The Relation Between InDel Genetic Distance and Heterosis in Rice

Xudong Zhu[1, *], Xueli Tang[2], Ju Zhao[1], Feng Wang[1], Qingming Zhou[1], Zhonghua Zhou[1, *]

[1]College of Agronomy, Hunan Agricultural University, Changsha, P.R. China

[2]Agricultural and Rural Bureau, Hengshan County, Hengyang, P.R. China

Email address:
zxdzhaoju@126.com (Xudong Zhu), 55960548@qq.com (Zhonghua Zhou)
[*]Corresponding author

Abstract: Heterosis is a common phenomenon in plant species, which has been extensively utilized in different crops. Genetic distance is one of important basis for selecting parents in hybrid breeding. There is a relation between the genetic distance and heterosis and hybrid rice combinations usually show a stronger heterosis with larger genetic distance between parental lines. In this study, 160 hybrid rice combinations were made by incomplete duallel cross using 10 sterile lines and 16 self-bred restorer lines. Moreover, the indica-japonica gene frequencies of 26 lines were identified by 20 pairs of the InDel markers. The relationships between the InDel genetic distance and the phenotype heterosis were analyzed. Here, the results showed that the InDel genetic distance between parents could be used to predict phenotype value of effective panicle per plant and control heterosis of yield per plant and seed setting rate of the tested rice combinations. The optimal distance between the parents was in the interval 0.1-0.55 for the control heterosis of several main yield traits including effective panicle per plant, filled grains per panicle and yield per plant of combinations. The present result will provide a useful reference for hybrid rice parents selection in breeding.

Keywords: Hybrid Rice, Parents, Indel Marker, Genetic Distance, Heterosis

1. Introduction

Rice is one of the most important food crops in the world, which supports the requirement of staple food for almost half of the world's population [16]. Heterosis is a universal phenomenon in which hybrids exhibit superior performance relative to parental phenotypes [27]. Hybrid breeding promotes to increase rice yield by use of heterosis, which lead to about 40% increase in rice yield per area during the past 30 years [1]. Since the three-line hybrid rice has been successfully developed in 1970s, the extensive research studies have been conducted on the utilization of heterosis [26, 27]. It is suggested that the heterosis in rice is closely related to the genetic distance between parents [7]. The heterosis is significantly positively correlated with the genetic distance, and the coincidence rate of heterosis prediction is up to 80% [24]. The genetic distance based on SSR markers was positively correlated with special combining ability but not significantly, while was significantly negatively correlated with the mid-parent heterosis [19]. A study of 3 maintainer (B) lines and 13 restorer (R) lines using 193 pairs SSR markers was concluded that the genetic distance was significantly positively correlated with panicles number, grain length and length-width ratio, but was significantly negatively correlated with grain width, yield per plant, seed setting rate, mid-parent heterosis, control heterosis and special combining ability [21]. The heterosis was predicted based on the correlation between the coefficient of marker polymorphism and the control heterosis of different traits for 40 F_1 hybrids, which showed that the coefficient of the marker polymorphism was significantly positively correlated with the yield per plant, effective tillering and weight per panicle [13]. The genetic distance based on SNPs was significantly correlated with the mid-parent heterosis of 1000-grain weight in wheat, but not significantly correlated with the high-parent heterosis, and weakly correlated with the mid-parent heterosis and super-

parent heterosis of effective panicle per plant, filled grain per panicl and the yield [12].

Furthermore, 21 pairs of Insertion/Deletion (InDel) primers were used to analyze a self-bred *japonica* Zhegeng4A and 22 *indica* restorer lines, to predict the heterosis of the combinations crossed by them. The results showed that the InDel genetic distance was significantly positively correlated with total grains per panicle, filled grains per panicle and the yield per plant, but not significantly correlated with seed setting rate and 1000-grain weight, which indicate that the InDel markers could be used for *indica-japonica* attribute identification and heterosis prediction [5, 30]. The *indica-japonica* recombinant inbred lines (Akihikari×Qishanzhan, RIL) and a *jponica* photo-thermo-sensitive male sterile line (GB028S) were used as materials to analyze the relationship between the *indica-japonica* components and heterosis. The genetic basis of these materials was analyzed by the methods of Cheng's index and the SSR markers. The results showed that the genetic distance of the parents had no significant relationship with the F_1 yield related traits and heterosis. However, the partial *japonica* coefficient of RIL had a significant or extremely significant curvilinear relationship with the F_1 yield and its heterosis, and when partial *japonica* coefficient of RIL was 0.55-0.65, the F_1 had the potential to perform high yield and heterosis [25].

In recent years, some major breakthroughs have emerged for researching of rice heterosis. The phenotypes of 10,074 F2 lines from 17 representative hybrid rice crosses were generated, sequenced and recorded, and classified modern hybrid rice varieties into three groups. found within each group, a small number of genomic loci from female parents explained a large proportion of the yield advantage of hybrids over their male parents. For some of these loci, they found support for partial dominance of heterozygous locus for yield-related traits and better-parent heterosis for overall performance when all of the grain-yield traits were considered together. These results inform on the genomic architecture of heterosis and rice hybrid breeding [6]. The apomixis system of rice was established by gene editing technology, and the hybrid rice seeds were successfully cloned, which realized the stable inheritance of hybrid rice traits to the next generation for the first time [8]. An important transcription factor OsDREB1C in rice was found. This transcription factor can improve the photosynthetic efficiency and nitrogen use efficiency of rice at the same time, and significantly improve crop yield [22].

In this study, we employed 10 sterile lines as female parents and 16 restorer lines as male parents to generate 160 hybrid rice combinations crossed by them according to incomplete dual cross. The relation between the InDel genetic distance of the parents and heterosis was explored to identify the suitable genetic distance interval of parents for selecting optimal combinations with strong heterosis. The result will provide a reference for parents selecting of hybrid rice breeding.

2. Materials and Methods

2.1. Material

In the present study, 10 sterile lines as female parents were labeled as A1, A2, A3, A4, A5, A6, A7, A8, A9 and A10, respectively (Table A1). Besides, 16 self-bred male parents were labeled as R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11, R12, R13, R14, R15, and R16, respectively (Table A2). Then 160 hybrid rice combinations were produced according to incomplete dual cross. Y Liangyou 1 was used as the control hybrid rice.

2.2. Methods

2.2.1. Experimental Location

The field experiment was designed as a comparative experiment in May 2017 at Ren Ziqiao Group, Chang'an Village, Huangxing Town, Changsha, China. All materials were planted in ten rows with ten plants each with a spacing of 16.7 cm ×16.7 cm. Total 50 plants were planted for each combination and the distance between combinations was 33 cm. Field management was accorded to the production requirements of conventional hybrid rice planting.

2.2.2. Sampling

To determine the growth performance of all hybrid rice combinations, we selected two spots for each material in the middle of the plot, and 10 plants of each plot were sampled to count effective panicle of unit area. According to the effective panicle of each spot to sample one plant, the traits of plant height, panicle length, total grain of per panicle, filled grains per panicle, seeding rate were also recorded. Randomly count out three 1000 grains, weigh them respectively, the average value was 1000 grains weight. Here, 50 plants were harvested at each plot and dried, then weighed and converted to the actual yield of unit area.

2.2.3. DNA Extraction

At rice tillering stage, 5 g of young leaves were sampled for DNA extraction according to the CTAB method [11], and the operation steps were carried out according to the method by Song et al [17]. The extracted DNA were dissolved into 100~200 μL by sterilized water and stored in a refrigerator at -20°C. The laboratory experiment was performed in the lab of Ramie Research Institute, Hunan Agricultural University, China.

2.2.4. PCR Amplification and Electrophoresis Detection

In this experiment, 25 pairs of InDel primers were selected according to the literature of Sahu and Cai [2, 16]. Total 20 pairs of InDel primers from 12 chromosomes of rice were selected through previous experiments to identify the *indica* or *japonica* frequencies of parental lines. The InDel primers used in this experiment were synthesized by Hunan Qingke Biotechnology Co., China. The PCR reaction system was consisted of 1μL DNA template, 1μL dNTPs, 1μL of 0.2μmol/L primers, 5μL of 2×Taq Mix and 2μL ddH$_2$O.

The PCR amplification reaction was as follows: pre-denaturation for 5mins at 95°C, denaturation for 35s at 95°C,

primer annealing at 55°C for 35s, elongation at 72°C for 1min, a total of 38 cycles, and finally elongation at 72°C for 7mins and preservation at 4°C. The PCR amplification products were separated by electrophoresis on 2% of agarose gel at a voltage of 130V, and electrophoresis gel images were taken using Bio-Rad gel imaging apparatus.

2.2.5. Analysis of Indica-Japonica Gene Frequency of the Tested Materials

An *indica* rice variety 9311 and a *japonica* variety Nipponbare were used as the control. At each InDel locus, samples with the same bands as 9311 were recorded as homozygous *indica* genotypes (II), or homozygous *japonica* genotypes as the same bands with Nipponbare (JJ), or *indica-japonica* heterozygous genotypes with both 9311 and Nipponbare bands (IJ). The average allele frequency (F) of *indica* and *japonica* was calculated at all InDel loci, and the *indica-japonica* gene frequency of the tested materials is calculated by Wang's method [21].

According to the frequency of *indica* or *japonica* genotypes at multiple InDel loci, the *indica* and *japonica* properties of the tested materials were determined according to Lu et al [10]. Moreover, the bands amplified by InDel primers with the same position of 9311, Nipponbare or heterozygous were counted as 1, 2 and 3, respectively. The genetic difference between each two materials is calculated by Nei's method for genetic similarity coefficient (GS) and genetic distance (GD) [15].

2.2.6. Date Preparation and Analysis

Microsoft Excel 2010 was used for calculating super-male parental heterosis and control heterosis, while DPS [18]7.5 software was used for correlation analysis between genetic distance and heterosis. The formula for super male heterosis and control heterosis were:

$$\text{Super male parent heterosis} = \frac{(Hp - F1)}{Hp} \times 100$$

Hp represents the phenotypic value of the male parent and F1 represents the phenotypic value of F1.

$$\text{Control heterosis} = \frac{(CK - F1)}{CK} \times 100$$

CK represents the phenotype value of the control, and F1 represents the phenotypic value of F1.

3. Results and Discussion

3.1. Results

3.1.1. Indica-Japonica Gene Frequencies and InDel Genetic Distance of Tested Rice Parents

Among these rice parents, the male parents of typical *indica* rice were R4, R5, R8, R11, R12 and R14, while *indica* rice were R1, R2, R6, R7, R9, R10, R13, R14 and R16 and the partial *indica* was R3 (Table 1). The tested female parents were distributed from *indica* type to intermediate type. Among them, typical *indica* rice female parent was A9, while *indica* rice female parents included A2, A4, A5, A7, A8, A9 and A10. The partial *indica* rice was A6 and intermediate types were A1 and A3 (Table 1). In short, the most of the tested parents were *indica* type or partial *indica* type.

***Table 1.** Indica-japonica gene frequencies of the tested rice parents.*

NO	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13
Indica gene frequency	0.800	0.775	0.700	0.925	1.000	0.900	0.900	1.000	0.900	0.850	1.000	1.000	0.850
Japonica gene frequency	0.200	0.175	0.300	0.075	0.000	0.100	0.100	0.000	0.100	0.150	0.000	0.000	0.150
NO	R14	R15	R16	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
Indica gene frequency	1.000	0.850	0.875	0.550	0.900	0.450	0.850	0.800	0.750	0.900	0.900	0.950	0.850
Japonica gene frequency	0.000	0.150	0.125	0.450	0.100	0.550	0.150	0.150	0.200	0.100	0.100	0.050	0.150

On the other hand, the InDel genetic distance between each two parents for 160 crossed combinations was ranged from 0.05 to 0.7 (Table 2). From the above analysis, it was found that the InDel genetic distance of the most combinations was between 0.1 and 0.3, accounting for 66.9% of the tested combinations.

***Table 2.** The InDel genetic distance for 26 rice parents.*

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
R1	0.45	0.10	0.45	0.25	0.20	0.25	0.10	0.30	0.15	0.15
R2	0.45	0.25	0.55	0.30	0.15	0.40	0.35	0.30	0.30	0.30
R3	0.45	0.40	0.55	0.45	0.35	0.45	0.40	0.40	0.35	0.45
R4	0.40	0.20	0.50	0.25	0.30	0.30	0.20	0.15	0.15	0.25
R5	0.45	0.10	0.55	0.15	0.20	0.25	0.10	0.10	0.05	0.15
R6	0.45	0.20	0.55	0.15	0.20	0.35	0.20	0.20	0.15	0.25
R7	0.55	0.20	0.55	0.15	0.20	0.35	0.20	0.20	0.15	0.25
R8	0.45	0.10	0.55	0.15	0.20	0.25	0.10	0.10	0.05	0.15
R9	0.35	0.20	0.45	0.25	0.30	0.35	0.20	0.20	0.15	0.25
R10	0.50	0.15	0.50	0.20	0.25	0.20	0.05	0.25	0.10	0.20
R11	0.45	0.10	0.55	0.15	0.20	0.25	0.10	0.10	0.05	0.15
R12	0.45	0.10	0.55	0.15	0.20	0.25	0.10	0.10	0.05	0.15
R13	0.30	0.25	0.40	0.20	0.25	0.40	0.25	0.25	0.20	0.30
R14	0.45	0.10	0.55	0.15	0.20	0.25	0.10	0.10	0.05	0.15

NO	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
R15	0.60	0.25	0.70	0.30	0.25	0.40	0.25	0.25	0.20	0.30
R16	0.45	0.25	0.55	0.20	0.25	0.35	0.25	0.20	0.20	0.30

3.1.2. The Correlation of Indica-Japonica Gene Frequency of Parents with Phenotypes and Heterosis

The result showed that the *indica* gene frequency of male parent was not significantly correlated with all phenotypic values (Table 3). However, the *indica* gene frequency of male parent was significantly or extremely significantly positively correlated with the super-parent heterosis of effective panicle per plant and seed setting rate traits, but was significantly negatively correlated with the super-parent heterosis of total grain number per panicle, 1000-grain weight and yield per plant traits (Table 3). Besides, the *indica* gene frequency of male parent was only significantly positively correlation with the control heterosis of total grain number per panicle traits (Table 3).

For the *japonica* gene frequency of male parent, it was significantly positive correlation with super paternal heterosis of total number of grain per panicle, seeding rate and yield per plant, but was negatively correlated with super paternal heterosis of effective panicle per plant (Table 3). The male paternal *japonica* gene frequency was significantly negatively correlated with the control heterosis of total grain per panicle (Table 3). Therefore, the male parental *indica* gene frequency could only predicted the phenotypic value of effective panicle per plant and the control heterosis of total grains per panicle, while the *japonica* gene frequency of male parent could only predict the super parental heterosis of total grain number per panicle and yield characters per plant.

Table 3. Correlations between indica-japonica gene frequency in male parent and phenotypic value, super paternal heterosis, control heterosis.

Male parent	Item	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
indica gene frequency	Phenotypic value	-0.0248	0.0316	0.1235	0.1410	0.1070	-0.1255
	Super paternal heterosis	0.2820**	-0.1864*	-0.1018	0.1698*	-0.1668*	-0.1669*
	Control heterosis	0.0213	0.0569	0.1641*	0.1371	0.0968	-0.1289
japonica gene frequency	Phenotypic value	-0.0010	-0.0222	-0.1074	-0.1356	-0.0900	0.1327
	Super paternal heterosis	-0.2737**	0.2009*	0.1163	-0.1996*	0.1182	0.2022*
	Control heterosis	-0.0180	-0.0572	-0.1565*	-0.1282	-0.0936	0.1309

Note: Critical values of correlation coefficient, a=0.05, r=0.1552, a=0.01, r=0.2031, the same as below.

We also analyzed the correlation of the *indica-japonica* gene frequency of female parent with phenotypes and heterosis. The *indica* gene frequency of female parent was significantly negatively correlated with the effective panicle per plant traits, but not with other traits (Table 4). Moreover, the *indica* gene frequency of female parent was also negatively correlated with the super paternal heterosis of effective panicle per plant traits. The *indica* gene frequency in female parent was significantly or extremely significantly negatively correlated with the control heterosis of effective panicle per plant, seed setting rate and yield per plant traits, while it was significantly positively correlated with the control heterosis of total grains per panicle trait (Table 4).

Based on the results of Table 4, it was suggested that the control heterosis of effective panicle could be predicted by the *indica* gene frequency of female parent, while the phenotypic value, super paternal and control heterosis of effective panicle per plant, control heterosis of seed setting rate and yield per plant traits could be predicted by the *japonica* gene frequency in female parent.

Table 4. Correlations between the indica-japonica gene frequency of female parent and phenotypic value, super paternal heterosis control heterosis.

female parent	Item	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
indica gene frequency	Phenotypic value	-0.2657**	0.0864	0.0835	-0.0060	0.0828	-0.0630
	Super paternal heterosis	-0.1596*	0.0398	0.0563	0.0177	0.0176	-0.0494
	Control heterosis	-0.1593*	0.2752**	-0.0574	-0.3759**	-0.0215	-0.2481**
japonica gene frequency	Phenotypic value	0.2641**	-0.0817	-0.0715	0.0182	-0.0721	0.0885
	Super paternal heterosis	0.1601*	-0.0369	-0.0502	-0.0127	-0.0563	0.0698
	Control heterosis	0.1597*	-0.2822**	0.0618	0.3890**	0.0288	0.2728**

3.1.3. The Correlation Between the InDel Genetic Distance of Parents and Phenolypes and Heterosis

We found that the InDel genetic distance was significantly positively correlated with the phenotypic value of effective panicle traits per plant (Table 5). The InDel genetic distance was extremely significant positive correlation with control heterosis of yield per plant and seed setting rate traits, but was significantly negatively correlated with control heterosis of the total grain number per panicle traits (Table 5). Based on the above analysis, the InDel genetic distance of parents could predict the phenotypic value of effective ear traits and the control heterosis of yield per plant and seed setting rate traits.

Table 5. *Correlations between phenotypic value, super male parents, control heterosis with InDel genetic distance of parents.*

Item	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 grain weight	Yield per plant
Phenotypic value	0.2316**	-0.1070	-0.1229	-0.0190	-0.0024	0.1278
Super paternal heterosis	-0.0420	0.0409	-0.0181	-0.0726	0.0957	0.0910
Control heterosis	0.1335	-0.2362**	-0.0116	0.2622**	0.0274	0.2522**

3.1.4. Analysis of the Optimum Interval of Male Paternal Indica-Japonica Gene Frequency for Phenotypes and Heterosis

The results showed that the phenotypic values of the yield related traits including effective panicle per plant, filled grains per panicle, 1000-grain weight and yield per plant traits of hybrid rice combinations were significantly positively correlated with the *indica* gene frequency with the interval of 0.85-0.9. There was a significant positive correlation between the super paternal heterosis of the traits of effective per plant, filled grains per panicle, 1000-grain weight and yield per plant and the *indica* gene frequency when it ranged from 0.775 to 0.90. When the *indica* gene frequency was 0.7-0.9, it had a significant positive correlation with the control heterosis of the traits of effective per plant, filled grains per panicle, 1000-grain weight and yield per plant traits of combinations (Table 6). Therefore, when the *indica* gene frequency of male parent was in the interval of 0.7-0.9, the combinations crossed by the tested parents had strong control heterosis.

Table 6. *Analysis of the optimum interval of male paternal indica gene frequency for phenotypic value, super parents heterosis and control heterosis.*

Item		Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
Phenotypic value	Interval	0.85-1	0.85-0.9	0.7-0.9	0.7-1	0.9-1	0.85-0.925
	Correlation significance	0.1979*	0.3511*	0.2103*	0.1621*	0.2857*	0.2365*
	r0.05	0.1975	0.3494	0.2096	0.1567	0.2335	0.2303
Super paternal heterosis	Interval	0.7-1	0.8-1	0.775-1	0.7-1	0.7-0.9	0.65-0.925
	Correlation significance	0.2820**	0.2117*	0.1689*	0.1698*	0.2631*	0.2227*
	r0.05	0.1552	0.1678	0.1642	0.1552	0.2461	0.2199
Control heterosis	Interval	0.7-0.9	0.775-0.85	0.7-1	0.7-1	0.7-1	0.7-1
	Correlation significance	0.2589*	0.3162*	0.1756*	0.1612*	0.1656*	0.1801*
	r0.05	0.2423	0.3081	0.1552	0.1562	0.1654	0.1710

For *japonica* gene frequency, the phenotypic values of the yield related traits of effective panicle per plant, filled grains per panicle, 1000-grain weight and yield per plant traits of combinations that the male parents crossed with test female parents were significantly positively correlated with the paternal *japonica* gene frequency in the interval of 0.1-0.125. The super paternal heterosis of the traits of effective panicle per plant, filled grains per panicle, 1000-grain weight and yield per plant were significantly positively correlated with the paternal *japonica* gene frequency with the interval of 0-0.175. The control heterosis of the main yield traits of effective panicle per plant, filled grains per panicle, 1000-grain weight and yield per plant traits of hybrid rice combinations were significantly positively correlated with the paternal *japonica* gene frequency with the interval of 0.125-0.15 (Table 7). Therefore, the male parent with the *indica* gene frequency of 0.75-0.85 and the female parent with the *japonica* gene frequency of 0.125-0.15, should be chose for combinations selection with better performance and stronger control heterosis.

Table 7. *Analysis of the optimum interval of paternal japonica gene frequency for phenotypic value, super parents heterosis and control heterosis.*

Item		Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
Phenotypic value	Interval	0-0.3	0-0.15	0.075-0.175	0.075-0.125	0.1-0.125	0-0.3
	Correlation significance	0.1926*	0.1933*	0.2161*	0.1836	0.5645*	0.1599*
	r0.05	0.2269	0.1918	0.2133	0.3440	0.4438	0.1562
Super paternal heterosis	Interval	0-0.175	0-0.3	0-0.3	0.075-0.3	0-0.3	0-0.3
	Correlation significance	0.1937*	0.1799*	0.1934*	0.1965*	0.1602*	0.1771*
	r0.05	0.1918	0.1552	0.1793	0.1927	0.1552	0.1552
Control heterosis	Interval	0-0.3	0.125-0.15	0.125-0.15	0.1-0.175	0.1-0.15	0-0.3
	Correlation significance	0.1880*	0.3307*	0.6420*	0.2680*	0.3748*	0.1692*
	r0.05	0.1918	0.3160	0.5529	0.2461	0.361	0.1552

3.1.5. Analysis of the Optimum Interval of Female Paternal Indica-Japonica Gene Frequency for Phenotypes and Heterosis

For the female parent with *indica* genes frequency between 0.8 and 0.9, a significant positive correlation was found between it and the phenotypic values of the main yield traits including per plant, grain number per panicle, 1000 grain weight and yield per plant. When the frequency of *indica* genes of female parent was 0.85-0.9, the super male parents

heterosis of the main yield traits including effective panicle per plant, filled grain per panicle, 1000 grain weight and yield per plant of the combination crossed with male parent was significantly positive correlation with it. The control heterosis of the main yield traits include effective panicle per plant, filled grains per panicle, 1000-grain weight and yield per plant traits of combinations that the female parents crossed with tested male parents were significantly positively correlated with the maternal *indica* gene frequency in the interval of 0.85-0.9 (Table 8).

Table 8. *Analysis on the optimum Interval of female paternal indica gene frequency for phenotypic value, super parents heterosis and control heterosis.*

Item	Traits	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
Phenotypic value	Interval	0.8-0.9	0.45-0.9	0.75-0.95	0.55-0.95	0.45-0.9	0.45-0.9
	Correlation significance	0.2283*	0.1997*	0.1854*	0.1947*	0.1814*	0.1827*
	r0.05	0.2242	0.1996	0.1736	0.1716	0.1714	0.1809
Super paternal heterosis	Interval	0.85-0.9	0.45-0.9	0.75-0.95	0.55-0.75	0.8-0.9	0.55-0.95
	Correlation significance	0.2562*	0.1979*	0.1821*	0.5555*	0.2262*	0.1880*
	r0.05	0.2441	0.1975	0.1801	0.5529	0.2199	0.1729
Control heterosis	Interval	0.80-0.9	0.45-0.95	0.85-0.95	0.8-0.95	0.75-0.9	0.75-0.95
	Correlation significance	0.2803*	0.2752**	0.2406*	0.2164*	0.2079*	0.2073*
	r0.05	0.2542	0.1552	0.2369	0.1986	0.2028	0.1809

For *japonica* gene frequency, the phenotypic values of the main yield traits of effective panicle per plant, filled grains per panicle and yield per plant traits of hybrid rice combinations were significantly positively correlated with the maternal *japonica* gene frequency in the interval of 0.2-0.5. The super paternal heterosis of main yield traits include effective panicle per plant, filled grains per panicle and yield per plant traits of combinations that the tested female parents crossed with tested male parents were significantly positively correlated with the female parents *japonica* gene frequency when it was in the interval of 0.15-0.55. When the *japonica* gene frequency of male parent was in the interval of 0.15-0.55, the control heterosis of the main yield traits include effective panicle per plant, filled grains per panicle and yield per plant traits of combinations that the tested female parents crossed with test male parents were significantly positively correlated with it (Table 9).

Based on the above analysis, we found that the combinations had stronger control heterosis when the *indica* gene frequency of the female parent was in the range of 0.85-0.9 and the *japonica* gene frequency of the male parent was between 0. 05 and 0.55.

Table 9. *Analysis on the optimum Interval of female paternal japonica gene frequency for phenotypic value, super parents heterosis and control heterosis.*

Item	Traits	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	Seeding rate	1000 graing weight	Yield per plant
Phenotypic value	Interval	0.05-0.55	0.2-0.55	0.2-0.5	0.1-0.55	0.1-0.2	0.1-0.55
	Correlation significance	0.2641**	0.4875*	0.4116*	0.2096*	0.2554*	0.3364**
	r0.05	0.1552	0.4683	0.3961	0.1848	0.2146	0.1764
Super paternal heterosis	Interval	0.05-0.55	0.05-0.45	0.15-0.55	0.1-0.55	0.1-0.2	0.1-0.55
	Correlation significance	0.1601*	0.1867*	0.2905*	0.1450	0.2346*	0.2352[**]
	r0.05	0.1552	0.1729	0.2816	0.2146	0.2172	0.1764
Control heterosis	Interval	0.05-0.55	0.05-0.2	0.05-0.55	0.05-0.55	0.05-0.2	0.05-0.55
	Correlation significance	0.1597*	0.2194*	0.1909*	0.3890[**]	0.1877*	0.2728[**]
	r0.05	0.1552	0.1824	0.1631	0.1552	0.1779	0.1552

3.1.6. Analysis of the Optimal Interval for InDel Genetic Distance of Parents for Phenotypes and Heterosis

The phenotypic values of the main yield traits including effective panicle per plant, filled grains per panicle and yield per plant traits of hybrid rice combinations were significantly positively correlated with the InDel genetic distance of the tested parents when they was in the interval of 0.35-0.45. When the InDel genetic distance of the parents was between 0.35 and 0.45, the super paternal heterosis of the main yield traits including effective panicle per plant, filled grains per panicle and yield per plant traits of combinations were significantly positively correlated with it. The control heterosis of the main yield traits of effective panicle per plant, filled grains per panicle and yield per plant of these combinations were significantly positively correlated with the InDel genetic distance of the parents in the interval of 0.25-0.4. When the indel genetic distance between the parents was in the range of 0.35-0.4, the phenotypic values and control heterosis of the yield and yield composing factors including effective panicles per plant, filled grains per panicle, 1000 grain weight and yield per plant were significantly positively correlated with the indel genetic distance (Table 10). Therefore, the optimal InDel genetic distance between parents should be 0.1-0.55 for hybrid rice combinations.

Table 10. *Analysis of the optimal interval for InDel genetic distance of parents for phenotypic value, super male parents and control heterosis.*

Item	Traits	Effective panicle per plant	Total grain of per panicle	Filled grains per panicle	seeding rate	1000 grain weight	Yield per plant
Phenotypic value	Interval	0.05-0.7	0.3-0.55	0.35-0.55	0.35-0.55	0.2-0.45	0.05-0.55
	Correlation significance	0.2316**	0.2904*	0.3106*	0.3636*	0.2356*	0.1635*
	R0.05	0.1552	0.2681	0.3081	0.3081	0.2159	0.1614
Super paternal heterosis	Interval	0.25-0.7	0.15-0.5	0.15-0.55	0.35-0.55	0.05-0.45	0.05-0.55
	Correlation significance	0.2566**	0.1949*	0.1856*	0.3720*	0.1878*	0.1650*
	R0.05	0.2387	0.1865	0.184	0.3610	0.1703	0.1598
Control heterosis	Interval	0.1-0.7	0.05-0.15	0.25-0.55	0.05-0.7	0.05-0.4	0.05-0.7
	Correlation significance	0.1652*	0.2597	2349*	0.2622**	0.1778*	0.2522**
	R0.05	0.1582	r0.1=0.251	0.2319	0.1552	0.1736	0.1552

3.2. Discussion

Since the inception of hybrid rice, the relations between heterosis and parental genetic distance have been extensively studied. It is believed that there is a close relationship between the heterosis and the parental genetic distance in rice [23]. Therefore, it is reliable to predict the heterosis of the yield per plant by using parental InDel genetic distance [30]. The InDel genetic distance was considered that it had a good ability for heterosis prediction among different types of varieties, which could be used to predict the heterosis of weight of per panicle [13, 20]. Some studies suggested that there was a significant linear relationship between yield heterosis and SSR marker genetic distance [14]. The RIL partial *japonica* coefficient had a significant or extremely significant conic curve relationship with the F_1 yield and its heterosis [25]. The molecular markers could be used as a useful tool for analyzing heterotic groups and genetic diversity, but the degree of correlation is not enough to predict the heterosis [29]. The relationship between *Mahalanobis* distance of parents [9], SSR markers distance [28] and the heterosis was not significant, so further studies are needed for the heterosis prediction. There was no significant correlation between RFLP genetic distance of maize parents and the heterosis of F_1 [3, 4]. Although it is generally believed that the genetic distance between the parents is closely related to heterosis, some studies have given different results. So it is needed to further study with more extensive materials for predicting the heterosis with genetic distance.

Based on the SSR genetic distance, the test combinations had been divided into 4 groups with genetic distance within 40-50, 50-60, 60-70 and 70-80. It was found that when the genetic distance within the group was 70-80, it was positively correlated with the yield performance, and the correlation coefficient was the largest, r =0.32 [19]. The F_1 had the potential to obtain higher yield and hybid heterosis when the RIL partial *japonica* coefficient was in the range of 0.55-0.65 [25]. These studies remind us that the correlation between the genetic distance of parents and the heterosis is not necessarily significant, but further analysis can find a genetic distance interval in which the correlation between the genetic distance of parents and heterosis is significant.

Yu considered that it was closely correlation for the yield and yield component factors with RIL partial *japonica* coefficient on male parental RIL of Chr.8, Chr.11, Chr.12 [25]. This study also found that *japonica* genes on the chromosome of parents were significantly or extremely significantly correlated with control heterosis of the yield per plant, but the location of these *japonica* genes on chromosome were different, which may be caused by the fact that the parents used in this study were *indica* type or partial *indica* type.

The possibility of obtaining a stronger heterosis combination is closely related to the selection of parents. Previous studies mainly focused on the genetic distance of parents. But there are few reports on what kind of female parents should be selected for different types of male parents, or what kind of male parents should be selected for different types of female parents. In this study, the *indica* gene frequency of the tested male parents could only predict the phenotypic value of effective panicle trait and the control heterosis of the total grains per panicle trait. The *japonica* gene frequency could only predict the super male parent heterosis of the total grains per panicle and the yield per plant traits for the tested combinations. The control heterosis of effective panicle per plant of the tested combinations could be predicted by the *indica* gene frequency of the female plants, and the phenotypic value, super male parents and control heterosis of effective panicle per plant, and control heterosis of seed setting rate and yield per plant traits of the tested combination could be predicted by *japonica* gene frequency of female parents. The InDel genetic distance between parents could be used to predict phenotype value of effective ear phenotype and control heterosis of yield per plant and seed setting rate. The optimal distance between the parents was in the interval 0.1-0.55 for the control heterosis of several main yield traits including effective panicle per plant, filled grains per panicle and yield per plant of combinations.

4. Conclusion

In this study, 10 sterile lines as female parents and 16 restorer lines as male parents were used to generate 160 hybrid rice combinations. Moreover, the relation between the InDel genetic distance of the 26 parents and the heterosis of the yield related traits was explored to find identify the suitable InDel genetic distance interval of parents for parents selecting optimal combinations with strongest heterosis. The

results showed that the InDel genetic distance between parents could be used to predict phenotype value of effective panicle per plant and control heterosis of yield per plant and seed setting rate. The optimal distance between the parents was in the interval 0.1-0.55 for the control heterosis of several main yield traits including effective panicle per plant, filled grains per panicle and yield per plant of combinations. This information will provide a useful reference for the parents selecting in hybrid rice breeding. It will be more prospective for producing the new hybrid rice combinations and improving the screening probability for strong heterosis combinations.

Author Contributions

Z. X., Z. Z. and Z. Q. conceived and designed the study. Z. X., Z. J., T. X., and W. F. performed the experiments. Z. Z. and Z. X. wrote the paper.

Acknowledgements

Appendix

Table A1. The information of female parents of rice.

Sterile line		The source of female parent	
A1	S1-3	Y58S/Erect panicle japonica	Self-bred
A2	15HS1	Y58S/ Erect panicle japonica	Self-bred
A3	S2-13	Y58S/ Erect panicle japonica	Self-bred
A4	Chuang 106A	Chuang xiang 28A/ Chuang 106B	Bred by Sichuan Academy of Agricultural Sciences, Rice and Sorghum Institute
A5	Chuang nogn 1A	D23A/Chuang nong 1B	Bred by Sichuan Agricultural university
A6	Jing 3A	K17A/Jing 3B	Bred by Sichuan Academy of Agricultural Sciences, Rice and Sorghum Institute
A7	Ye xiang A	You IA/Ye xiang B	Bred by Guangxi Lvhai seed co., Ltd
A8	65A	Xie qin zao A/65B	Bred by An Hui Academy of Agricultural Sciences, Rice and Sorghum Institute
A9	Guang he A	Tian feng A/Guang he B	Bred by Guangxi Zhao He seed co., Ltd
A10	Guang 8A	325A/Guang 8B	Bred by Guan Dong Academy of Agricultural Sciences, Rice and Sorghum Institute

Table A2. The information of male parents of rice.

Male parents	The source of male parents
R1	Tetep//Guang qinhui/Digu B
R2	Tetep//8006/yan gen2108
R3	Tetep//Guang qinhui
R4	Guang qinhui /Cong sheng rice//1126
R5	Duo Xi 1/Material of resistance to bacterial blight
R6	Huang Huazhan /005
R7	R025/ Guang qinhui
R8	Digu B//R005/ Guang qinhui
R9	GuMei 4//8006/Yan Jing 2108
R10	R025/ Guang qinhui
R11	8006/ Yan Jing 2108
R12	Guang qinhui / Cong sheng rice///1126///R025/014
R13	Guang qinhui // Cong sheng rice //1126
R14	Tetep //3898/Jing chao 1
R15	Huang hua zhan // Guang jing 1/9311
R16	Huang hua zhan //Guang jing 1/9311

Table A3. The information of 20 InDel markers in rice.

Primer	Chromosome	Forward primer sequence (5´-3´)	Reverse primer sequence (5´-3´)
R1M7	1	ATTGTGGTTCTACATTAGTTA	CGCCTCACTAGAATATCGGA
R1M30	1	AAGGGGCCCTAATTTATCTAG	TGTTTACTTTGTTCTTGGACTG
R2M26	2	GCAGCAAAGTGCGGAGTA	CAGGTGAATTGCCAATTT
R2M50	2	CCTGAAGGAAATGATAGCAATAG	GTTTTGTATGCTCTTCACTTGTC
R3M10	3	CCGAGTACCATTGCTTTC	CTGCCATAGTTACTGCTCTGTT
R3M53	3	ACACTGGCTACGGCAAAG	TTTGTTCGGGAATAATGATGC
R4M13	4	TACACGGTAGACATCCAACA	ATGATTTAACCGTAGATTGG
R4M17	4	AGTGCTCGGTTTTGTTTTC	GTCAGATATAATTGATGGATGTA
R5M30	5	CTCAATTTCACCCATCCC	CGCTCCGTCTCCAACCTC
R6M14	6	AAATGTCCATGTGTTTGCTTC	CATGTGTGGAATGTGGTTG
R6M44	6	TTAGGAATAAAGGCTGGATA	TTACCGTTAATAGGTGGAA

Primer	Chromosome	Forward primer sequence (5´-3´)	Reverse primer sequence (5´-3´)
R7M7	7	ACCTTCCCTCCCCCTTTTGAT	AACTTGGTCTTCCTGTTTTATTG
R7M37	7	CAGCCCTAAATCTAAATACCC	ACGTTGAGACAGGCGAGC
R8M33	8	CGAAAGAGGAGAGGGGTAGT	CGAAAACGAGAAACAAATA
R9M10	9	CTTTGGATTCAGGGGGA	AACTTGAAACGGAGGCAG
R9M42	9	ACTGCTTTGATGGCTTGTG	CTCCCCAAACTGAATCC
R10M30	10	CCCTAAAAATAGAGCAACCT	ACCCATAATACTACCAATCAAC
R11M23	11	AAGGTTGACAAGGACAGAAG	TGGGAGGAATGGATAAAA
R11M40	11	AAGAAAAATATCTATTGAGGAGTG	GGAGGACCATAAATGACGG
R12M27	12	ATTTCATTGCCATCAGTT	GTAATCTTCTATCCGTTCA

Table A4. *The classification standards of indica and japonica gene frequencies based on the InDel markers.*

Gene frequency		Type of identification
Indica	*Japonica*	
>0.90	<0.10	Typical *indica* rice
0.75~0.90	0.11~0.25	*Indica* rice
0.60~0.75	0.25~0.40	Partial *indica* rice
0.40~0.60	0.40~0.60	Intermediate type
0.25~0.40	0.60~0.75	Partial *japonica* rice
0.10~0.25	0.75~0.90	*Japonica* rice
≤0.10	≥0.90	Typical *japonica* rice

References

[1] Beukert U., Zuo L., Liu G., Zhao Y. S., Ramachandra N., Mirdita V., Pita F., Pillen K., Reif J. C. (2017), Genome-Based Identification of Heterotic Patterns in Rice. Rice (New York, NY), 10: 22.

[2] Cai, X. X., Liu, J. (2007). Differentiation of Indica-Japonica Rice Revealed by Insertion/deletion Fragments Obtained from Comparative Genomic Study of DNA Sequences between 93-11 (Indica) and Nipponbare (Japonica). Journal of Fudan University (Natural Science), 045 (003): 309-315.

[3] Dudley, J. W., Saghai Maroof, M. A., Rufener, G. K. (1991). Molecular markers and grouping of parents in maize breeding programs. Crop Science, 31: 718-723.

[4] Godshalk, E. B., Lee, M., Lamkey, K. R. (1990). Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. Theoretical and Appllied Genetics, 80: 273-280.

[5] Gui, J. M., Wang, L. Y., Fan, X. J., Qi, Y. B., Zhang, L. X., Fan, H. H., Jin, Q. S., Wang, J. J. (2016). Comparison the Heterosis o Indica-Japonica Hybrids and Japonica-Japonica Hybrids Using InDel Markers. Scientia Agricultura Sinica., 49 (2): 219-231.

[6] Huang X. H., Yang S. H., Gong J. Y., Zhao Q., Feng Q., Zhan Q. L., Zhao Y., Li W. J., Cheng B. Y., Xia J. H., Chen N., Huang T., Zhang L., Fan D. L., Chen J. Y., Zhou C. C., Lu Y. Q., Weng Q. J., Han B. (2016). Genomic architecture of heterosis for yield traits in rice. Nature, 537: 629–633.

[7] Ji, Q., Zhang, M. J., Lu, J., Wang, H. M., Lin B., Liu, Q Q., Chao, Q., Zhang, Y., Liu, C. X., Gu, M. D., Xu M. L. (2012). Molecular basis underlying the S5-dependent reproductive isolation and compatibility of indica/japonica Rice Hybrids. Plant physiology, 3 (3): 1319-1328.

[8] Li L. Y. (2019). Chinese scientists successfully cloned hybrid rice seeds [J]. Agricultural Engine, 9 (2): 1-1.

[9] Liao, F. M., Zhou, K. L., Yang, H. H., Xu, Q. H. (1998). Genetic Difference of Parents and Its Relation to Heterosis in Hybrid Rice [J]. Chinese journal of rice science, 12 (4): 193-199.

[10] Lu, B. R., Cai, X. X., Jin, X. (2009). The method of highly efficient molecular identification and its signidicance for researching of Indica and Japonica rice breeding and evolution. Advances in natural Science, 19 (6): 628-638.

[11] Murray, M., Thompson, W. F. (1980). Rapid isolateon of high moleculer weight plant DNA. Nucleic Acids Research, 8: 4321-4325.

[12] Nie, Y. B., Ji, W. Q., Ma, S. M. (2019). Songmei Ma. Assessment of Heterosis Based on Genetic DistanceEstimated Using SNP in Common Wheat. Agronomy, 9 (2), 66-75.

[13] Pavani, M., Sundaram, R. M., Ramesha, M. S., Kavi Kishor, P. B., Kemparaju, K. B. (2018). Prediction of heterosis in rice based on divergence of morphological and molecular markers. Journal of Genetics, 97 (5): 1263–1279. https://link.springer.com/article/10.1007/s12041-018-1023-8.

[14] Phetmanyseng, Xangsayasane, Xie, F. M., Jose E., Hernandez, Teresita H., Boirromeo. (2010). Hybrid rice heterosis and genetic diversity of IRRI and Lao rice [J]. Field Crops Research, 117: 18-23.

[15] Rohlf, F. J. (1998) Numerical taxonomy and multivariate analysis system, version 2.0. NTSYS-pc 2.1.

[16] Sahu P. K, Mondal S., Sharma D., Vishwakarma G., Kumar V., Das B. K. (2017). InDel marker based genetic differentiation and genetic diversity in traditional rice (Oryza sativa L.) landrace of Chhattisgarh, India. PLoS One, 12: e0188864.

[17] Song, Z. P., Xu, X., Wang, B., Chen, J. K. Lu, B. R. (2003). Genetic diversity in the northernmost Oryza rufipogon populations estimated by SSR markers. Theoretical and applied genetics, 107: 1492-1449.

[18] Tang Q. Y. (2007). DPS data processing system: experiment design, statistical analysis and data mining [M]. Science Press.

[19] Vikas, K. S., Priti, U., Pallavi, S., Ashish, K. M., Ranjith, K. E., Atul, S., Sanjay, K. J., Sunil, B., Ramakrishna, S., Sundaram, R. M., Ilyas, A., Viraktamath, B. C., Kole, C., Sukhpal, S. (2011). Prediction of hybrid performance based on the genetic distance of parental lines in two-line rice (Oryza satival) hybrids. Journal of Crop Science & Biotechnology, 14 (1): 1-10.

[20] Wang, L. Y., Zhang, L. X., Gou, X. X., Fan, H. H., Jin, Q. H., Wang, J. J. (2014). Identification of indica-japonic attribute and prediction of heterosis of Zheyou hybrids rice using InDel molecular markers. Scientia Agricultura Sinica, 47: 1243-1255.

[21] Wang, Y. H., Cai, Q. H, Xie, H. G., Wu, F. G, Lian, L. L,, He, W. L., Chen, L. P., Xie, H. A., Zhang J. F. (2018). Determination of Heterotic Groups and Heterosis Analysis of Yield Performance in indica Rice. Rice Science, 25 (05): 31-40.

[22] Wei S. B., Li X., Lu Z. F., Zhang H., Ye X. Y., Zhou Y. J., Li J., Yan Y. Y., Pei H. C., Duan F. Y., Wang D. Y., Chen S., Wang P., Zhang C., Shang L. G., Zhou, Yan P., Zhao M., Huang J. R., Bock R., Qian Q., Zhou W. B. (2022). A transcriptional regulator that boosts grain yields and shortens the growth duration of rice. Science. 377, 382-383.

[23] Xu J. f., Wang, L. Y. (1980), Heterosis in rice and. genetic distance. Hunan Agricultural Sciences, (6): 11-15.

[24] Yang, J., Wang, J., Cao, Q., Chen, Z. D., and Zhong, W. G. (2009). Development and application of a functional marker for wide compatibility gene s5-n of rice. Acta Agronomica Sinica, (11): 2000-2007.

[25] Yu, Y. H., Liu, Y., Li, Z. Y., Chen, G. H., Xu, Z. J., Tang, L., Mao, T., Xu, H. (2016). Relationship between Indica-Japonica Index of Parents and Heterosis of Hybrid and Its Genetic Basis in Japonica Two Line Hybrid Rice. Acta Agronomica Sinica, 42 (5): 648-657.

[26] Yuan L. P. (2016) Research and development of Hybrid rice in China [J]. Science & Technology Review, 34 (20): 64–65.

[27] Zhai, R. R., Feng, Y., Wang, H. M., Wang, X. D., Zhan, X. H., Shen, W. M., Zhang Y. X., Chen D. B., Dai G. X., Yang Z. L, Cao L. Y., Cheng S. H. (2013) Transcriptome analysis of rice root heterosis by RNA-Seq. BMC Genomics, 14, 19. https://doi.org/10.1186/1471-2164-14-19.

[28] Zhang, T., Han, L., Xu, J. D., Jiang, K. F., Wu, X. J., Wang, X. D., Zheng, J. K. (2006). Correlation Between Genetic Distance and Yield Heterosis of Hybrid Aromatic Rice [J]. Scientia Agricultura Sinica, 39 (04): 831-835.

[29] Zhao, Q. Y., Zhu, Z., Zhang, Y. D, Zhao, L., Chen, T., Zhang, Q. F., Wang, C. L. (2009). Analysis on Correlation Between Heterosis and Genetic Distance Based on Simple Sequence Repeat Markers in japonica Rice. Chinese journal of rice science, 23 (2): 141-147.

[30] Zhao, X. Y., Ye, S. H., Li, X. H., Zhai, R. G., Yu, P., Jin, Q. S., Zhang, X. M. (2015). Identification of indica-japonica attribute and prediction of heterosis using InDel markers in rice. Acta Agriculturae Zhe jiangensis, 27 (8): 1309-1316.