



Correlation analysis between heterosis and genetic distance evaluated by 15K SNP array in hybrid wheat

Mingliang Ding^{1,2†}, Shaoxiang Li^{2†}, Jia Liu³, Hongsheng Li², Mujun Yang², Kun Liu², Zhonghui Yang², Jian Yin², Sedhom Abdelkhalik^{2,4}, Ahmed M. S. Elfanah^{2,4}, Zhanhong Ma¹

¹Department of Plant Pathology, College of Plant Protection, China Agricultural University, Beijing, China

²Food Crops Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China

³Tuber Crops Research Institute, Yuxi Academy of Agricultural Sciences, Kunming, China

⁴Wheat Research Department, Field Crops Research Institute, Agricultural Research Center, Giza, Egypt

Hybrid breeding is a promising technology for improving wheat yield potential and stability, especially in marginal environments. However, creating hybrids with strong heterosis is challenging. This study screened T307 restorer lines and seven thermo-photosensitive genic male sterile (TPGMS) lines genotyped by a 15K SNP array. Subsequently, based on the genetic distances (GDs) among parental lines, two male sterile lines and 16 restorer lines representing close, middle, and distant genetic distances relative to each sterile line were selected for production of 18 hybrids. The hybrids were subjected to a two-year field trial for assessment of 27 agronomic and end-use quality characteristics. We found that the GDs among the selected 2 sterile lines and 16 restorer lines ranged from 0.04 to 0.66, with an average of 0.47. The 18 parental lines are divided into five groups based on the phylogenetic tree. The relationship between GDs and the mid-parent heterosis showed a weak negative correlation for a majority of characteristics. Likewise, for the grain yield per plant, non-significant negative correlations were observed between GDs and the mid-parent heterosis (MPH), better parent heterosis (BPH), and commercial check heterosis (CCH). However, hybrid S003×R080 with the genetic distance of 0.36 and hybrid S005×R084 with the genetic distance of 0.44, demonstrated superior performance across all tested attributes. This study suggests that SNP marker based genetic distance is a poor predictor of heterosis and high heterosis can be obtained in hybrids of moderately distant parental lines.

Keywords: Genetic distance, GYT biplots, Heterosis, Hybrid wheat, SNP marker, TPSGMS

INTRODUCTION

Wheat is a vital food crop, supplying approximately 20% of daily caloric intake and 21% of protein needs (Tadesse et al., 2019; Li et al., 2020). Wheat productivity should be increased based on genetic improvements to cope with the increasing demand globally and enhance food security, particularly under climate change (e.g., global warming) (Godfray et al., 2010; Melonek et al., 2021; El Hanafi et al., 2022). Among various

breeding strategies for genetic improvement of wheat, hybrid wheat is a promising choice for improving yield potentiality and stability across different environments. Although the exploitation of hybrid wheat remains a global challenge, wheat researchers worldwide continue to strive towards this goal, making this breeding strategy one of the hot topics in global research. During the past sixty years, hybrid wheat systems based on cytoplasmic male sterility (CMS), chemical hybridization agent (CHA), photoperiod-sensitive

ARTICLE HISTORY

Submitted: January 7, 2025

Accepted: July 3, 2025

CORRESPONDANCE TO:

- Sedhom Abdelkhalik

Food Crops Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China; Wheat Research Department, Field Crops Research Institute, Agricultural Research Center, Giza, Egypt

Email: Sedhom_aiad@yahoo.com

- Zhanhong Ma

Department of Plant Pathology, College of Plant Protection, China Agricultural University, Beijing, China

Email: mazh@cau.edu.cn

†These authors have contributed equally to this work and share first authorship.

DOI: 10.21608/ejbo.2025.350896.3153

EDITED BY: Reda Gaafar

cytoplasmic male sterility (PCMS), and thermo-photo sensitive genic male sterility (TPSGMS) have been established worldwide for utilizing heterosis in wheat (Hoagland et al., 1953; Godfray et al., 2010; Basnet et al., 2018; Tadesse et al., 2019; Melonek et al., 2021; El Hanafi et al., 2022). At present, among these hybrid wheat breeding strategies, the “two-line hybrid wheat” based on thermo-photo sensitive or photo-thermo sensitive male sterile lines has been implemented on a large scale and has effectively increased yield per unit area, making China the only country to commercialize hybrid wheat using sterile lines (Li et al., 2022; Zhao, 2022). This strategy comprises male sterile lines cultivated under short-day and low-temperature conditions for hybrid seed generation, whereas the same male sterile line subjected to long-day and high-temperature conditions performs as fertile lines in the event of self-crossing. This strategy is considered a main technology for the utilization of wheat heterosis in the future (Zhao, 2022). Despite the release of 18 hybrid wheat varieties since 2002 employing this strategy, the efficiency of generating strong heterosis hybrid wheat combinations remain relatively low compared to other crops such as rice, corn, and rapeseed. One important reason is that the method for selecting strong heterosis combinations in hybrid wheat still relies on randomly creating a large number of hybrid combinations for yield evaluation. This breeding strategy is not only time-consuming and labor-intensive but also has a low efficiency. Conducting heterosis prediction and subsequent classification of heterotic groups in the current predicament is an effective approach to overcoming the aforementioned challenges.

In recent years, numerous researchers have employed molecular markers to ascertain the genetic distances (GDs) of wheat parents and predict wheat heterosis, yielding some promising research outcomes. For instance, Al-Ashkar et al. (2020) calculated the GDs among 16 wheat parents using 60 SSR markers and discovered that heterosis in terms of grain filling period, thousand-grain weight, kernel number per spike, harvest index, grain yield, and filling rate was significantly positively correlated with the GDs between parents. Conversely, there was a significant negative correlation with the growth period, indicating the potential to predict wheat hybrids with high grain yield and early maturity. Liu et al. (1999) calculated the GDs among 20 wheat parents based on a limited number of RAPD

markers, but they found no significant correlation between the heterosis for grain yield and the GDs among the parents. As stated by Nie et al. (2019), the wheat genome is vast, and the use of a limited number of molecular markers to assess the GDs in wheat is inaccurate; further employment of high-density molecular markers is necessary for precise evaluation. As wheat genome sequencing efforts continue to advance, an increasing number of high-density wheat SNP arrays with various application objectives are being developed and utilized, such as: 9KiSelect, 90KiSelect, 820K Axiom array, 35K Axiom array, 50K Trait Breed, 15K, 660K Axiom array, 55K, etc. (Sun et al., 2020). These SNP arrays have been extensively applied in the fields of wheat genetic diversity and population structure analysis (Joukhadar et al., 2017), QTL mapping (Li et al., 2021; Jia et al., 2024), and trait association analysis (Ye et al., 2024). However, the application of SNP array technology for predicting heterosis in hybrid wheat is rarely reported. Nie et al. (2019) utilized the 90K SNP chip to assess the GDs among 20 wheat genotypes and further analyzed the relationship between GDs and the F_1 heterosis. The study found that there was a non-significant positive correlation between grain yield's mid-parent heterosis and better-parent heterosis with the GDs among parents; a non-significant negative correlation was observed for the mid-parent heterosis and better-parent heterosis of kernel number per spike and effective spike number with the GDs among parents; a significant positive association was found between the mid-parent heterosis of thousand-kernel weight and the GDs among parents, while the better-parent heterosis of thousand-kernel weight showed a significant correlation. Chen et al. (2022) utilized a wheat 660K SNP chip to assess GDs among 32 wheat genotypes, and further analyzed the relationship between GDs and heterosis in their F_1 hybrids. It was found that both mid-parent heterosis and over-parent heterosis in single-plant grain yield were not significantly positively correlated with the GDs between parents. In summary, the varying research outcomes concerning the relationship between genetic distance and heterosis in wheat parents exhibit certain discrepancies, which may be associated with the types and quantities of molecular markers utilized, as well as the types and numbers of wheat parental resources studied.

In this study, we collected 307 restorer lines from different geographical origins and seven thermo-photoperiod-sensitive genic male sterile

(TPSGMS) lines, and genotyped these lines using the 15K SNP chip to determine the genetic distances among lines. Subsequently, we selected two representative TPSGMS lines were selected to cross with three groups of restorers with far, middle, and near genetic distances, resulting in 18 hybrid combinations. In the field trials conducted during the 2021/22 and 2022/23 seasons, we evaluated 27 agronomic and quality traits of these hybrids and their parents. This study aimed to assess the feasibility of heterosis prediction by using SNP marker based genetic distance in two-line hybrid wheat breeding program.

MATERIALS AND METHODS

The wheat germplasm

Seven wheat thermo-photosensitive genetic male sterility lines, including S001(K456S), S002 (K43S), S003 (K64S), S004 (K78S), S005 (K47S), S006 (K66S), and S007 (K63S) were bred by the Food Crops Research Institute, Yunnan Academy of Agricultural Sciences (YAAS), Yunnan, China. Restorer 307 lines have diverse sources, mainly from Yunnan province and others from different provinces in China and CIMMYT (Supplementary Tables 1).

Genotyping of restorer and sterile lines by single nucleotide polymorphism (SNP)

During the 2019/20 cropping season, whole wheat genotypes comprising seven male sterile lines and 307 restorer lines were sown at the Songming experimental farm (altitude 1882.2m, east longitude 103° 6' 41", north latitude 25° 21' 18") of

the Yunnan Academy of Agricultural Sciences. The genotypes were employed to obtain DNA samples prior to the jointing growth stage by individually collecting fresh leaves. The DNA extraction procedure utilizing CTAB was conducted as per the methodology outlined by Rogers & Bendich (1985). The DNA purity was estimated by UV spectrophotometer at the absorbance wavelengths of $A_{260-280}$ nm. The purity ranged between 1.8-2.0, and the DNA concentration adjusted to approximately 100 ng/μL. The Zhongyujin Biotechnology laboratory (Co., Ltd., Beijing, China) labeled DNA samples from male sterile and restorer lines. The Axiom platform genotyped all lines using the Zhongmaixin No.1, a 15K SNP array for wheat breeding in Affimatrix®. Zhongmaixin 1 chip consists of 13,702 SNP markers, including 1272 functional markers, with a uniform distribution of markers, and the average physical distance between the two markers on each chromosome is about 1M (Table 1).

The primary genotypic data analysis using Axiom Analysis Suite software was performed. First, SNP site quality control is performed on samples with a data quality control (DQC) of > 0.82 and a marker detection rate (CR) of > 95%. The remaining 11,198 markers were then filtered to remove the markers with a deletion rate (Miss) of < 10% and a minimum allele frequency (MAF) of > 0.05%, then the total remaining were 10198 markers. Two restorer lines, R085 and R181, were excluded from the SNP study due to inferior DNA purity.

Table 1. SNP Marker distribution and the average physical distance per marker in 21 chromosomes of common wheat

Chromosome code	Number of SNP markers	Average physical distance	Chromosome code	Number of SNP markers	Average physical distance
1A	625	950563	5B	795	897044
2A	932	837767	6B	802	898988
3A	606	1239016	7B	657	1142497
4A	782	952159	1D	370	1339063
5A	713	995475	2D	634	1028159
6A	480	1287665	3D	536	1148419
7A	760	969350	4D	265	1923989
1B	685	1007083	5D	550	1029238
2B	755	1061267	6D	428	1106525
3B	1025	810566	7D	697	916336
4B	605	1113417			

Genetic distance

Computing whole genetic distances (GDs) between 307 restorers and seven male sterile lines, selected genotypes according to near, moderate, and far GDs. Two sterile lines were pointed on the same cluster, whereas the 16 selected restorers were distributed on several evolutionary tree branches and divided into three groups according to their GDs from sterile lines (Supplementary Figure S1).

Focusing on GD between S003 (awned type) and S005 (awnless type) male sterile lines and their mating restorer lines, including The 16 selected restorer lines, R065, R068, R080, R084, R129, R139, R155, R159, R162, R192, R199, R202, R220, R228, R269, and R299, were calculated. The relationship between far, middle, and near genetic distances with heterosis performance was separately calculated (GDs) to discover their relationship.

Each sterile line was pollinated with nine restorer lines to produce 18 hybrids, i.e., three near lines, three moderate, and three lines far from GDs. These restorer lines, male sterile lines, and their

crossing combinations (F_1 hybrids) were evaluated in the open field trials to examine the hybrid yield superiority in appropriate conditions for two-line hybrid wheat production.

Field experiments and estimated traits

The hybrid seeds were harvested and replanted in October 2021/22 and 2022/23 cropping seasons at the Jiajing experimental farm (1615.2m above sea level, 102 ° 17 ' 52 " E, 24 ° 10 ' 3 " N) of the Yuxi Academy of Agricultural Sciences in Yanhe Town, Yuxi City, Yunnan Province. The experimental soil type is sandy loam with flat terrain and uniform soil fertility, and the previous crop was rice.

Thirty-seven-wheat genotypes comprised 16 restorer lines, two male sterile lines (parents), 18 hybrids, and the commercial check Yunmai 56 are shown in Table 2. Each genotype was planted in one row of 105cm with plants spaced 15cm (e.g., sown seven plants per row) and rows spaced 25cm. Experiments were designated with three replicates in a randomized complete block design (RCBD) in both cropping seasons.

Table 2. Genotype number, hybrid combination, parent name, and check.

Genotype	Hybrid	Genotype	Parent/check name
1	S005/R065	19	R065
2	S005/R068	20	R068
3	S005/R084	21	R080
4	S005/R129	22	R084
5	S005/R155	23	R129
6	S005/R159	24	R139
7	S005/R192	25	R155
8	S005/R220	26	R159
9	S005/R299	27	R162
10	S003/R068	28	R192
11	S003/R080	29	R199
12	S003/R129	30	R202
13	S003/R139	31	R220
14	S003/R162	32	R228
15	S003/R199	33	R269
16	S003/R202	34	R299
17	S003/R228	35	S003 (K64S)
18	S003/R269	36	S005 (K47S)
-	-	37	Yunmai 56 (Check)

Agronomic traits were determined, such as growth period (GP, days), plant height (PH, cm), tillers number per plant (TNP), effective spikes per plant (ESP), spikelets number (SN), number of degenerated spikelets (NDS), number of filling spikelets (NFS), spike length (SL, cm), grains number per spike (GNS), single spike weight (SSW, g), grains number per plant (GNP), thousand-grain weight (TGW, g), and three represented plants (without the bordered plants) were randomly harvested from each row to investigate grain yield per plant average (GYP, g).

The grain shape attributes were tested by an automatic seed analysis system (Wanshen SC-G, Hangzhou, China). More than 200 grains were measured for each genotype, including grain area (GA, mm²), grain circumference (GC, mm), grain length (GL, mm), grain width (GW, mm), and length-to-width ratio (LWR, %). For more details for each trait, the technical regulations for regional trials of crop varieties (wheat) were followed, and the traits of sterile lines were estimated for their outcrossing plants (Ding et al., 2022).

The grain quality characteristics were determined using a DA7200 near-infrared quality analyzer (Pertin Instruments, Huddinge, Sweden). The estimated traits were water absorption rate (WAR, %), protein content (PC, %), bulk density (BD, g/l), wet gluten (WG, %), stability time (ST, min), formation time (FT, min), hardness value (HV, %), settlement value (SV, ml) and flour yield (FY, %). For more details, refer to Ling et al. (2020).

The hybrid heterosis evaluation

The mid-parent heterosis, better-parent heterosis, and commercial check heterosis for all estimated traits in each hybrid combination verify the accuracy of the heterosis group divided by the genetic distances computed. The calculation formulas are listed as follows: Mid-parent heterosis (%) = $(F_1 \text{ mean} - \text{mid-parents mean}) / (\text{mid-parents mean} \times 100)$; High parent heterosis (%) = $(F_1 \text{ mean} - \text{high parent mean}) / (\text{high parent average} \times 100)$; Commercial check heterosis (%) = $(F_1 \text{ mean} - \text{commercial check mean}) / (\text{commercial check mean} \times 100)$.

Statistical analyses

The evolutionary tree of all studied genotypes (307 restorers plus seven sterile lines) was inferred by a hierarchical neighbor-joining method utilizing DARwin 6.0.21 software (Dissimilarity Analysis and Representation for Windows) (Perrier et al., 2003).

Genetic distances (GDs) between the parental lines of hybrids were calculated based on the SNP markers using Pairwise Distance from section/ribbon of Distances in MEGA 11 software Tamura et al. (2021). The model was used according to the methods of Tajima & Nei (2000), and the Neighbor Jointed Tree from the section of Phylogeny was utilized to compute and construct the phylogenetic tree. The bootstrap (1000 replicates) and p-distance methods were used to identify the evolutionary distances (Nei & Kumar, 2000).

The agronomic and quality characteristics data of field trials were subjected to analysis of variance (ANOVA) of randomized complete block design (RCBD). For each season, then mutual (combined) analysis over 2021/22 and 2022/23 seasons using GenStat 23rd edition (VSN International Ltd., Hemel Hempstead, UK).

Moreover, the mean performance of predicted characteristics for all parents and hybrids, aggregated across seasons, was employed in genotyping yield*trait (GYT) biplot methodologies, as per Yan & Frégeau-Reid (2018). The GenStat program finalized the data standardization and visualization processes. To ascertain the optimal genotype based on trait combinations. The data of studied traits of parents and hybrids were averaged over seasons and subjected to calculate the heatmap dendrogram. In contrast, the correlation plot employed all studied traits' mid-parent heterosis (MPH) with GDs. Additionally, to generate a radar chart, traits' data were normalized by minimum and maximum values to convert their values to unitless traits. The heatmap dendrogram, correlation plot, and radar chart were computed by Origin (Pro), version 2021 (Origin Lab Corporation, Northampton, MA, USA).

RESULTS

Parental selection and genetic distance of hybrid parents

Based on genetic distances (GDs) obtained from 15K SNP markers analysis for 307 restorers and seven sterile lines, we selected two sterile lines and 16 restorers according to the near, middle, and far GDs between sterile lines and restorers from the evolutionary tree (Supplementary Figure S1). Two sterile lines include S003 (awned type) and S005 (awnless type) in the same cluster, while 16 restorers R065, R068, R080, R084, R129, R139, R155, R159, R162, R192, R199, R202, R220, R228, R269, and R299 are distributed on five

clusters (Figure 1). Meanwhile, the GDs between these selected parents ranged from 0.04 to 0.66, with an average of 0.47 (Supplementary Table 2). As listed in Table 3, hybrids S003×R139, S003×R080, S003×R068, and S005×R159, S005×R068, S005×R192 have closer GDs between their male and female parents, with GDs ranged from 0.18 to 0.36, while S003×R269, S003 × R129, S003 × R199, and S005 × R129, S005 × R299 are distant GDs combinations, with GDs ranged from 0.54 to 0.64. Others are middle GDs combinations.

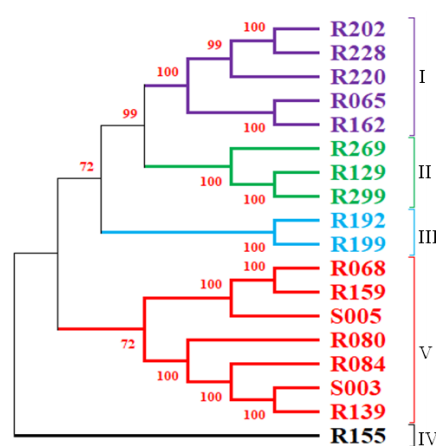


Figure 1. The phylogenetic tree of two sterile lines, S003, S005, and 16 restorers based on SNP markers

Mid-parent, better Parent, and commercial check heterosis

Figure 2 represents the mean performance of 18 wheat hybrids in mid-parent, better parents, and commercial heterosis, relying on the estimated traits over both seasons. The mid-parent heterosis for all hybrids was common and the values were higher than those of better parents heterosis, specifically for hybrids 11 (S003×R080) and 3 (S005×R084). Among all 18 hybrids, hybrid 11 (S003×R080) had the highest performance of three types of heterosis, it has a near parental GD at 0.36. However, hybrid 4 has the smallest mid-parent heterosis, and better parents heterosis, and hybrid 5 has the smallest commercial check heterosis.

Correlation between genetic distance and three heterosis types

The Pearson's correlation plot for genetic distances and mid-parent heterosis of the 18 hybrids for agronomic and quality traits are shown in Figure 3. Except for traits such as tillers number per plant (TNP) and protein content (PC), a non-significant negative correlation is observed between GDs

among parents and mid-parent heterosis for most traits. The Pearson's correlation plot for genetic distances and mid-parent heterosis (MPH), better parents (BPH) and commercial check heterosis (CCH) of the 18 hybrids for grain yield are shown in Figure 4. Likewise, for the significant trait of per plant grain yield, non-significant negative correlations are observed between its mid-parent heterosis (MPH), better parent heterosis (BPH), commercial check heterosis (CCH), and the genetic distances among parents.

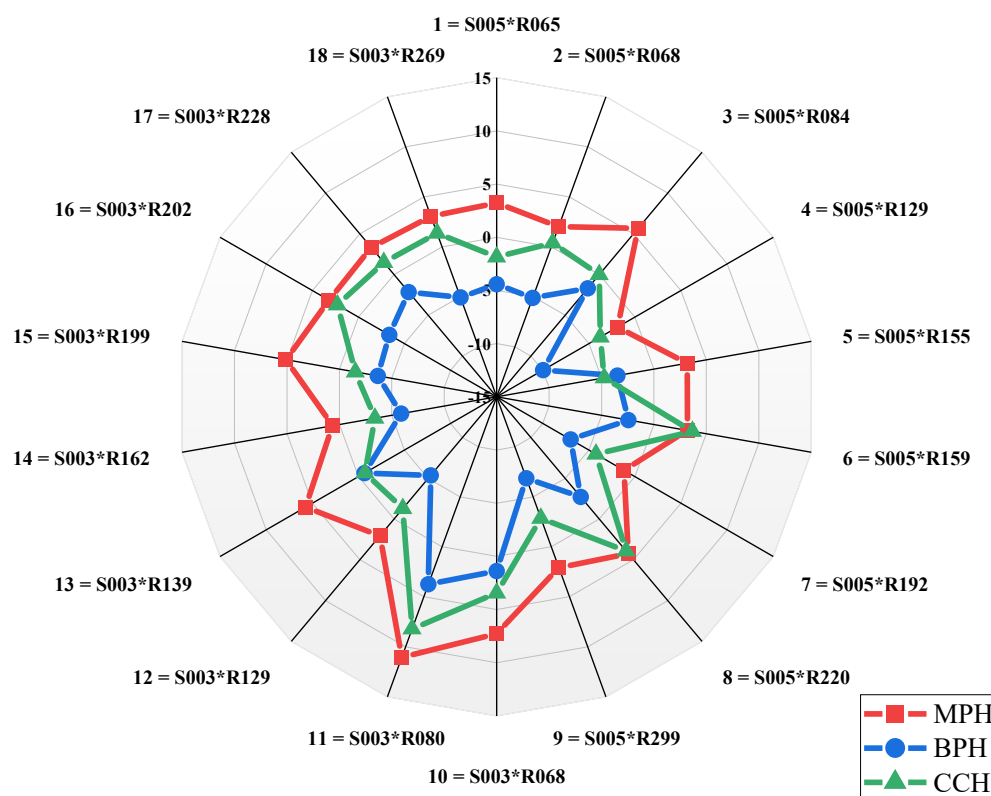
Genotype by yield*traits (GYT) biplot and rankings

The genotype by yield traits (GYT) analysis of 37 genotypes and their grain yield, multiplied by various attributes (combinations), is depicted in Figure 5. The principal component analysis (PC1 and PC2) accounts for 89.6% of the overall variation in the GYT biplot. Genotype 11 is a hybrid wheat (S003×R080) situated on the vertices of the polygon containing most of the trait's combinations, and hybrid 3 is located in the other sector containing combinations such as GYP*NDS, GYP*GNP, and GYP*ESP. These results indicated the superiority of hybrid S003×R080 and S005×R084.

Figure 6 displays the genotype ranking of all assessed genotypes and trait combinations in the GYT biplot. Genotype 11 (S003×R080) is regarded as the superior hybrid, succeeded by genotypes 3 (S005×R084). The lowest rating was attributed to genotype 25 (restorer line). Moreover, genotypes 11 and 3 are proximate to the average tester coordination (ATC) line (i.e., a line with an arrow) and exhibited brief projections. The brief projection indicates that the stable genotypes in both examined environments are oriented away from the ATC line in contrast to genotypes 1 and 33. Therefore, Among all evaluated hybrid wheat combinations and their parents, the two hybrid combinations with the best comprehensive performance are Genotype 11 (S003×R080) and Genotype 3 (S005×R084), which demonstrates that the heterosis in hybrid wheat is indeed real. Interestingly, the genetic distance between the parents of the best-performing hybrid wheat combinations, S003×R080 (0.36) and S005×R084 (0.44), is less than or equal to the average GD of S003 (0.48) and S005 (0.44) with all their respective restorer lines. This suggests that a moderate genetic distance between parents is conducive to the expression of heterosis when formulating hybrid wheat combinations using the thermo-photoperiod sensitivity two-line system.

Table 3. 18 hybrids and their parental genetic distance of selected sterile lines and restorer based on SNP marker analysis

Genotype	Hybrid	Genetic distance	Genotype	Hybrid	Genetic distance
1	S003 × R139	0.18	10	S005 × R068	0.29
2	S003 × R068	0.32	11	S005 × R159	0.29
3	S003 × R080	0.36	12	S005 × R192	0.35
4	S003 × R162	0.54	13	S005 × R084	0.44
5	S003 × R202	0.54	14	S005 × R155	0.48
6	S003 × R228	0.55	15	S005 × R065	0.51
7	S003 × R269	0.60	16	S005 × R220	0.51
8	S003 × R129	0.63	17	S005 × R129	0.54
9	S003 × R199	0.64	18	S005 × R299	0.54

**Figure 2.** Radar chart of mid-parent (MPH), better parents (BPH), and commercial check heterosis (CCH) (estimated heterosis values) of all agronomic and grain quality characteristics for 18 hybrids estimated during two cropping seasons of 2021/22 and 2022/23

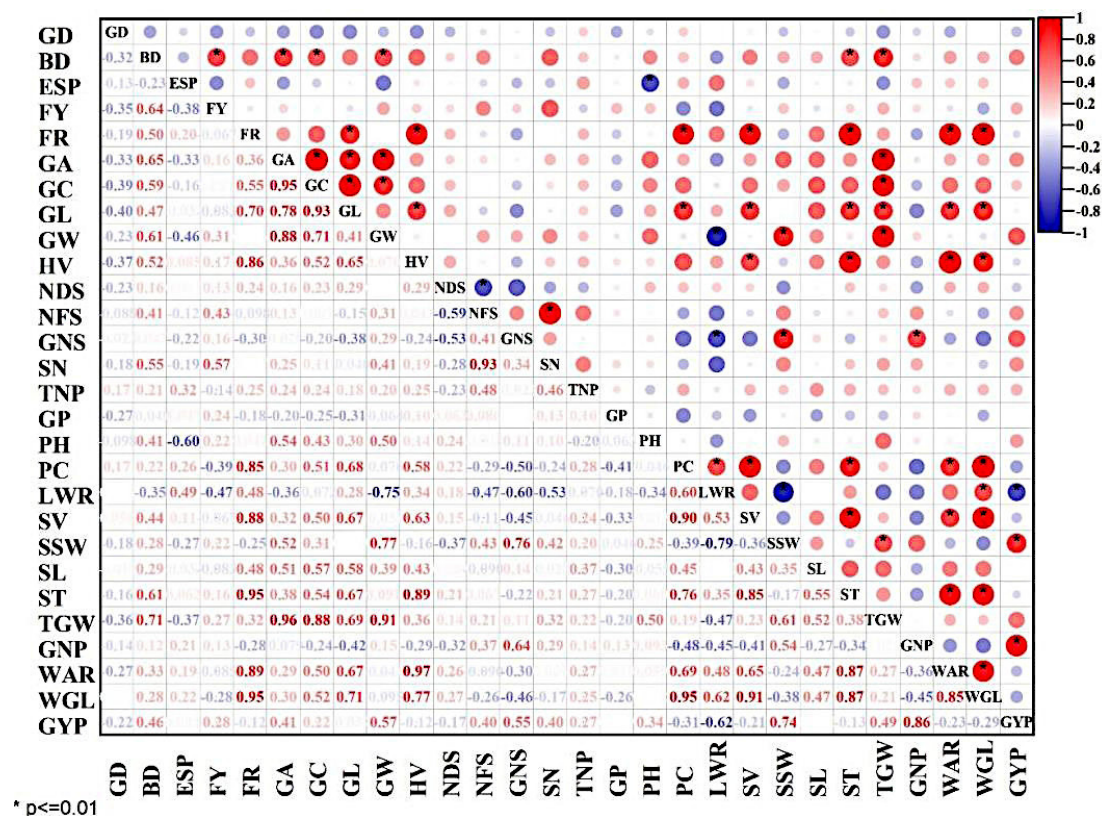


Figure 3. Person correlation plot for mid-parent heterosis performance and genetic distance (GD) for 18 hybrid wheat of agronomic and quality traits [e.g., GP, growth period; PH, plant height; TNP, tillers number per plant; ESP, effective spikes per plant; SN, spikelets number; NDS, number of degenerated spikelets; NFS, number of filling spikelets; SL, spike length; GNS, grains number per spike; SSW, single spike weight; GNP, grains number per plant; TGW, thousand-grain weight; GYP, grain yield per plant; GA, grain area; GC, grain circumference; LWR, length to width ratio; GL, grain length; TGW, total grain width; WAR, water absorption rate; PC, protein content; BD, bulk density; WG, wet gluten; ST, stable time; FT, formation time; HV, hardness value; SV, settlement value; FY, flour yield estimated from the two seasons average]

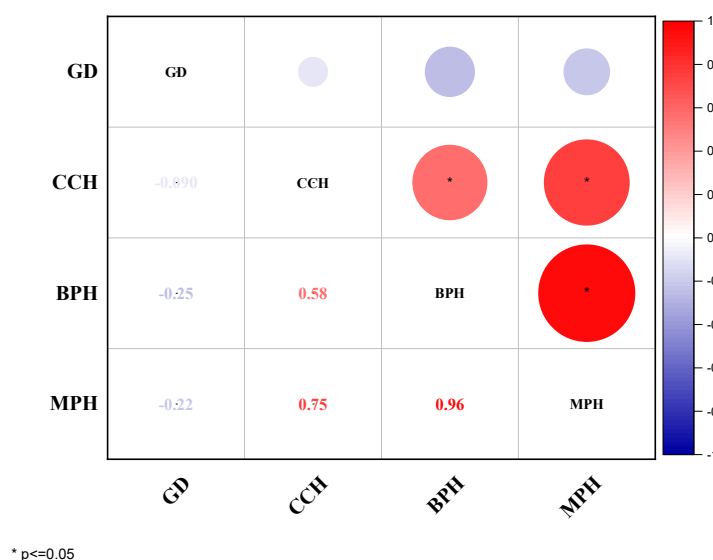


Figure 4. Pearson's correlation plot for mid-parent heterosis (MPH), better parents (BPH) and commercial check heterosis (CCH) performance and genetic distance (GD) for 18 hybrid wheat of grain yield

DISCUSSION

Based on the near, middle, and far genetic distances (GDs) of two TPGMS lines from 16 restorer lines utilizing 15K of SNP markers, hybrid combinations were produced to evaluate the heterosis in field multi-trials. This selection reflects the hybrid breeding programs' main target: identifying genotypes with broad genetic diversity. Hence, it is expected to generate strong heterosis performance, precise DNA level, and field experimental verification (Boeven et al., 2016; Al-Ashkar et al., 2020; Wu et al., 2021). This study utilized various amounts of GDs, as presented in Table 2 and Figure 1, to develop hybrid wheat combinations. This selection criterion and approach align with other prior studies, including those on wheat (Chen et al., 2010; Nie et al., 2019; Al-Ashkar et al., 2020; Semagn et al., 2021), maize (Masuka et al., 2017; Demail et al., 2020; Jiang et al., 2023), and pearl millet (Sattler et al., 2019). Most of these studies utilized different genotyping techniques in computing the GDs and predicting F_1 performance for several agronomic, physiological, and end-use quality characteristics.

Moreover, the GDs among the selected 2 sterile lines and 16 restorer lines varied from 0.04 to 0.66 in our study, with a mean value of 0.47 as presented in Table 1S. This range indicates the extent of variation in genetic distances, despite the different geographical origins of the wheat germplasm. Similarly, our estimated genetic dissimilarities are less than those studied by SSR markers on wheat, i.e., Huang et al. (2002) stated that the genetic diversity ranged from 0.43-0.94 with an average of 0.77, applied for 998 bread wheat genotypes. However, our estimated GDs are greater than the GDs determined by Nie et al. (2019) using a wheat 90K SNP chip for genotyping among 20 wheat parents, where the GD among all parents ranged from 0.008 to 0.276, with an average of 0.212. This indicates that the wheat parents selected in our study have a favorable GDs, and the combinations prepared are representative for evaluating the relationship between GD and heterosis. Consequently, we choose close, moderate, and far GDs of restorers from sterile lines to forecast superior heterotic performance. Indeed, our study has produced hybrid combinations with varying levels of heterosis, with genotype 11 (S003/R080) and genotype 3 (S005/R084) being the most advantageous. However, the former has a closer GD, while the latter has a middle GD.

Genotypes can be categorized into heterotic and genetic groups based on phylogenetic tree and cluster analysis. In the current research, the parental lines are divided into five distinct groups, especially the sterile lines categorized in Group V in Figure 1. These findings align with the prior results reported by Hussain et al. (2022) in rice that identified heterotic groups. Additionally, in wheat and based on SSR markers coupled with an evolutionary tree, three main groups and six subsets of eight parental genotypes are reported by Al-Ashkar et al. (2020). However, in the same way, Sang et al. (2022) divided 41 inbred lines into five heterotic groups in maize. Besides, cluster analysis was performed to classify 20 wheat genotypes into five groups matching their pedigree (characteristics) (Boeven et al., 2016). It is noteworthy that the parents of the two best-performing combinations for both mid-parent heterosis and over-parent heterosis are concentrated in Group V, which is part of the heterosis group. This also indicates that the parents of thermo-photo-sensitive two-line hybrid wheat with relatively closer GDs may be more likely to produce high heterosis.

Plant breeders have devised methods to predict the performance of heterosis, including the assessment of genetic distance between parents using various DNA technologies, thereby clarifying the relationship between GDs and heterosis, and subsequently guiding the formulation of hybrid combinations. This method has been effectively applied in crops such as hybrid corn (Jiang et al., 2023), hybrid rice (Zhang et al., 2022), and upland cotton (Geng et al., 2021). Our research indicates that a negative correlation was observed between GDs and most traits except for traits such as tiller number per plant (TNP) and protein content (PC). Concurrently, for the significant trait of per plant grain yield, non-significant negative correlations are observed between its mid-parent heterosis (MPH), better parent heterosis (BPH), commercial check heterosis (CCH), and the genetic distances among parents. Our research results are consistent with previous studies by Nie et al. (2019) and Liu et al. (1999) on wheat, both considering that there is a very weak correlation between GDs among parents and heterosis. Although the correlation between GDs calculated based on SNP chip technology and heterosis is weak and cannot well predict heterosis, it can still be used to classify the heterosis groups of thermo-photo-sensitive two-line hybrid wheat parents and guide the testing and matching of hybrid combinations.

The GYT biplot model enables graphical selection of genotypes in plant breeding. This technique facilitates the collection of grain yield and associated parameters, along with genotypes shown in a visible biplot, thereby aiding plant breeders in the selection of superior genotypes (Yan & Frégeau-Reid, 2018). For example, researchers Elfanah et al. (2023a, b) utilized agronomic and physiological traits and spectral reflectance indices in the GYT model to identify and select wheat genotypes' salinity tolerance. Similarly, other authors applied the GYT biplot to select the wheat drought tolerance based on agronomic and physiological attributes and spectroscopy indices (Darwish et al., 2023). Our study hybrid combinations (S003×R080 and S005×R084) are superior hybrid wheat from the GYT biplot view (Figure 6). However, the genetic distances between the parents of the two combinations are neither the maximum nor the minimum. Despite the combinations (S003×R080 and S005×R084) exhibiting genetic distances that are less than or equal to the average genetic distance between the two sterile lines and all tested restorer lines, they demonstrate the best performance in mid-parent, better-parent, and commercial check heterosis. Therefore, an appropriate level of genetic distance may more readily result in high heterosis. Although genetic distance is necessary for the occurrence of heterosis, it is not the case that the greater the genetic distance, the higher the heterosis will be. In fact, there exists an optimal genetic distance, beyond which heterosis may decline due to an increase in genetic incompatibilities (Würschum et al., 2023). This relationship is sometimes described as “hump-shaped,” meaning that at a certain optimal point of genetic distance, heterosis reaches its maximum value, and it decreases before and after this optimal point (Würschum et al., 2023). Therefore, breeders need to find a balance, that is, a moderate genetic distance, to maximize heterosis when conducting hybrid breeding.

CONCLUSIONS

In this study, we conclude that SNP chips can be used for wheat genotyping and can accurately assess the genetic distance between wheat parents. The GDs among the selected 2 sterile lines and 16 restorer lines ranged from 0.04 to 0.66, with an average of 0.47. Except for traits such as tillers number per plant (TNP) and protein content (PC), a non-significant negative correlation is observed between genetic distance (GD) among parents and mid-parent heterosis for most traits. Concurrently,

for the significant trait of per plant grain yield, non-significant negative correlations are observed between its mid-parent heterosis (MPH), better parent heterosis (BPH), commercial check heterosis (CCH), and the genetic distances among parents. However, hybrid S003×R080 with the genetic distance of 0.36 and hybrid S005×R084 with the genetic distance of 0.44, demonstrated superior performance across all tested attributes. This study suggests that SNP marker based genetic distance is a poor predictor of heterosis and high heterosis can be obtained in hybrids of moderately distant parental lines.

Conflict of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' contributions: MD and JL performed most of the phenotyping and molecular experiments. HL, MY, KL, ZY, JY and SA undertook part of the field and laboratory work. MD, SL, and AM wrote the manuscript. AM and ZM critically revised the manuscript for important intellectual content and approved the final version to be published. All authors contributed to the article and approved the submitted version.

Funding: This research was funded by the Natural Science Foundation of Yunnan Province (202301AT070004), the Major Science and Technology Project in Yunnan Province (202302AE090011), the Central-Guided Local Science and Technology Development Fund (202407AB110015) and the National Key Research and Development Program of China (2016YFD0101603).

Supplementary material: Supplementary Tables 1, supplementary Tables 2 and supplementary Figure 1 can be found in the supplementary material.

Data availability statement: The raw data supporting the conclusions of this article will be available from the authors, without condition.

Ethical approval: Not applicable.

REFERENCE

- Al-Ashkar, I., Alotaibi, M., Refay, Y., Ghazy, A., Zakri, A., Al-Doss, A. (2020). Selection criteria for high-yielding and early-flowering bread wheat hybrids under heat stress. *Plos One*, 15(8): e0236351.
- Basnet, B.R., Crossa, J., Dreisigacker, S., Pérez-Rodríguez, P., Manes, Y., Singh, R.P., Rosyara,

- U.R., Camarillo-Castillo, F., Murua, M. (2018). Hybrid wheat prediction using genomic, pedigree, and environmental covariables interaction models. *Plant Genome*, 12. doi.org/10.3835/plantgenome2018.07.0051
- Boeven, P.H.G., Longin, C.F.H., Würschum, T. (2016). A unified framework for hybrid breeding and the establishment of heterotic groups in wheat. *Theoretical and Applied Genetics*, 129: 1231-1245.
- Chen, X., Sun, D., Rong, D.-f., Sun, G., Peng, J. (2010). Relationship of genetic distance and hybrid performance in hybrids derived from a new photoperiod-thermo sensitive male sterile wheat line 337S. *Euphytica*, 175: 365-371.
- Chen, X., Wu, X., Fang, F., Song, L., Dong, N., Hu, T., Ru, Z. (2022). Heterosis prediction of wheat based on SNP genetic distance and combining ability. *Chinese Science Bulletin*, 67: 3221-3232 (in Chinese).
- Darwish, M.A., Elkot, A.F., Elfanah, A.M., Selim, A.I., Yassin, M.M., Abomarzoka, E.A., et al. (2023). Evaluation of wheat genotypes under water regimes using hyperspectral reflectance and agro-physiological parameters via genotype by yield trait approaches in Sakha station, Delta, Egypt. *Agriculture*, 13(7): 1338. doi.org/10.3390/agriculture13071338
- Dermail, A., Suriharn, B., Chankaew, S., Sanitchon, J., Lertrat, K. (2020). Hybrid prediction based on SSR-genetic distance, heterosis and combining ability on agronomic traits and yields in sweet and waxy corn. *Scientia Horticulturae*, 259: 108817. doi.org/10.1016/j.scienta.2019.108817
- Ding, M., Liu, J., Li, S., Yang, Z., Bai Y., Zhang, Y., et al. (2022). Response of yield and quality traits of new wheat variety 'Yunmai 110' to different planting densities. *Chinese Agricultural Science Bulletin*, 38(24): 1-7(in Chinese).
- El Hanafi, S., Cherkaoui, S., Kehel, Z., Sanchez-Garcia, M., Sarazin, J.B., Baenziger, S., Tadesse, W. (2022). Hybrid seed set in relation with male floral traits, estimation of heterosis and combining abilities for yield and its components in wheat (*Triticum aestivum* L.). *Plants*, 11(4): 19. doi.org/10.3390/plants11040508.
- Elfanah, A.M., Darwish, M.A., Selim, A.I., Elmoselhy, O.M., Ali, A.M., El-Maghraby, et al. (2023a). Hyperspectral reflectance and agro-physiological traits for field identification of salt-tolerant wheat genotypes using the genotype by yield* trait biplot technique. *Frontiers in Plant Science*, 14. doi.org/10.3389/fpls.2023.1165113.
- Elfanah, A.M., Darwish, M.A., Selim, A.I., Shabana, M.M., Elmoselhy, O.M., Khedr, R.A., et al. (2023b). Spectral reflectance indices' performance to identify seawater salinity tolerance in bread wheat genotypes using genotype by yield* trait biplot approach. *Agronomy*, 13(2): 353. doi.org/10.3390/agronomy13020353
- Geng, X., Qu, Y., Jia, Y., He, S., Pan, Z., Wang, L., Du, X. (2021). Assessment of heterosis based on parental genetic distance estimated with SSR and SNP markers in upland cotton (*Gossypium hirsutum* L.). *BMC Genomics*, 22: 1-11.
- Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L.J., Lawrence, D., Muir, J.F., et al. (2010). Food security: The challenge of feeding 9 billion people. *Science*, 327: 812-818.
- Hoagland, A.R., Elliott, F., Rasmussen, L. (1953). Some histological and morphological effects of maleic hydrazide on a spring wheat. *Agronomy Journal*, 45(10): 468-472.
- Huang, X. Q., Börner, A., Röder, M.S., Ganal, M.W. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics*, 105: 699-707.
- Hussain, I., Ali, S., Liu, W., Awais, M., Li, J., Liao, Y., et al. (2022). Identification of heterotic groups and patterns based on genotypic and phenotypic characteristics among rice accessions of diverse origins. *Frontiers in Genetics*, 13: 811124. doi.org/10.3389/fgene.2022.811124
- Jia, Y., Zhang, Y., Sun, Y., Ma, C., Bai, Y., Zhang, H., et al. (2024). QTL mapping of yield-related traits in tetraploid wheat based on wheat 55K SNP array. *Plants*, 13(10): 1285. doi.org/10.3390/plants13101285
- Jiang, F., Yin, X., Li, Z.W., Guo, R., Wang, J., Fan, J., et al. (2023). Predicting heterosis via genetic distance and the number of SNPs in selected segments of chromosomes in maize. *Frontiers in Plant Science*, 14: 1111961. doi.org/10.3389/fpls.2023.1111961
- Joukhadar, R., Daetwyler, H.D., Bansal, U.K., Gendall, A.R., Hayden, M.J. (2017). Genetic diversity, population structure and ancestral origin of Australian wheat. *Frontiers in Plant Science*, 8: 2115. doi.org/10.3389/fpls.2017.02115
- Li, H., Li, S., Abdelkhalik, S., Shahzad, A., Gu, J., Yang, Z., et al. (2020). Development of thermo-photo sensitive genic male sterile lines in wheat using

- doubled haploid breeding. *BMC Plant Biology*, 20: 1-10.
- Li, L., Zhang, Y., Zhang, Y., Li, M., Xu, D., Tian, X., et al. (2021). Genome-wide linkage mapping for preharvest sprouting resistance in wheat using 15K single-nucleotide polymorphism arrays. *Frontiers in Plant Science*, 12: 749206. doi.org/10.3389/fpls.2021.749206
- Li, S.X., Ding, M.L., Li, H.S., Liu, K., Yang, Z.H., Gu, J., Yang, M.J. (2022). Research progress and reflection of two-line hybrid wheat based on thermo-photo sensitive genic male sterility in Yunnan. *Chinese Science Bulletin*, 67: 3197-3206 (in Chinese).
- Ling, C., Wenwen, W., Yue, T., Tong H. (2020). Accuracy analysis of wheat quality determination using near infrared spectroscopy. *Hans Journal of Agricultural Sciences*, 10: 1020-1024 (in Chinese).
- Liu, Z.Q., Pei, Y., Pu, Z.J. (1999). Relationship between hybrid performance and genetic diversity based on RAPD markers in wheat, *Triticum aestivum* L. *Plant Breeding*, 118(2): 119-123.
- Masuka, B., Biljon, A., Cairns, J.E., Das, B., Labuschagne, M.T., Macrobert, J.F., et al. (2017). Genetic diversity among selected elite CIMMYT maize hybrids in East and Southern Africa. *Crop Science*, 57: 2395-2404.
- Melonek, J., Duarte, J., Martin, J.J.J., Beuf, L., Murigneux, A., Varenne, P., et al. (2021). The genetic basis of cytoplasmic male sterility and fertility restoration in wheat. *Nature Communications*, 12: 1036. doi.org/10.1038/s41467-021-21225-0
- Nei, M., Kumar, S. (2000). "Molecular Evolution and Phylogenetics". Oxford University Press.
- Nie, Y., Ji, W., Ma, S. (2019). Assessment of heterosis based on genetic distance estimated using SNP in common wheat. *Agronomy*, 9(2): 66. doi.org/10.3390/agronomy9020066
- Perrier, X., Flori, A., Bonnot, F. (2003). Data analysis methods. In: "Genetic Diversity of Cultivated Tropical Plants", P. Hamon, M. Seguin, X. Perrier, J.C.E. Glaszmann (Eds.) (Montpellier: Enfield Science Publishers), pp. 43-76.
- Rogers, S.O., Bendich, A.J. (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Molecular Biology*, 5: 69-76.
- Sang, Z., Zhang, Z., Yang, Y.-x., Li, Z.-W., Liu, X., Xu, Y., Li, W.-h. (2022). Heterosis and heterotic patterns of maize germplasm revealed by a multiple-hybrid population under well-watered and drought stress conditions. *Frontiers in Plant Science*, 13. doi.org/10.1016/j.fpls.2022.07.006
- Sattler, F.T., Pucher, A., Ango, I.K., Sy, O., Ahmadou, I., Hash, C.T., Hausmann, B.I.G. (2019). Identification of combining ability patterns for pearl millet hybrid breeding in West Africa. *Crop Science*, 59(4): 1590-1603.
- Semagn, K., Iqbal, M., Alachiotis, N.S., N' Diaye, A., Pozniak, C.J., Spaner, D.M. (2021). Genetic diversity and selective sweeps in historical and modern Canadian spring wheat cultivars using the 90K SNP array. *Scientific Reports*, 11. doi.org/10.1038/s41598-021-02666-5
- Sun, C., Dong, Z., Zhao, L., Ren, Y., Zhang, N., Chen, F. (2020). The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnology Journal*, 18(6): 1354-1360.
- Tadesse, W., Sanchez-Garcia, M., Assefa, S.G., Amri, A., Bishaw, Z., Ogbonnaya, F., et al. (2019). Genetic gains in wheat breeding and its role in feeding the world. *Crop Breeding, Genetics and Genomics*, 1: 1-28.
- Tajima, F., Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution*, 1(3): 269-285.
- Tamura, K., Stecher, G., Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38: 3022-3027.
- Wu, X., Liu, Y., Zhang, Y., Gu, R. (2021). Advances in research on the mechanism of heterosis in plants. *Frontiers in Plant Science*, 12. doi.org/10.3389/fpls.2021.745726
- Würschum, T., Zhu, X., Zhao, Y., Jiang, Y., Reif, J.C., Maurer, H.P. (2023). Maximization through optimization? On the relationship between hybrid performance and parental genetic distance. *Theoretical and Applied Genetics*, 136(9): 186. doi.org/10.1007/s00122-023-04436-5
- Yan, W., Frégeau-Reid, J. (2018). Genotype by yield* trait (GYT) biplot: A novel approach for genotype selection based on multiple traits. *Scientific Reports*, 8(1), 8242. doi.org/10.1038/s41598-018-26688-8
- Ye, X., Li, J., Cheng, Y., Yao, F., Long, L., Wang, Y., et al. (2024). Genome-wide association study reveals new loci for yield-related traits in Sichuan wheat germplasm under stripe rust stress. *BMC Genomics*, 20: 1-17.

Zhang, T., Ni, X.L., Jiang, K.F., Deng, H.F., Qing, H.E., Yang, Q.H., et al. (2022). Relationship between heterosis and parental genetic distance based on molecular markers for functional genes related to yield traits in rice. *Rice Science*, 30: 1-12.

Zhao, C.P. (2022). Research progress and prospects of two-line hybrid wheat in China. *Chinese Science Bulletin*, 67: 3119-3128(in Chinese).