

1
2
3
4
5
6
7

8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Minireview Article
**Prediction of Heterosis in Crop Plants – Status
and Prospects**

ABSTRACT

Heterosis has been commercially exploited in many crops; however, its genetic and molecular basis is still not clear. The availability of advanced genomic tools has encouraged several research groups to dissect heterosis into genetic components. Prediction of heterosis using molecular markers is the first step towards this direction, as it may lead to the identification of markers and genomic regions associated with heterosis. Practically, it increases the efficiency of hybrid breeding by marker-assisted selection of superior parental combinations without large-scale field evaluation. Extensive studies in important crops like rice, wheat and maize have identified the genes / genomic regions associated with heterosis and also demonstrated the potential of molecular prediction. Recent trends in whole transcriptome and metabolome analyses will help in the identification of genes or metabolic networks involved in the expression of heterosis. Various genomic tools such as heterotic loci, molecular markers associated with heterotic gene blocks, ESTs and heterosis-related gene databases provide better opportunities for understanding the molecular basis of heterosis and to develop reliable methodologies for a realistic prediction of heterosis. This review provides critical analysis of the prediction of heterosis in crop plants and its implications in hybrid development.

Keywords: Heterosis; molecular markers; genetic distance; prediction methods; gene expression

1. INTRODUCTION

Development of hybrids through the exploitation of heterosis is considered as one of the greatest breakthroughs in plant breeding, as evident from the dramatic increases in the productivity of many crops. In crops like maize, sugarbeet, sunflower, sorghum, pearl millet and many vegetable crops, hybrids are grown predominantly, whereas in many others, like rice, cotton, safflower and rapeseed the area under hybrids is rapidly increasing. Even though the genetic and molecular basis of heterosis is still unclear, plant breeders continue to exploit heterosis extensively to develop superior hybrids. The development of molecular markers and saturated linkage maps has offered useful tools to dissect the complex trait of yield and other agriculturally important traits into genetic components through quantitative trait loci (QTL) analysis, which may lead to a better understanding of the genetic basis of heterosis. Some of the recent research is also focused on the identification of heterotic loci [1, 2, 3] and the molecular markers associated with them thereby helping in understanding the phenomenon of heterosis at the molecular level. In addition, complementary approaches involving genomic, transcriptomic, proteomic and metabolomic data coupled with mathematical modeling [4, 5, 6, 7] have resulted in a better prediction of heterosis in comparison to a single approach. The growing interest for the prediction of heterosis in different crops necessitates the prediction of heterosis in crop plants using phenotypic as well as molecular data including DNA markers (SSR and SNPs), gene expression, profiling of proteins and metabolites. The progress made during the last two decades in improving heterosis prediction in plants using various molecular tools has been discussed critically in this review.

2. PREDICTION OF HETEROSIS

Heterosis refers to the superiority of F_1 hybrids with respect to biomass, size, yield, rate of development, fertility, biotic and abiotic stress tolerance compared to the average of their parental inbreds [8, 9]. Breeders have to evaluate a large number of parental lines for their combining ability, make thousands of test crosses and evaluate the F_1 hybrids for yield heterosis to identify the best hybrid. This consumes a lot of resources and time. Moreover, the field trials can accommodate only a limited number of crosses in all possible combinations between the parental lines, and the results are subjective. Therefore, hybrid breeders have always been interested in the selection of potential lines among the available parental lines which are expected to give heterotic hybrids by-passing the field evaluation of a large number of crosses involving these parental lines. Over the years, several methods were employed to predict heterosis, such as, *per se* performance of parental lines, mitochondrial complementation, combining ability and genetic diversity estimated from geographical origin, coefficient of parentage, multivariate analysis of morphological traits and isozyme and molecular marker analysis. Among these methods, mitochondrial complementation based heterosis prediction is unpopular since the results were not reproducible. Hence, this method will not be discussed further in this review. Apart from these methods, gene expression is being used in recent studies to predict heterosis.

2.1 Phenotypic Data-based Prediction of Heterosis

Heterosis prediction can be done through *per se* performance of parental lines, combining ability and genetic diversity studies using the phenotypic data collected from field evaluation of genotypes. Several studies in different crops reported contrasting conclusions regarding the effectiveness of *per se* performance in the prediction of heterosis. Studies in maize [10] and sugarcane [11] concluded that there was no association between *per se* performance of parental lines and heterosis in F_1 hybrids. Selection of potential parents based on *per se* performance was found to be effective in Triticale for traits other than grain yield [12], while in tropical maize the performance of hybrid progenies under excessive moisture stress can be predicted, to some extent, based on *per se* performance of their inbred parents that were selected and improved for excessive moisture stress [13]. Therefore, it can be concluded that heterosis prediction based on *per se* performance of parents may not be a reliable indicator of heterosis.

Identification of superior parental lines for developing heterotic hybrids was generally done by employing combining ability tests such as top-cross test, poly-cross test, single cross test, diallel mating, and line $\times$ tester analysis, though with variable levels of success [14]. In general, selection of parental lines with high general combining ability (GCA) effects resulted in the development of heterotic hybrids in any crop. However, heterotic combinations could also be derived from parents exhibiting low GCA effects as noticed in rice [15] and such combinations could not be derived from parents with high GCA [16]. Studies in crops like maize [17], wheat [18], Ethiopian mustard [19] and rapeseed [20] have shown moderate to strong correlation between combining ability and *per se* performance as well as mid-parent heterosis, but in varying degrees. It is important to note that the strong relationship between the mean performance and GCA of inbred lines may be due to the presence of additive genetic variance. SCA is related to heterosis but the association is more tricky than assumed. However, the non-additive genetic variance or specific combining ability (SCA) is ignored here, which will drastically reduce this correlation. Therefore, SCA has to be given due consideration since it has a direct impact on heterosis.

The extent of genetic diversity between the two parents has been proposed as a possible indicator for the prediction of heterosis [9, 21]. However, heterosis and genetic diversity

between parents rarely exhibit a strong correlation [22, 23, 24]. More reports have been published in maize [25] and rice [21, 26, 27], but the extent of correlation varied widely from one trait to another and from one data set to another. In Ethiopian mustard, Teklewold and Becker [28] reported that it was possible to predict the heterosis and parental combining ability to some extent based on genetic distance calculated over phenotypic traits. However, in sorghum, a perfect correspondence between heterosis and genetic diversity between parental lines was lacking [29]. In a study on pigeonpea, hybrids derived by crossing genotypes from high diversity groups exhibited significant and positive heterosis for seed yield, with some exceptions. Moreover, higher yield, heterosis and specific combining ability were observed in inter-group crosses as compared to intra-group crosses, suggesting that genetic diversity could be used as a reliable parameter for the prediction of heterosis in pigeon pea [30]. Due to lack of consistency in the prediction of heterosis based on genetic diversity and combining ability among parental lines through field evaluation, there is an immense need for the prediction of heterosis based on the molecular marker polymorphism from a large pool of parental lines without field evaluation.

Heterotic groups refers to genetically diverse groups of genotypes / parental lines, crossing of which, may result in heterotic hybrids. Involving the genotypes from different heterotic groups in crossing program will lead to heterosis and yield stability of the new cultivars [31, 32]. Utilization of the estimates of genetic distances and specific combining ability will help in the identification of prospective parental lines from a larger gene pool for crossing that would lead to the development of superior hybrids [28, 31, 33, 34]. Heterotic groups among inbred parents in maize for the identification of superior hybrid combinations was performed using morphological / field evaluation data [35, 36, 37] and DNA markers [38]. Recently, four heterotic patterns within two heterotic groups were identified in tropical *indica* rice based on SSR marker data and field trials using IRRI B- and R-line germplasm resulting in the enhancement of hybrid breeding efficiency [39]. The present review discusses mainly on the prediction of heterosis based on molecular marker data and gene expression data, such as transcript abundance and expression levels/patterns that are becoming popular in the prediction of heterosis.

2.2 Molecular Marker-based Prediction of Heterosis

Genetic divergence of parental lines is thought to be related to heterosis or hybrid vigor of the hybrid as discussed earlier. Thus, biochemical and molecular marker-assayed genetic variation of the parental lines may potentially be useful for predicting heterosis. Prior to molecular markers, isozymes were the commonly used biochemical markers for the prediction of heterosis. Many of the studies in maize used only fewer isozyme loci as well as limited parental lines. As a result, low and non-significant correlations were reported in these studies between isozymes diversity and specific combining ability. Even in a study involving 100 maize hybrids derived from 37 parental lines and 31 isozyme loci, an R^2 value of only 0.36 was reported [40]. Furthermore, in a study on rice, no association was observed between isozyme diversity and grain yield of hybrid using six isozyme loci in 75 F1 hybrids [41]. Distances computed from isozyme data were in some cases significantly correlated to heterosis but their utility in the prediction of heterosis was limited primarily to lines with similar pedigrees [42]. Isozymes were considered to be unpopular for heterosis predictions since they could sample only a limited number of loci contributing to the hybrid response and it is unlikely that these loci will have a direct effect on the phenotypic expression of the targeted trait [43].

One of the important challenges that have encouraged research in employing molecular markers and genomics in heterosis breeding relates to the understanding of genetic and molecular basis of heterosis and prediction of heterosis. The use of molecular markers led

heterosis prediction into a new phase since it appeared in 1980s, but the conclusions were not consistent with each other. High positive correlation of yield heterosis with genetic distance based on RAPD and SSR markers was reported for *indica* x *indica* and *japonica* x *japonica* crosses but not for *indica* x *japonica* crosses [44]. Significant and effective correlations between genetic distance and heterosis were reported in rice [26, 45], maize [25], wheat [46], sunflower [47] and rapeseed [48]. On the other hand, weak or non-significant correlations have also been reported between genetic distance and heterosis in some studies. Table 1 summarizes the studies conducted on the prediction of heterosis using molecular markers in different crops and their conclusions.

With the prominence and popularity of single locus markers, especially microsatellite markers, several efforts were made in rice in assessing the utility of SSR markers for heterosis prediction based on the relationship between molecular diversity of parental lines and heterosis of the resulting hybrid [53, 54, 55, 83]. Consistency in the prediction was lacking in these studies, and this may be because prediction of heterosis was based on genetic diversity of anonymous molecular markers, which may not be associated with functional regions of the genome. Therefore, prediction based on functional markers, especially the markers associated with genes controlling heterosis for yield traits, might be more powerful than that based on anonymous markers. Markers used for prediction of heterosis should be differentiated into positive markers, since every marker is not linked to a quantitative trait locus (QTL) and therefore, with heterosis [27]. It was also suggested to use QTL-linked markers, thus distinguishing between general and specific heterozygosity. An attempt was made in rice to associate morphological traits, pedigree records and SSR polymorphism with heterosis [53]. While no significant correlation was observed for morphological traits and pedigree records, a low correlation value was noticed between SSR polymorphism and heterosis. They also reported that molecular markers would be useful for predicting hybrid performance only when a significant portion (>50%) of the selected markers is linked to QTL. In a study on rice, genomic and EST-SSR marker polymorphism was compared for its correlation with heterosis and concluded that the latter exhibited a significant positive correlation with heterosis resulting in the identification of 10 informative EST-SSR markers having the potential for the prediction of heterosis [56]. Once such informative markers are identified, they should be tested in different populations of parental lines varying in their genetic background to ascertain their consistency in the prediction of heterosis. Moreover, it is also important to analyze the expression of these genes in hybrids and their parents to understand their role in heterotic expression.

Molecular markers are widely used in QTL mapping of quantitative traits [84, 85, 86] and exploration of genetic diversity, with the latter often employed in the prediction of heterosis [27, 52, 87]. Methods using the genetic distance between the parental lines estimated using random / anonymous molecular markers failed to predict reliably the hybrid performance and hence prediction methods using markers linked to QTL were successfully developed. The advantage of using such markers is that it can reveal the locus and/or locus pair that could be targeted for improvement in further breeding programs by manipulating the related QTL [88]. Unfortunately, studies on the prediction of heterosis based on analysis of polymorphism of yield related QTL are very limited. Neither heterozygosities nor QTL for yield traits were found to be suitable to predict hybrid performance and heterosis in *Brassica napus* [78]. On the contrary, a study on rice indicated that the QTL identified for 1000-grain weight showed a high contribution to heterosis. The heterosis over mid-parents of QTL could be predicted based on QTL effects [89]. A total of 13 heterotic loci in maize [1], genome regions harboring numerous heterosis-related QTL in oilseed rape [3] and ten genomic positions involved in biomass heterosis during early development in *Arabidopsis* were identified [2]. Very recently, Ben-Israel et al. [90] identified three heterotic trait loci with synergistic intra-locus effects on heterosis for grain yield in sorghum, which accounts for 19.0% of the heterotic variation.

Table 1 Heterosis prediction studies in different crops using molecular markers

Crop Species	Marker type	Plant material	Conclusions
Rice	RFLP	11 elite lines in a diallel set involving <i>indica</i> and <i>japonica</i>	Complex relationship was observed between marker heterozygosity and hybrid performance [49].
	RAPD, ISSR, RFLP and STMS	Two CMS lines, four restorers and 28 hybrids	Molecular markers may not always be useful for selection of parents [50].
	SSR and RAPD	13 'Tongil'-type rice cultivars and 78 hybrids	DNA marker-based GD may not be useful for heterosis prediction [51].
	AFLP	Eight CMS lines, 10 restorers and 80 hybrids	Heterosis was related to GD between the parents, with a few loci contributing to it [52].
	Pedigree record, quantitative traits and SSR	37 maintainers, 43 restorers and 34 hybrids	Prediction is difficult through SSR and pedigree-based diversity of complex traits [53].
	RAPD and SSR	41 hybrids from a half diallel with 10 <i>japonica</i> cultivars	Proposed the role of 'key' DNA markers in the prediction of heterosis [54].
	SSR	13 CMS lines, 19 restorers and 151 hybrids	Prediction of heterosis based on effect-increasing loci was more effective [55].
	SSR and EST-SSR	Nine CMS lines, 32 restorers and 20 hybrids	Prediction of heterosis is better using EST-SSRs [56].
Maize	RAPD	18 inbred lines	DNA marker is useful in aiding the choice of superior crosses to be made [57].
	RFLP	Eight and 10 S ₃ selected inbred lines in partial diallel	RFLP-based GDs are not useful for the prediction of heterosis [58].
	AFLP	53 inter-heterotic hybrids of a 13 × 13 half-diallel	Some QTL of grain yield were located in the vicinity of loci having quantitative effects on grain yield [59].
	Morphological data and RAPD	28 open pollinated varieties in a diallel scheme and 378 hybrids	Low and positive correlation was observed between RAPD-based GD and SCA for yield [60].
	RAPD	13 inbred lines and 78 hybrids	RAPDs are not suitable for the prediction of yield performance of hybrids [61].
	AFLP and SSR	18 S ₃ inbred lines in a partial diallel mating design	Single cross performance can be predicted through AFLP-based GD [62].
	SSR	15 elite inbred lines and 105 hybrids	Prediction of yield heterosis using SSR markers is difficult [63].

	AFLP	52 maize elite inbred lines and 270 hybrids	Prediction efficiency ranged from 0.46 to 0.86 for grain yield and from 0.59 to 0.96 for grain dry matter content [64].
	AFLP	52 elite inbred lines and 270 hybrids	Prediction of heterosis improved by variable haplo-type block length than with AFLP markers [65].
	AFLP	21 inbreds and 225 hybrids	Integration of molecular and phenotypic data will lead to better prediction of heterosis [66].
	AFLP	26 inbred lines were crossed with six testers	Significant positive correlations with low magnitude between GDs and mid-parent heterosis can be of predictive value [67].
	AFLP	400 hybrids from nine factorial crosses and 79 inbred parents	Even without testcross data of parents, the hybrid performance was predicted efficiently by molecular markers [68].
Wheat	RFLP and RAPD	Eight-parent diallel cross and top cross (4 males × 25 females)	A weak correlation was observed between parental diversity and hybrid performance [69].
	RAPD	10 CMS lines, 10 restorers and 41 hybrids	No significant correlation was observed between RAPD marker-based GD with heterosis [70].
	RAPD	18 parental lines and 76 F ₂ hybrids	GDs between parents can be a potential predictor of hybrid performance for selected traits [71].
Sorghum	RFLP	162 hybrids derived from 70 inbred lines	A very weak association of GD with yield is of no value for identifying superior hybrids [72].
	SSR	210 hybrids from 37 parental lines	A significant, positive and moderate correlation between coefficient of marker polymorphism of parental lines with grain yield heterosis of their respective hybrids indicated the potential of molecular marker-assisted prediction of heterosis [73].
	AFLP, SSR and SRAP	90 cross combinations five sterile lines and 18 Sudan grass lines.	An average correlation was observed between molecular marker value and grain yield heterosis (r = 0.65), which will help breeders to exploit the markers associated with yield QTL for a reliable prediction of heterosis [74]
Triticale	SSR and AFLP	206 F ₁ crosses and 61 parental lines	Low correlations between GD and SCA hampered the prediction of SCA from molecular data [75].

<i>Brassica napus</i>	SRAP	10 maintainer lines and 12 restorers	Correlation coefficient between GD and yield (0.64) indicated a moderately strong relationship [76].
Oilseed rape	Morphology, isozymes, proteins and RAPD	Six male-sterile lines, five restorers and 30 hybrids	GD based on significant and favorable markers significantly correlated with plant height, seeds per silique and seed yield [77].
Rapeseed	AFLP and SSR	Three double low (erucic acid and glucosinolates) self-incompatible lines, 22 varieties and 66 hybrids	Neither heterozygosities nor QTLs for yield traits were suitable for predicting heterosis [78].
Faba Bean	AFLP	18 inbreds, 62 hybrids and their F ₂ progenies	AFLP-based genetic similarities are not useful to predict heterosis [79].
Pepper	Morphology and AFLP	Seven parents and 21 hybrids	Correlations of AFLP-based GDs with mid- and high-parent heterosis were non-significant and proved to be of no predictive value [80].
Cotton	RAPD and SSR	Three CMS lines, 10 restorers and 22 hybrids	The relationship between SSR marker heterozygosity and hybrid performance can be used to predict the fiber length during inter-specific hybrid cotton breeding [81].
Chickpea	RAPD and SSR	Eight genotypes in diallel set	Two heterotic groups could be identified on the basis of the general marker heterozygosity [82].

Abbreviations: RFLP, restriction fragment length polymorphism; RAPD, randomly amplified polymorphic DNA; ISSR, inter-simple sequence repeats; AFLP, amplified fragment length polymorphism; SSR, simple sequence repeats; STMS, sequence tagged microsatellite sites; SRAP, sequence related amplified polymorphism; EST, expressed sequence tag; GD, genetic distance; QTL, quantitative trait loci; SCA, specific combining ability

194 Also, a significant 21% increase in grain yield was observed for one of these loci (hDPW4.1)
195 in a consecutive F_2 population. Such studies could be helpful in the identification of
196 QTL/genomic regions associated with heterosis and markers associated with them may help
197 in a better prediction of heterosis. If successful, such prediction methods may lead to the
198 selection of the limited number of parental combinations for synthesizing experimental
199 hybrids for field evaluation for the identification of highly heterotic hybrids, thus increasing
200 the efficiency of hybrid development.

201

202 3. METHODS FOR IMPROVING THE PREDICTION OF HETEROSIS

203

204 Even though molecular marker systems such as restriction fragment length polymorphism
205 (RFLP), randomly amplified polymorphic DNA (RAPD), inter-simple sequence repeats
206 (ISSR) and amplified fragment length polymorphism (AFLP) were used earlier for the
207 prediction of heterosis, the studies ended up with limited success. This may be because
208 each marker system differs in their ability in predicting heterosis. The RFLP data was found
209 promising for the prediction of hybrid performance for crosses between lines from the same
210 heterotic group, practically zero for crosses between lines from divergent heterotic groups,
211 and medium for mixtures of both [91, 92]. The correlation between RAPD-based genetic
212 distances and hybrid performance were very low to predict heterosis [50]. A significant
213 improvement in the correlations between genetic distance and hybrid performance was
214 observed in maize by using AFLP markers as compared to RFLP markers due to the
215 analysis of a large number of pre-selected markers linked to quantitative trait loci [93].
216 Significantly higher correlations between the coefficient of parental polymorphism and
217 heterosis were obtained using EST-SSR markers in rice [56] in comparison to genomic SSR
218 markers.

219 Heterosis is presumed to be due to heterozygosity and hence the selection of the
220 appropriate type of marker is very important for assessing the genetic diversity among the
221 parents. Codominant and single locus markers such as SSRs and SNPs are advantageous,
222 since they could estimate the heterozygosity based on the parental alleles. This will also
223 help in the identification of specific alleles associated with heterosis of various traits that can
224 be employed in the improvement of inbred parents. Moreover, with the sequencing of crop
225 genomes, these markers can be easily located in the physical map. In contrast, the dominant
226 markers are not informative, since they possess information on the allele of only one parent
227 at a particular locus. Such lack of complete allelic information will be a great setback for the
228 parental line improvement aimed at enriching of superior alleles.

229 Several efforts were made to improve the power of heterosis prediction. A study based on
230 computer simulations concluded that at least 30-50% of the molecular markers should be
231 linked to QTL while not more than 20-30% of the molecular markers should be randomly
232 dispersed [94]. Later, TM-BLUP (trait marker-best linear unbiased prediction) was
233 developed, which takes into account the molecular marker data together with the trait data
234 for the prediction of hybrid performance. This approach slightly improved the quality of the
235 prediction [95, 96]. A linear regression model, namely TCSM (total sum of selected markers)
236 with a promising predictive power was developed for the prediction of hybrid performance,
237 which calculates genotypic value of hybrids by considering joint effect of genetic factors [59].
238 This approach was later improved by involving multiple regression as well as haplo-block
239 organization of markers [64, 65, 68]. With the availability of the heterosis related genes of
240 rice, maize and wheat in the public database (Heterosis related genes database,
241 <http://hrgd.big.ac.cn/index.html>) [97], PCR-based molecular markers can be developed and
242 used for the prediction of heterosis in other closely related crop species. Development of
243 SSR or INDEL markers from heterosis related genes is not explored until now and the use of

such molecular markers associated with heterosis will help in the reliable prediction of heterosis in comparison to prediction using random/anonymous markers.

4. EMERGING CONCEPTS IN MARKER-BASED PREDICTION OF HETEROSIS

The concept of heterosis prediction using molecular markers has been refined for more than two decades since it appeared in 1980s. Initial predictions with isozymes did not gain importance due to the availability of a limited number of loci and inconsistent results. Later, DNA markers such as RFLPs, RAPDs and ISSRs were predominantly used during the 1990s with very limited success since the correlation between genetic diversity based on these markers with heterosis was very low to predict heterosis. The popularity of SSR markers during the late 1990s resulted in several studies in different crop species towards the prediction of heterosis with the help of SSR-based genetic distance. To start with, anonymous markers (genomic SSR markers) were used, which led to inconsistent results. Then, the focus was shifted to the selection of specific markers based on their effects with heterosis or QTL information instead of using random markers.

A series of studies in rice for the past two decades resulted in the development of the concept of heterosis prediction. The concept has evolved from anonymous or random markers to specific or informative markers. Specific heterozygosity of positive markers (markers with significant effects on the target traits) showed a high correlation with heterosis for grain yield in rice [21]. The importance of significant correlations of specific heterozygosity based on the small number of positive markers (markers associated with heterotic yield response) with hybrid performance than those with general heterozygosity in the prediction of heterosis was also reported [27]. Favorable and unfavorable SSR alleles significantly affecting yield heterosis were identified with a possible use in hybrid rice breeding [26]. Another study revealed that the presence of loci with positive or negative effects, and their counteraction **during the prediction of heterosis** [87]. Markers associated with hybrid vigor and hybrid weakness were identified in rice [54], and they proposed a new concept of 'key DNA markers' for the prediction of heterosis. In line with this, the concept of 'effect-increasing' loci was proposed recently for the prediction of hybrid performances in *indica* rice [55]. Comparative analysis of genomic and EST-SSR markers in rice revealed that the latter exhibited a significant positive correlation with heterosis. A set of informative markers was identified for the prediction of heterosis by a careful and critical analysis of EST-SSR markers [56]. The concept of heterosis is summarized in Fig. 1.

In sorghum, recent studies have highlighted the potential of molecular marker-assisted prediction of heterosis. Correlation of molecular marker polymorphism between parental lines with the grain yield heterosis of their respective hybrids resulted in the identification of a set of 30 SSR markers exhibiting the potential for the prediction of heterosis [73]. Using yield-related QTL, a moderate correlation ($r=0.65$) was observed between the marker value of these loci and yield heterosis, which will help the molecular breeders to exploit such loci for a reliable and accurate prediction of heterosis [74]. In all the above discussed reports, it is very clear that molecular marker-based prediction of heterosis should not involve the total markers, but, selective or informative markers, which will increase the reliability and accuracy of prediction.

5. PREDICTION OF HETEROSIS USING TRANSCRIPTOME AND METABOLOME DATA

Besides molecular marker data, the transcriptomic and metabolomic data also have the potential for the prediction of heterosis. This can be attributed to the fact that heterosis is a genome-wide phenomenon involving a network of genes and proteins leading to changes in

the plant metabolism. Rapid advances in this area provide new avenues for the expression-based prediction of heterosis. Several projects investigating the phenomenon of heterosis in different crop species have contributed to the identification and characterization of differentially expressed genes (DEGs) related to heterosis. Among the various techniques applied to identify genes possibly associated with heterosis, differential display analysis, cDNA-AFLP and microarrays have been more popularly used. Analysis of the differential expression of genes at different developmental stages of hybrids and their parental lines have proven to be a useful methodology to identify the genes associated with heterosis.

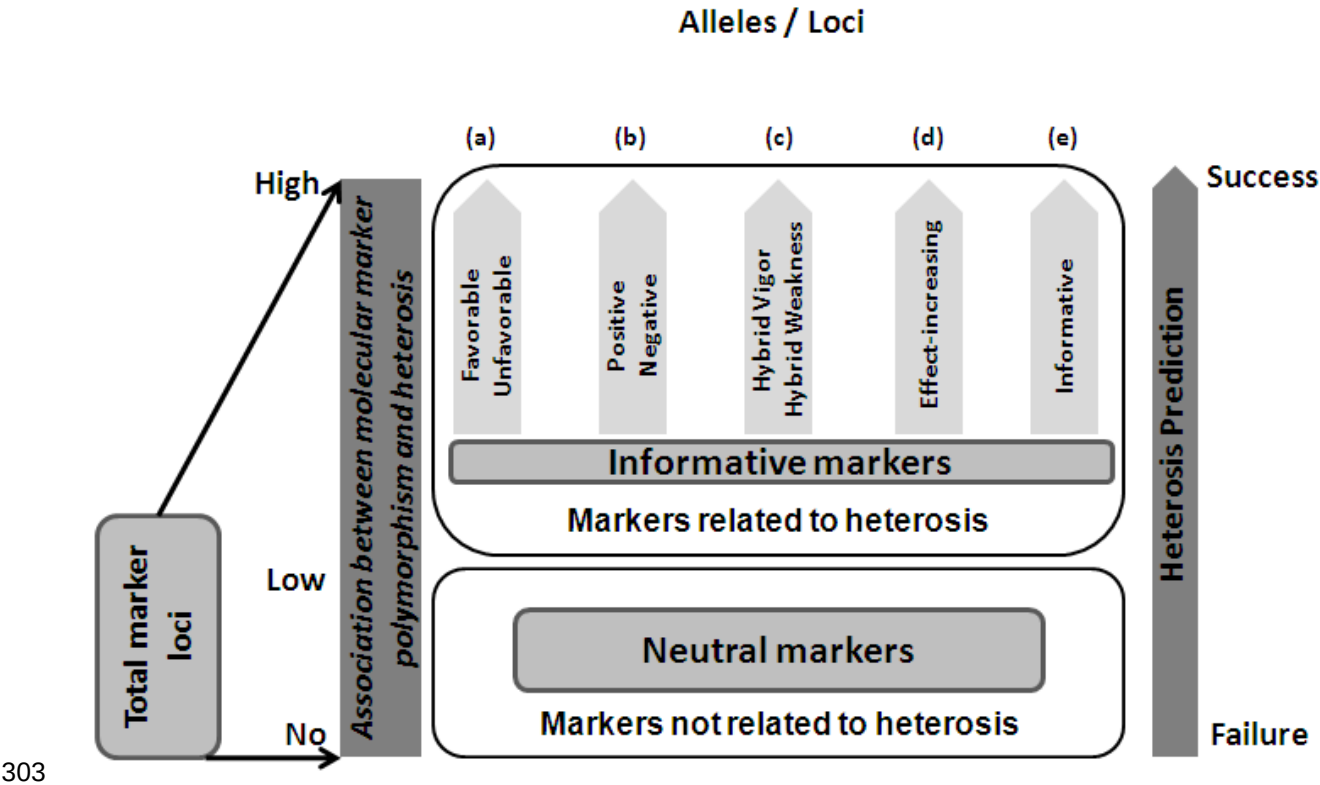


Fig. 1 Concept of heterosis prediction in crop plants. The concept can be summarized based on few important studies that led to the identification of specific marker loci / alleles for the prediction of heterosis viz., (a) 'favorable and unfavorable' SSR alleles [26], (b) loci with 'positive or negative' effects [87], (c) 'hybrid vigor and hybrid weakness' alleles [54], (d) 'effect-increasing' loci [55], (e) 'informative' markers [56].

Earlier studies have indeed indicated that the differential gene expression patterns were correlated with heterosis [98, 99]. The levels of expression of some differentially expressed genes in the parents exhibited highly significant correlations with heterosis and hybrid performance for grain yield and also with grain dry matter content [100]. Such a strong correlation provided the evidence on the prediction potential of genes and their expression levels. Recent studies on the transcriptome-based prediction of heterosis showed promising results. Based on a study in maize, it was concluded that the prediction of hybrid performance was more precise with transcriptome-based distances using selected markers than earlier prediction models involving DNA markers or the estimates of general combining ability [101]. A novel method of predicting heterosis involving a combination of transcriptomic technology and mathematical modeling was developed [102] that could identify a sub-group

of genes based on the correlation of their transcript abundance with the magnitude of heterosis in hybrids using linear regression algorithms. Moreover, the allelic patterns of inbreds can be used as potential markers for the prediction of heterosis in hybrids developed from these lines. This approach was successfully demonstrated in *Arabidopsis thaliana* and maize, which can also be extended to other crops for a reliable prediction of heterosis for important traits, including yield. In addition, development of molecular markers for heterosis using transcriptome data was found to be promising even with a limited number of hybrids as well as field performance data [6]. However, the use of such approach needs further development and evaluation before being applied in hybrid breeding programs.

The applicability of morphological characters, isozymes, proteins, expression profiles, SNP markers, and metabolites in different combinations for prediction of heterosis was investigated by various research groups. Genetic markers were employed in combination with morphological characters, isozymes, and proteins for the same purpose in oilseed rape [77]. In *Arabidopsis thaliana*, it was demonstrated that parental genetic markers in combination with metabolic markers, identified via feature selection and regression methods, significantly improves the prediction of biomass heterosis in the hybrid [4]. Though genetic markers provide very little information, especially in the case of genes with minor effect, more integrated markers such as the levels of plant metabolites, will contribute a decisive part of information, since they represent the combined effect of many genes. A proof of concept of a new complementary approach for the prediction of hybrid biomass in *Arabidopsis thaliana* was presented that involves metabolite profiles in addition to SNP markers and feature selection procedure instead of the whole set of SNPs and metabolites [5]. Recently, whole genome prediction (WGP) was suggested to be a powerful complementary approach in hybrid breeding for highly polygenic traits as revealed prediction accuracies in the range of 0.72 - 0.81 for SNPs and 0.60 - 0.80 for metabolites thereby a reliable screening of a large pool of diverse inbred lines is possible to identify potential parental lines for the development of superior hybrids [7].

Gene expression based approaches are expensive and demands sophisticated infrastructure and expertise. Hence, they may not be suitable for the routine screening of the large number of parental lines for the identification of superior parental combinations. So, there is an immense need for the development of easy, cheap, rapid and routinely usable assays that will help those involved in hybrid development to predict heterosis in different crops. In this direction, a better and pragmatic approach would be the use of PCR-based markers targeting the sequence polymorphism responsible for the differential gene expression in the hybrid as compared to the parents for the prediction of heterosis.

6. CONCLUSION

An accurate and reliable prediction of heterosis will depend on the exploitation of metabolic networks involving several genes or genomic regions since heterosis is a complex phenomenon. Recent studies discussed in this review highlighted a careful selection of molecular markers is important for a better prediction of heterosis and no single category such as anonymous or QTL linked markers could serve the purpose, however, a major proportion of the markers should be linked to QTLs for the trait of interest. Therefore, it would be better to select a judicious mix of anonymous markers, EST/gene-based markers and QTL-associated markers for heterosis prediction. Advances in transcriptomics, proteomics and metabolomics have helped to focus research towards the identification of the complete network of genes responsible for different pathways, including those of transcription factors. The popular approach involving whole transcriptome analysis in a hybrid in relation to its parents helps in the identification of genes associated with heterosis. Hence, the future prediction strategies should integrate the valuable information from these

data along with molecular marker information to develop a rapid, cheap, efficient and easy-to-use assay to achieve a realistic prediction of heterosis. The current focus in many crops should be on the identification of sequence polymorphisms in the regulatory regions of genes involved in heterotic expression will help to develop a practically viable molecular marker-based approach that targets such a polymorphism responsible for the heterotic expression of different traits. Once genes/genomic regions/QTL associated with heterosis are identified as a consequence of current heterosis projects in important crops, genetic transformation and marker-assisted introgression can be employed for the enrichment of heterotic genes/alleles in the parental lines so that the resultant hybrid will exhibit high heterosis.

REFERENCES

1. Tang J, Yan J, Ma X, Teng W, Wu W, Dai J, et al. Dissection of the genetic basis of heterosis in an elite maize hybrid by QTL mapping in an immortalized F₂ population. *Theor. Appl. Genet.* 2010;120:333-340.
2. Meyer RC, Kusterer B, Lisec J, Steinfath M, Becher M, Scharr H, et al. QTL analysis of early stage heterosis for biomass in *Arabidopsis*. *Theor. Appl. Genet.* 2010;120:227-237.
3. Basunanda P, Radoev M, Ecke W, Friedt W, Becker HC, Snowdon RJ. Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 2010;120:271-281.
4. Gartner T, Steinfath M, Andorf S, Lisec J, Meyer RC, Altmann T, et al. Improved heterosis prediction by combining information on DNA and metabolic markers. *PLoS ONE* 2009;4:e5220.
5. Steinfath M, Gartner T, Lisec J, Meyer RC, Altmann T, Willmitzer L, et al. Prediction of hybrid biomass in *Arabidopsis thaliana* by selected parental SNP and metabolite markers. *Theor. Appl. Genet.* 2010;120:239-247.
6. Stokes D, Fraser F, Morgan C, O'Neill CM, Dreos R, Magusin A, et al. An association transcriptomics approach to the prediction of hybrid performance. *Mol. Breed.* 2010;26:91-106.
7. Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisec J, Technow F, Sulpice R, et al., Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics* 2012;44:217-221.
8. Shull G. Beginnings of the heterosis concept. In: Gowen JW, editor. *Heterosis: A record of researches directed toward explaining and utilizing the vigor of hybrids*. Iowa State College Press, Ames; 1952.
9. Falconer DS, Mackay TFC. *Introduction to quantitative genetics*, 4th edn. Longman, England; 1996.
10. Lee EA, Ash MJ, Good B. Re-examining the relationship between degree of relatedness, genetic effects, and heterosis in maize. *Crop Sci.* 2007;47:629-635.
11. Verma PS, Singh SB. Heterosis in relation to *per se* performance and effects of general combining ability in sugarcane. *Sugar Tech.* 2004;6:181-185.

- 414 12. Oettler G, Tams SH, Utz HF, Bauer E, Melchinger AE. Prospects for hybrid breeding
415 in winter triticales. I. Heterosis and combining ability for agronomic traits in European
416 elite germplasm. *Crop Sci.* 2005;45:1476-1482.
- 417 13. Zaidi PH, Mani Selvan P, Sultana R, Srivastava A, Singh AK, Srinivasan G, et al.
418 Association between line *per se* and hybrid performance under excessive soil
419 moisture stress in tropical maize (*Zea mays* L.). *Field Crops Res.* 2007;101:117-126.
- 420 14. Virmani SS. Heterosis and hybrid rice breeding. Monographs on Theor Appl. Genet.
421 22, Springer-Verlag, Berlin; 1994.
- 422 15. Kumar I, Saini SS. Diallel analysis in rice. Combining ability for various quantitative
423 characters. *Genet. Agric.* 1981;35:243-252.
- 424 16. Shrivastava MN, Seshu DV. Combining ability for yield and associated characters in
425 rice. *Crop Sci.* 1983;23:741-744.
- 426 17. Betran FJ, Ribaut JM, Beck D, Gonzalez de Leon D. Diversity, specific combining
427 ability, and heterosis in tropical maize under stress and non-stress environments.
428 *Crop Sci.* 2003;43:797-806.
- 429 18. Dreisigacker S, Melchinger AE, Zhang P, Ammar K, Flachenecker C, Hoisington D, et
430 al. Hybrid performance and heterosis in spring bread wheat, and their relations to
431 SSR-based genetic distances and coefficients of parentage. *Euphytica* 2005;144:51-
432 59.
- 433 19. Teklewold A, Becker HC. Heterosis and combining ability in a diallel cross of Ethiopian
434 mustard inbred lines. *Crop Sci.* 2005;45:2629-2635.
- 435 20. Qian W, Sas, O, Meng J, Li M, Frauen M, Jung C. Heterotic patterns in rapeseed
436 (*Brassica napus* L.): I. Crosses between spring and Chinese semi-winter lines. *Theor.*
437 *Appl. Genet.* 2007;115:27-34.
- 438 21. Zhang QF, Gao YJ, Yang SH, Ragab RA, Saghai Maroof MA, Li ZB. A half-diallel
439 analysis of heterosis in elite hybrid rice based on RFLP and microsatellites. *Theor.*
440 *Appl. Genet.* 1994;89:185-192.
- 441 22. Melchinger AE. Genetic diversity and heterosis. In: Coors JG, Pandey S, editors. The
442 genetics and exploitation of heterosis in crops. CSSA, Madison, WI; 1999.
- 443 23. Dixit UN, Swain D. Genetic divergence and heterosis in sesame. *Indian J. Genet.*
444 2000;60:213-219.
- 445 24. Singh SP, Singh M. Multivariate analysis in relation to genetic improvement in *Cuphea*
446 *procumbens*. *J. Genet. Breed.* 2004;58:105-112.
- 447 25. Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ. Similarities among a group of
448 elite maize inbreds as measured by pedigree, F_1 grain yield, grain yield heterosis and
449 RFLPs. *Theor. Appl. Genet.* 1990;80:833-840.
- 450 26. Liu XC, Wu JL. SSR heterogenic patterns of parents for marking and predicting
451 heterosis in rice breeding. *Mol. Breed.* 1998;3:263-268.

- 452 27. Zhang QF, Zhou ZQ, Yang GP, Xu CG, Liu KD, Saghai Maroof MA. Molecular marker
453 heterozygosity and hybrid performance in *indica* and *japonica* rice. Theor. Appl.
454 Genet. 1996;93:1218-1224.
- 455 28. Teklewold A, Becker HC. Comparison of phenotypic and molecular distances to
456 predict heterosis and F₁ performance in Ethiopian mustard (*Brassica carinata* A.
457 Braun). Theor. Appl. Genet. 2006;112:752-759.
- 458 29. Rani KJ, Rao SS. Relationship between heterosis and genetic divergence in rabi
459 sorghum [*Sorghum bicolor* (L.) Moench]. Res. Crops 2009;10:319-322.
- 460 30. Pandey P, Pandey VR, Kumar A, Yadav S, Tiwari D, Kumar R. Relationship between
461 heterosis and genetic diversity in Indian pigeonpea [*Cajanus cajan* (L.) Millspaugh]
462 accessions using multivariate cluster analysis and heterotic grouping. Aust. J. Crop
463 Sci. 2015;9(6):494-503.
- 464 31. Fan XM, Zhang YM, Yao WH, Chen HM, Tan J, Xu CX, Han XL, Luo LM, Kang MS.
465 Classifying maize inbred lines into heterotic groups using a factorial mating design.
466 Agron J. 2009;101:106-112.
- 467 32. Esposito MA, Gatti I, Cravero VP, Anido FSL, Cointy EL. Combining abilities and
468 heterotic groups in *Pisum sativum* L. Aust J. Crop Sci. 2013;7(11):1634-1641.
- 469 33. Riday H, Brummer EC, Campbell TA, Luth D, Cazcarro PM. Comparisons of genetic
470 and orphological distance with heterosis between *Medicago sativa* subsp. *sativa* and
471 subsp. *falcata*. Theor. Appl. Genet. 2003;131:37-45.
- 472 34. Bhatnagar S, Betran FJ, Rooney LW. Combining abilities of quality protein maize
473 inbreds. Crop Sci. 2004;44:1997-2005.
- 474 35. Gonzalez S, Cordova H, Rodrigue S, de Leon H, Serrato VM. Determinacion de un
475 patron heterotico a partir de la evaluacion de un dialelo de diez lineas de maiz
476 subtropical. Agron. Mesoamericana 1997;8:1-7.
- 477 36. Fan XM, Zhang YM, Yao WH, Chen HM, Tan J, Xu CX, Han XL, Luo LM, Kang MS.
478 Classifying maize inbred lines into heterotic groups using a factorial mating design.
479 Agron. J. 2009;101(1);106-112.
- 480 37. Bidhendi MZ, Choukan R, Darvish F, Mostafavi K, Majidi E. Classifying of maize
481 inbred lines into heterotic groups using diallel analysis. World Acad. Sci. Engineer.
482 Tech. 2012;6;1059-1162.
- 483 38. Choukan R, Hossainzadeh A, Ghannadha MR, Warburton ML, Talei AR, Mohammadi
484 SA. Use of SSR data to determine relationships and potential heterotic groupings
485 within medium to late maturing Iranian maize inbred lines. Field Crops Res.
486 2006;95:212-222.
- 487 39. Xie F, He Z, Esguerra MQ, Qiu F, Ramanathan V. Determination of heterotic groups
488 for tropical Indica hybrid rice germplasm. Theor. Appl. Genet. 2014;127:407-417.
- 489 40. Smith JSC, Smith OS. Comparison of heterosis among hybrids as a measure of
490 relatedness with that to be expected on the basis of pedigree. Maize Genet. Coop.
491 Newsl. 1989;63:86-87.

- 492 41. Peng JY, Glaszmann JC, Virmani SS. Heterosis and isozymes diversions in *indica*
493 rice. Crop Sci. 1988;28:561-563.
- 494 42. Frei OM, Stuber CW, Goodman MM. Use of allozymes as genetic markers for
495 predicting performance in maize single cross hybrids. Crop Sci. 1986;26:37-42.
- 496 43. Stuber CW, Polacco M, Senior ML. Synergy of empirical breeding, marker-assisted
497 selection, and genomics to increase crop yield potential. Crop Sci. 1999;39:1571-
498 1583.
- 499 44. Xiao J, Li J, Yuan L, McCouch SR, Tanksley SD. Genetic diversity and its
500 relationships to hybrid performance and heterosis in rice as revealed by PCR-based
501 markers. Theor. Appl. Genet. 1996;92:637-643.
- 502 45. Zhang T, Ni XL, Jiang KF, Deng HF, He Q, Yang QH, et al. Relationship between
503 heterosis and parental genetic distance based on molecular markers for functional
504 genes related to yield traits in rice. Rice Science 2010;17:288-295.
- 505 46. Corbellini M, Perenzin M, Accerbi M, Vaccino P, Borghi B. Genetic diversity in bread
506 wheat, as revealed by coefficient of parentage and molecular markers, and its
507 relationship to hybrid performance. Euphytica 2002;123:273-285.
- 508 47. Cheres MT, Miller JF, Crane JM, Knapp SJ. Genetic distance as a predictor of
509 heterosis and hybrid performance within and between heterotic groups in sunflower.
510 Theor. Appl. Genet. 2000;100:889-894.
- 511 48. Diers BW, McKetty PBE, Osborn TC. Relationship between heterosis and genetic
512 distance based on restriction fragment length polymorphism markers in oilseed rape
513 (*Brassica napus* L). Crop Sci. 1996;36:79-83.
- 514 49. Zhao MF, Li XH, Yang JB, Xu CG, Hu RY, Liu DJ, et al. Relationship between
515 molecular marker heterozygosity and hybrid performance in intra- and inter-
516 subspecific crosses of rice. Plant Breeding 1999;118:139-144.
- 517 50. Joshi SP, Bhavé SG, Chowdari KV, Apte GS, Dhonukshe BL, Lalitha K, et al. Use of
518 DNA markers in prediction of hybrid performance and heterosis for a three-line hybrid
519 system in rice. Biochemical Genetics 2001;39:179-200.
- 520 51. Kwon SJ, Ha WG, Hwang HG, Yang SJ, Choi HC, Moon HP, et al. Relationship
521 between heterosis and genetic divergence in 'Tongil'-type rice. Plant Breeding
522 2002;121:487-492.
- 523 52. Liu X, Ishiki K, Wang W. Identification of AFLP markers favorable to heterosis in
524 hybrid rice. Breeding Science 2002;52:201-206.
- 525 53. Xu W, Virmani SS, Hernandez JE, Sebastian LS, Redoña ED, Li Z. Genetic diversity
526 in the parental lines and heterosis of the tropical rice hybrids. Euphytica
527 2002;127:139-148.
- 528 54. Cho YI, Park CW, Kwon SW, Chin JH, Ji HS, Park KJ, et al. Key DNA markers for
529 predicting heterosis in F_1 hybrids of *japonica* rice. Breed. Sci. 2004;54:389-397.

- 530 55. Renming Z, Yinghua L, Zhenglin Y, Fangming Z, Bingqiang Z, Rong X, et al.
531 Prediction of hybrid grain yield performances in *indica* rice (*Oryza sativa* L.) with
532 effect-increasing loci. Mol. Breed. 2008;22:467-476.
- 533 56. Jaikishan I, Rajendrakumar P, Ramesha MS, Viraktamath BC, Balachandran SM,
534 Neeraja CN, et al. Prediction of heterosis for grain yield in rice using 'key' informative
535 EST-SSR markers. Plant Breeding 2010;129:108-111.
- 536 57. Lanza LLB, de Souza Jr CL. Genetic distance of inbred lines and prediction of maize
537 single-cross performance using RAPD markers. Theor. Appl. Genet. 1997;94:1023-
538 1030.
- 539 58. Benchimol LL, de Souza Jr CL, Garcia AAF, Kono PMS, Mangolin CA, Barbosa AMM,
540 et al. Genetic diversity in tropical maize inbred lines: heterotic group assignment and
541 hybrid performance determined by RFLP markers. Plant Breeding 2000;119:491-496.
- 542 59. Vuylsteke M, Kuiper M, Stam P. Chromosomal regions involved in hybrid performance
543 and heterosis: their AFLPs® -based identification and practical use in prediction
544 models. Heredity 2000;85:208-218.
- 545 60. Parentoni SN, Magalhaes JV, Pacheco CAP, Santos MX, Abadie T, Gama EEG, et al.
546 Heterotic groups based on yield-specific combining ability data and phylogenetic
547 relationship determined by RAPD markers for 28 tropical maize open pollinated
548 varieties. Euphytica 2001;121:197-208.
- 549 61. Shieh GJ, Thseng FS. Genetic diversity of Tainan-white maize inbred lines and
550 prediction of single cross hybrid performance using RAPD markers. Euphytica
551 2002;124:307-313.
- 552 62. Barbosa AMM, Geraldi IQ, Benchimol LL, Garcia AAF, Souza Jr CL, Souza AP.
553 Relationship of intra- and inter-population tropical maize single cross hybrid
554 performance and genetic distances computed from AFLP and SSR markers.
555 Euphytica 2003;130:87-99.
- 556 63. Xu SX, Liu J, Liu GS. The use of SSRs for predicting the hybrid yield and yield
557 heterosis in 15 key inbred lines of Chinese maize. Hereditas 2004;141:207-215.
- 558 64. Schrag TA, Melchinger AE, Sorensen AP, Frisch M. Prediction of single-cross hybrid
559 performance for grain yield and grain dry matter content in maize using AFLP markers
560 associated with QTL. Theor. Appl. Genet. 2006;113:1037-1047.
- 561 65. Schrag TA, Maurer HP, Melchinger AE, Piepho HP, Peleman J, Frisch M. Prediction
562 of single-cross hybrid performance in maize using haplotype blocks associated with
563 QTL for grain yield. Theor. Appl. Genet. 2007;114:1345-1355.
- 564 66. Kiula BA, Lyimo NG, Botha AM. Association between AFLP-based genetic distance
565 and hybrid performance in tropical maize. Plant Breeding 2008;127:140-144.
- 566 67. Legesse BW, Myburg AA, Pixley KV, Afriyie ST, Botha AM. Relationship between
567 hybrid performance and AFLP based genetic distance in highland maize inbred lines.
568 Euphytica 2008;162:313-323.

- 569 68. Schrag TA, Mohring J, Maurer HP, Dhillon BS, Melchinger AE, Piepho HP, et al.
570 Molecular marker-based prediction of hybrid performance in maize using unbalanced
571 data from multiple experiments with factorial crosses. *Theor. Appl. Genet.*
572 2009;118:741-751.
- 573 69. Perenzin M, Corbellini M, Accerbi M, Vaccino P, Borghi B. Bread wheat: F₁ hybrid
574 performance and parental diversity estimates using molecular markers. *Euphytica*
575 1998;100:273-279.
- 576 70. Liu ZQ, Pei Y, Pu ZJ. Relationship between hybrid performance and genetic diversity
577 based on RAPD markers in wheat, *Triticum aestivum* L. *Plant Breeding* 1999;118:119-
578 123.
- 579 71. Krystkowiak K, Adamski T, Surma M, Kaczmarek Z. Relationship between phenotypic
580 and genetic diversity of parental genotypes and the specific combining ability and
581 heterosis effects in wheat (*Triticum aestivum* L.). *Euphytica* 2009;165:419-434.
- 582 72. Jordan DR, Tao Y, Godwin ID, Henzell RG, Cooper M, McIntyre CL. Prediction of
583 hybrid performance in grain sorghum using RFLP markers. *Theor. Appl. Genet.*
584 2003;106:559-567.
- 585 73. Rajendrakumar P, Hariprasanna K, Jaikishan I, Madhusudhana R, Patil JV. Potential
586 of microsatellite marker polymorphism in the prediction of grain yield heterosis in
587 sorghum [*Sorghum bicolor* (L.) Moench]. In: Rakshit S, Das IK, Shyamprasad G,
588 Mishra JS, Ratnavathi CV, Chapke RR, Tonapi VA, Dayakar Rao B and Patil JV (eds)
589 Compendium of papers and abstracts: Global consultation on millets promotion for
590 health and nutritional security, 18–20 Dec 2013. Society of Millets Research, 11-127,
591 Directorate of Sorghum Research, Rajendranagar, Hyderabad 500030, India, p 329,
592 2013.
- 593 74. Lu XP, Liu DD, Wang SY, Mi FG, Han PA, Lu ES. Genetic effects and heterosis
594 prediction model of *Sorghum bicolor* × *S. sudanense* grass. *Acta Agron. Sin.*
595 2014;40(3):466-475.
- 596 75. Tams SH, Bauer E, Oettler G, Melchinger AE, Schon CC. Prospects for hybrid
597 breeding in winter triticale: II. Relationship between parental genetic distance and
598 specific combining ability. *Plant Breeding* 2006;125:331-336.
- 599 76. Riaz A, Li G, Quresh Z, Swati MS, Quiros CF. Genetic diversity of oilseed *Brassica*
600 *napus* inbred lines based on sequence-related amplified polymorphism and its relation
601 to hybrid performance. *Plant Breeding* 2001;120:411-415.
- 602 77. Yu CY, Hu SW, Zhao HX, Guo AG, Sun GL. Genetic distances revealed by
603 morphological characters, isozymes, proteins and RAPD markers and their
604 relationships with hybrid performance in oilseed rape (*Brassica napus* L.). *Theor.*
605 *Appl. Genet.* 2005;110:511-518.
- 606 78. Shen JX, Fu TD, Yang GS, Tu JX, Ma CZ. Prediction of heterosis using QTL for yield
607 traits in rapeseed (*Brassica napus* L.). *Euphytica* 2006;151:165-171.
- 608 79. Zeid M, Schon CC, Link W. Hybrid performance and AFLP-based genetic similarity in
609 faba bean. *Euphytica* 2004;139:207-216.

- 610 80. Geleta LF, Labuschagne MT, Viljoen CJ. Relationship between heterosis and genetic
611 distance based on morphological traits and AFLP markers in pepper. *Plant Breeding*
612 2004;123:467-473.
- 613 81. Zhang XQ, Wang XD, Jiang PD, Hua SJ, Zhang HP, Dutt Y. Relationship between
614 molecular marker heterozygosity and hybrid performance in intra- and inter-specific
615 hybrids of cotton. *Plant Breeding* 2007;126:385-391.
- 616 82. Sant VJ, Patankar AG, Sarode ND, Mhase LB, Sainani MN, Deshmukh RB, et al.
617 Potential of DNA markers in detecting divergence and in analyzing heterosis in Indian
618 elite chickpea cultivars. *Theor. Appl. Genet.* 1998;98:1217-1225.
- 619 83. Riday H, Brummer EC, Campbell TA, Luth D, Cazcarso PM. Comparisons of genetic
620 and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and
621 subsp. *falcata*. *Euphytica* 2003;131:37-45.
- 622 84. Xie X, Song MH, Jin F, Ahn SN, Suh JP, Hwang HG, et al. Fine mapping of a grain
623 weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived
624 from a cross between *Oryza sativa* and *Oryza rufipogon*. *Theor. Appl. Genet.*
625 2006;113:885-894.
- 626 85. Parh DK, Jordan DR, Aitken EAB, Mace ES, Junai P, McIntyre CL, et al. QTL analysis
627 of ergot resistance in sorghum. *Theor. Appl. Genet.* 2008;117:369-382.
- 628 86. Wu Y, Huang Y. Molecular mapping of QTL for resistance to the green bug *Schizaphis*
629 *graminum* (Rondani) in *Sorghum bicolor* (Moench). *Theor. Appl. Genet.*
630 2008;117:117-124.
- 631 87. He GH, Hou L, Li DM, Luo XY, Niu GQ, Tang M, et al. Prediction of yield and yield
632 components in hybrid rice by using molecular markers. *Acta Genet. Sin.* 2002;29:438-
633 444.
- 634 88. Yang J, Zhu J. Methods for predicting superior genotypes under multiple
635 environments based on QTL effects. *Theor. Appl. Genet.* 2005;110:1268-1274.
- 636 89. Gao YM, Zhu J. Mapping QTL with digenic epistasis under multiple environments and
637 predicting heterosis based on QTL effects. *Theor. Appl. Genet.* 2007;115:325-333.
- 638 90. Ben-Israel I, Kilian B, Nida H, Fridman E. Heterotic trait locus (HTL) mapping identifies
639 intra-locus interactions that underlie reproductive hybrid vigor in *Sorghum bicolor*.
640 *PLoS ONE* 2012;7:e38993.
- 641 91. Melchinger AE, Boppenmaier J, Dhillon BS, Pollmer WG, Herrmann RG. Genetic
642 diversity for RFLPs in European maize inbreds. II. Relation to performance of hybrids
643 within versus between heterotic groups for forage traits. *Theor. Appl. Genet.*
644 1992;84:672-681.
- 645 92. Boppenmaier J, Melchinger AE, Brunklaus-Jung E, Geiger HH, Herrmann RG.
646 Genetic diversity for RFLPs in European maize inbreds. I. Relation to performance of
647 flint x dent crosses for forage traits. *Crop Sci.* 1992;32:895-902.

- 648 93. Ajmone-Marson P, Castiglioni P, Fusar F, Kuiper M, Motto M. Genetic diversity and its
649 relationship to hybrid performance in maize as revealed by RFLP and AFLP markers.
650 Theor. Appl. Genet. 1998;96:219-227.
- 651 94. Bernardo R. Relationship between single cross performance and molecular marker
652 heterozygosity. Theor. Appl. Genet. 1992;83:628-634.
- 653 95. Bernardo R. Predicting the performance of untested single crosses: trait and marker
654 data. In: Lamkey KR, Staub JE, editors. Concepts and breeding of heterosis in crop
655 plants. Madison: Crop Science Society of America; 1998.
- 656 96. Bernardo R. Marker-assisted best linear unbiased prediction of single-cross
657 performance. Crop Sci. 1999;39:1277-1282.
- 658 97. Song S, Huang Y, Wang X, Wei G, Qu H, Wang W, et al., HRGD: a database for
659 mining potential heterosis-related genes in plants. Plant Mol. Biol. 2009;69:255-260.
- 660 98. Xiong LZ, Yang GP, Xu CG, Zhang Q, Saghai-Maroo MA. Relationships of differential
661 gene expression in leaves with heterosis and heterozygosity in a rice diallel cross.
662 Mol. Breed. 1998;4:129-136.
- 663 99. Wu LM, Ni ZF, Wang ZK, Lin Z, Sun QX. Relationship between differential gene
664 expression patterns of multigene families and heterosis in a wheat diallel crosses.
665 Acta Genetica Sinica 2001;28:256-266.
- 666 100. Thiemann A, Fu J, Schrag TA, Melchinger AE, Frisch M, Scholten S. Correlation
667 between parental transcriptome and field data for the characterization of heterosis in
668 *Zea mays* L. Theor. Appl. Genet. 2010;120:401-413.
- 669 101. Frisch M, Thiemann A, Fu J, Schrag TA, Scholten S, Melchinger AE. Transcriptome-
670 based distance measures for grouping of germplasm and prediction of hybrid
671 performance in maize. Theor. Appl. Genet. 2010;120:441-450.
- 672 102. Stokes D, Morgan C, O'Neill CM, Fraser F, Bancroft I. Transcriptome-based predictive
673 modeling of heterosis. Abstract International Symposium on Heterosis in Plants. 18-20
674 May, 2006 Potsdam-Golm.