of covariance model for the environment Willstätt in the year 2011. Therefore, it was not possible to separate the genotype-by-environment interaction variance from the error variance. In order to estimate stability variances in models (3a, 3b, 3c) to (5a), we fixed the residual variance ($\sigma_{\varepsilon\,(\text{tiny value})}^2$) to a tiny value (0.0001). Thus, the variance estimates obtained for the interaction effects comprised both the true interaction as well as the error of adjusted means. We fixed the error variance not to the average error variance of the individual trials as in model (2), because we wanted to avoid that for highly stable genotypes a stability variance of zero is estimated. For simplicity and better readability we kept the notation from model (2) although the genotype-by-environment interaction variance (σ^2_{ge}) is of course no longer the pure genotype-by-environment interaction variance but a combination of the genotype-by-environment interaction variance and the error of adjusted means.

Differences in stability variance between groups were investigated by three different approaches. In the first approach the adjusted genotype yields were used for each group (hybrids, female lines, male lines, commercial lines) separately. The group-specific dataset was analyzed with the following model:

$$y_{ij.} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ij.} \tag{3a}$$

where $y_{ij.}$ was the adjusted yield of the ith genotype at the jth environment. The effect μ denoted the intercept and g_i the effect of the ith genotype, e_j the effect of the jth environment, and $(ge)_{ij}$ the genotype-by-environment interaction effect of the ith genotype with the jth environment. Intercept and genotype main effects were assumed to be fixed. The other effects were assumed to be random with independent distribution, zero mean and variances σ^2_e, σ^2_{ge}, and $\sigma^2_{\varepsilon\,(\text{tiny value})}$, where σ^2_e is the environmental variance, σ^2_{ge} the genotype-by-environment interaction variance, and $\sigma^2_{\varepsilon\,(\text{tiny value})}$ the error variance.

In the second approach, the whole dataset was used and the genotype-by-environment interaction variance was assumed to be heterogeneous for groups, i.e.,

$$\text{the variance of } (ge)_{ij} \text{ was } \sigma^2_{ge\,(t)} \text{ for the } t\text{th group} \tag{3b}$$

In the third approach group-by-environment interaction effects were added:

$$y_{tij.} = \mu + g_i + e_j + (ae)_{tj} + (ge)_{ij} + \varepsilon_{tij.}\,, \tag{3c}$$

where $y_{tij.}$ was the adjusted yield of the ith genotype at the jth environment belonging to the tth group and $(ae)_{tj}$ the group-by-environment interaction effect of the tth group with the jth environment. Group-by-environment interaction effects were assumed to be random with independent distribution, zero mean and variance σ^2_{ae}.

Pairwise differences between group-specific stability variances were investigated in models 3b and 3c with likelihood-ratio tests, where in the full model a separate stability for each group was assumed, but in the reduced model a common stability variance for the two groups under comparison. For the models 3b and 3c the AIC was calculated. The AIC was defined as minus two times the REML log-likelihood plus two times the number of variance parameters (Müller et al. 2010). Parameters that had an estimated variance of zero were included, but the residual parameter with the fixed variance, was not taken into account. A smaller AIC value denoted a better model fit.

Mean grain yield of individual genotypes across locations was estimated with the following model:

$$y_{tij.} = \mu + g_i + e_j + (ae)_{tj} + (ge)_{ij} + \varepsilon_{tij.} \tag{4a}$$

The model was identical to model (3c) with the exception that the genotype-by-environment interaction variance was homogeneous (σ^2_{ge}).

For estimating yield stability of individual genotypes we modified model (4a) to model (4b) by assuming a separate genotype-by-environment interaction variance for each genotype, i.e., the variance of $(ge)_{ij}$ was $\sigma^2_{ge_{(i)}}$. The genotype-by-environment interaction variance of the ith genotype ($\sigma^2_{ge_{(i)}}$) was also termed 'stability variance' of genotype i.

Estimating rank-correlation-heritability of mean grain yield and of stability variance in dependence of the number of test environments was performed following Mühleisen et al. (2014b). In order to display the variation in heritability among the 1,000 resampling runs (see Mühleisen et al. 2014b) we estimated in addition quantiles of the heritability by using the 0.25th and 0.75th quantile of the Spearman correlation coefficients.

Midparent heterosis was calculated as $\text{MPH} = \text{HYB} - (P_{\text{female}} + P_{\text{male}})/2$, better parent heterosis as $\text{BPH} = \text{HYB} - P_{\max}$, and commercial heterosis as $\text{CH} = \text{HYB} - \text{CL}_{\max}$, where HYB denotes the trait value of the hybrid, P_{female} the trait value of the female parent, P_{male} the trait value of the male parent, $P_{\max}$ the trait value of the parent with the higher trait value and $\text{CL}_{\max}$ the trait value of the inbred line with the highest trait value. For the stability variance the lower parent heterosis (LPH) and the lower commercial heterosis (LCH) were calculated using the minimum values as reference, since a small stability variance indicates the desired high yield stability.

In order to obtain estimates of GCA and SCA effects we fitted the following model:

$$\begin{aligned} y_{tipqj.} = {} & \mu + a_t + d_{\text{hybrid}} \times f_p + d_{\text{hybrid}} \times m_q + d_{\text{hybrid}} \times g_i \\ & + d_{\text{female}} \times g_i + d_{male} \times g_i + d_{\text{commercial line}} \times g_i \\ & + e_j + (ae)_{tj} + d_{\text{hybrid}} \times (fe)_{pj} \\ & + d_{\text{hybrid}} \times (me)_{qj} + d_{\text{hybrid}} \times (ge)_{ij} \\ & + d_{\text{female}} \times (ge)_{ij} + d_{\text{male}} \times (ge)_{ij} \\ & + d_{\text{commercial line}} \times (ge)_{ij} + \varepsilon_{tipqj.}. \end{aligned} \tag{5a}$$

This model is identical to model (2), with the exception that a_t was assumed to be fixed and the error variance was fixed to a small value (0.0001).

For estimating GCA and SCA effects of yield stability, we fitted the following model:

$$\begin{aligned} y_{tipqj.} = {} & \mu + g_i + e_j + (ae)_{tj} + d_{\text{hyf01}} \times (fe)_{01j} \\ & + d_{\text{hyf02}} \times (fe)_{02j} + \ldots d_{\text{hyf13}} \times (fe)_{13j} \\ & + d_{\text{hym01}} \times (me)_{01j} + d_{\text{hym02}} \times (me)_{02j}, \\ & + \ldots d_{\text{hym07}} \times (me)_{07j} + (ge)_{ij} + \varepsilon_{tipqj.}. \end{aligned} \tag{5b}$$

where d_{hyf01} to d_{hyf13} are dummy variables that are one for hybrids being offspring of the respective female parent. The effect $d_{\mathrm{hyf01}} \times (fe)_{01j}$ is the GCA effect of the first female parent of the stability variance, $d_{\mathrm{hym01}} \times (me)_{01j}$ the GCA effect of the first male parent of the stability variance, and $(ge)_{ij}$ are the SCA effects of the hybrids and the per se effects of female, male and commercial lines of yield stability.

The effects μ and g_i are assumed to be fixed. The other effects are assumed to be random with independent normal distribution, zero mean and variances σ_e^2, $\sigma_{a\times e}^2$, $\sigma^2_{\mathrm{GCA(female01)}\times e}$, …, $\sigma^2_{\mathrm{GCA(female13)}\times e}$, $\sigma^2_{\mathrm{GCA(male01)}\times e}$, …, $\sigma^2_{\mathrm{GCA(male07)}\times e}$, $\sigma^2_{\mathrm{SCA(hybrid01)}\times e}$, …, $\sigma^2_{\mathrm{SCA(hybrid91)}\times e}$, $\sigma^2_{g\mathrm{(female01)}\times e}$, …, $\sigma^2_{g\mathrm{(female13)}\times e}$, $\sigma^2_{g\mathrm{(male01)}\times e}$,…, $\sigma^2_{g\mathrm{(male07)}\times e}$, $\sigma^2_{g\mathrm{(commercial\ 01)}\times e}$,…, $\sigma^2_{g\mathrm{(commercial\ 30)}\times e}$, and $\sigma^2_{\varepsilon\mathrm{(tiny\ value)}}$.

Spearman's rank correlation coefficient was calculated between mean grain yield of hybrids (BLUEs from model 4a) and midparent performance (BLUEs from model 4a) as well as between mean grain yield of hybrids (BLUEs from model 4a) and the sum of parental GCA effects (BLUPs from model 5a).

For stability variance, Spearman's rank correlation coefficient was calculated between stability variance of hybrids (variance components from model 4b) and average stability variance of the two parents (variance components from model 4b) as well as between stability variance of hybrids (variance components from model 4b) and the sum of stability variances of parental GCA effects (variance components from model 5b).

The correlation between GCA effects and hybrid performance is potentially overestimated if the hybrid to be predicted is included in the estimation set. Therefore, for each hybrid, the GCA effect of its female and of its male parent, was estimated based on a reduced data set, excluding the respective hybrid following the suggestion of Schrag et al. (2009). For hybrids, female lines, male lines and commercial lines, the Spearman rank correlation coefficient was calculated between grain yield and stability variance. The association between grain yield and stability variance was also visually shown in a scatter plot. All statistical models were analyzed using the software package ASReml-R (Butler et al. 2009). Heritability, correlation coefficients, and heterosis were calculated using the statistical software R 3.1.0 (R Development Core Team 2014).

Results

The stability variance of particular genotypes depends on the reference population. As estimates of σ_g^2 and σ_{ge}^2 varied strongly among the four different genotype groups, we first analyzed the individual groups separately. Hybrids had the highest stability variance and were therefore less stable than female and male parental lines and commercial lines. Among the three groups of inbred lines, female parents showed the highest yield stability. One disadvantage of modeling the four groups separately is that significance tests of group comparisons are difficult to implement. Therefore, we studied the stability variance of the four groups in one model. In approach two, where solely the stability variance was assumed to be heterogeneous for groups, the results deviate severely from the split approach. The AIC was 1041.9. In approach three, where in addition, a group-by-environment interaction term was included, the AIC fell to 985.3 and the estimated stability variances were very close the estimates of the split approach. Therefore, in all further models we included group-by-environment interaction effects.

Rank-correlation-heritability of mean grain yield was already with three test environments high with correlations between observed and predicted mean values being larger than 0.6 (Fig. 1). For stability variance, observed mean values of rank-correlation-heritability varied between 0.03 and 0.06 for three to ten test environments. An extrapolation revealed that even with 20 environments average rank-correlation-heritability is not expected to exceed 0.2.

The average midparent heterosis of grain yield amounted to 3.0 % with a range from −15.0 to 11.5 % (Table 3). Better parent heterosis was on average negative amounting to −4.2 % with a maximum value of 8.4 %. The best inbred line included in the present study was the breeding line 'TIW841' with a grain yield performance of 9.93 Mg ha^{-1} (Supplementary Table S2). The best hybrid 'hybrid f09m02' yielded slightly less than 'TIW841' with a mean grain yield of 9.65 Mg ha^{-1}. Therefore, commercial heterosis was negative for all hybrids, ranging from −26.0 to −2.8 %.

The average midparent heterosis for stability variance was 30.2 % ranging from −76.3 to 268.5 %, i.e., dynamic yield stability was lower for hybrids than for lines. Better parent heterosis was even more in favor of the more yield stable lines in contrast to the hybrids. The line with the highest dynamic yield stability was the 'female f06' with a stability variance of 0.165 being less yield stable as the most stable hybrid 'hybrid f13m07' with a stability variance of 0.089 (supplementary Table S2). The most stable hybrid was also better in yield performance compared to the most stable line (Fig. 2).

For grain yield, midparent performance was moderately correlated with hybrid performance ($r_S = 0.51$; $P < 0.001$) (Fig. 3). Correlation between GCA-predicted and observed hybrid performance was not substantially higher with $r_S = 0.52$ ($P < 0.001$). For the stability variance, correlation between midparent performance and hybrid performance was not significantly ($P > 0.05$) larger than zero, but

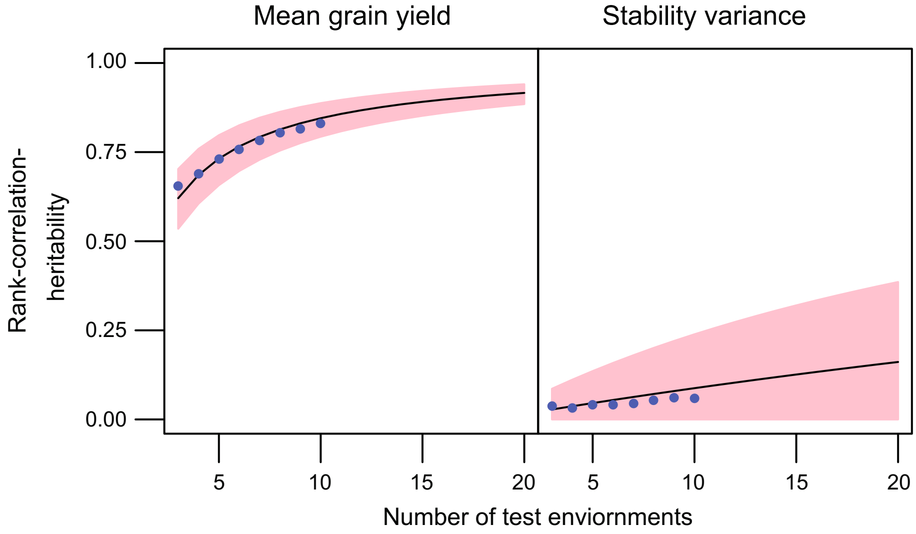

Fig. 1 Observed (*blue dots*) and predicted (*black lines*) values of the rank-correlation-heritability for mean grain yield and stability variance as well as the 0.25 and 0.75 quantiles (*rose area*) of the predicted heritability (color figure online)

GCA-predicted and observed hybrid performance was significantly correlated ($r_S = 0.23$, $P < 0.05$).

Spearman's rank correlation coefficient between mean grain yield and stability variance was for hybrids, female lines, male lines, and commercial lines not significantly ($P > 0.05$) larger than zero. The scatter plot between mean grain yield and stability variance revealed that the hybrids varied strongly in mean yield and yield stability (Fig. 2). Female parental lines had the lowest mean yields. In contrast, male and commercial lines had high yields. In yield stability, there were no obvious differences between the four groups, except that there were several hybrids exhibiting low yield stability.

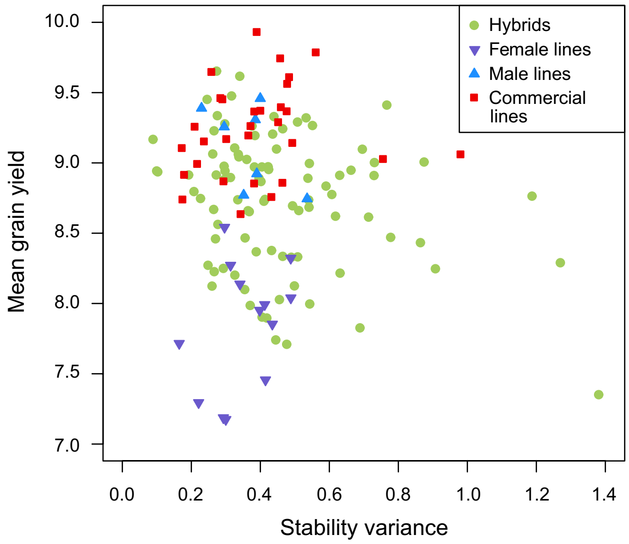

Fig. 2 *Scatter plot* between grain yield and stability variance for hybrids, female lines, male lines, and commercial lines evaluated for grain yield (Mg ha^{-1}) in up to 20 European environments

Discussion

Amount of heterosis was not associated with the yield level of environments

It has been hypothesized that the amount of heterosis is more pronounced in marginal environments (Jordaan 1996). We observed only a moderate range of average midparent heterosis for grain yield between −4.6 and 7.2 % among individual environments (Table 2) despite the use of diverse set of locations with naturally occurring abiotic stress (Mühleisen et al. 2013a). In addition, there was no significant correlation between yield level and average midparent heterosis (Table 2). Hence, our data did not support the hypothesis that heterosis is more pronounced in marginal environments.

CMS system potentially impairs heterosis for grain yield

Previous studies suggested that average midparent heterosis of CMS-based triticale hybrids (Gowda et al. 2013) is substantially lower than for CHA-based hybrids (Oettler et al. 2005; Fischer et al. 2010). Our results confirmed the finding that MPH in CMS-based hybrids is low and amounted on average only to 3 % (Table 3). Possible reasons are negative effects of the CMS cytoplasm and/or incomplete restoration. In the field, visible differences between isogenic triticale genotypes with normal versus CMS cytoplasm are limited to higher ability of tillering and around 2 days faster development of genotypes backcrossed to the CMS cytoplasm (Sigrid Weissmann, Saatzucht Dr. Hege, unpublished data 2014). Severe grain yield penalties of the *Triticumtimopheevii* Zhuk.cytoplasmin triticale have to our knowledge not yet been described elsewhere. Consequently, a comprehensive study is required examining the

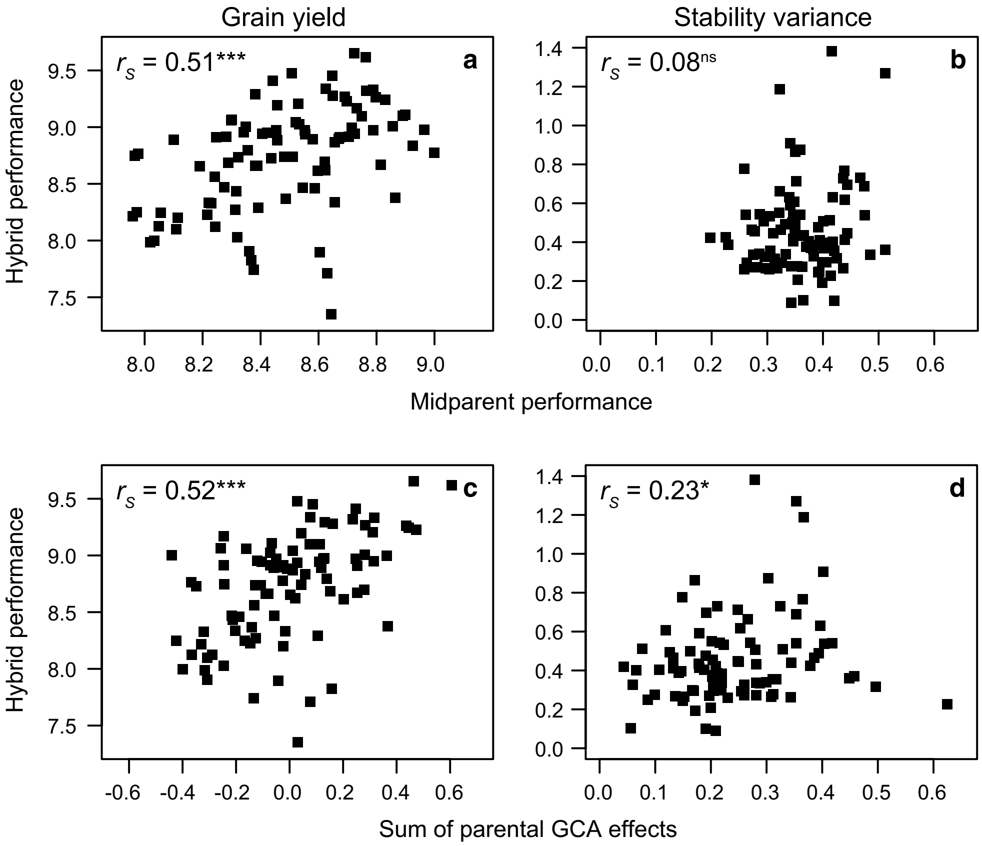

Fig. 3 Association between performance of 91 triticale hybrids and their parents' performance for mean grain yield (**a**) and stability variance (**b**). Association between hybrid performance and sum of the general combining ability (GCA) effects of their parents for mean grain yield (**c**) and stability variance (**d**). GCA effects for each hybrid were estimated in a reduced data set, excluding the respective hybrid

effects of the *Triticumtimopheevii* Zhuk.cytoplasm on grain yield performance in more detail.

Incomplete restoration was frequently observed in CMS originating from *Triticumtimopheevii* Zhuk. (Wilson and Driscoll 1983; Curtis and Lukaszewski 1993). This was also confirmed by previous experiments conducted in the hybrid triticale program of the State Plant Breeding Institute of the University of Hohenheim. Out of 3,000 European elite triticale inbred lines converted into CMS cytoplasm originating from *Triticumtimopheevii* Zhuk. Only 15 genotypes achieved full restoration (Hans Peter Maurer, unpublished data 2011). The above mentioned restoration tests were conducted in at least four locations and represent reliable estimates. The detected restorer lines have been used to build up the male pool. It is important to consider, however, that restoration ability is complex and depends also on the genetic background of the female lines (Wilson and Driscoll 1983). In the restoration tests at the State Plant Breeding Institute, however, male lines were in general only crossed with one female line and the amount of restoration of other females is largely unknown. Therefore, incomplete restoration cannot be ruled out in our experiments. A detailed survey investigating the effect of incomplete restoration on the seed set has not been undertaken but seems necessary to complement our understanding of the competiveness of CMS-based triticale hybrids.

In the European triticale germplasm, the frequency of partial restorer is higher than 95 % (Hans Peter Maurer, unpublished data, 2014). Therefore the development of complete restorer and maintainer lines is subject to a strong selection pressure. Furthermore, the conversion of maintainers into the CMS cytoplasm requires additional time. These aspects slow down the breeding progress in the development of parental lines and resulted in a negative commercial heterosis (Table 4). Systematic breeding of hybrid varieties including the development of heterotic groups might enhance the amount of commercial and of midparent heterosis (Fischer et al. 2010).

Estimation of stability variance in presence of groups

The dynamic yield stability of a genotype depends on the underlying reference population (Becker and Leon 1988). Despite this, dynamic yield stability of hybrids and lines have previously been contrasted ignoring the specificity of the reference populations (e.g., Borghi and Perenzin 1990; Peterson et al. 1997; Bruns and Peterson 1998; Oury et al. 2000; Koemel et al. 2004; Mühleisen et al. 2014a, 2014b) leading to potentially biased results. In line with this, we observed that the implementation of the estimation of yield stability severely influenced the contrasts of the four germplasm groups (Table 4). Our results clearly suggested that including group and group-by-environment interactions effects are crucial to properly analyze the different germplasm groups together in one biometrical model (Table 4). Data analyses have benefited from modeling group-by-environment interaction variance (Table 4 and results). Therefore, we discussed differences of yield stability among groups based on results obtained from Approach 3 considering group and group-by-environment interactions effects.

Table 2 Variance components [group, σ^2_a; hybrids, $\sigma^2_{g(\mathrm{hybrids})}$; general combining ability (GCA) of female parents, $\sigma^2_{\mathrm{gca\,(females)}}$; GCA of male parents, $\sigma^2_{\mathrm{gca\,(males)}}$; specific combining ability (SCA), σ^2_{sca}; females, $\sigma^2_{g\,(\mathrm{females})}$; males, $\sigma^2_{g\,(\mathrm{males})}$; commercial lines, $\sigma^2_{g\,(\mathrm{clines})}$; environment, σ^2_e; environment × group interaction, $\sigma^2_{a\times e}$; environment × hybrid interaction, $\sigma^2_{g(\mathrm{hybrids})\times e}$; environment × GCA of female parents interaction, $\sigma^2_{\mathrm{gca\,(females)}\times e}$; environment × GCA of male parents interaction, $\sigma^2_{\mathrm{gca\,(males)}\times e}$; environment × SCA interaction, $\sigma^2_{\mathrm{sca}\times e}$; environment × females interaction, $\sigma^2_{g\,(\mathrm{females})\times e}$; environment × male interaction, $\sigma^2_{g\,(\mathrm{males})\times e}$; environment × commercial lines interaction, $\sigma^2_{g\,(\mathrm{clines})\times e}$; pooled error of mean yields, $\sigma^2_{\varepsilon m\,(\mathrm{pooled})}$; pooled error of plot yields, $\sigma^2_{\varepsilon p\,(\mathrm{pooled})}$] estimated with model (2) for 141 winter triticale genotypes evaluated for grain yield (Mg ha^{-1}) in up to 20 environments

Source	Component	Standard error	*P* value
σ^2_a	0.383	0.323	<0.001
$\sigma^2_{g(\mathrm{hybrids})}$ [a]	0.207		
$\sigma^2_{\mathrm{gca\,(females)}}$	0.061	0.035	<0.001
$\sigma^2_{\mathrm{gca\,(males)}}$	0.029	0.026	0.039
σ^2_{sca}	0.117	0.022	<0.001
$\sigma^2_{g\,(\mathrm{females})}$	0.182	0.082	<0.001
$\sigma^2_{g\,(\mathrm{males})}$	0.072	0.052	<0.001
$\sigma^2_{g\,(\mathrm{clines})}$	0.09	0.029	<0.001
σ^2_e	1.606	0.526	<0.001
$\sigma^2_{a\times e}$	0.03	0.011	<0.001
$\sigma^2_{g(\mathrm{hybrids})\times e}$ [b]	0.366		
$\sigma^2_{\mathrm{gca\,(females)}\times e}$	0.124	0.015	<0.001
$\sigma^2_{\mathrm{gca\,(males)}\times e}$	0.122	0.019	<0.001
$\sigma^2_{\mathrm{sca}\times e}$	0.12	0.01	0.186
$\sigma^2_{g\,(\mathrm{females})\times e}$	0.231	0.034	<0.001
$\sigma^2_{g\,(\mathrm{males})\times e}$	0.248	0.048	<0.001
$\sigma^2_{g\,(\mathrm{clines})\times e}$	0.268	0.024	<0.001
$\sigma^2_{\varepsilon m\,(\mathrm{pooled})}$ [c]	0.124		
$\sigma^2_{\varepsilon p\,(\mathrm{pooled})}$ [d]	0.232		

[a] Sum of $\sigma^2_{\mathrm{gca\,(females)}}$, $\sigma^2_{\mathrm{gca\,(males)}}$, and σ^2_{sca}

[b] Sum of σ^2 $\sigma^2_{\mathrm{gca\,(females)}\times e}$, $\sigma^2_{\mathrm{gca\,(males)}\times e}$, and $\sigma^2_{\mathrm{sca}\times e}$

[c] Average error variance of means yields

[d] Average error variance of plot yields

Table 3 Relative values of mean, minimum and maximum heterosis of 91 triticale hybrids evaluated for grain yield (Mg ha^{-1}) in up to 20 European environments

	Mean (%)	Minimum (%)	Maximum (%)
Mean grain yield			
Midparent heterosis	3.0	−15.0	11.5
Better parent heterosis	−4.2	−18.2	8.4
Commercial heterosis	−12.1	−26.0	−2.8
Stability variance			
Midparent heterosis	30.2	−76.3	268.5
Lower parent heterosis	58.3	−71.7	366.0
Lower commercial heterosis	178.6	−46.1	737.8

Yield stability of hybrid triticale

Yield stability of hybrids is expected to outperform that of inbred lines because of general vigor and fitness advantages. This was confirmed in several studies for allogamous and partially allogamous crops (for review see Léon 1994). For autogamous crops, such as wheat, barley, and triticale, contrasting results were reported, ranging from higher yield stability of hybrids (Oury et al. 2000; Mühleisen et al. 2014a) to no differences in yield stability (Peterson et al. 1997; Bruns and Peterson 1998; Koemel et al. 2004). Mühleisen et al. (2014a) hypothesized that the latter three studies might have also found a higher yield stability of hybrids using a proper definition of the environmental index.

Table 4 Stability variance of four groups of winter triticale genotypes evaluated for grain yield (Mg ha^{-1}) in 20 European environments

	Single-cross hybrids	Female parental lines	Male parental lines	Commercial lines
Approach 1	0.455	0.346	0.387	0.392
Approach 2	0.463^{a}	0.451^{a}	0.369^{a}	0.417^{a}
Approach 3	0.455^{b}	0.353^{a}	0.370^{ab}	0.392^{a}

Stability variance was estimated with four different approaches. Stability variances within the same row with no common letter were significantly different ($P < 0.05$)

We found a significantly lower yield stability of hybrids compared to female parents and commercial lines (Table 3). This can be explained by the potential negative effects of the CMS system. Assuming that the lower yields were caused by incomplete restoration of CMS cytoplasm, it is likely that the amount of partial restoration was not constant across environments (Wilson and Driscoll 1983). This would reduce the yield stability of hybrids. Therefore, the presented results should not be regarded as proof of lower yield stability of hybrids in contrast to lines. Our results rather point to the strong need to optimize the CMS system enabling to produce high yielding and yield stable triticale hybrids.

Prediction of hybrid yield and yield stability

Selection of superior hybrid combinations is crucial for efficient hybrid breeding. Evaluation of all possible hybrid combinations between the parents exceeds the capacity of a single hybrid triticale breeding program and breeders need a reliable method for predicting the performance of untested hybrids. Hybrid performance and midparent performance were moderately correlated with $r_S = 0.51$ (Fig. 3), which is lower than the correlation of 0.66 reported by Fischer et al. (2010) for CHA-based hybrids. GCA-based hybrid prediction is assumed to be more precise than midparent prediction because GCA effects consider additionally dominance effects (Mühleisen et al. 2013a). In accordance with the expectation, Fischer et al. (2010) reported a higher correlation of 0.84. In contrast, we observed only a slight increase of 0.01 (Fig. 3). One reason for the observed discrepancy can be that we used a leave-one-out cross-validation which was not applied by Fischer et al. (2010). A re-analysis of the GCA prediction revealed a non-cross-validated correlation of 0.72 in our study which is substantially biased compared to the cross-validated accuracy of 0.52. This clearly underlines the strong need to use cross-validations in order to obtain reliable estimates of the potential to predict hybrid performance based on GCA effects.

Since yield and stability variance were independent (Fig. 2), separate selection for yield stability is required. Yield stability of the hybrids was not associated with the average stability of the parents, but with the sum of the parental GCA effects (Fig. 3). The correlation was with 0.23, however, much lower than observed for hybrid performance. Despite this our results indicate that pre-selection of parents with yield stability should be performed based on their GCA effects rather than on their per se performance.

Conclusions

Yield performance and stability of CMS-based triticale hybrids lag behind the expectations based on comparable studies investigating CHA-based triticale hybrids. The major challenge in breeding CMS-based hybrid triticale will be to identify and eliminate potential disadvantages of the CMS system. For this purpose a detailed study on the reasons for the drawback of the currently existing CMS system in triticale is urgently required comprising also the search of potentially alternative hybridization systems.

Author contribution statement J Mühleisen was the leading researcher, including statistical analyses, literature review, and writing of the manuscript. JC Reif and HP Maurer designed the project. HP Piepho provided statistical support throughout the analyses. HP Piepho assisted in writing the manuscript. JC Reif participated in writing the manuscript and reviewed recent literature.

Acknowledgments We thank Jens Möhring for statistical consulting. Jonathan Mühleisen was supported by BMELV/BLE withinthe „Züchtung von Triticalesorten für extreme Umwelten—eine Frage des Sortentyps?“ project (Grant ID: 2814502410).

Conflict of interest The authors declare that they have no conflict of interest.

References

Ahokas H (1979) Cytoplasmic male sterility in barley. Acta Agric Scand 29:219–224

Becker HC (1987) Zur Heritabilität statistischer Maßzahlen für die Ertragssicherheit. VortrPflanzenzüchtg 12:134–144

Becker HC, Leon J (1988) Stability analysis in plant breeding. Plant Breed 101:1–23

Borghi B, Perenzin M (1990) Yield and yield stability of conventional varieties and F1 bread wheat hybrids. J Genet Breed 44:307–310

Bruns R, Peterson CJ (1998) Yield and stability factors associated with hybrid wheat. Euphytica 100:1–5

Butler D, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R, reference manual. Version 3, Queensland department of primary industries and fisheries, Brisbane

Coors JG, Pandey S (1999) The genetics and exploitation of heterosisin crops, ASA, CSSA, and SSSA, Madison

Curtis CA, Lukaszewski AJ (1993) Localization of genes in rye that restore male fertility to hexaploid wheat with timopheevi cytoplasm. Plant Breed 111:106–112

Eberhart ST, Russell WA (1966) Stability parameters for comparing varieties. Crop Sci 6:36–40

Fischer S, Maurer HP, Würschum T, Möhring J, Piepho HP, Schön CC, Thiemt EM, Dhillon BS, Weissmann EA, Melchinger AE, Reif JC (2010) Development of heterotic groups in triticale. Crop Sci 50:584–590

Gowda M, Kling C, Würschum T, Liu W, Maurer HP, Hahn V, Reif JC (2010) Hybrid breeding in durum wheat: heterosis and combining ability. Crop Sci 50:2224–2230

Gowda M, Zhao Y, Maurer HP, Weissmann EA, Würschum T, Reif JC (2013) Best linear unbiased prediction of triticale hybrid performance. Euphytica 191:223–230

Jalaluddin MD, Harrison SA (1993) Repeatability of stability estimators for grain yield in wheat. Crop Sci 33:720–725

Jordaan JP (1996) Hybrid wheat: advances and challenges. In: Reynolds MP, Rajaram S, and McNabA (eds) Increasing yield potential in wheat: breaking the barriers, CIMMYT, Mexico, pp 66–75

Kempe K, Rubtsova M, Gils M (2014) Split-gene system for hybrid wheat seed production. Proc Natl Acad Sci USA 111:9097–9102

Koemel JE, Guenzi AC, Carver BF, Payton ME, Morgan GH, Smith EL (2004) Hybrid and pureline hard winter wheat yield and stability. Crop Sci 44:107–113

Kumar S, Singh O, Rheenen HV, Rao KVS (1998) Repeatability of different stability parameters for grain yield in chickpea. Plant Breed 117:143–146

Léon J (1994) Mating system and the effect of heterogeneity and heterozygosity on phenotypic stability. In: van Ooijen JW, Jansen J (eds) Biometrics in plant breeding: applications of molecular markers. Proceedings of the 9th meeting of the EUCARPIA section biometrics in plant breeding, Wageningen, pp 19–31

Léon J, Becker HC (1988) Repeatability of some statistical measures of phenotypic stability—correlation between single year results and multi years results. Plant Breed 100:137–142

Lin CS, Binns MR, Lefkovitch LP (1986) Stability analysis: where do we stand? Crop Sci 26:894–900

Longin CFH, Mühleisen J, Maurer HP, Zhang H, Gowda M, Reif JC (2012) Hybrid breeding in autogamous cereals. Theor Appl Genet 125:1087–1096

Longin CFH, Gowda M, Mühleisen J, Ebmeyer E, Kazman E, Schachschneider R, Schacht J, Kirchhoff M, Zhao Y, Reif JC (2013) Hybrid wheat: quantitative genetic parameters and consequences for the design of breeding programs. TheorAppl Genet 126:2791–2801

Longin CF, Reif JC, Würschum T (2014) Long-term perspective of hybrid versus line breeding in wheat based on quantitative genetic theory. Theor Appl Genet 127:1635–1641

McMaster GS, Wilhelm WW (1997) Growing degree-days: one equation, two interpretations. Agric For Meteorol 87:291–300

Mühleisen J, Maurer HP, Stiewe G, Bury P, Reif JC (2013a) Hybrid breeding in barley. Crop Sci 53:819–824

Mühleisen J, Reif JC, Maurer HP, Möhring J, Piepho HP (2013b) Visual scorings of drought stress intensity as covariates for improved variety trial analysis. J Agron Crop Sci 199:321–330

Mühleisen J, Piepho HP, Maurer HP, Longin CFH, Reif JC (2014a) Yield stability of hybrids versus lines in wheat, barley, and triticale. Theor Appl Genet 127:309–316

Mühleisen J, Piepho HP, Maurer HP, Zhao Y, Reif JC (2014b) Exploitation of yield stability in barley. Theor Appl Genet 127:1949–1962

Müller BU, Kleinknecht K, Möhring J, Piepho HP (2010) Comparison of spatial models for sugar beet and barley trials. Crop Sci 50:794–802

Oettler G, Tams SH, Utz HF, Bauer E, Melchinger AE (2005) Prospects for hybrid breeding in winter triticale: I. Heterosis and combining ability for agronomic traits in European elite germplasm. Crop Sci 45:1476–1482

Oury FX, Brabant P, Berard P, Pluchard P (2000) Predicting hybrid value in bread wheat: biometric modeling based on a top-cross design. Theor Appl Genet 100:96–104

Peterson CJ, Moffatt JM, Erickson JR (1997) Yield stability of hybrid vs. pureline hard winter wheats in regional performance trials. Crop Sci 37:116–120

Pham HN, Kang MS (1988) Interrelationships among and repeatability of several stability statistics estimated from international maize trials. Crop Sci 28:925–928

Piepho HP, Möhring J, Schulz-Streeck T, Ogutu JO (2012) A stagewise approach for the analysis of multi-environment trials. Biom J 54:844–860

R Core Team (2014) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna. URL: http://www.R-project.org/

Robert N (2002) Comparison of stability statistics for yield and quality traits in bread wheat. Euphytica 128:333–341

Schrag TA, Möhring J, Maurer HP, Dhillon BS, Melchinger AE, Piepho HP, Sørensen AP, Frisch M (2009) Molecular marker-based prediction of hybrid performance in maize using unbalanced data from multiple experiments with factorial crosses. Theor Appl Genet 118:741–751

Shukla GK (1972) Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29:237–245

Sneller CH, Kilgore-Norquest L, Dombek D (1997) Repeatability of yield stability statistics in soybean. Crop Sci 37:383–390

Whitford R, Fleury D, Reif JC, Garcia M, Okada T, Korzun V, Langridge P (2013) Hybrid breeding in wheat: technologies to improve hybrid wheat seed production. J Exp Bot 64:5411–5428

Wilson P, Driscoll CJ (1983) Hybrid wheat. In: Frankel R (ed) Heterosis: reappraisal of theory and practice. Monogr Theor Appl. Genet, vol 6. Springer, New York, pp 94–123

Zhao Y, Mette MF, Reif JC (2014) Genomic selection in hybrid breeding. Plant Breed. doi:10.1111/pbr.12231