Check for updates

Received: 30 November 2023 | Accepted: 9 April 2024

DOI: 10.1002/tpg2.20464

The Plant Genome

ORIGINAL ARTICLE

Genomic prediction of synthetic hexaploid wheat upon tetraploid *durum* and diploid *Aegilops* parental pools

Susanne Dreisigacker[1] | **Johannes W. R. Martini**[2] | **Jaime Cuevas**[3] | **Paulino Pérez-Rodríguez**[4] | **Nerida Lozano-Ramírez**[1] | **Julio Huerta**[1] | **Pawan Singh**[1] | **Leonardo Crespo-Herrera**[1] | **Alison R. Bentley**[5] | **Jose Crossa**[1,4]

[1]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, México

[2]Aardevo B.V., Nagele, The Netherlands

[3]Universidad Autónoma del Estado de Quintana Roo, Chetumal, México

[4]Colegio de Postgraduados, Campus Montecillos, Texcoco, México

[5]Australian National University, Research School of Biology, Canberra, Australia

Correspondence
Jose Crossa, International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Edo. de México, México.
Email: j.crossa@cgiar.org

Funding information
Bill and Melinda Gates Foundation, FCDO, Grant/Award Number: INV-003439

Abstract

Bread wheat (*Triticum aestivum* L.) is a globally important food crop, which was domesticated about 8–10,000 years ago. Bread wheat is an allopolyploid, and it evolved from two hybridization events of three species. To widen the genetic base in breeding, bread wheat has been re-synthesized by crossing durum wheat (*Triticum turgidum* ssp. *durum*) and goat grass (*Aegilops tauschii* Coss), leading to so-called synthetic hexaploid wheat (SHW). We applied the quantitative genetics tools of "hybrid prediction"—originally developed for the prediction of wheat hybrids generated from different heterotic groups — to a situation of allopolyploidization. Our use-case predicts the phenotypes of SHW for three quantitatively inherited global wheat diseases, namely tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB). Our results revealed prediction abilities comparable to studies in 'traditional' elite or hybrid wheat. Prediction abilities were highest using a marker model and performing random cross-validation, predicting the performance of untested SHW (0.483 for SB to 0.730 for TS). When testing parents not necessarily used in SHW, combination prediction abilities were slightly lower (0.378 for SB to 0.718 for TS), yet still promising. Despite the limited phenotypic data, our results provide a general example for predictive models targeting an allopolyploidization event and a method that can guide the use of genetic resources available in gene banks.

Plain Language Summary

Bread wheat is an allopolyploid from two hybridization events of three species. To widen the genetic base in breeding, bread wheat has been re-synthesized by

Abbreviations: CIMMYT, International Maize and Wheat Improvement Center; CV1, cross-validation 1; CV2, cross-validation 2; CV3, cross-validation 3; DW, durum wheat; GBLUP, genomic best linear unbiased predictor; GCA, general combining ability; GK, Gaussian kernel; GM, genomic model; PGM-1, pedigree + genomic model; PGM-2, pedigree + genomic model; PGM-3, pedigree + genomic model; PM, pedigree model; SB, spot blotch; SCA, specific combining ability; SHW, synthetic hexaploid wheat; SNB, septoria nodorum blotch; SNP, single nucleotide polimorphism; TS, tan spot.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2024 International Maize and Wheat Improvement Center (CIMMYT). *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

crossing durum wheat and goat grass, leading to so-called synthetic hexaploid wheat (SHW). We applied the "hybrid prediction" to a situation of allopolyploidization. We performed genomic prediction phenotypes of SHW for three quantitatively inherited global wheat diseases, tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB). Our results revealed comparable genomic prediction abilities to studies in "traditional" elite or hybrid wheat.

1 | INTRODUCTION

Polyploidy, a prevalent occurrence in the Plantae kingdom, refers to the condition in which a cell has more than two copies of each homologous chromosome. While the evolutionary advantage of polyploidy remains a topic of discussion, it is commonly associated with amplified growth vigor and is believed to enhance the adaptive potential of species (Van de Peer et al., 2017, 2021). Several major food crops are polyploids, either as allopolyploids such as cotton, rapeseed, and wheat or as autopolyploids, as seen in the majority of commercial potato varieties (Chen, 2007). Allopolyploidy refers to the merging of the genomes of two (or more) different species, and it involves intergenomic interaction, which may produce new phenotypic variation and heterosis (Chen, 2007; Ni et al., 2009).

Bread wheat (*Triticum aestivum* L.) is an allohexaploid that was formed between 8000 and 10,000 years ago as a hybrid between the early cultivated allotetraploid *Triticum turgidum* ssp. *durum* Desf. MacKey ($2n = 4x = 28$, AABB) and the diploid goat grass *Aegilops tauschii* Coss ($2n = 2x = 14$, DD), followed by spontaneous chromosome doubling (Kihara, 1944). After its domestication, hexaploid wheat emerged as a superior form, surpassing its tetraploid and diploid ancestors. Hexaploid wheat grows more robustly, confers broader adaptability, provides an enhanced ability against pathogens, and offers versatile end products (Dubcovsky & Dvorak, 2007). Through centuries of cultivation and selection, hexaploid wheat has become a staple food crop.

Global wheat production amounts to ~770 metric tons annually (FAO, 2020. Over two-thirds of its production is devoted to food consumption. This significant contribution establishes wheat as a vital dietary component for an estimated 35% of the global population (Grote et al., 2021; Shiferaw et al., 2013). Considering the expected increase in the demand for wheat—to keep up with the world's increasing population and to counter global crises and unrest that trigger volatile markets—accelerated wheat improvement and enhanced productivity are required. Moreover, global warming has already become a worldwide threat, with more frequent and extreme weather events, including severe droughts or heat stress (Gaupp et al., 2020; Kornhuber et al., 2020; Sarhadi et al., 2018). New pathogen strains are emerging more frequently, posing difficult challenges for breeding new advanced wheat varieties and global wheat cultivation (Nnadi & Carter, 2021).

The polyploidization event that bread wheat underwent in its domestication history can be repeated ("synthesized") by interspecific hybridization followed by spontaneous chromosome doubling (Mujeeb-Kazi & Hettel, 1995; Figure 1). This artificially formed nascent bread wheat is named "synthetic hexaploid wheat" (SHW). SHW can be designed to capture novel genetic diversity from tetraploid and diploid relatives of wheat that is not found in "traditional" bread wheat. With no significant reproduction barrier, SHW lines are useful to introduce agronomically needed traits into bread wheat from available genetic resources (Singh et al., 2018).

SHW often exhibits increased resilience to biotic stresses and thus provides an opportunity to discover a new resistance to many pathogens (Aberkane et al., 2020). In particular, wheat foliar diseases have become increasingly importance in recent years due to various factors, such as the threat of climate change and the widespread cultivation of susceptible wheat varieties on a commercial scale (Lozano-Ramirez et al., 2022a; Singh et al., 2010). The causal agents of the diseases infect multiple wheat tissues, such as root, stem, leaf, spike, and grain. Based on the frequency and severity levels of disease epidemics, the diseases that infect leaf and spike are considered more important (Lozano-Ramirez et al., 2022a). Some of the most important foliar diseases are "tan spot" (TS), caused by *Pyrenophora tritici-repentis* Drechs (anamorph *Drechslera tritici-repentis* Shoem.); "septoria nodorum blotch" (SNB), caused by *Parastagonospora nodorum* (syn. ana, *Stagonospora*; teleo. Phaeosphaeria) (Berk.) Quaedvlieg, Verkley, and Crous; and "spot blotch" (SB), caused by *Bipolaris sorokiniana*.

While many SHWs are characterized by the resistance to diverse pathogens, it has been difficult to forecast the resistance phenotype based on the relative performance of the parents. Possible contributing factors reported in previous studies are distinct levels of genetic and epigenetic alternations, changing gene expression patterns, and aneuploidy during the synthesis of SHW (Chen et al., 2023; Eagles et al., 2011; Li et al., 2018; Mestiri et al., 2010; Vasudevan et al., 2023; Yu et al., 2017). In addition, the nature and growth of the wild *Ae. tauschii* species inhibit reliable phenotypic screening, which

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

means that the degree of resistance/susceptibility of the relevant parent itself cannot be characterized. The main criteria to select *Ae. tauschii* as a parent to synthesize a novel SHW have therefore been focused on capturing diverse material from accessions in gene bank collections, facing the key challenge of countless cross-combination options (Aberkane et al., 2020).

The prediction of the performance of hybrid combinations of different heterotic groups, such as the one used in maize (Technow et al., 2012), involves a similar task of predicting the most promising combinations of parents, only within and not between species. Genomics-enabled prediction, including pedigree and marker information, has recently shown reliable results in the prediction of hybrid performance from the wheat parents forming the single crosses (Basnet et al., 2019; Rembe et al., 2019; Zhao et al., 2013). Genomic-enabled prediction methods could also help harness the genetic diversity provided by tetraploid wheat and *Ae. tauschii* collections as a special case of the general attempt to use predictive breeding methods to harness the potential of germplasm collections (Martini et al., 2021; Yu et al., 2016). In this research, we aimed to model the variation of SHW quantitatively, with genomic hybrid prediction considering durum wheat (DW) and *Ae. tauschii* parents as two different parental pools.

Core Ideas

- Genomic prediction of synthetic hexaploid wheat (SHW) from tetraploid durum wheat and diploid *Aegilops* parental pools was investigated.
- To widen the genetic base in breeding, bread wheat has been re-synthesized by crossing durum wheat and goat grass.
- We applied the "hybrid prediction" to a situation of allopolyploidization.
- Our use-case predicts the phenotypes of SHW for three quantitatively inherited wheat diseases.

2 | MATERIALS AND METHODS

2.1 | Plant materials inoculation procedures and phenotypic data measurements

A total of 443 SHW lines were available (Table S1), derived from an interspecific cross between DW ($2n = 4x = 28$, AABB, *T. turgidum* L.) and goat grass ($2n = 2x = 14$, DD, *A. tauschii* Coss.) (Figure 1). The lines were created by the International Maize and Wheat Improvement Center

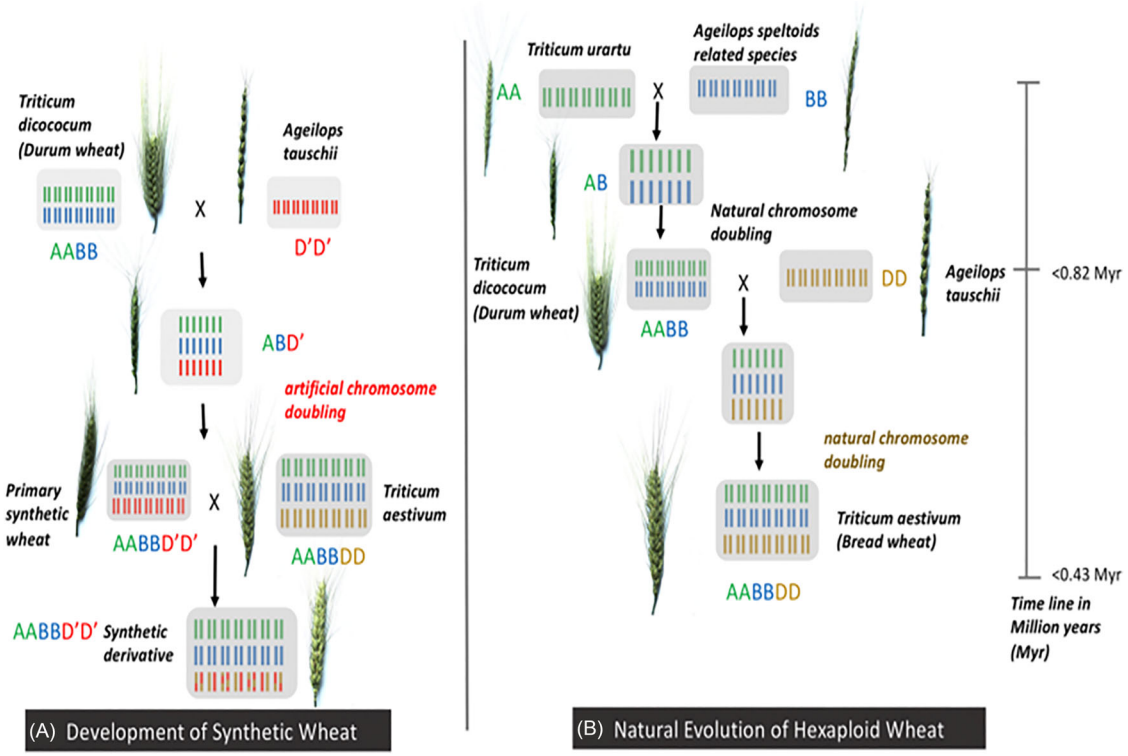


FIGURE 1 Development of synthetic hexaploidy wheat in comparison with the natural evolution of bread wheat (adopted from Rosyara et al., 2019).

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(CIMMYT) in its "Wheat Wide Crosses Program" between 1988 and 2010. The SHW included 57 different DW parents and 281 *Ae. tauschii* accessions (Table S1).

The SHW, along with the DW parents, was evaluated for four globally important diseases: (1) TS, caused by *P. tritici-repentis*, a foliar disease that has become a major threat within the complex of leaf spotting diseases and has increased worldwide over the last several decades; (2) SNB, caused by *P. nodorum*, which occurs in wheat-growing areas worldwide, and especially in areas with warm and moist weather; and (3) SB, from the fungal pathogen *B. sorokiniana*, that causes enormous losses in areas with climates characterized by high temperatures and humidity at late growth stages such as in eastern India, Southeast Asia, Latin America, Nepal, China, and Africa.

The assessment of the leaf-spotting diseases (TS, SNB, and SB) was performed under greenhouse conditions at the seedling stage in the CIMMYT headquarters in El Batan, Mexico (19°31′N, 98°50′W; elevation of 2249 m above sea level) in 2018–2019. The SHW and DW seeds were scarified to break dormancy and obtain an even germination. The experiments were arranged in a randomized complete block design with 12 replicates for TS and SNB and six replicates for SB. For each entry, four plants were grown in plastic containers as experimental units to derive mean values for subsequent analysis. The seedlings were grown under controlled environmental conditions with an air temperature of 22–25°C/16–18°C (day/night) and a 16-h photoperiod. Checks for TS and SNB evaluation included the bread wheat cultivars Erik (resistant), 6B-662 (moderately resistant), 6B-365 (moderately susceptible), and Glenlea (susceptible). For SB, checks used were Chirya (resistant), Francolin (moderately susceptible), Ciano F79 (susceptible), and Sonalika (susceptible).

For TS, the Mexican PTR isolate CIMFU 531-PTR1 (race 1) was applied. The isolate was grown in a V8-PDA medium (Lamari & Bernier, 1989), and conidia concentration for inoculation was adjusted to 4×10^3 spores mL^{-1} using a Fuchs-Rosenthal counting chamber. One drop of Tween 20 (a surfactant reagent) was added to every 100 mL of the spore suspension. The Mexican isolate CIMFU463-SN4 was used to evaluate SNB resistance. The isolate was adjusted to a concentration of 10^7 conidia per milliliter in a Neubauer counting chamber. One milliliter of Tween 20 was added to every 100 mL of isolate suspension for inoculation. Similarly, the Mexican isolate CIMFU483-BSG4M2 was used for the inoculation of SB. This isolate was adjusted to a concentration of 7500 conidia per milliliter in the same Neubauer chamber, and 1 mL Tween 20 was added to every 100 mL isolate suspension for inoculation.

The seedlings were inoculated with the respective conidial suspension for every disease in the two-leaf stage when the second leaf was fully expanded, or 2 weeks after sowing. Isolates were induced until the leaves were at a dew point, and this process was carried out four times using a hand sprayer. When the leaves were dry, the trays were moved to a mist chamber (relative humidity 100%, 21–22°C) to facilitate infection. After 24 h (48 h in the case of SB), the plants were transferred back to the greenhouse bench. Seedling response was evaluated 7 days after inoculation by following the 1–5 lesion rating scale developed by Lamari and Bernier (1989), which is based on the lesion type shown on the secondary leaf where 1 indicates resistant, 2 indicates moderately resistant, 3 indicates moderately resistant to moderately susceptible, 4 indicates susceptible, and 5 indicates highly susceptible.

2.2 | DNA isolation and sequencing

Fresh young leaves bulked from five individual plants per SHW were used for DNA extraction using a modified cetyltrimethylammonium bromide method (Dreisigacker et al., 2016). The DNA was quantified with a NanoDrop 8000 spectrophotometer V 2.1.0, and all entries were genotyped with the DArTseq technology at the Genetic Analysis Service for Agriculture (SAGA) laboratory in CIMMYT in Mexico (Sansaloni et al., 2011). For this study, we excluded single nucleotide polymorphisms (SNPs) with >15% of missing values and minor allele frequencies of 5%; missing genotypes were imputed with samples from the marginal distribution of observed marker genotypes based on the sample mean of each marker already recoded to its corresponding numeric values (0,1,2). The final 4169 high-quality SNPs were used for further analyses. Seeds from the DW and *Ae. tauschii* parents were, in many cases, no longer accessible. Searching for historical genotypic data generated by the CIMMYT gene bank revealed SNP data for only a limited number of 32 DW parents and 27 *Ae. tauschii* accessions, which formed a total of 147 SHW. SNP data of parental lines were therefore not used.

2.3 | Statistical models

Two stages were used in this study. In the first stage, a phenotypic ordinal linear model was fitted with the aim of estimating the effects of the SHW cultivars (l_i) (Alvarado et al., 2020). In the second stage, models for the genomic adjustment and prediction of missing cultivars were applied. Both phenotypic and genomic models presented were fitted using the BGLR package (Pérez & de los Campos, 2014), comprising 50,000 iterations. The initial 5000 iterations were discarded, and a thinning of 5 was applied to minimize random errors.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

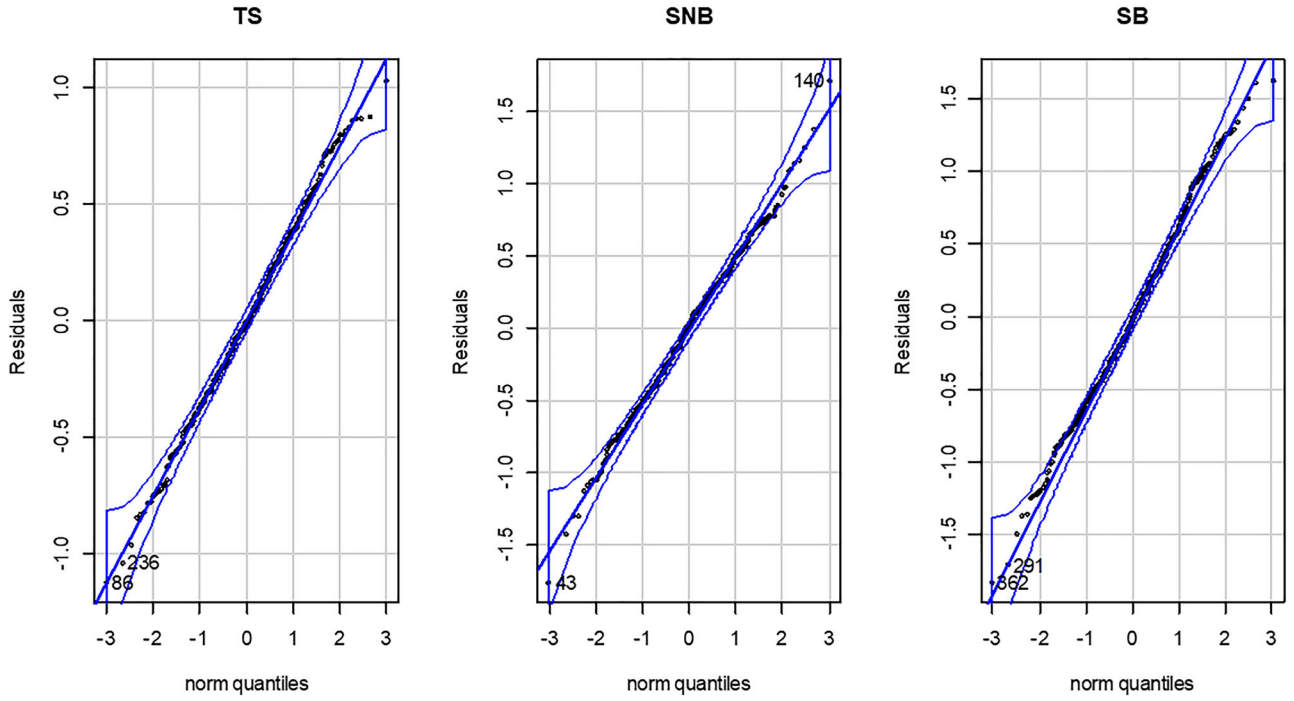


FIGURE 2 Residuals with a normal distribution obtained when fitting genomic model (GM) (model 2) for traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB).

2.4 | Phenotypic model for disease traits TS, SNB, and SB

We estimated the effects of the SHW cultivars (l_i), using a linear model for ordinal traits, where the response can take values on C ordered values, $y_{ij} \in \{1, \ldots, C\}$. We used the probit link, and the probability of each observation belonging to each category is given by:

$$P\left(y_{ij} = c\right) = \Phi\left(\eta_{ij} - \gamma_c\right) - \Phi\left(\eta_{ij} - \gamma_{c-1}\right), \quad (1)$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable, $\eta_{ij} = r_j + l_i$, which corresponds to the linear predictor, which includes the effect of replicates (r_j) and SHW lines (l_i); γ_c are threshold parameters, with $\gamma_0 = -\infty$, $\gamma_c \geq \gamma_{c-1}$, $\gamma_C = \infty$. The model was fitted using the BGLR package (Pérez & de los Campos, 2014) using data augmentation; that is, the latent variables (liabilities) L_{ij} are introduced in the model as follows: $L_{ij} = \eta_{ij} + \varepsilon_{ij}$, where ε_{ij} are independent and identically distributed random variables with mean 0 and variance 1. For more details about the threshold model, see Gianola (1982).

Note that the effects of cultivars (l_i) computed in stage 1 are considered response variables to be used in stage 2 when performing the genomic prediction. In stage 2 (see the description of the various genomic models below), the response for traits TS, SNB, and SB obtained from stage 1 is considered a continuous variable with a symmetric distribution that guarantees a normal distribution of the residuals (Figure 2).

2.5 | Genomic prediction models with pedigree and markers

Pedigree information was available for the DW parents and the SHW lines. The accession names of the *Ae. tauschii* parents were adopted assuming no pedigree relationship. We created prediction models that included normally distributed random terms with covariance defined by pedigree or genomic information. Thus, models implemented in this study attempt to capture (1) additive gene effects from the two parents, DW and *Ae. tauschii* (general combining ability, GCA), and (2) nonadditive effects due to the DW × *Ae. tauschii* cross to represent variability due to intra-locus interaction (dominance) and inter-loci interaction (epistasis) (specific combining ability, SCA).

The metric used to assess the model's prediction ability is the average Pearson's correlation between observed and predicted values. In this study, a fivefold random cross-validation (CV) was used to obtain the Pearson's correlation. Genomic prediction was assessed considering three main random CVs explained below. Model comparisons were carried out for the mean differences between the various pedigree model (PM), genomic model (GM), PGM-1, PGM-2, and PGM-3 models.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.5.1 | Pedigree model (PM)

Only the pedigree information represented by the numerical relationship matrix ($\boldsymbol{A}$) was used in this model to determine the genetic additive relationship between DW parents, *Ae. tauschii* parents, and their crosses forming the SHW. The relationship matrix was derived by the coefficient of co-ancestry computed using the BROWSE software (McLaren et al., 2000, 2005). The full genetic model to assess the GCA of the DW and *Ae. tauschii* parents and their SCA is described by the following linear mixed model:

$$\boldsymbol{y} = \mu 1 + \boldsymbol{Z}_{d_p}\boldsymbol{d}_p + \boldsymbol{Z}_{t_I}\boldsymbol{t}_I + \boldsymbol{Z}_{c_p}\boldsymbol{c}_p + \boldsymbol{e}, \quad (2)$$

where matrices $\boldsymbol{Z}_{d_p}$, $\boldsymbol{Z}_{t_I}$, and $\boldsymbol{Z}_{c_p}$ are the incident matrices that relate the values of random vectors $\boldsymbol{d}_p, \boldsymbol{t}_I$, and $\boldsymbol{c}_p$ to the response variable $\boldsymbol{y}$, μ is an intercept (fixed effect), $\boldsymbol{d}_p \sim \text{MN}(0, \sigma^2_{d_p}\boldsymbol{A}_d)$ is the random vector for the GCA of the DW (d subscript) obtained from the numerator relationship matrix ($\boldsymbol{A}_d$) derived from the coefficient of coancestry (the subscript p denotes pedigree) and assumed to be distributed following a multivariate normal (MN) random variable with $\sigma^2_{d_p}$ as the additive parameter variance estimated from the PM. The term $\boldsymbol{t}_I \sim \text{MN}(0, \sigma^2_{t_I}\boldsymbol{I})$ is a random vector for the GCA of the *Ae. Tauschii* parents (t subscript) assuming that they are normal, independent, and identically distributed (*niid*) with a mean of 0 and a variance of $\sigma^2_{t_I}$ being the additive variance for the *Ae. tauschii* parents with no relationship ($\boldsymbol{I}$ denotes the identity matrix). Furthermore, the term $\boldsymbol{c}_p \sim \text{MN}(0, \sigma^2_{c_p}\boldsymbol{A}_c)$ is the random vector of the SCA of the DW × *Ae. tauschii* crosses forming SHW. The relationship matrix ($\boldsymbol{A}_c$) is derived from the coefficient of co-ancestry (subscript p denotes pedigree and subscript c, the cross). The variance component associated with the cross is identified as $\sigma^2_{c_p}$. The random residual term $\boldsymbol{e}$ is distributed as $\text{MN}(0, \sigma^2_e\boldsymbol{I})$ with σ^2_e being the variance of the residuals. We also assume that the random terms $\boldsymbol{d}_p$, $\boldsymbol{t}_I$, $\boldsymbol{c}_p$, and $\boldsymbol{e}$ are independent.

2.5.2 | Genomic model (GM)

The information in this model only uses the SNP data of the SHW lines because only limited SNP data for parental lines were available. The linear model is given by:

$$\boldsymbol{y} = \mu 1 + \boldsymbol{Z}_{c_m}\boldsymbol{c}_m + \boldsymbol{e}, \quad (3)$$

where $\boldsymbol{c}_m \sim \text{MN}(0, \sigma^2_{c_m}\boldsymbol{K}_c)$ is the random vector of the SHWs, and $\boldsymbol{Z}_{c_m}$ is the incidence matrix that relates the random vector $\boldsymbol{c}_m$ with the phenotype data. The genomic relationship matrix ($\boldsymbol{K}_c$) is derived from the DArTSeq markers (subscript m denotes markers and subscript c, the cross). $\sigma^2_{c_m}$ is the variance of the cross. Let $\boldsymbol{X}$ be a $c \times p$ matrix of standardized markers (mean zero and variance one for each marker), where c is the number of crosses (individuals) and p is the number of markers.

We defined two kernel matrices of genomic relationship: (1) the first kernel matrix is linear $\boldsymbol{K} = \boldsymbol{X}\boldsymbol{X}'/p$ (Lopez-Cruz et al., 2015), and it is the relationship matrix used in the genomic best linear unbiased predictor (GBLUP) (GB) model and (2) the second kernel matrix is nonlinear and is constructed with each element computed as $\boldsymbol{K}_{ii'} = \exp(-h \times \text{dist}(\boldsymbol{x}_i, \boldsymbol{x}_{i'})$ where $\text{dist}(\boldsymbol{x}_i, \boldsymbol{x}_{i'})$ is the squared distance between two rows $\boldsymbol{x}_i, \boldsymbol{x}_{i'}$, $i,\ i' = 1, \ldots, n$ (number of records with genotypic information) and h is two times the reciprocal of the median of the squared distances; this nonlinear kernel is named Gaussian kernel (GK) (Gianola & van Haam, 2008) or reproducing kernel Hilbert spaces (Crossa et al., 2010).

2.5.3 | Pedigree + genomic model (PGM-1)

By combining PM (Equation 2) and GM (Equation 3), we obtain model PGM-1

$$\boldsymbol{y} = \mu 1 + \boldsymbol{Z}_{d_p}\boldsymbol{d}_p + \boldsymbol{Z}_{t_I}\boldsymbol{t}_I + \boldsymbol{Z}_{c_p}\boldsymbol{c}_p + \boldsymbol{Z}_{c_m}\boldsymbol{c}_m + \boldsymbol{e}, \quad (4)$$

where all the components are already defined above.

2.5.4 | Pedigree + genomic model (PGM-2) (single-step approach developing the $\boldsymbol{H}$ matrix)

This model considers the information of the markers and pedigree using the single step approach by Legarra et al. (2009) and Aguilar et al. (2010). The PGM-2 model is given by:

$$\boldsymbol{y} = \mu 1 + \boldsymbol{Z}_{h_d}\boldsymbol{h_d} + \boldsymbol{Z}_{h_t}\boldsymbol{h_t} + \boldsymbol{Z}_{h_c}\boldsymbol{h_c} + \boldsymbol{e}, \quad (5)$$

where $\boldsymbol{Z}_{h_d}$, $\boldsymbol{Z}_{h_t}$, and $\boldsymbol{Z}_{h_c}$ are incidence matrices that relate observations with the random vectors $\boldsymbol{h_d}$, $\boldsymbol{h_t}$, and $\boldsymbol{h_c}$, respectively; $\boldsymbol{h}_d \sim \text{MN}(0, \sigma^2_{h_d}\boldsymbol{H}_d)$ is a random vector for the general combining ability of the DW (d) parents, where $\sigma^2_{H_d}$ is the variance of the DW parents that is the parameter associated with the matrix $\boldsymbol{H}_d$ that includes pedigree and genomic information for the DW lines; $\boldsymbol{h}_t \sim \text{MN}(0, \sigma^2_{h_t}\boldsymbol{H}_t)$ is a random vector for the general combining ability of the *Ae. tauschii* (t) parents, where $\sigma^2_{H_t}$ is the variance of the *Ae. tauschii* parents that is the parameter associated with the matrix $\boldsymbol{H}_t$ for *Ae. tauschii* lines; $\boldsymbol{h}_c \sim \text{MN}(0, \sigma^2_{h_c}\boldsymbol{H}_c)$ is the random vector of specific combining ability of the DW × *Ae. tauschii* crosses (c), where $\sigma^2_{c_H}$ is the variance of the cross (DW × *Ae. tauschii*) that is the parameter attached to the joined pedigree and genomic similarity matrix $\boldsymbol{H}_c$ for the cross.

Matrices $\boldsymbol{H}$ were generated using the relationship matrix derived from pedigree ($\boldsymbol{A}$) and a relationship matrix derived

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from markers ($\boldsymbol{K}$) according to Legarra et al. (2009) and Aguilar et al. (2010), and is given by:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix}, \tag{6}$$

where $\boldsymbol{H}_{11} = \boldsymbol{A}_{11} + \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}(\boldsymbol{K} - \boldsymbol{A}_{22})\ \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$, $\boldsymbol{H}_{12} = \boldsymbol{A}_{12}\ \boldsymbol{A}_{22}^{-1}\boldsymbol{K}$, $\boldsymbol{H}_{22} = \boldsymbol{K}$, $\boldsymbol{H}_{21} = \boldsymbol{H}'_{12}$. The matrix is divided according to whether the individuals are genotyped or not; 1 denotes the individuals that were not genotyped, and 2 denotes the individuals genotyped, so that, for example, $\boldsymbol{A}_{21}$ represents the relationship between genotyped and non-genotyped individuals (Pérez-Rodríguez et al., 2017).

According to Christensen (2012), matrix $\boldsymbol{K}$ should be adjusted in order to have the same scale as matrix $\boldsymbol{A}$, that is:

$$\boldsymbol{K}_a = \beta\boldsymbol{K} + \alpha, \tag{7}$$

where β and α are obtained by solving the following system of equations:

$$\text{Avg}\left(\text{diag}\left(\boldsymbol{K}\right)\right)\beta + \alpha = \text{Avg}\left(\text{diag}\left(\boldsymbol{A}_{22}\right)\right), \tag{8}$$

$$\text{Avg}\left(\boldsymbol{K}\right)\beta + \alpha = \text{Avg}\left(\boldsymbol{A}_{22}\right). \tag{9}$$

2.5.5 | Pedigree + genomic model (PGM-3) (interactions DW × *Ae. tauschii*)

This PGM-3 model considers the information of pedigree and markers based on the single step of Legarra et al. (2009) and Aguilar et al. (2010) in combination with the perspective provided by Acosta-Pech et al. (2017) for the hybrid prediction from the interaction combination of the parents.

$$\boldsymbol{y} = \mu 1 + \boldsymbol{h}_{\boldsymbol{dxt}} + \boldsymbol{Z}_{h_c}\boldsymbol{h}_{\boldsymbol{c}} + \boldsymbol{e}, \tag{10}$$

where $\boldsymbol{h}_{\boldsymbol{d\times t}} \sim \text{MN}(0, \sigma^2_{d\times t}(\boldsymbol{Z}_{h_t}\boldsymbol{H}_t\boldsymbol{Z}_{h_t}{}')\#(\boldsymbol{Z}_{h_d}\boldsymbol{H}_d\boldsymbol{Z}_{h_d}{}')$ is the random vector of interaction between the parents from DW and *Ae. tauschii*, $\sigma^2_{d\times t}$ is the variance component and matrices $\boldsymbol{Z}_{h_t}$, $\boldsymbol{Z}_{h_d}$, $\boldsymbol{H}_t$, $\boldsymbol{H}_d$ were defined earlier, with # being the Hadamard product. Acosta et al. (2017) suggested a Kronecker product, but since the data of this study are not balanced, the equivalent operation is the Hadamard product (Martini et al., 2020); $\boldsymbol{h}_c \sim \text{MN}(0, \sigma^2_{h_c}\boldsymbol{H}_c)$ is the random vector of SCA of the DW × *Ae. tauschii* crosses, where $\sigma^2_{h_c}$ is the variance of the cross (DW × *Ae. tauschii*) that is the scaled parameter associated with the joint pedigree and genomic similarity matrix $\boldsymbol{H}_c$ for the cross.

2.6 | Cross-validation schemes

We studied several genomic prediction problems related to approaches breeders or scientists might deploy when evaluating or generating new SHW. Three types of prediction problems were assessed, similar or related to prediction problems used in hybrid breeding, as in Basnet et al. (2019), as shown in Figure 3. The three CVs include different prediction situations.

2.6.1 | CV1: Assign SHW to five folds

CV1 studies the prediction of the SHW when both parents, DW and *Ae. tauschii*, are known. We performed a random CV analysis using a scheme that is known as cross-validation CV1 (Burgueño et al., 2012), which considers the prediction of a certain proportion of untested cultivars. In our study, we assume several SHW are assessed for disease resistance, while for other SHWs, phenotypic values are unobserved (missing). Overall, more than 1000 SHW have been generated at the CIMMYT since the 1980s, but only a subset of SHW has been evaluated for individual traits; thus, CV1 reflects the problem breeders face of usually not having the full capacity to evaluate all possible cultivars (germplasm) for all type of target traits.

2.6.2 | CV2: Assign DW parent to five folds

CV2 investigates the prediction of SHW when one parent, *Ae. tauschii*, is known, but a proportion (20%) of the DW parent is unobserved. We performed CV assigning the DW parents to five folds; therefore, if a set of DW lines is assigned to the training (testing) set, then all the corresponding SHW lines were assigned to the training (testing) set. The training set includes all the SHW lines obtained when using 80% of the DW and predicting the remaining 20% of SHW (from unobserved DW). This CV scheme mimics the problem of predicting SHW whose DW parent has not yet been observed in any SHW combination. Since the number of SHW was not equal across DW, the SHW lines in training and testing sets vary.

2.6.3 | CV3: Assign *Ae. tauschii* parents to five folds

CV3 comprises the prediction of SHW when one parent, DW, is known, but a proportion (20%) of the parent *Ae. tauschii* is unobserved. Similar to CV2, we performed CV assigning the *Ae. tauschii* wheat parents to folds. In general, a larger number of *Ae. tauschii* parents are used when generating new SHW. Here, the training set includes all the SHW obtained when

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

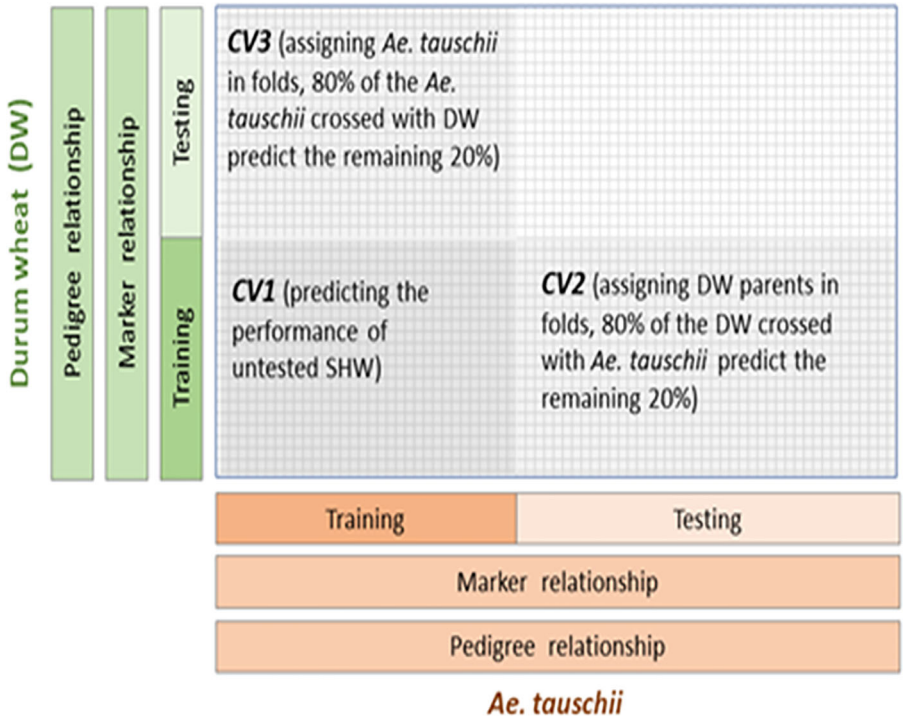


FIGURE 3 Three prediction problems of tested and untested synthetic hexaploid wheat (SHW) and their parents. CV1 tested SHW (DW × *Ae. tauschii*) using fivefold random cross-validation with 80% of data in the training set and 20% in the testing set. CV2 leave 80% of the DW parents out and predict the SHW with the remaining 20% DW parents. CV3 leave 80% of the *Ae. tauschii* parents out and predict the SHW with the remaining 20% *Ae. tauschii* parents. DW, durum wheat.

using 80% of the *Ae. tauschii* crossed with durum wheat and predicting the remaining 20% of the SHW. This CV scheme reflects the problem of predicting SHW using DW parents whose crosses with any of the *Ae. tauschii* accessions were not yet observed.

2.7 | Heritability based on genomic model

To compute heritability based on genomics, we considered the proposal by Feldman et al. (2022) that attempts to improve the precision of a heritability estimate by using the concept of average semi-variance. It is assumed that the SHW cultivars represent a random effect with identical and independent distributions (*iid*) $g_i \overset{iid}{\sim} N(0,\ \boldsymbol{K}_{\text{ASV}}\sigma_g^2)$, where $\boldsymbol{K}_{\text{ASV}}$ is the relationship matrix computed following Feldman et al. (2022); the variance of the SHW is represented as σ_g^2. Then, the heritability is

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\varepsilon^2}, \tag{11}$$

where the variance component σ_ε^2 denotes the residual error variance when assuming residuals being normally independent and identically distributed.

3 | RESULTS AND DISCUSSION

Data analyses were based on the phenotypic effects of the SHW across the 12 replicates for disease traits TS, SNB, and the six replicates for the disease trait SB. Note that the replicates were established in the greenhouse and sown at different planting dates.

For each of the models described above, the two kernels (GB and GK) and three types of CVs (CV1, CV2, and CV3) were used. For the SHW and DW parents, the pedigree information was available, while the *Ae. tauschii* accessions were assumed to be unrelated. SNP data for all SHW were generated in the study by Lozano-Ramirez et al. (2022a). Table 1 shows that the SNP data on the parental lines were only available for 32 DW parents and 127 *Ae. tauschii* accessions forming a total of 147 out of the 386 SHW used. The limited data availability for the parental lines justifies the use of the single step models (PGM-2 and PGM-3).

For genomic adjustment and prediction, two stages were carried out. In the first stage, the adjusted means of the cultivars were estimated using an ordinal response model. Table 1 displays the results of this adjustment, showing the means and standard deviations. Additionally, skewness coefficients of the adjusted means were calculated, resulting in 0.20, 0.15, and 0.25 for TS, SNB, and SB, respectively, indicating

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Information from adjusted means, genomic heritability, and sample size for each trait: tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB).

				Available sample size					
	Adjusted means			SHW		DW		*Ae. tauschii*	
Trait	Mean	SD	Heritability	Pedigree	Markers	Pedigree	Markers	Identity	Markers
TS	1.53	0.68	0.64	386	386	39	32	264	127
SNB	1.70	0.71	0.67	386	386	39	32	264	127
SB	1.51	0.58	0.57	386	386	39	32	264	127

Note: Details of the inoculation procedures for each disease are given in Section 2.
Abbreviations: DW, durum wheat; SHW, synthetic hexaploid wheat.

only small asymmetries. Similarly, variances of the adjusted means were calculated, with mean variances of 0.11, 0.12, and 0.16, as well as standard deviations of 0.03, 0.05, and 0.05 for TS, SNB, and SB, respectively. Based on the results, it can be assumed that the adjusted means follow a symmetric distribution with homogeneous variances.

Additionally, the homogeneity of residual variances was tested using the proposal by Breusch and Pagan (1979). In a simplified form, this involves fitting a regression model with the squared residuals from the original model as the response variable and the covariates from the original model as predictors. With this adjusted model, a statistic is calculated, equal to the number of observations (n) multiplied by the coefficient of determination. Subsequently, a p-value is determined using the χ^2 distribution with degrees of freedom equal to the number of covariates in the model. In all cases, the p-value of this test was close to 1, indicating no evidence to reject the null hypothesis of a constant variance existing in the residuals.

We reiterate that for stage 2, genomic prediction models are fitted for traits TS, SNB, and SB assuming the responses are continuous, and data are from a normal distribution such that the residuals of the models have a normal distribution as shown in Figure 2, which shows that all the residuals are contained in the corresponding confidence intervals. Furthermore, for the three diseases (TS, SNB, and SB), the Shapiro test for normality indicated evidence of not rejecting the null hypothesis that the residuals come from normal distribution.

3.1 | Resistance scores and variance components

The 386 SHW evaluated in the greenhouse displayed high levels of resistance and moderate resistance for all four foliar diseases (Table 1), and reaction types were comparable to the disease-specific resistance checks (data not shown). The means for the TS, SNB, and SB disease scores were similarly high. Heritability estimates for the traits were similar except for trait SB with a heritability of 0.53. The DW parents that were evaluated for the same diseases in parallel showed overall lower levels of resistance (data not shown), while the *Ae. tauschii* parents could not be screened. The lower resistance scores in the DW parents suggest that the resistance in the SHW lines was derived from both the DW and *Ae. tauschii* parents. Genome-wide association studies conducted on the same data sets confirmed this result and showed that genomic loci from all three genomes played a role in conferring resistance within the SHW lines (Lozano-Ramirez et al., 2022a; Lozano-Ramirez et al., 2022b). Higher levels of Fusarium crown rot resistance in SHW compared with their tetraploid parents were also observed by Chen et al. (2023). The authors validated that phenylalanine ammonia lyase genes, involved in the biosynthesis of lignin and salicylic acid, displayed a higher level of expression to Fusarium crown rot infection in the SHWs, leading to improved resistance.

Variance components of the fitted models (PM, GM, PGM-1, PGM-2, and PGM-3) for all three diseases were estimated (Table 2). The characterization of the variance components due to GCA and SCA is of central importance in hybrid prediction. GCA is directly related to the breeding value of a parent and is associated with additive genetic effects, while SCA is the relative performance of cross that is associated with nonadditive gene action, predominantly contributed by dominance and epistasis, suggesting the use of a nonlinear GK that enables accounting for complex interaction (Crossa et al., 2010). The variance explained by the sub-genomes of the DW and *Ae. tauschii* parents and the respective SHW in our study is comparable to modeling GCA and SCA. Both are fundamental sources for prediction based on either pedigree or marker information of the lines forming the interspecific hybridization (SHW).

Similar proportions of variance were explained by the DW and *Ae. tauschii* parents (GCA : $\sigma^2_{d_p}$, $\sigma^2_{t_I}$) and of the SHW (SCA : $\sigma^2_{c_p}$) for all traits using the PM model (Table 2). For GM, PMG-1, PGM-2, and PGM-3, a much larger proportion of variance was explained by genomic SCA ($\sigma^2_{c_m} = 0.727$ for GM, $\sigma^2_{c_m} = 0.542$ for PGM $-$ 1, $\sigma^2_{c_m} = 0.594$ for PGM $-$ 2, and $\sigma^2_{H_c} = 0.668$ for PGM-3). Trends for all traits were similar.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Variance components for models pedigree model (PM), marker model (GM), pedigree + marker model 1 (PGM-1), PGM-2, PGM-3 for all diseases traits, tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB).

Trait	Model	Variance components								
		$\sigma^2_{d_p}$	$\sigma^2_{t_I}$	$\sigma^2_{c_p}$	$\sigma^2_{c_m}$	$\sigma^2_{H_d}$	$\sigma^2_{H_t}$	$\sigma^2_{H_c}$	$\sigma^2_{H_{dxt}}$	σ^2_e
	PM	0.254	0.197	0.281						0.424
TS	GM				0.555					0.330
	PGM-1	0.121	0.096	0.109	0.412					0.213
	PGM-2					0.192	0.120	0.453		0.265
	PGM-3							0.497	0.207	0.252
	PM	0.220	0.265	0.194						0.453
SNB	GM				0.727					0.403
	PGM-1	0.118	0.129	0.100	0.542					0.280
	PGM-2					0.179	0.134	0.594		0.342
	PGM-3							0.668	0.213	0.325
	PM	0.190	0.211	0.184						0.582
SB	GM				0.604					0.536
	PGM-1	0.134	0.138	0.124	0.362					0.395
	PGM-2					0.228	0.130	0.438		0.472
	PGM-3							0.534	0.204	0.473

The general higher proportion of variance by SCA indicates that dominance and epistasis play a large role in a situation of allopolyploidization. Epistatic interaction between the A, B, and D genomes was also already suggested for TS resistance by Lozano-Ramirez et al. (2022a). Genomic loci from all three genomes were contributing to the resistance in the SHW while the DW parents showed moderately susceptible to susceptible reaction types, suggesting possible favorable epistatic interaction (activation) between genomes. Epistatic interaction activating or suppressing disease resistance have also been reported in other studies on wheat (Chen et al., 2023; Chu et al., 2008; Hiebert et al., 2020; Nelson et al., 1997). Genes may be activated or suppressed in SHW after hybridization. For instance, the resistance to stem rust pathogen (*Puccinia graminis* f. sp. *tritici*) is suppressed in the hexaploid state by Med15, a component of the Mediator complex encoded by the D sub-genome (Hiebert et al., 2020). These findings are also aligned with the observed alternations in gene expression during allopolyploidization (Kenan-Eichler et al., 2011; Li et al., 2014; Vasudevan et al., 2023; M. Yu et al., 2017). Early research studies documented that the onset of both allotetra- and allohexa-polyploidization events in wheat is associated with rapid and extensive structural changes at the genetic level (Ozkan et al., 2001; Shaked et al., 2001), as well as the gene expression and/or epigenetic level in wheat (Adams & Wendel et al., 2005). However, more recent work suggested that most structural changes may have taken place during allotetraploidization (Mestiri et al., 2010; Zhang et al., 2013). Compared with their parental lines, changes in genome structures and copy numbers of certain genes have been observed in SHW (Zhao et al., 2011; Zhao et al., 2011), while once SHW archive euploid chromosome sets, they are genetically stable (Mestiri et al., 2010; Zhao et al., 2011). The possible changes in genome structure during allohexaploidization indicate that pedigree information might not be the best predictor for crosses forming the SHW. Furthermore, some genetic variation remaining within *Ae. tauschii* accessions derived from gene banks is likely, and thus the identity of each accession might not always reflect the same genotype in a cross.

In addition to the early reports of gene expression and epigenetic changes, Rapp et al. (2009) defined the concept of "expression-level dominance" evaluating gene expression patterns during allopolyploidization in cotton. Expression-level dominance means that the total expression level for all homeologs of a gene in a polyploid is statistically equivalent to the expression level of one of the polyploid parents instead of additivity between homoeologs. Such a pattern suggests that there has been sub-functionalization of genes from the two parental lines to some extent in SHW. Li et al. (2014) identified functional dominance of genes in SHW with similar expression levels to those in tetraploid parent enriched for plant development, and genes with similar expression levels to those in *Ae. tauschii* were found to be involved in adaptation. Expression-level dominance in SHW could contribute to SCA variance.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Genomic prediction

It has been proven in wheat (Crossa et al., 2007; Pérez-Rodríguez et al., 2012) that large linkage blocks allow for relatively stable epistatic effects within and between the sub-genomes A, B, and D. The use of nonlinear kernels such as the GK enables the capturing of cryptic small epistatic effects. In this study of SHW, the better predictive models were mostly those using the nonlinear GK that is supposed to capture gene × gene interaction, especially those existing between sub-genomes AB and D.

To our knowledge, this is the first study that uses genomic models developed for hybrid breeding to predict the phenotypes of allopolyploids as a hybrid of different species. The prediction problems applied in this study are graphically displayed in Figure 3. The three examples of CV (CV1, CV2, and CV3) represent major problems CIMMYT pre-breeders and breeders face when testing SHW lines for any agronomic trait or when planning a new generation of SHW lines. CV1 predicts the performance of new SHW, based on the subset of SHW evaluated in the field, while other SHW are unobserved (missing). CV1 mirrors the problem, not having the full capacity to evaluate all available SHW phenotypically in the field or in the greenhouse, especially for time-consuming and cost-intensive phenotypes. CV2 and CV3 perform CV on the DW and *Ae. tauschii* parents, respectively. If a set of DW or *Ae. tauschii* parents were assigned to training and testing sets, all the corresponding SHW synthesized with these parents were also assigned to training-testing sets. For example, the training set includes all the SHW obtained when using 80% of the DW parents crossed with any *Ae. tauschii* accession and predicting the remaining 20% of the SHW. Both prediction problems, CV2 and CV3, mimic the problem of predicting SHW using parents (DW or *Ae. tauschii* accessions) that were not necessarily tested in combination.

The prediction abilities obtained varied with correlations between the observed and predicted values. These differences were caused by the model applied and the individual prediction problems tested (Table 3 and Figure 4 for random cross-validation CV1 scheme, Table 4 and Figure 5 for random cross-validation CV2 scheme, and Table 5 and Figure 6 for random cross-validation CV3). In general, for all three testing schemes, it was found that the predictive ability was the lowest when PM was used and slightly increasing when using GK compared to GB. The highest prediction abilities for CV1 were achieved with GK models; for example, the highest correlations were 0.730 (GM-GK) and 0.720 (PGM-2-GK) for TS and 0.645 (GM-GK) for SNB (Table 3 and Figure 4). Prediction abilities were generally lower for models PM and PGM-3.

The highest prediction abilities for CV2 were also obtained with GK, except for the model PGM-1 for trait SNB (0.428) (Table 4). Similarly, the GK resulted in slightly higher prediction abilities for all the traits except for SB for cross-validation CV3 with prediction accuracies of 0.483 for traits with the model PGM-2 (Table 5).

While the models including SNP marker data showed no significant differences, the GM and PGM-2 models showed slightly higher prediction abilities. The PM model revealed the lowest prediction ability for all diseases, up to twofold. Furthermore, prediction abilities tend to be higher for CV1 than for CV2 (Table 4, Figure 5) and CV3 (Table 5, Figure 6), regardless of the model deployed. Higher prediction abilities were observed for CV3, which assigned the *Ae. tauschii* parent into folds, compared to CV2 for almost all the models and traits. A much larger number of *Ae. tauschii* accessions were used for the synthesis of SHW compared to a smaller number of DW parents, resulting in more crosses made per DW parent. A favorable genetic component due to any closer relationship between the fewer DW parents (including sister lines) was therefore not apparent in CV2.

Across all predictions, the standard deviations of the predictive correlations are relatively large, leading to the PM model being the only model with significantly lower results. Overall, levels of prediction abilities for the diseases presented were similar or higher in comparison to previous studies. SB was the trait with the lowest prediction ability. Daetwyler et al. (2014) predicted leaf rust resistance in a set of diverse bread wheat landraces and reported average prediction abilities of 0.35. More recently, Semagn et al. (2022) described prediction accuracies for TS with values between 0.11 and 0.41 in an association mapping panel of Canadian spring wheat cultivar within and among environments. Our results were also comparable to those that explore predictions of grain yield in hybrid wheat (Basnet et al., 2019; Y. Zhao et al., 2013). In general, CV schemes predicting the performance of untested hybrids, equivalent to the prediction of the performance of untested SHW (CV1) in our study, showed better results than CV schemes predicting untested female or male parents.

In general, when using the nonlinear kernel (GK), in the GM, PGM-1, PGM-2, PGM-3 models, the prediction average consistently increases compared to the use of linear GB (Tables 3–5, Figures 4–6) for all diseases but is more noticeable in the SNB. This result is consistent with what was previously discussed regarding SCA manifesting nonadditive effects (epistasis) that can be captured better by the nonlinear kernel (GK) than by the linear kernel (GB). Clearly, the PM is overcome, in terms of prediction accuracy, by all the other model-kernel types. Genomic prediction for models and kernel type seems to be different for the different diseases investigated in this study. For example, for TS and SNB diseases, the best model-kernel options were GM-GK and PGM-3-GK. For the disease SB, models PMG-1 and PMG-2

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Average Pearson correlation between observed and predicted values for fivefold random cross-validation validation for prediction from missing crosses (CV1) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB). For the pedigree model (PM), marker model (GM), pedigree + marker model 1 (PGM-1), PGM-2, and PGM-3 models.

		Models								
		PM	GM		PGM-1		PGM-2		PGM-3	
Traits		A	GB	GK	GB	GK	GB	GK	GB	GK
TS	Mean	0.416	0.707	**0.730**	0.700	0.719	0.693	0.720	0.709	0.708
	SD	0.06	0.05	0.05	0.04	0.05	0.05	0.06	0.06	0.06
SNB	Mean	0.399	0.606	**0.645**	0.587	0.617	0.596	0.629	0.618	0.603
	SD	0.09	0.09	0.08	0.08	0.07	0.07	0.07	0.10	0.10
SB	Mean	0.282	0.467	0.470	0.468	0.471	**0.477**	0.473	0.469	0.461
	SD	0.08	0.12	0.12	0.07	0.07	0.07	0.07	0.11	0.11

Note: Models with the best genomic prediction ability for each trait are shown in bold.
Abbreviations: GB, GBLUP matrix; GK, Gaussian kernel matrix.

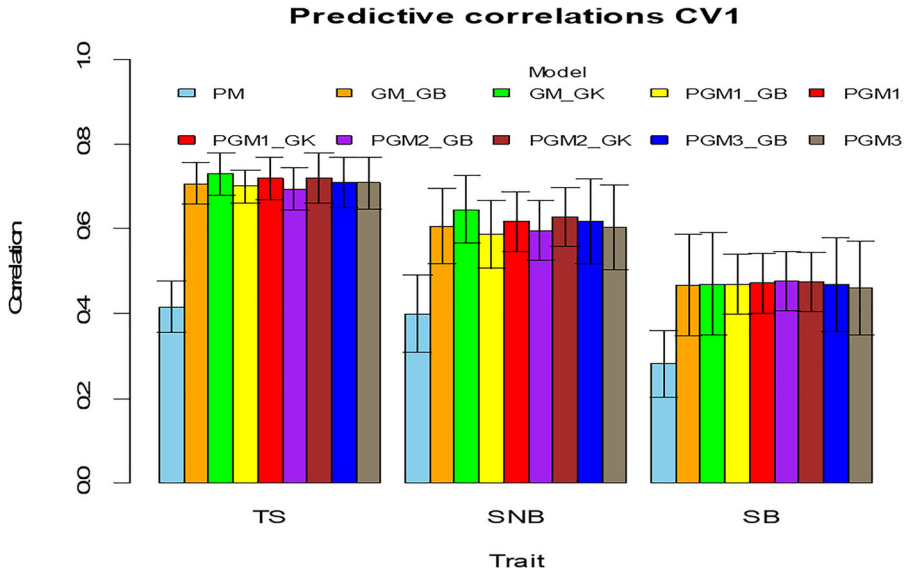


FIGURE 4 Average Pearson correlation (bars) and standard deviation (whiskers) between observed and predicted values for fivefold random cross-validation (CV) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB) for the prediction of crosses (CV1) for the PM, GM.GB, GM.GK, PGM-1.GB, PGM-1.GK, PGM-2.GB, PGM-2.GK, PGM-3.GB, and PGM-3.GK models.

TABLE 4 Average Pearson correlation between observed and predicted values for fivefold random cross-validation for prediction from missing durum wheat (DW) parents (CV2) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB). For the pedigree model (PM), marker model (GM), pedigree + marker model 1 (PGM-1), PGM-2, and PGM-3 models.

		Models								
		PM	GM		PGM-1		PGM-2		PGM-3	
Traits		A	GB	GK	GB	GK	GB	GK	GB	GK
TS	Mean	0.217	0.612	**0.646**	0.597	0.626	0.603	0.635	0.620	0.626
	SD	0.15	0.06	0.05	0.07	0.06	0.06	0.05	0.06	0.05
SNB	Mean	0.188	0.421	0.417	**0.428**	0.423	0.399	0.401	0.423	0.417
	SD	0.11	0.17	0.14	0.16	0.13	0.18	0.15	0.17	0.17
SB	Mean	0.140	0.342	0.353	0.367	0.370	0.360	**0.378**	0.359	0.354
	SD	0.22	0.13	0.09	0.14	0.13	0.13	0.11	0.12	0.11

Note: Models with the best genomic prediction ability for each trait are shown in bold.
Abbreviations: A, pedigree matrix; GB, GBLUP matrix; GK, Gaussian kernel matrix.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

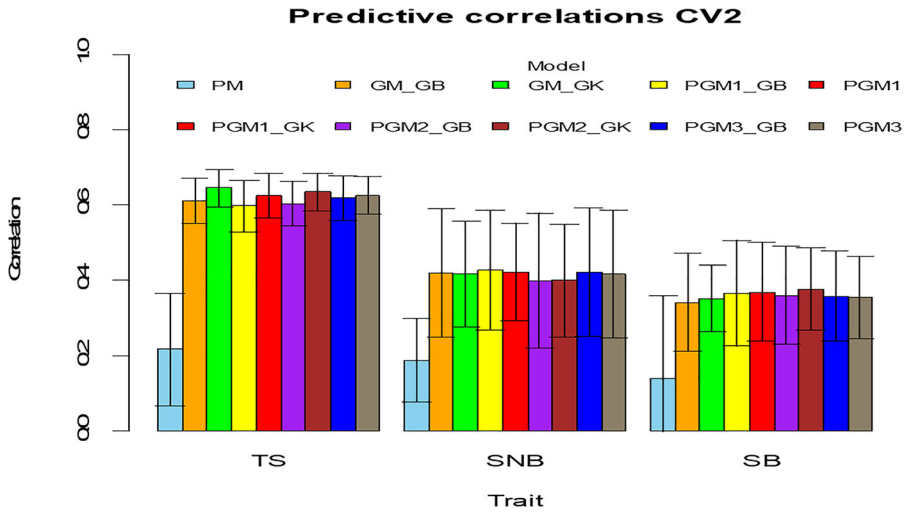


FIGURE 5 Average Pearson correlation (bars) and standard deviation (whiskers) between observed and predicted values for fivefold random cross-validation (CV) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB) for the prediction of crosses (CV2) for the PM, GM.GB, GM.GK, PGM-1.GB, PGM-1.GK, PGM-2.GB, PGM-2.GK, PGM-3.GB, and PGM-3.GK models.

TABLE 5 Average Pearson correlation between observed and predicted values for fivefold random cross-validation for prediction from missing *Ae. tauschii* parents (CV3) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB). For the pedigree model (PM), marker model (GM), pedigree + marker model 1 (PGM-1), PGM-2, and PGM-3 models.

		Models								
		PM	GM		PGM-1		PGM-2		PGM-3	
Traits		A	GB	GK	GB	GK	GB	GK	GB	GK
TS	Mean	0.389	0.694	**0.718**	0.680	0.699	0.675	0.702	0.694	0.692
	SD	0.09	0.07	0.06	0.07	0.07	0.08	0.07	0.06	0.07
SNB	Mean	0.382	0.596	**0.617**	0.591	0.610	0.591	0.611	0.600	0.596
	SD	0.071	0.036	0.05	0.03	0.06	0.03	0.05	0.04	0.03
SB	Mean	0.235	0.463	0.473	0.472	0.468	**0.483**	0.472	0.464	0.454
	SD	0.11	0.10	0.12	0.10	0.10	0.10	0.10	0.10	0.10

Note: Models with the best genomic prediction ability for each trait are shown in bold.
Abbreviations: A, pedigree matrix; GB, GBLUP matrix; GK, Gaussian kernel matrix.

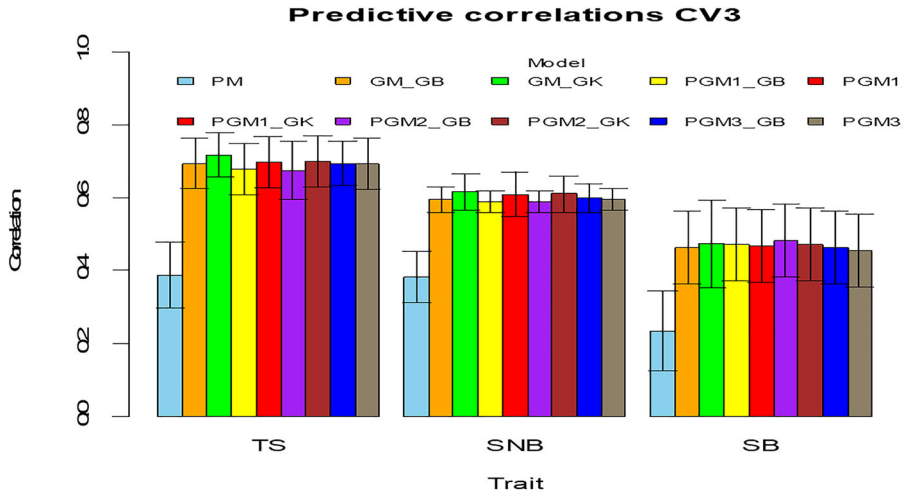


FIGURE 6 Average Pearson correlation (bars) and standard deviation (whiskers) between observed and predicted values for fivefold random cross-validation (CV) for disease traits tan spot (TS), septoria nodorum blotch (SNB), and spot blotch (SB) for the prediction of crosses (CV3) for the PM, GM.GB, GM.GK, PGM-1.GB, PGM-1.GK, PGM-2.GB, PGM-2.GK, PGM-3.GB, PGM-3.GK models.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are slightly superior in genomic prediction ability to the others yet with few differences between kernel types.

The CIMMYT produced ~1600 SHW since the 1980s and subsequently generated thousands of crosses with elite bread wheat. Synthetic derivative lines have been selected as parents for mainstream breeding, with rigorous selection resulting in advanced lines with excellent performance for yield and other traits. Synthetic derivative lines have been selected as candidates in international yield trials, which are disseminated by the CIMMYT globally every year (Dreisigacker et al., 2008; Rosyara et al., 2019). Many of the CIMMYT SHW lines have been evaluated by different research groups and for a series of diverse traits including diseases, quality characteristics, and agronomic traits. However, only smaller numbers of subsets of SHW that allow manageable field or greenhouse trials have been performed. Our results therefore provide a new approach to guide breeders assessing additional untested SHW.

3.3 | Model comparison

The results of the hypothesis tests for the mean differences between model PM and each of the GM, PGM-1, PGM-2, PGM-3 models indicate that the differences in their means are significant at the 0.05 level. However, when comparing the means of GM with the mean of each of the PGM-1, PGM-2, PGM-3 models, it was found that the differences in their means are not significant, despite some noticeable differences in the means of PGM-3 compared to the other models, as observed in the graphs. The contribution of genomic information from the parents in prediction appears to be positive in terms of increasing the predictions, although this increase does not meet the levels of statistical significance. This is likely due to the lack of information about the markers of the parents, and in the case of *Ae. tauschii*, there were 127 markers available out of a total of 264, and the pedigree matrix was not available (see Table 1).

4 | CONCLUSION

The genomic prediction approach for this type of material is the main novelty of this study. We could draw sound and robust conclusions despite a limited phenotype data. We successfully applied the quantitative genetics tools, originally developed for the prediction of intraspecific hybrids to a situation of allopolyploidization. We predicted the phenotypes of SHW for four major quantitatively inherited diseases in wheat. Our study describes a predictive approach that could further guide the use of genetic resources available in gene banks to increase the extent of useful genetic diversity available to breeders. Results indicated that the genomic prediction of several wheat diseases in the SHW developed from the cross of DW with *Ae. tauschii* can be successfully achieved. In general, genomic prediction models including pedigree and nonlinear kernels overcome the prediction ability of genomic models that use only linear kernels. These results indicated that nonlinear kernels explained and predicted the complex genome interaction existing when crossing DW and *Ae. tauschii* better than the linear kernel.

AUTHOR CONTRIBUTIONS

Susanne Dreisigacker: Conceptualization; investigation; methodology; project administration; supervision; validation; writing—original draft; writing—review and editing. **Johannes Martini**: Conceptualization; investigation; visualization; writing—original draft; writing—review and editing. **Jaime Cuevas**: Data curation; formal analysis; investigation; writing—original draft. **Paulino Perez-Rodriguez**: Data curation; formal analysis; investigation; methodology. **Nerida Lozano Ramirez**: Data curation; investigation; resources; validation; visualization; writing—original draft. **Julio Huerta-Espino**: Data curation; funding acquisition; investigation; project administration. **Pawan Kumar Singh**: Data curation; investigation; methodology; project administration. **Leonardo Crespo**: Conceptualization; data curation; investigation. **Alison Bentley**: Conceptualization; investigation; methodology; supervision; validation. **Jose Crossa**: Conceptualization; investigation; methodology; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

We thank all scientists, field workers, and lab assistants from the National Programs, CIMMYT, and ICRISAT who collected the data used in this study. Authors are thankful to Dr. Masahiro Kishii for assistance in doing the crosses between durum wheat and *Ae. Tauschii* parents. We are thankful for the financial support provided by the Bill & Melinda Gates Foundation and FCDO (INV-003439, Accelerating Genetic Gains in Maize and Wheat for Improved Livelihoods [AG2MW]), the USAID projects (USAID Amend. No. 9 MTO 069033, USAID-CIMMYT Wheat/AGGMW, AGG-Maize Supplementary Project, AGG [Stress Tolerant Maize for Africa]), and the CIMMYT CRP (maize and wheat). We are also thankful for the financial support provided by the Foundation for Research Levy on Agricultural Products (FFL) and the Agricultural Agreement Research Fund (JA) through the Research Council of Norway for grants 301835 (Sustainable Management of Rust Diseases in Wheat) and 320090 (Phenotyping for Healthier and more Productive Wheat Crops).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DATA AVAILABILITY STATEMENT

The phenotypic (TS, SNB and SB) and genotypic data sets can be found in the following link: https://hdl.handle.net/11529/10548948.

ORCID

Susanne Dreisigacker https://orcid.org/0000-0002-3546-5989
Paulino Pérez-Rodríguez https://orcid.org/0000-0002-3202-1784
Pawan Singh https://orcid.org/0000-0003-4610-3120
Alison R. Bentley https://orcid.org/0000-0001-5519-4357
Jose Crossa https://orcid.org/0000-0001-9429-5855

REFERENCES

Aberkane, H., Payne, T., Kishi, M., Smale, M., Amri, A., & Jamora, N. (2020). Transferring diversity of goat grass to farmers' fields through the development of synthetic hexaploid wheat. *Food Security*, *12*(5), 1017–1033. https://doi.org/10.1007/s12571-020-01051-w

Acosta-Pech, R., Crossa, J., De Los Campos, G., Teyssèdre, S., Claustres, B., Pérez-Elizalde, S., & Pérez-Rodríguez, P. (2017). Genomic models with genotype × environment interaction for predicting hybrid performance: An application in maize hybrids. *Theoretical and Applied Genetics*, *130*(7), 1431–1440. https://doi.org/10.1007/s00122-017-2898-0

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, *93*(2), 743–752. https://doi.org/10.3168/jds.2009-2730

Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez, P., & Lopez-Cruz, M. A. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *Crop Journal*, *8*(5), 745–756. https://doi.org/10.1016/j.cj.2020.03.010

Basnet, B. R., Crossa, J., Dreisigacker, S., Pérez-Rodríguez, P., Manes, Y., Singh, R. P., Rosyara, U. R., Camarillo-Castillo, F., & Murua, M. (2019). Hybrid wheat prediction using genomic, pedigree, and environmental covariables interaction models. *Plant Genome*, *12*(1), 1–13. https://doi.org/10.3835/plantgenome2018.07.0051

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, *47*, 1287–1294. https://doi.org/10.2307/1911963

Burgueño, J., De Los Campos, G., Weigel, K., & Crossa, J. (2012). Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Science*, *52*(2), 707–719. https://doi.org/10.2135/cropsci2011.06.0299

Chen, Y., Wang, Y., Guan, F., Long, L., Wang, Y., Li, H., Deng, M., Zhang, Y., Pu, Z., Li, W., Jiang, Q., Wang, J., Wei, Y., Ma, J., Xu, Q., Kang, H., Qi, P., Yuan, Z., Zhang, L., … Jiang, Y. (2023). Comparative analysis of Fusarium crown rot resistance in synthetic hexaploid wheats and their parental genotypes. *BMC Genomics*, *24*(1), 178. https://doi.org/10.1186/s12864-023-09268-7

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variaton in plant polyploids. *Annual Review of Plant Biology*, *58*(1), 377–406. https://doi.org/10.1146/annurev.arplant.58.032806.103835

Christensen, O. F. (2012). Correction: Compatibility of pedigree-based and marker-based relationship matrices for single-step genetic evaluation. *Genetics Selection Evolution*, *44*, 37. https://doi.org/10.1186/1297-9686-46-20

Chu, C.-G., Friesen, T. L., Xu, S. S., & Faris, J. D. (2008). Identification of novel tan spot resistance loci beyond the known host-selective toxin insensitivity genes in wheat. *Theoretical and Applied Genetics*, *117*(6), 873–881. https://doi.org/10.1007/s00122-008-0826-z

Crossa, J., Burgueño, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S. A., Lillemo, M., Singh, R. P., Trethowan, R., Warburton, M., Franco, J., Reynolds, M., Crouch, J. H., & Ortiz, R. (2007). Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics*, *177*, 1–25. https://doi.org/10.1534/genetics.107.078659

Crossa, J., Campos, G. D. L., Pérez, P., Gianola, D., Burgueño, J., Araus, J. L., Makumbi, D., Singh, R. P., Dreisigacker, S., Yan, J., Arief, V., Banziger, M., & Braun, H.-J. (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*, *186*(2), 713–724. https://doi.org/10.1534/genetics.110.118521

Daetwyler, H. D., Bansal, U. K., Bariana, H. S., Hayden, M. J., & Hayes, B. J. (2014). Genomic prediction for rust resistance in diverse wheat landraces. *Theoretical and Applied Genetics*, *127*(8), 1795–1803. https://doi.org/10.1007/s00122-014-2341-8

Dreisigacker, S., Kishii, M., Lage, J., & Warburton, M. (2008). Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. *Australian Journal of Agricultural and Research*, *59*(5), 413–420. https://doi.org/10.1071/AR07225

Dreisigacker, S., Sehgal, D., Reyes Jaimez, A., Luna Garrido, B., Muñoz Savala, S., Núñez-Ríos, C., Mollins, J., & Mall, S. (2016). *CIMMYT wheat molecular genetics: Laboratory protocols and applications to wheat breeding*. CIMMYT.

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, *316*(5833), 1862–1866. https://doi.org/10.1126/science.1143986

Eagles, H. A., Cane, K., & Trevaskis, B. (2011). Veery wheats carry an allele of Vrn-A1 that has implications for freezing tolerance in winter wheats. *Plant Breeding*, *130*(130), 413–418. https://doi.org/10.1111/j.1439-0523.2011.01856.x

FAO. (2020). *FAO—Food and Agricultural Organization of the United Nations, 2020*. http://www.fao.org/faostat/

Feldmann, M. J., Piepho, H.-P., & Knapp, S. J. (2022). Average semivariance directly yields accurate estimates of the genomic variance in complex trait analyses, 2022. *G3 Genes|Genomes|Genetics*, *12*(6), jkac080. https://doi.org/10.1093/g3journal/jkac080

Gaupp, F., Hall, J., Hochrainer-Stigler, S., & Dadson, S. (2020). Changing risks of simultaneous global breadbasket failure. *Nature Climate Change*, *10*, 54–57. https://doi.org/10.1038/s41558-019-0600-z

Gianola, D. (1982). Theory and analysiss of the threshold characters. *Journal of Animal Science*, *54*(5), 1979–1996. https://doi.org/10.2527/jas1982.5451079x

Gianola, D., & Van Kaam, J. B. C. H. M. (2008). Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics*, *178*, 2289–2303. https://doi.org/10.1534/genetics.107.084285

Grote, U., Fasse, A., Nguyen, T. T., & Erenstein, O. (2021). Food security and the dynamics of wheat and maize value chains in Africa and Asia.

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Frontiers in Sustainable Food Systems, *4*(February), 1–17. https://doi.org/10.3389/fsufs.2020.617009

Hiebert, C. W., Moscou, M. J., Hewitt, T., Steuernagel, B., Hernández-Pinzón, I., Green, P., Pujol, V., Zhang, P., Rouse, M. N., Jin, Y., Mcintosh, R. A., Upadhyaya, N., Zhang, J., Bhavani, S., Vrána, J., Karafiátová, M., Huang, L., Fetch, T., Doležel, J., … Spielmeyer, W. (2020). Stem rust resistance in wheat is suppressed by a subunit of the mediator complex. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-14937-2

Kenan-Eichler, M., Leshkowitz, D., Tal, L., Noor, E., Melamed-Bessudo, C., Feldman, M., & Levy, A. A. (2011). Wheat hybridization and polyploidization results in deregulation of small RNAs. *Genetics*, *188*(2), 263–272. https://doi.org/10.1534/genetics.111.128348

Kihara, H. (1944). Discovery of the DD-analyzer, one of the analyzers of *Triticum vulgare*. *Agricuture and Horticulture*, *19*(1), 889–890.

Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., & Horton, R. M. (2020). Amplified Rossby waves enhance risk of concurrent heatwaves in major breadbasket regions. *Nature Climate Change*, *10*, 48–53. https://doi.org/10.1038/s41558-019-0637-z

Lamari, L., & Bernier, C. C. (1989). Evaluation of wheat lines and cultivars to tan spot [*Pyrenophora tritici-repentis*] based on lesion type. *Canadian Journal of Plant Pathology*, *11*, 49–56. https://doi.org/10.1080/07060668909501146

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, *92*(9), 4656–4663. https://doi.org/10.3168/jds.2009-2061

Li, A., Liu, D., Wu, J., Zhao, X., Hao, M., Geng, S., Yan, J., Jiang, X., Zhang, L., Wu, J., Yin, L., Zhang, R., Wu, L., Zheng, Y., & Mao, L. (2014). mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell*, *26*(5), 1878–1900. https://doi.org/10.1105/tpc.114.124388

Li, A., Liu, D., Yang, W., Kishii, M., & Mao, L. (2018). Synthetic hexaploid wheat: Yesterday, today, and tomorrow. *Engineering*, *4*(4), 552–558. https://doi.org/10.1016/j.eng.2018.07.001

Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J.-L., Singh, R. P., Autrique, E., & De Los Campos, G. (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3 Genes, Genomes, Genet*, *5*(4), 569–582. https://doi.org/10.1534/g3.114.016097

Lozano-Ramirez, N., Dreisigacker, S., Sansaloni, C. P., He, X., Sandoval-Islas, S., Pérez-Rodríguez, P., Carballo, A. C., Nava-Díaz, C., Kishii, M., & Singh, P. K. (2022a). Genome-wide association study for resistance to tan spot in synthetic hexaploid wheat. *Plants*, *11*(3), 433. https://doi.org/10.3390/plants11030433

Lozano-Ramirez, N., Dreisigacker, S., Sansaloni, C., He, X., Sandoval-Islas, J., Pérez-Rodríguez, P., Carballo, A. C., Diaz, C. N., Kishii, M., & Singh, P. K. (2022b). Genome-wide association study for spot blotch resistance in synthetic hexaploid wheat. *Genes*, *13*(8), 1387. https://doi.org/10.3390/genes13081387

Martini, J. W. R., Crossa, J., Toledo, F. H., & Cuevas, J. (2020). On Hadamard and Kronecker products in covariance structures for genotype × environment interaction. *Plant Genome*, *13*(3), 1–12. https://doi.org/10.1002/tpg2.20033

Martini, J. W. R., Molnar, T. L., Crossa, J., Hearne, S. J., & Pixley, K. V. (2021). Opportunities and challenges of predictive approaches for harnessing the potential of genetic resources. *Frontiers in Plant Science*, *12*, 10–13. https://doi.org/10.3389/fpls.2021.674036

Mclaren, C. G., Bruskiewich, R. M., Portugal, A. M., & Cosico, A. B. (2005). The International Rice Information System. A platform for meta-analysis of rice crop data. *Plant Physiology*, *139*(2), 637–642. https://doi.org/10.1104/pp.105.063438

McLaren, C., Ramos, L., Lopez, C., & Eusebio, W. (2000). Applications of the geneaology manegment system. In C. McLaren, J. White, & P. Fox (Eds.), *International crop information system* (Technical development manual, version VI) (pp. 8–13). CIMMYT.

Mestiri, I., Chagué, V., Tanguy, A.-M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Chalhoub, B., & Jahier, J. (2010). Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. *New Phytologist*, *186*(1), 86–101. https://doi.org/10.1111/j.1469-8137.2010.03186.x

Miller, D., Ball, E., & Technau, U. (2005). Novel patterns of gene expression in polyploid plants. *Trends in Genetics*, *21*(10), 539–543. https://doi.org/10.1016/j.tig.2005.08.002

Mujeeb-Kazi, A., & Hettel, G. (1995). *Utilizing wild grass biodiversity in wheat improvement: 15 years of wide cross research at CIMMYT*. CIMMYT.

Nelson, J. C., Singh, R. P., Autrique, J. E., & Sorrells, M. E. (1997). Mapping genes conferring and suppressing leaf rust resistance in wheat. *Crop Science*, *37*(6), 1928–1935. https://doi.org/10.2135/cropsci1997.0011183x003700060043x

Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., & Chen, Z. J (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature*, *457*(7227), 327–331. https://doi.org/10.1038/nature07523

Nnadi, N. E., & Carter, D. A. (2021). Climate change and the emergence of fungal pathogens. *PLoS Pathogens*, *17*(4), 1–6. https://doi.org/10.1371/journal.ppat.1009503

Ozkan, H. (2001). Allopolyploidy-induced rapid genome evolution in the wheat (Aegilops-Triticum) group. *Plant Cell*, *13*(8), 1735–1747. https://doi.org/10.1105/tpc.13.8.1735

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

Pérez-Rodríguez, P., Crossa, J., Rutkoski, J., Poland, J., Singh, R., Legarra, A., Autrique, E., Campos, G. D. L., Burgueño, J., & Dreisigacker, S. (2017). Single-step genomic and pedigree genotype × environment interaction models for predicting wheat lines in international environments. *Plant Genome*, *10*(2), plantgenome 2016.09.0089. https://doi.org/10.3835/plantgenome2016.09.0089

Pérez-Rodríguez, P., Gianola, D., González-Camacho, J. M., Crossa, J., Manès, Y., & Dreisigacker, S. (2012). Comparison between linear and non-parametric regression models for genome-enabled prediction in wheat. *G3 Genes, Genomes, Genet*, *2*(12), 1595–1605. https://doi.org/10.1534/g3.112.003665

Rapp, R. A., Udall, J. A., & Wendel, J. F. (2009). Genomic expression dominance in allopolyploids. *BMC Biology*, *7*, 1–10. https://doi.org/10.1186/1741-7007-7-18

Rembe, M., Zhao, Y., Jiang, Y., & Reif, J. C. (2019). Reciprocal recurrent genomic selection: An attractive tool to leverage hybrid wheat breeding. *Theoretical and Applied Genetics*, *132*(3), 687–698. https://doi.org/10.1007/s00122-018-3244-x

Rosyara, U., Kishii, M., Payne, T., Sansaloni, C. P., Singh, R. P., Braun, H.-J., & Dreisigacker, S. (2019). Genetic contribution of synthetic hexaploid wheat to CIMMYT's spring bread wheat breeding germplasm. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-47936-5

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., & Kilian, A. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings*, *5*(S7), P54. https://doi.org/10.1186/1753-6561-5-s7-p54

Sarhadi, A., Ausín, M. C., Wiper, M. P., Touma, D., & Diffenbaugh, N. S. (2018). Multidimensional risk in a nonstationary climate: Joint probability of increasingly severe warm and dry conditions. *Science Advances*, *4*(11), eaau3487. https://doi.org/10.1126/sciadv.aau3487

Semagn, K., Iqbal, M., Jarquin, D., Crossa, J., Howard, R., Ciechanowska, I., Henriquez, M., Randhawa, H., Aboukhaddour, R., Mccallum, B., Brûlé-Babel, A., Navabi, A., N'diaye, A., Pozniak, C., & Spaner, D. (2022). Genomic predictions for common bunt, FHB, stripe rust, leaf rust, and leaf spotting resistance in spring wheat. *Genes*, *13*(4), 565. https://doi.org/10.3390/genes13040565

Shaked, H. (2001). Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell*, *13*(8), 1749–1759. https://doi.org/10.1105/tpc.13.8.1749

Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, *5*(3), 291–317. https://doi.org/10.1007/s12571-013-0263-y

Singh, P. K., Singh, R. P., Duveiller, E., Mergoum, M., Adhikari, T. B., & Elias, E. M. (2010). Genetics of wheat-*Pyrenophora tritici-repentis* interactions. *Euphytica*, *171*(1), 1–13. https://doi.org/10.1007/s10681-009-0074-6

Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., Sansaloni, C. P., Joynson, R., Brabbs, T., Ortiz, C., Solis-Moya, E., Govindan, V., Gupta, N., Sidhu, H. S., Basandrai, A. K., Basandrai, D., Ledesma-Ramires, L., Suaste-Franco, M. P., Fuentes-Dávila, G., … Pixley, K. V. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-30667-4

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

Van De Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2021). Polyploidy: An evolutionary and ecological force in stressful times. *Plant Cell*, *33*(1), 11–26. https://doi.org/10.1093/plcell/koaa015

Van De Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*(7), 411–424. https://doi.org/10.1038/nrg.2017.26

Vasudevan, A., Lévesque-Lemay, M., Edwards, T., & Cloutier, S. (2023). Global transcriptome analysis of allopolyploidization reveals large-scale repression of the D-subgenome in synthetic hexaploid wheat. *Communications Biology*, *6*(1), 426. https://doi.org/10.1038/s42003-023-04781-7

Yu, M., Guan, L.-L., Chen, G.-Y., Pu, Z.-E., & Hou, D.-B (2017). Allopolyploidy-induced rapid genomic changes in newly generated synthetic hexaploid wheat. *Biotechnology and Biotechnological Equipment*, *31*(2), 236–242. https://doi.org/10.1080/13102818.2016.1273797

Yu, X., Li, X., Guo, T., Zhu, C., Wu, Y., Mitchell, S. E., Roozeboom, K. L., Wang, D., Wang, M. L., Pederson, G. A., Tesso, T. T., Schnable, P. S., Bernardo, R., & Yu, J. (2016). Genomic prediction contributing to a promising global strategy to turbocharge gene banks. *Nature Plants*, *2*, 16150. https://doi.org/10.1038/nplants.2016.150

Zhang, H., Bian, Y., Gou, X., Dong, Y., Rustgi, S., Zhang, B., Xu, C., Li, N., Qi, B., Han, F., Von Wettstein, D., & Liu, B. (2013). Intrinsic karyotype stability and gene copy number variations may have laid the foundation for tetraploid wheat formation. *PNAS*, *110*(48), 19466–19471. https://doi.org/10.1073/pnas.1319598110

Zhao, N., Xu, L., Zhu, B., Li, M., Zhang, H., Qi, B., Xu, C., Han, F., & Liu, B. (2011). Chromosomal and genome-wide molecular changes associated with initial stages of allohexaploidization in wheat can be transit and incidental. *Genome*, *54*(8), 692–699. https://doi.org/10.1139/g11-028

Zhao, N., Zhu, B., Li, M., Wang, L., Xu, L., Zhang, H., Zheng, S., Qi, B., Han, F., & Liu, B. (2011). Extensive and heritable epigenetic remodeling and genetic stability accompany allohexaploidization of wheat. *Genetics*, *188*(3), 499–509. https://doi.org/10.1534/genetics.111.127688

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 can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dreisigacker, S., Martini, J. W. R., Cuevas, J., Pérez-Rodríguez, P., Lozano-Ramírez, N., Huerta, J., Singh, P., Crespo-Herrera, L., Bentley, A. R., & Crossa, J. (2024). Genomic prediction of synthetic hexaploid wheat upon tetraploid *durum* and diploid *Aegilops* parental pools. *The Plant Genome*, *17*, e20464. https://doi.org/10.1002/tpg2.20464

19403372, 2024, 2, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.20464 by Cochrane Mexico, Wiley Online Library on [03/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License