



HETEROSIS AND COMBINING ABILITY FOR GRAIN YIELD AND EARLINESS IN ACCESSIONS OFA RICE CORE COLLECTION

Mariana Rodrigues Feitosa Ramos¹, João Antônio Mendonça², Rosana Pereira Vianello²,
Odilon Peixoto de Moraes Júnior¹, José Manoel Colombari², Tereza Cristina de Oliveira Borba²,
Adriano Pereira de Castro², Claudio Brondani²

¹ Universidade Federal de Goiás (UFG), 74.690-900, Goiânia, GO, Brazil

² Embrapa Arroz e Feijão, 75.375-000, Santo Antônio de Goiás, GO, Brazil

Corresponding author: Claudio Brondani (claudio.brondani@embrapa.br).

Abstract: Rice has extensive collections of germplasm, but a small fraction of genotypes are used in breeding programs worldwide. This aimed to assess the heterosis and the combining ability for higher grain yield and earliness, and their correlations to genetic distance among accessions belonging to the Embrapa's Rice Core Collection. Crosses were made in complete diallel, without the reciprocals, being phenotyped in generations F_2 and F_7 for grain yield (GY) and days to flowering (DTF). It were estimated the varietal effect, mean heterosis, varietal heterosis, specific combining ability and general combining ability (GCA) of the parents in each generation. For the selection of new lines, the combinations involving the genitors with the highest magnitudes and positive effects for GCA were considered more promising for the GY, of which Canela Curta, Maninjau, Epagri 108 and Diamante stood out, and were also the most productive. For the DTF, the genitors with the largest earliness estimates were Araguaia, CT11632, Irat 122, Pratinha Branco and Tox 503. The recommended parameters for estimating GY in F_7 generation from the data of the F_2 generation were the mean of GY, GCA and the mean Rogers-W genetic distance coefficient. Crosses with Rogers-W distance coefficients above 0.9 showed higher GY and earliness in F_7 generation, whereas in F_2 the superior crosses were those that showed Rogers-W between 0.7 and 0.89. The above genotypes and the specific combinations identified may lead to new rice cultivars.

Keywords: *Oryza sativa*, diallel analysis, genetic divergence, SSR markers, molecular breeding

Introduction

Most breeding programs of self-pollinating species, such as rice, generate segregating populations of biparental crosses between elite genitors, usually genetically related, for the development of cultivars. This strategy aims to ensure success in obtaining new superior inbred lines through recombination among parents who already accumulate many favorable alleles, since successive generations of self-fertilization from the F_2 generation do not promote sufficient recombination to increase the transgressive frequency of favorable gene combinations (Dennis et al., 2008). However, this uniformity, in addition to limiting the annual genetic gain in grain yield, makes the elite germplasm of these programs genetically uniform, and more likely to be susceptible to pests and diseases (Huang et al., 2014).

An alternative to increase the exploration of the potential variability to be worked by breeding programs is the better organization of the genetic resources available in genebanks, through core collections formed by a smaller set of accessions representing the largest fraction of genetic variability of the total collection (Brown, 1989). For rice, core collections have already been developed for the germplasm stored in countries such as Brazil (Abadie et al., 2005), the U.S.A. (Yan et al., 2007) and China (Zhang et al., 2011). However, there would be a multitude of possible crossings to be made with such accessions with intent to broaden the genetic base useful to the breeding program. The most widely used method for selecting parents is diallel analysis, which allows for estimating combining ability, a term introduced by Sprague and Tatum (1942).

Diallel analysis permits to infer the potential of each genitor to provide superior hybrid combinations, the gene action that controls the traits involved, and the existence of heterosis, providing the opportunity to identify and develop superior lines (Huang et al., 2015). The general combining ability generates information regarding a set of genes of predominantly additive effects, and the specific combining ability is related to a set of genes of dominance and epistatic effects (Asfaliza et al., 2012; Fasahat et al., 2016).

The use of molecular markers for genetic characterization of germplasm offers real benefits to increase efficiency at various stages of breeding programs and can be adopted with high accuracy and at reduced costs for, as examples, cross validation, certification of cultivars and determination of seed lot genetic uniformity. The objective of this study was to assess the heterosis and the combining ability for higher grain yield and earliness, and their correlations to genetic distance among accessions belonging to the Embrapa Rice Core Collection.

Material and Methods

Among the 550 accessions of the Embrapa Rice Core Collection, the 12 genotypes selected were the most productive in a previous experiment (Bueno et al., 2012) carried out at Palmital Farm (Goianira, GO), which belongs to Embrapa Arroz e Feijão. Those included five landraces, two Brazilian cultivars and five genotypes from abroad (Table 1). The 12 genotypes constituted the group of parents of the complete diallelic crosses without reciprocals, which resulted in 66 combinations of crosses. During the experiment in generation F_2 , and during the process of advance of generations, five of the 66 combinations (Bico Roxo x Lageado, Canarinho x Lageado, Canarinho x Bico Roxo, Diamante x Bico Roxo and Diamante x Irat 122) were lost. The F_1 hybrids were self-fertilized to obtain generation F_2 , which was evaluated through field experiment, at Palmital Farm, using a randomized complete block design, with three replications in the 2006/2007 crop season. The plot consisted of 4 rows, each 4 meters in length, and a plot useful area of 2.40 m² (2 central rows). Afterwards, the diallelic cross group was advanced in *bulk*, from generation F_2 to F_7 . The F_7 generation was evaluated in the 2013/2014 crop season, using the randomized complete block design with two replications, with the same size of the plot and in the same location of the F_2 generation.

The traits evaluated in the experiments of generations F_2 and F_7 were number of days to flowering (DTF), corresponding to the number of days between sowing and mean flowering of the plot (50% of flowered panicles), and grain yield (GY), corresponding to the dried grain weight, in kg ha⁻¹. The phenotypic means of the parents and the crosses in the two generations, adjusted for the year and the block effects, were estimated by the

joint analysis of the two experiments using the following fixed linear model:

$$y_{ijk} = \mu + \frac{(v_i + v_j)}{2} + \theta(\bar{h}_k + h_{ki} + h_{kj} + s_{ijk}) + \bar{e}_{ijk} \quad (\text{Eq. 1})$$

where Y_{ijk} is the observation of the treatment i , in block k , in year j ; μ is the constant, inherent to all observations; a_j is the effect of the year j ; $b_{k(j)}$ is the effect of block k within the year j ; t_i is the effect of the treatment i , being i equal to parents or crosses; and e_{ijk} is the experimental error associated with the observation ijk , assuming $NID \cap (0, \sigma^2)$, where NID stands for normally and independently distributed. The interaction treatments *versus* years were considered as a component of the experimental error, a necessary condition for the estimation of the adjusted means of the treatments (Searle, 1971). In addition, treatments and years were treated as fixed effects for the following reasons: (i) both locations, represented by the years, consisted of environments with their own characteristics that were targets of interest for the rice breeding program; (ii) treatments included lines or parental cultivars selected for desirable traits in the breeding program; thus, they were not randomly selected. For each generation (year) was obtained the vector estimation \hat{Y} , corresponding to the means of treatments adjusted for effects described in the linear model, and also the covariance matrix [$\hat{V}(\hat{Y})$] of the adjusted means for the joint analysis of the genotype group, in accordance with Morais Júnior et al. (2017a).

The combining ability of the accessions was estimated using the Gardner and Eberhart (1966) method, based on the generational model proposed by Pereira et al. (2008), with necessary adaptation to the generalized method of diallelic analysis of unbalanced data, as proposed by Silva et al. (2000). The justification for using this method is based on the existence of unbalance among the experimental data, due to the different number of repetitions between some treatments. The Gardner and Eberhart method (1966) was used in the diallelic analysis involving the crosses and their parents in the generations F_2 and F_7 , thus allowing for the decomposition of the total heterosis effects in mean heterosis, heterosis of parents and specific heterosis (or specific combining ability – SCA). In addition, Cruz and Regazzi (1994) point out that this method is recommended when there are significant differences between diallelic mean values due to heterogeneity among parents, and also in the presence of heterosis. Such conditions are present in the set of crosses due to the different origins of the parents. To estimate the effects of genitors and heterosis components for each generation, the following linear model was used:

$$y_{ijk} = \mu + \frac{(v_i + v_j)}{2} + \theta(\bar{h}_k + h_{ki} + h_{kj} + s_{ijk}) + \bar{e}_{ijk}$$

where y_{ijk} is the estimated mean of crosses between i and j genitors, in the F_k generation; μ is the general mean;

v_i and v_j are the effects of the i and j genitors, respectively; θ is the indicator variable, where $\theta = 0$ if $i=j$, otherwise $\theta = 1$; \bar{h}_k is the mean heterosis in the F_k generation; h_{ki} and h_{kj} represent the heterosis attributed to the i or j genitors, respectively, in generation F_k ; S_{ijk} is the specific heterosis (or SCA) of the cross between the i and j genitors in the F_k generation; and \bar{e}_{ijk} is the mean experimental error associated with the mean y_{ijk} (non-homogeneous and non-orthogonal).

In order to perform a decomposition of the sum of squares (SS) of treatments, which are non-orthogonal necessarily in this case, adjustments were made in SS (v), SS (\bar{h}_k), SS (h_{ki}), SS (h_{kj}), SS (S_{ijk}). For each k , the adapted equation of the general sum of squares hypothesis in linear models (Searle, 1971) was used for each variation component:

$$SS = (C\beta)'(CQC')^{-1}C\beta$$

where C it is a matrix of estimate contrasts; β is the solution vector of the diallelic equation; Q is the covariance matrix of β [$\hat{V}(\hat{\beta})$], divided by the mean square of residue from the joint variance analysis ($\hat{\sigma}^2$), i.e., ...

$Q = \hat{V}(\hat{\beta})/\hat{\sigma}^2$. Information on the structure of $\hat{V}(\hat{\beta})$ are detailed in Silva et al. (2000) and Morais Júnior et al. (2017b). The constraint matrix was structured, admitting the following restrictions in the solution space: $\sum_i \hat{v}_i = 0$; $\sum_j \hat{h}_{kj} = 0$, for each k ; and $\sum_j \hat{S}_{ijk}$, for each i and k . From the effects of genitor (\hat{v}_i or \hat{v}_j) and of genitor heterosis \hat{h}_{ki} in the mean of the generations, estimates of the general combining ability (GCA) of each genitor (\hat{g}_i) were obtained by the expression:

$$g_i = \frac{1}{2} v_i + h_i; \left(\sum_i g_i = 0 \right).$$

The genitors and diallelic crosses for each generation (F_2 and F_7) were compared with the purpose of obtaining the varietal effect, in addition to the estimates of mean heterosis for all crosses, heterosis of crosses and specific heterosis in F_2 . All statistical analyses to obtain adjusted means of treatments and effects were performed using platform R (R Core Team, 2013). Pearson coefficient correlation analyses between the Rogers genetic distance modified by Wright (RW) (Wright, 1978), specific heterosis (\hat{S}_{ij}) and the means of the two characters evaluated were obtained by the Genes program (Cruz, 2013). The RW genetic distance estimates were determined from the genotyping data of the parents through 24 SSR markers described by Borba et al. (2009).

Results and discussion

The coefficients of variation (CV%) obtained by the joint analysis of the experiments, in generations F_2 and F_7 , were 16.98% and 3.65% for GY and DTF, respectively. These values are considered satisfactory for the experimental design and plot size adopted for the evaluated traits, in accordance with Silva et al. (2000); thus indicating a high experimental precision. The control group displayed the highest GY mean, followed by the group of crosses in generation F_7 , that were superior to the respective genitors and the crosses in generation F_2 , also surpassing the general mean for the trait (Tables 1 and 2). In the generation advances by the bulk method, the average of crosses in F_7 should have been equal to the average of the genitors, since there is no dominance effects. In addition, the GY F_2 average should have been higher than that of the parents due to the presence of half of heterosis. According to Charlesworth et al. (2017) these results can be explained by the fact that over the course of the inbreeding generations, the most adapted individuals can generate more descendants due to greater grain yield. For the GY trait, the mean heterosis referring to generations F_2 and F_7 was not significant ($p > 0.05$), which indicates that crosses in both F_2 and F_7 generation do not differ from the mean of the genitors. However, specific heterosis (or SCA) was highly significant, indicating the occurrence of certain crosses that stood out in relation to the respective parents (positive heterosis) (Table 3). For DTF, there was significance for mean heterosis ($p < 0.01$) in both generations F_2 and F_7 , indicating that the set of crosses in both generations differs from the mean of the parents. Although the magnitude of the mean specific heterosis (or SCA), has been reduced from F_2 to F_7 , it remained significant, indicating the presence of epistasis between loci involving additive effects (Bhullar et al., 1979).

Grain yield (GY)

The parameters varietal effect (\hat{v}), varietal heterosis (\hat{h}_v) and the combining ability of the parents (\hat{g}_i) in generations F_2 and F_7 were estimated (Table 4). The varietal heterosis of a genitor is the difference between the mean of the heterosises of all their crosses and the mean heterosis of the diallel in the same generation (Cruz and Regazzi, 1994). Therefore, with the significance of varietal heterosis, it can be inferred that the parent differ statistically from each other in terms of their mean allele frequencies. Canela Curta, CT11632, Irat 122, Lageado, Maninjau and Pratinha Branco genitors showed significant estimates of varietal heterosis in generation F_2 ($\hat{h}_v F_2$), indicating that these genitors have a greater divergence in relation to the group of genitors of the diallel for the GY trait, while CT11632, Irat 122 and Pratinha Branco showed negative effects for this trait. Varietal heterosis in generation F_7 ($\hat{h}_v F_7$) was positive and significant for

Lageado and Maninjou (repeating the result of generation F_2), along with the Canarinho, and was negative and significant for CT11632 and Irat 122 (repeating the result of generation F_2), together with Epagri 108.

Epagri 108, Diamante, Irat 122, Maninjou and Canela Curta showed the highest estimates ($\alpha = 0.05$) of varietal effect (i.e. the *per se* effect) for GY, while the genitors Lageado, Canarinho, Bico Roxo and Tox 503 showed negative estimates (Table 4). The Maninjou genitor, in addition to having a prominent varietal effect, increased, on mean, the GY of the crosses in which it participated, considering the two generations evaluated. On the other hand, despite having a negative varietal effect, on average, Lageado also increased the GY in both generations. This demonstrates the importance of conducting studies to select potential genitors for breeding programs through diallelic crosses, because, despite the negative and significant varietal effect, Lageado participated in specific crosses with potential to generate lines with high GY. The genitors that showed the highest varietal effect also showed the highest values for general combining ability in F_2 and F_7 , with the exception of Irat 122 (Table 4).

The specific heterosis or specific combining ability - SCA (\hat{S}_{ij}) is an important parameter for the identification of the best combinations of pair-on-pair parents (Asfaliza et al., 2012). Among the 61 combinations evaluated for GY, eight were positive and significant in generation F_2 and two in F_7 , while the negative and significant estimates were eight in F_2 and three in F_7 (Table 5). The trend of reduction of specific heterosis, resulting from the advances in inbreeding, was already expected, because in F_7 generation the dominance effects are practically null. Thus, contributions in medium and specific heterosis in combinations in advanced generations, as in F_7 , are the result of epistatic effects involving additive effects that have been accumulated with generations of self-fertilization. In addition, the importance of specific heterosis will be manifested if there are considerable differences between the average allele frequencies of at least part of the parents. Therefore, differences between degrees of complementation of frequencies may also be responsible for the manifestation of this type of combining ability.

Considering the mean SCA values of the cross groups classified as Indica x Indica (acronym I x I), Indica x Japonica (I x J), and Japonica x Japonica (J x J), in F_2 the highest mean SCA was between I x I crossings, and the lowest I x J, while in F_7 , the largest was between J x J, and the smallest was, once again, for I x J crosses. It is expected that crosses of type I x J show higher heterosis, as reported by Najeeb et al. (2013). However, the broad genetic diversity of the diallelic genitors of the present

study resulted in the greater heterosis within the Indica and Japonica groups.

Estimates of varietal heterosis in the early F_2 generation were used to predict SCA of crosses in advanced F_7 generation of the diallel. Considering the three genitors that showed positive and significant varietal heterosis in F_2 , Maninjou participated in a cross with positive and significant SCA (with Pratinha Branco) and one with negative and significant SCA (with Araguaia) in F_7 , while Canela Curta participated in a cross with negative SCA (with Canarinho) (Table 5). Considering the three genitors with negative and significant varietal heterosis in F_2 , CT11632 and Pratinha Branco were present in a cross with positive SCA in F_7 (with Araguaia and Maninjou, respectively), whereas CT11632 and Irat 122 were genitors of a cross with a negative value of SCA in F_7 , as result of depression by endogamy. These results indicate that estimates of varietal heterosis were not a good predictor for SCA in F_7 , and this has implications for the use of generation F_2 to predict the potential crosses with higher GY in F_7 . According to Gärtner et al. (2009), specifically regarding GY, heterosis can vary very strongly depending on the specific combination, and in this case the selection of the best parental combination may be erroneously determined. Additionally, one of the ways to improve the heterosis prediction would be the use of molecular markers to infer the complementarity between the genitors.

The genitors with the highest general combining ability (GCA) values (\hat{g}_i) can be included in breeding programs of self-pollinating crops in order to develop genotypes that exceed their genitors in advanced generations (Fasahat et al., 2016). In the present study, of the 12 genitors evaluated, four showed high and positive in both generations (Canela Curta, Diamante, Epagri 108 and Maninjou). At least one of these four genitors was involved in the 20 most productive combinations, and in only four of the 20 least productive combinations in F_2 , while in F_7 they were present in 19 and four combinations, respectively. In addition, these genitors also had outstanding *per se* productivity, as they were among the five most productive in the joint analysis of generations F_2 and F_7 , demonstrating the importance of the contribution of its additive effects (Table 1). Huang et al. (2015) also observed that parents with higher GCA showed higher productivity performance. Of the four parents with high \hat{g}_i , only Maninjou was positive and significant in the five estimates evaluated for GY (Table 4). In self-pollinating species, defining the combining ability of a given genotype in very early generations may lead to an erroneous conclusion about their capacity to derive promising inbred lines, because for these species is not important the effects of dominance, but the additive effects of homozygous loci (Lorencetti et al., 2006; Bertan

et al., 2009). GCA is usually related to genes in additivity (Huang et al., 2015), suggesting that selection for the GY trait could be performed in early generations (Asfaliza et al., 2012). However, non-additive interactions are markedly present in early generations, such as dominance, over-dominance, and epistasis, which are difficult to work with in the progeny due to low heritability and interactions with the environment (Liu et al., 2015; Morais Júnior et al., 2017b), which may induce an error in the choice of genitors with greater productive potential.

Earliness analysis

As described in this study, the Epagri 108, Diamante, Maninjau and CanelaCurta accessions were recommended as potential sources for obtaining lines with greater productive potential. With the exception of CanelaCurta, these genitors had a varietal effect (*per se*) that was favorable and significant for earliness in flowering. However, for the other estimated parameters, only Maninjau showed earliness, and only as GCA in generation F_7 . In addition, it was observed that GY specific heterosis correlated positively with DTF specific heterosis, indicating that the most productive combinations were also the latter ones ($p < 0.01$). Late flowering in rice allows higher biomass accumulation, and consequently higher productivity (Jung and Müller, 2009; Gao et al., 2014), while very early varieties have comparatively lower productivity (Shahriar et al., 2014). On the other hand, the parents that generated earlier specific combinations, in both F_2 and F_7 , were BicoRoxo, Canarinho, Irat 122 and Lageado. Of these, only Irat 122 had a significant GCA for earliness (in F_2 and F_7), along with Araguaia, CT11632, PratinhaBranco and Tox503 (Table 6). For DTF in generation F_2 , all the parents were present in at least one cross that showed positive and significant SCA (increase in the cycle), with particular emphasis on Pratinha Branco (present in four crossings). The Maninjau accession, on the other hand, was present in three crosses that showed GCA which significantly decreased the cycle (Table 7). In F_7 , however, Canela Curta participated in three crosses with negative and significant SCA, while Bico Roxo, Lageado, Araguaia, Irat 122 and Maninjau did not participate in crosses with negative and significant SCA for DTF. In the case of Maninjau, a prominent genitor for GY, none of the three crosses in which it participated showed SCA for earliness in F_7 (Table 7).

SSR markers to predict GY and earliness

One of the main applications of combining ability estimation is the possibility of defining heterotic groups which, for example, form the basis of hybrid programs of rice (Li et al., 2008) and corn (Fasahat et al., 2016). Molecular markers with high PIC (Polymorphism Information Content) can be used to predict the performance of hybrids (Soni et al., 2018), but the correlation is highly dependent on the predictive capacity of the model and the genetic structure of individuals (Dan et al., 2016). In the present study, no significant correlation was found between the genetic distance of the parents, obtained by the Rogers-W coefficient (RW acronym), and the specific heterosis in F_2 and SCA in F_7 generations (data not shown). The mean RW distance of the group of genitors, obtained by the panel of 24 SSR markers, was high (0.88; Table 8), and is usually reported in groups of genotypes with great genetic variability, which is indicative of unadapted accessions (Borba et al., 2009).

From the mean genetic distance data of the parents at each crossing, the overall mean for each parent was obtained. In accordance with this criteria, the most divergent were Epagri 108 and Irat 122 (both with $RW = 0.93$), and the least divergent was Lageado landrace (0.71). The diallelic combination with the lowest RW distance occurred between Pratinha Branco x Canela Curta landraces (0.45), while 21 combinations showed maximum distance (1.0). Considering the groups formed in accordance with the origin of the materials, the largest mean distance was in crosses between improved varieties from Brazil (IVB x IVB) and between improved varieties from other countries (IVC x IVC) ($RW = 0.92$). Crossings between landraces (LRC x LRC) showed smaller mean distance ($RW = 0.74$) in relation to those two groups. This result indicates that the materials developed by rice breeding programs analyzed in this study have good genetic diversity. With the mean RW values (RW_m), calculated from the RW of each genitor of a given cross, classes comprising high genetic distance (classes 1 and 2, RW_m 0.9 to 1), intermediate (classes 3 and 4, RW_m 0.7 to 0.89) and low (class 5, RW_m less than 0.7) were obtained (Table 9). Class 1, with 21 crosses, resulted in the highest GY average in generation F_7 , while class 3, with 16 crosses, showed the highest GY average in F_2 . Class 4, with 11 crosses, showed the highest earliness average in generation F_2 , while class 2, with 9 crosses, had the highest earliness average in generation F_7 . Crosses belonging to the two classes with the greatest RW_m (classes 1 and 2) resulted in the combinations with higher GY average and earliness in F_7 , while in F_2 , the combinations of genitors with the highest mean GY and earliness belonged to the classes with intermediate RW_m genetic distances (classes 3 and 4). Based on these results, in theory, in

a breeding program that uses routinely the germplasm genotyping with molecular markers, pairs of genitors with intermediate RWm values should be chosen, in order to generate superior individuals early in F₂ generation.

If in such programs the objective would be to proceed the generation advancement of progenies (e.g., by bulk method), for selection of lines in advanced generations, pairs of genitors resulting in highRWm genetic distance between them should be chosen.

Table 1. Adjusted means (considering F₂ and F₇) for grain yield (GY, in kg ha⁻¹) and days to flowering (DTF) of the diallel genitors and the check cultivars used in the experiments. I: *Oryza sativa* subspecies Indica, J: *O. sativa* subspecies Japonica.

Genotypes	Code	I/J	Country	Group*	Grain Yield**	DTF
Genitors						
Bico Roxo	BGA012924	I	Brazil	LRC	4441.0	e 128.53 a
Canarinho	BGA013974	I	Brazil	LRC	3054.0	f 125.93 b
Canela Curta	BGA013191	I	Brazil	LRC	6645.0	b 107.22 c
Lageado	BGA013263	I	Brazi.	LRC	2864.33	f 128.70 a
Pratinha Branco	BGA013619	J	Brazil	LRC	5348.50	d 90.72 e
Araguaia	BGA004206	J	Brazil	IVB	4814.50	e 85.39 g
Diamante	BGA004899	I	Brazil	LRC	7523.83	a 99.55 d
Epagri108	BGA010688	I	Brazil	IVB	7772.66	a 100.55 d
CT11632	BGA009115	J	Colombia	IVC	5881.83	c 89.22 f
Irat 122	BGA004752	J	France	IVC	7067.66	a 101.89 d
Maninjau	BGA006910	I	Indonesia	IVC	6991.66	a 92.89 e
Tox 503	BGA004788	J	Nigeria	IVC	4579.66	e 89.55 f
Check cultivars						
BRS Alvorada	-	I	Brazil		7337.87	a 101.33 a
BRS Irga 409	-	I	Brazil		7204.10	a 94.20 b
BRS Taim	-	I	Brazil		6887.32	b 99.33 a
BRS Irga 424	-	I	Brazil		6720.32	b 87.33 c

* LRC: landrace; IVB: improved varieties from Brazil; IVC: improved varieties from other countries.

** Means followed by equal letters in the column do not differ from each other by Scott-Knott test at 5% probability.

Table 2. Joint analysis of variance involving the two generations (F₂ and F₇) for the traits of grain yield (GY) and days to flowering (DTF).

FV	GY			DTF		
	DF	SS	MS	DF	SS	MS
Generation (K)	1	37,192,043	37,192,043**	1	12	12 ^{ns}
Block/K	4	6,620,976	1,655,244 ^{ns}	4	98	24 ^{ns}
Group (G)	3	55,990,222	18,663,407**	3	1,400	466**
Treatment/G	134	1,072,307,950	8,002,298**	134	32,311	241**
Diallel/G (D)	131	1,070,999,237	8,175,567**	131	31,592	241**
Genitors/G (P)	11	172,678,855	15,698,078**	11	14,924	1,356**
F ₂ Crossings/G (C ₂)	60	700,049,229	11,667,487**	60	12,664	211**
F ₇ Crossings/G (C ₇)	60	198,266,541	3,304,442**	60	4,007	66**
Checks/G (T)	3	1297,684	432,561 ^{ns}	3	717	239 ^{ns}
Residues	307	341,332,399	1,111,832	313	4,196	13

General Means	6.210,95			100,41		
Genitors	5.617,66 c #			102,36 a #		
F ₂ Crossings	4.834,02 d			94,77 b		
F ₇ Crossings	6.674,49 b			91,23 c		
Checkers	7.030,24 a			95,56 b		
F ₂ Generation	4.972,29 b			95,91 a		
F ₇ Generation	6.726,97 a			93,25 b		

CV%	16,98			3,65		

** , ns: significant at 1% probability and not significant by the F test, respectively.

#Means followed by equal letters in the column do not differ from each other by Scott-Knott test at 5% probability.

Table 3. Variance analysis of the non-orthogonal decomposition of the sum of squares for grain yield (GY) in kg ha⁻¹ and days to flowering (DTF).

FV	GY			DTF		
	DF	SS	MS	DF	SS	MS
Treatments	131	1.07	817,556,669 **	131	315.92	241.2**
Varieties	11	1.79	163,162,693 **	11	148.66	1351.4**
Het. in F ₂	60	4.24	706,088,028 **	60	191.02	318.4**
Mean Het. F ₂	1	2.76	2,759,158 ^{ns}	1	288.30	2883.0**
$\hat{h}_v F_2$	11	7.41	673,308,136 **	11	339	308.2**
$\hat{s}_{ij} F_2$	48	3.56	742,306,008 **	48	295.45	61.6**
Het. in F ₇	60	144,001,440	2,400,024 **	60	825.35	137.6**
Mean SCA F ₇	1	3,299,152	3,299,152 ^{ns}	1	214.28	2142.8**
$\hat{h}_v F_7$	11	5,388,1976	489,836,145 **	11	261.53	237.8**
$\hat{s}_{ij} F_7$	48	982,276,130	20,464,086 **	48	1,546,116	32.2**
Residue	307	3,41	111.183.192	313	4,196,924	134,057,508

CV%	18.20			3.54		

** , ns: significant at 1% probability and not significant by the F test, respectively.

Table 4. Estimates of the effects of varieties (\hat{v}), varietal heterosis in the generations F_2 ($\hat{h}_v F_2$) and F_7 ($\hat{h}_v F_7$), general combining ability (\hat{g}_{F_2} and \hat{g}_{F_7}) and mean heterosis, for the grain yield among twelve genitors, in kg ha⁻¹.

Genitors	\hat{v}	$\hat{h}_v F_2$	$\hat{h}_v F_7$	\hat{g}_{F_2}	\hat{g}_{F_7}
Canela Curta	1,053.22	* 1,148.21	** 302.58	ns 1,674.82	829.19
Araguaia	-777.27	ns -75.54	ns -178.71	ns -464.19	-567.36
Bico Roxo	-873.70	* -536.19	ns -8.42	ns -973.05	-445.27
Canarinho	-2,265.12	** 556.16	ns 932.54	** -576.39	-200.01
CT11632	290.05	ns -782.37	** -859.18	** -637.35	-714.16
Diamante	1,932.05	* -494.11	ns -322.66	ns 471.91	643.36
Irat 122	1,475.88	** -886.56	** -1,002.4	** -148.62	-264.46
Lageado	-3160.5	** 886.744	** 1,055.79	* -693.51	-524.46
Manin jau	1,399.88	* 944.74	** 1,000.34	* 1,644.69	1,700.29
Pratinha Branco	-243.27	ns -577.90	* -204.41	ns -699.55	-326.06
Tox 503	-1,012.11	* 20.13	ns -724.70	ns -485.92	-496.81
Epagri 108	2,180.88	** -203.29	ns 3,57	* 887.15	365.74
Mean Heterosis	-	-287.88	ns 148.93	-	-
General Mean	5,791.9				
CV (%)	18.20				

ns: not significant. * and **: significant by the t-test, 5% and 1% probability, respectively.

Table 5. Estimates of specific heterosis (\hat{S}_{ij}) of the crosses in F_2 (above the diagonal) and specific combining ability (in \hat{S}_{ij}) F_7 (below the diagonal) for the grain yield, in kg ha⁻¹.

F_2	F_7	C. Curta	Araguaia	B. Roxo	Canarinho	CT11632	Diamante	Irat 122	Lageado	Manin jau	P. Branco	Tox 503	Epagri 108
Canela Curta	-	3,600.2**	-263.30	175.11	-636.92	-1,928.17**	46.84	1,615.15	-1,121.22	-384.97	145.89	-1,248.68	
Araguaia	3.56	-	662.78	-1,451.88**	1,071.83	-1,734.93**	-421.89	-1,025.01	-1,573.7**	950.28	711.40	-789.16	
Bico Roxo	267.48	-497.46	-	NA	-883.80	NA	-473.03	NA	1,677.25**	-1,284.11	-632.48	1,196.69	
Canarinho	-1,932.28**	434.77	NA	-	-1,724.47**	3,265.04**	-277.93	NA	9,005.7	-2,175.52**	616.60	667.28	
CT11632	79.36	1,618.41*	403.83	202.07	-	-329.267	-148.73	-1,168.1	-153.29	2,866.19**	1,470.81*	-364.25**	
Diamante	180.84	208.89	NA	-33.94	413.69	-	NA	2,121.15	-584.05	-544.06	-766.69	500.98	
Irat 122	1,639.66	379.71	144.63	-627.12	-2,208.98**	NA	-	-775.32	-257.51	452.46	185.59	1,669.52**	
Lageado	206.66	793.21	NA	NA	-26.48	1,042.0	41.81	-	1,839.38**	-2,202.65**	-1,107.77	703.15	
Manin jau	-69.58	-2,598.03**	-67.11	1,571.62	-400.23	-633.25	-964.43	-111.93	-	1,116.90	-1,290.47	-559.04	
Prat. Branco	-295.73	-246.18	304.51	-745.02	-164.38	-288.40	-100.08	-1,061.08	2,592.66**	-	1,824.51**	-619.05	
Tox 503	-535.48	204.06	-27.51	531.72	137.86	-31.65	1,178.16	-1,016.83	249.91	-148.23	-	-1,157.43	
Epagri 108	455.46	-300.98	-528.56	598.17	-55.18	-858.20	516.61	132.61	430.36	151.71	-542.03	-	

NA: missing crosses. * and **: significant by the t-test, 5% and 1% probability, respectively.

Table 6. Estimates of the effects of varieties (\hat{v}), varietal heterosis in the generations F_2 (\hat{h}_{F_2}) and F_7 (\hat{h}_{F_7}), overall combining ability (\hat{g}_{F_2} and \hat{g}_{F_7}) and mean heterosis, for the trait days to flowering among twelve rice genitors.

GENITORS	\hat{v}	\hat{h}_{F_2}	\hat{h}_{F_7}	\hat{g}_{F_2}	\hat{g}_{F_7}
Canela Curta	3.87**	2.96**	1.74 ^{ns}	4.90	3.68
Araguaia	-17.95**	2.98**	3.96**	-5.99	-5.02
Bico Roxo	25.18**	-7.61**	-8.31**	4.98	4.28
Canarinho	22.59**	-4.48**	-7.13**	6.81	4.16
CT11632	-14.12**	1.64 ^{ns}	3.49**	-5.42	-3.57
Diamante	-3.79**	2.59**	4.48**	0.70	2.59
Irat 122	-1.45 ^{ns}	-5.46**	-2.69*	-6.19	-3.43
Lageado	25.35**	-6.69**	-8.51**	5.98	4.16
Maninjabu	-10.45**	8.28**	2.36*	3.06	-2.87
Pratinha Branco	-12.62**	2.17*	3.79**	-4.14	-2.52
Tox 503	-13.79**	0.95 ^{ns}	3.53**	-5.94	-3.37
Epagri 108	-2.79**	2.65**	3.28*	1.26	1.88
Mean Heterosis	-	-8.25**	-11.18**	-	-
General Mean	103.35				
CV(%)	3.54				

ns: not significant. * and **: significant by the t-test, 5% and 1% probability, respectively.

Table 7. Estimates of specific heterosis (\hat{s}_{ij}) of crosses in F_2 (above the diagonal) and specific combining ability (\hat{s}_{ij}) in F_7 (below the diagonal) for the trait days to flowering.

F_2	C. Curta	Araguaia	B. Roxo	Canarinho	CT11632	Diamante	Irat 122	Lageado	Maninjabu	P. Branco	Tox 503	Epagri 108
Canela Curta	-	7,87**	2,54	5,07**	-2,95	-2,56	-1,17	5,65**	-2,42	-2,47	-3,92**	-5,62**
Araguaia	2,09	-	-0,96	-3,28	2,93	-0,42	-1,28	-2,26	-6,28**	0,16	2,71	0,76
Bico Roxo	3,79	-1,00	-	NA	-3,28	-2,25	5,74	-2,00	-4,01**	4,29	-0,61**	3,77**
Canarinho	1,41	-1,88	NA	-	-0,61	3,77**	-2,58	NA	-0,33	-4,69**	-0,08	2,71
CT11632	-4,86**	0,84	-2,95	0,16	-	-1,49	1,14	-3,71**	-3,61*	6,33**	7,38**	-2,06
Diamante	8,48**	-1,31	NA	1,51	0,73	-	NA	4,60**	-0,47	-0,52	-0,72	-2,17
Irat 122	-6,00	-0,80	1,40	6,02**	-2,25	NA	-	-3,75**	0,41	1,86**	1,16	6,46**
Lageado	3,91	-3,38	NA	NA	2,66	1,01	-2,97	-	4,99**	-5,80	-3,00	3,29
Maninjabu	-1,56	-1,86	-1,65	-3,53	0,19	-3,46	1,05	-1,03	-	5,11**	-0,58	-2,53
Prat. Branco	-6,41**	2,79	-1,50	-3,88	4,84**	-2,31	1,20	-1,88	8,14**	-	4,11**	-2,08
Tox 503	6,44**	1,14	-0,15	-4,53**	4,69**	-6,96**	0,05	-0,53	4,49	-0,36**	-	-3,03
Epagri 108	-7,31**	3,39	2,09	4,71**	-4,06	2,28	2,30	2,21	-0,76	-0,61	-4,26	-

NA: missing crosses. **: significant by the t-test, 5% and 1% probability, respectively.

Table 8. Genetic distance (RW coefficient) among pairs of diallel parents. (*) - distances of the crosses that were not evaluated for grain yield and days to flowering. Mean RW: Mean of the distances between a genitor and all other genitors.

Mean RW	Genitors	Lageado	BicoRoxo	Canarinho	CanelaCurta	PratinhaBranco	Epagri 108	Diamante	Araguaia	Maninjavau	Irat 122	CT11632
0.72	Lageado	0.00										
0.91	BicoRoxo	0.66 (*)	0.00									
0.95	Canarinho	0.61 (*)	0.91 (*)	0.00								
0.84	CanelaCurta	0.59	0.82	1.00	0.00							
0.87	PratinhaBranco	0.60	0.77	1.00	0.45	0.00						
0.93	Epagri 108	0.84	1.00	1.00	1.00	0.00	0.00					
0.90	Diamante	0.79	0.89 (*)	0.89	1.00	1.00	0.80	0.00				
0.88	Araguaia	0.73	0.89	0.89	0.89	0.87	0.97	1.00	0.00			
0.91	Maninjavau	0.81	1.00	0.89	1.00	1.00	0.79	0.71	0.95	0.00		
0.93	Irat 122	0.79	0.89	1.00	1.00	1.00	0.91	0.91 (*)	0.94	0.90	0.00	
0.83	CT11632	0.60	0.87	0.87	0.74	0.83	0.97	0.94	0.71	0.94	0.87	0.00
0.92	Tox 503	0.75	1.00	1.00	0.77	1.00	1.00	0.96	0.88	1.00	1.00	0.79

Table 9. Mean grain yield (GY, in kg ha⁻¹) data and number of days to flowering (DTF) in RW genetic distance classes.

Class	RW in range*	Number of Crosses	GY average F2	GY average F7	DTF average F2	DTF average F7	DTF average F7
1	1 (1)	21	5602.1	6346.5	95.5	91.8	91.8
2	0.9-0.99 (0.94)	9	5650.3	5864.1	90.3	86.9	86.9
3	0.8-0.89 (0.87)	16	5901.4	5874.9	94.4	90.7	90.7
4	0.7-0.79 (0.76)	11	5053.6	5480.6	89.6	88.0	88.0
5	<0.7 (0.56)	4	4777.0	5326.3	91.1	89.3	89.3

* Between parenthesis: Average of mean RW within each class

Conclusions

On average, the Lageado and Maninjou genitors contributed to increase the grain yield of the combinations that they were part of, in both generations, while the CT11632 and Irat 122 genitors showed the opposite behavior. The genitors that showed the highest varietal effect and general combining ability for GY in generations F_2 and F_7 were Epagri 108, Diamante, Maninjou and Canela Curta; thus they can be recommended for the

breeding program as important genetic assets to obtain lines with greater grain yield potential. Regarding days to flowering, the genitors with the largest earliness estimates were Araguaia, CT11632, Irat 122, Pratinha Branco and Tox 503. The most suitable parameters for estimating grain yield in generation F_7 , from the data attained from F_2 , were the average grain yield, general combining ability and the RW genetic distance between pairs of genitors, obtained through genotyping using SSR markers.

References

- ABADIE, T.; CORDEIRO, C.M.T.; FONSECA, J.R.; ALVES, R.B.N.; BURLE, M.L.; BRONDANI, C.; RANGEL, P.H.N.; CASTRO, E.M.; SILVA, H.T.; FREIRE, M.S.; ZIMMERMANN, F.J.P.; MAGALHÃES, J.R. 2005. Construção de uma coleção nuclear de arroz para o Brasil. **Pesquisa Agropecuária Brasileira**, 40:129–136.
- ASFALIZA, R.; RAFII, M.Y.; SALEH, G.; OMAR, O.; PUTEH, A. 2012. Combining ability and heritability of selected rice varieties for grain quality traits. **Australian Journal of Crop Sciences**, 6:1718–1723.
- BERTAN, I.; CARVALHO, I.F.I.F.; OLIVEIRA, A.C.; SILVA, J.A.G.; BENIN, G.; HARTWIG, I.; SCHMIDT, D.M.; VALÉRIO, I.P.; FONSECA, D.R.; SILVEIRA, G. 2009. Efeitos da heterose e endogamia em caracteres de importância agrônômica em trigo. **Revista Ceres**, 56:753–763.
- BHULLAR, G.S.; GILL, K.S.; KHEHRA, A.S. 1979. Combining ability analysis over F_1 - F_5 generations in diallel crosses of bread wheat. **Theoretical and Applied Genetics**, 55:77–80.
- BORBA, T.C.O.; BRONDANI, R.P.V.; RANGEL, P.H.N.; BRONDANI, C. 2009. Microsatellite marker-mediated analysis of the Embrapa Rice Core Collection genetic diversity. **Genetica**, 137:293-304.
- BROWN, A.H.D. 1989. Core collections: A practical approach to genetic resources management. **Genome**, 31:818–824.
- BUENO, L.G.; VIANELLO, R.P.; RANGEL, P.N.; UTUMI, M.M.; CENTENO, A.C.; PEREIRA, J.A.; FRANCO, D.F.; NETO, F.M.; MENDONÇA, J.A.; COELHO, A.S.; OLIVEIRA, J.P.; BRONDANI, C. 2012. Adaptabilidade e estabilidade de acessos de uma coleção nuclear de arroz. **Pesquisa Agropecuária Brasileira**, 47:216–226.
- CHARLESWORTH, D.; BARTON, N.H.; CHARLESWORTH, B. 2017. The sources of adaptive variation. **Proceedings of the Royal Society B**, 284:20162864.
- CRUZ, C.D.; REGAZZI, A.J. 1994. **Modelos biométricos aplicados ao melhoramento genético**. UFV, Viçosa, 390p.
- CRUZ, C.D. 2013. GENES - a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum**, 35:271–276.
- DAN, Z.; HU, J.; ZHOU, W.; YAO, G.; ZHU, R.; ZHU, Y.; HUANG, W. 2016. Metabolic prediction of important agronomic traits in hybrid rice (*Oryza sativa* L.). **Scientific Reports**, 6:21732.

-
- DENNIS, E.S.; ELLIS, J.; GREEN, A.; LLEWELLYN, D.; MORELL, M.; TABE, L.; PEACOCK, W.J. 2008. Genetic contributions to agricultural sustainability. **Philosophical Transactions of the Royal Society B**, 363: 591–609.
- FASAHAT, P.; RAJABI, A.; RAD, J.M.; DERERA, J. 2016. Principles and utilization of combining ability in plant breeding. **Biometrics & Biostatistics International Journal**, 4:00085
- GAO, H.; JINA, M.; ZHENG, X.M.; CHEN, J.; YUAN, D.; XIN, Y.; WANG, W.; HUANG, D.; ZHANG, Z.; ZHANG, Z.; ZHOU, K.; SHENG, P.; MA, J.; MA, W.; DENG, H.; JIANG, L.; LIU, S.; WANG, H.; WUA, C.; YUAN, L.; WAN, J. 2014. *Days to heading 7*, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. **Proceedings of the National Academy of Sciences of the United States of America**, 111:16337–16342.
- GARDNER, C.O.; EBERHART, S.A. 1966. Analysis and interpretation of the variety cross diallel and related populations. **Biometrics**, 22:439–452.
- GÄRTNER, T.; STEINFATH, M.; ANDORF, S.; LISEC, J.; MEYER, R.C. 2009. Improved heterosis prediction by combining information on DNA and metabolic markers. **PLoS ONE**, 4:e5220.
- HUANG, M.; CHEN, L.Y.; CHEN, Z.Q. 2015. Diallel analysis of combining ability and heterosis for yield and yield components in rice by using positive loci. **Euphytica**, 205:37–50.
- HUANG, S., WANG, L.; LIU, L.; FU, Q.; ZHU, D. 2014. Nonchemical pest control in China rice: a review. **Agronomy for Sustainable Development**, 34:275–291
- JUNG, C.; MÜLLER, A.E. 2009. Flowering time control and applications in plant breeding. **Trends in Plant Science**, 14:563–573.
- LI, L.; LU, K.; CHEN, Z.; MU, T.; HU, Z.; LI, X. 2008. Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. **Genetics**, 180:1725–1742.
- LIU, C.; SONG, G.; ZHOU, Y.; QU, X.; GUO, Z.; LIU, Z.; JIANG, D.; YANG, D. 2015. *OsPRR37* and *Ghd7* are the major genes for general combining ability of DTH, PH and SPP in rice. **Scientific Reports**, 5:12803.
- LORENCETTI, C.; CARVALHO, F.I.F.; OLIVEIRA, A.C.; VALÉRIO, I.P.; BENIN, G.; ZIMMER, P.D.; VIEIRA, E.A. 2006. Distância genética e sua associação com heterose e desempenho de híbridos em aveia. **Pesquisa Agropecuária Brasileira**, 41:591–598.
- MORAIS JÚNIOR, O.P.; BRESEGHELLO, F.; DUARTE, J.B.; MORAIS, O.P.; RANGEL, P.H.N.; COELHO, A.S.G. 2017a. Effectiveness of recurrent selection in irrigated rice breeding. **Crop Science**, 57:3043–3058.
- MORAIS JÚNIOR, O.P.; PEREIRA, J.A.; MELO, P.G.S. GUIMARÃES, P.H.R.; MORAIS, O.P. 2017b. Gene action and combining ability for certain agronomic traits in red rice lines and commercial cultivars. **Crop Science**, 57:1295–1307.
- NAJEEB, S.; AHANGAR, A.; DAR, S.H. 2013. An analysis of hybrid sterility in rice (*Oryza sativa* L.) using genetically diverse germplasm under temperate ecosystem. **African Journal Agricultural Research**, 8:3820–3827.
- PEREIRA, J.A.; MORAIS, P.O.; BRESEGHELLO, F. 2008. Análise da heterose de cruzamentos entre variedades de arroz-vermelho. **Pesquisa Agropecuária Brasileira**, 43:1135–1142.

-
- R CORE TEAM. 2013. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0.
- SEARLE, S.R. 1971. **Linear models**. John Wiley & Sons, New York, 532p.
- SHAHRIAR, M.D.; ROBIN, A.H.K.; HOQUE, A. 2014. Diversity assessment of yield, yield contributing traits, and earliness of advanced T-aman rice (*Oryza sativa* L.) lines. **Journal of Bioscience and Agriculture Research**, 1:102–112.
- SILVA, S.A.G.; MORAIS, O.P.; RAVA, C.A.; COSTA, J.G.C. 2000. Método generalizado de análise de dialelos desbalanceados. **Pesquisa Agropecuária Brasileira**, 35:1999–2005.
- SONI, S.K.; TIWARI, S.; NEWMAN, J.T.; DOSSOU-AMINON, I.; SUNDARAM, R.M. 2018. Prediction of hybrid performance in crop plants: molecular and recent approaches. **International Journal of Current Microbiology and Applied Sciences**, 7:98–108.
- SPRAGUE, G.F.; TATUM, L.A. 1942. General versus specific combining ability in single crosses of corn. **Journal of American Society of Agronomy**, 34:923–932.
- WRIGHT, S. 1978. **Evolution and the genetics of populations. Variability within and among natural populations**. University of Chicago Press, Chicago, 590p.
- YAN, W.; RUTGER, J.N.; BRYANT, R.J.; BOCKELMAN, H.E.; FJELLSTROM, R.G.; CHEN, M.H.; TAI, T.H.; MCCLUNG, A.M. 2007. Development and evaluation of a core subset of the USDA rice (*Oryza sativa* L.) germplasm collection. **Crop Science**, 47:869–878.
- ZHANG, P.; LI, J.; LI, X.; LIU, X.; ZHAO, X.; LU, Y. 2011. Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers. **PLoS ONE**, 6:e27565.