Theor Appl Genet (2015) 128:421–430
DOI 10.1007/s00122-014-2441-5

ORIGINAL PAPER

Heterotic groups of tropical *indica* rice germplasm

Kai Wang · Fulin Qiu · Wenceslao Larazo · Madonna Angelita dela Paz · Fangming Xie

Received: 13 July 2014 / Accepted: 6 December 2014 / Published online: 16 December 2014

Abstract

Key message **Four heterotic groups were identified for tropical *indica* rice germplasm to develop hybrid rice in the tropics based on two studies.**

Abstract Heterotic groups are of fundamental importance in hybrid crop breeding. This study investigated hybrid yield, yield heterosis and combining ability within and among groups based on genetic distance derived from single-nucleotide polymorphism markers. The main objectives of the study were to (1) evaluate the magnitude of yield heterosis among marker-based groups, (2) identify possible heterotic groups for tropical *indica* hybrid rice, and (3) validate heterotic patterns concluded from a previous study. Seventeen rice parents selected from improved *indica* germplasm from the tropics with high genetic divergence and 136 derived hybrids were evaluated in five environments. The hybrids had more yield than their parents with an average of 24.1 % mid-parent heterosis. Genotype × environment interaction was the major factor affecting variations in yield and yield heterosis, which raised a necessity and a challenge to develop heterotic rice hybrid adapted to different regions and seasons in the tropics. Yield, yield heterosis and combining ability were significantly increased in inter-group than in intra-group hybrids. Four heterotic groups and three promising hybridization patterns, which could be used in tropical hybrid rice breeding, were identified based on marker-based grouping, yield and yield heterosis analyses in the two studies. The study reveals that molecular markers analysis can serve as a basis for assigning germplasm into heterotic groups and to provide guidelines for parental selection in hybrid rice breeding.

Communicated by Peter Langridge.

K. Wang · F. Qiu · W. Larazo · M. A. dela Paz · F. Xie (✉)
International Rice Research Institute, DAPO Box 7777,
1301 Metro Manila, Philippines
e-mail: f.xie@irri.org

K. Wang
China National Rice Research Institute, Hangzhou 310006,
Zhejiang, China

K. Wang
Yahua Seed Research Institute, Changsha 410001, Hunan, China

F. Qiu
Liaoning Rice Research Institute, Shenyang 110101,
Liaoning, China

Introduction

Recognition and determination of heterotic groups and patterns are fundamentally important for breeding hybrid crops as indicated in many studies for maize (Melchinger and Gumber 1998; Menkir et al. 2003, 2004; Reif et al. 2003; Akinwalea et al. 2014; Suwarno et al. 2014), rye (Fischer et al. 2010a), sunflower (Reif et al. 2013), sorghum (Menz et al. 2004), triticale (Fischer et al. 2010b) and rice (Sun et al. 2000; Wang and Lu 2006; Xie et al. 2014).

Parents of heterotic hybrids are usually derived from different heterotic groups with high genetic divergence. Grouping of germplasm in divergent pools is advantageous to maximize the expected heterosis (Reif et al. 2005). The success of the hybrid rice that resulted from the utilization of heterosis is based on the genetic divergence of germplasm, basically on geographic divergence for three-line hybrid rice and subspecific genetic divergence for two-line hybrid rice. Heterotic groups of a crop have been assessed by information, such as phenotypic differences, germplasm origins, pedigree and combining ability; however, those traditional evaluation methods are time-consuming and usually impractical to

breeders due to numerous hybrid combinations and tremendous field work. Furthermore, extensive germplasm exchange among breeding programs without detailed information on the pedigrees and genetic background, as well as effects of environment and genotype × environment interaction, makes parental assessment more complicated. Molecular markers have provided an efficient way and effective tool for genetic diversity evaluation at the DNA level (Semagn et al. 2012; Frascaroli et al. 2013; Li et al. 2014). Heterotic groups for hybrid crops could possibly be determined by marker-based groups as studied in maize (Choukan et al. 2006; Xie et al. 2008; Lu et al. 2009; Romay et al. 2013), as well as in other crops (Tams et al. 2004; Du et al. 2011). In a study conducted by He et al. (2012), genetic diversity of hybrid rice parents developed at the International Rice Research Institute (IRRI) was evaluated with simple sequence repeats (SSR) and single-nucleotide polymorphism (SNP) markers where it was demonstrated that heterotic hybrids could be formed based on marker-based parent groups to increase efficiency of hybrid rice breeding (Xie et al. 2014). To further investigate heterotic groups of tropical hybrid rice with more germplasm, a subset of samples from a previous study (Wang et al. 2013), involving 736 SNP-grouped rice samples collected from major *indica*-growing tropical regions, was selected to examine the heterotic response of parents with the objectives of (1) evaluating the magnitude of yield heterosis among divergent tropical *indica* germplasm grouped with SNP markers, (2) identifying possible heterotic groups from marker-based genotypes, and (3) validating heterotic patterns concluded from a previous study.

Materials and methods

Plant materials

Seventeen parents (Table 1) were selected from a diverse panel of 736 improved, semi-dwarf *indica* rice varieties and elite breeding lines genotyped with 384 SNPs in a previous study (Wang et al. 2013). The parents' selection was based on (1) representing the original population structure with six subgroups, (2) possessing a maximum allelic variation and (3) commonly cultivated in tropical regions with normal maturity and fertility. The 17 parents were crossed with a diallel mating design without reciprocals to develop 136 F_1 hybrids at IRRI in Los Baños, Philippines in wet season (WS) of 2012 and dry season (DS) of 2013.

Field experiments

A total of 155 entries, including 136 hybrids, 17 parents and two checks (inbred, IRRI123 and hybrid, IR75217H), were

Table 1 Parents and their yield means (g plant^{-1}) and GCA across environments

Code	Name	Group	Yield mean over environments	GCA	Origin
Parent					
P12	TOX3416-170-2-1-1	B4	28.4 a	0.30	AfricaRice
P1	IR65623-94-3-1-3-3R	B1	27.3 ab	0.82	IRRI
P4	IRRI105	B2	26.8 ab	0.43	IRRI
P16	IR77801B	B6	26.4 abc	−0.17	IRRI
P6	IR80285-34-3-3-2	B2	26.2 abc	0.61	IRRI
P15	CT22147-6P-9SR-2P-3SR-1P	B5	25.7 abc	−0.30	CIAT
P7	IR64	B3	25.1 abcd	0.63	IRRI
P9	IRRI135	B3	24.5 bcde	−0.23	IRRI
P10	IR79228-67-1-1-3	B4	23.9 bcde	0.42	IRRI
P2	IR02A127	B1	23.7 bcdef	1.58	IRRI
P8	IRRI110	B3	23.6 bcdef	0.26	IRRI
P3	SILUGONGGO	B1	23.0 cdef	−0.50	Indonesia
P5	IR79203-105-1-1-3	B2	22.0 def	0.87	IRRI
P14	CT22139-4P-2SR-4P-3SR-1P	B5	21.2 efg	−0.44	CIAT
P18	IR80157B	B6	21.0 efg	−1.76	IRRI
P11	CT18920-1-10-4-1SR-3P	B4	20.1 fg	−0.78	CIAT
P17	IR78365B	B6	18.1 g	−1.75	IRRI
	Mean		23.9		
Check					
	IRRI123		25.0		
	IR75217H		27.6		

Different letters mean significant difference at 0.05 level by Duncan's grouping test

AfricaRice Africa Rice Center, *CIAT* International Center for Tropical Agriculture

evaluated in field trials under five environments (year, season and location): (1) 2013DS and 2013WS, Los Baños, Philippines, 14°10′N/121°13′E; (2) 2013DS and 2013WS, Muñoz, Philippines, 15°42′N/120°54′E; and (3) 2013WS, San Mateo, Isabela, Philippines, 16°52′N/121°35′E. All entries were grown in a randomized complete block design with two replications. Forty 21-day-old seedlings were transplanted in four rows with ten plants per row with spacing of 20 cm × 20 cm. The field management was carried out following local recommendations for each cropping season. At ripening stage, five plants in central rows of each plot were randomly harvested and measured for plant height and yield components. The rest of the 11 plants in central rows of each plot excluding the plants at row ends were also harvested and assessed for grain yield, wherein the grain yield per plant was calculated based on the total grain weight of the harvested 16 plants.

Statistical analyses

Genetic diversity between each pair of the 17 parents was measured as Cavalli-Sforza and Edwards (1967) genetic distance (GD) using PowerMarker version 3.25 (Liu and Muse 2005). Clusters based on GDs were generated using DARwin 5 (Perrier and Jacquemoud-Collet 2006). Analyses of variances were conducted among parents and among hybrid combinations for each and across all environments with the PROC GLM procedure (SAS Institute Inc 2012) following the model:

$$Y = \mu + E + R + G + \mathrm{GEI} + e,$$

where Y is a observed value of parent or hybrid yield from each test unit, μ is a population mean, E is a environmental (Env) effect, R is a replication effect within each environment, G is a genotype (parent or hybrid) effect, GEI is a interaction effect between each genotype and environment, and e is a residual effect. The environment and genotype were treated as fixed factors and the replication within environment was considered to be a random factor. The significance of environmental variance was tested against replication within environment entity. General combining ability (GCA) effects of the parents and specific combining ability (SCA) of the hybrids were estimated for each and across environments following Griffing's (1956) method 2 model 1 using the R program (R Development Core Team 2011). Yield heterosis for each hybrid was calculated as (1) mid-parent heterosis (MPH) = 100 × (F_1 − MP)/MP, (2) better parent heterosis (BPH) = 100 × (F_1 − BP)/BP, and (3) standard heterosis over inbred check (SDHI) or over hybrid check (SDHH) = 100 × (F_1 − CK)/CK, where F_1 is the hybrid yield, MP is the yield mean of both parents, BP is the yield of the better-yielding parent, and CK is the yield of the check variety, either the inbred (IRRI123) or the hybrid check (IR75217H). Correlations between GD estimates and means of hybrid yield, yield heterosis and combining ability were calculated using PROC CORR of SAS (SAS Institute Inc 2012).

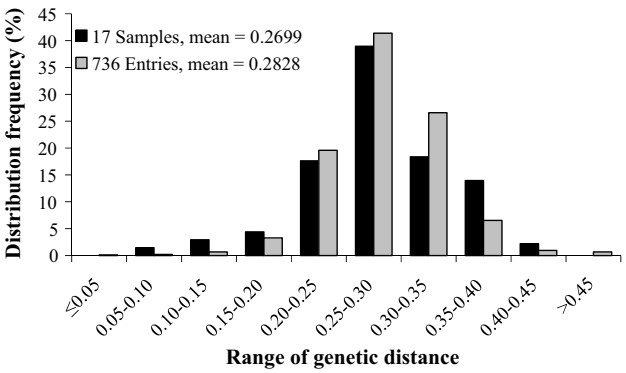

Fig. 1 Distribution of genetic distances of the original 736 lines and the 17 parents sampled for the study based on 384 SNP markers

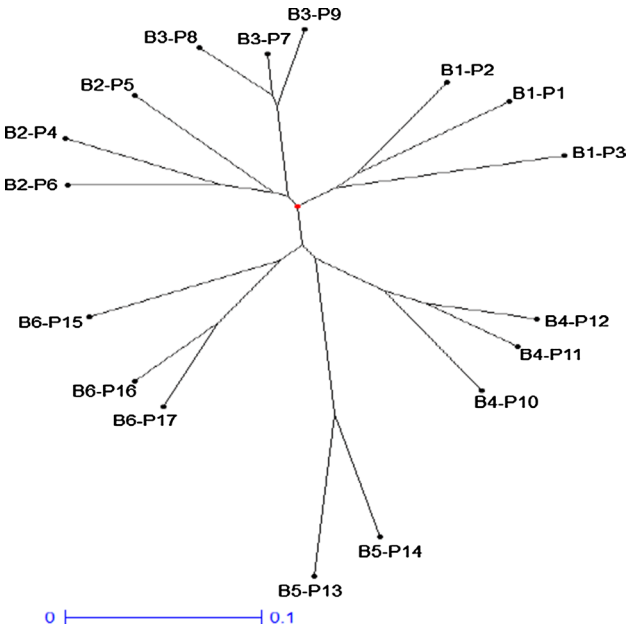

Fig. 2 Cluster of the 17 samples based on genetic distance with 384 SNP markers. *Letter* and *number* combination refers to parent as marker-based group-parent code

Results

Parental genetic diversity

The GD values for the 17 parents ranged from 0.0760 (IR64 vs IRRI110) to 0.3983 (IRRI135 vs CT22147-6P-9SR-2P-3SR-1P) with an average of 0.2699 which was slightly lower than that of the original panel of 736

Table 2 Genetic distance among groups and performance of hybrid yield and yield heterosis

Hybrid group	Genetic distance	Hybrid yield	SCA	MPH	BPH	SDHI	SDHH
B1 × B2	0.2704	30.5 a	0.97	28.2	14.7	22.5	11.3
B2 × B5	0.3730	29.8 ab	1.39	26.0	10.6	20.1	8.8
B1 × B5	0.3535	29.8 ab	1.37	25.9	12.2	20.3	9.0
B1 × B3	0.2399	29.7 ab	0.62	24.8	12.3	19.3	8.1
B2 × B3	0.2284	29.5 ab	0.41	23.0	8.7	18.4	7.4
B1 × B4	0.3097	29.4 ab	0.63	27.2	11.9	19.2	7.8
B2 × B2	0.1978	29.4 ab	−0.06	24.7	10.5	18.2	7.2
B2 × B4	0.2770	29.4 ab	0.58	26.1	11.5	18.7	7.6
B1 × B1	0.2085	29.1 abc	−0.32	19.8	6.1	15.8	5.0
B3 × B6	0.2700	29.1 abc	1.94	30.3	13.8	17.0	6.1
B3 × B4	0.2812	29.1 abc	0.67	24.5	9.4	17.6	6.4
B3 × B3	0.0926	28.9 abc	0.28	19.3	8.4	16.4	5.2
B5 × B6	0.3389	28.6 abc	1.89	28.7	10.4	14.7	3.8
B4 × B4	0.1472	28.4 abcd	0.24	22.4	5.3	14.0	3.4
B4 × B5	0.3363	28.4 abcd	0.55	21.1	6.0	14.5	3.8
B1 × B6	0.2960	28.0 bcd	0.30	24.8	8.7	11.7	1.4
B2 × B6	0.2800	27.7 bcd	0.14	23.3	8.3	10.8	0.6
B4 × B6	0.2706	27.2 cd	0.26	23.6	7.1	9.4	−0.8
B3 × B5	0.4002	26.5 d	−1.51	12.3	1.8	7.1	−3.1
B5 × B5	0.2068	24.5 e	−1.06	4.0	−7.2	−3.2	−12.1
B6 × B6	0.1868	23.8 e	−1.95	9.9	−5.2	−6.3	−15.0
Summarized by inter- or intra-group							
Inter-group	0.2855	28.9 a	0.68	24.9	10.0	16.2	5.3
Intra-group	0.1643	27.7 b	−0.49	18.3	4.3	10.7	0.3
Summarized by environment							
2013DS Muñoz		35.6 a		16.6	5.4	13.4	7.5
2013DS Los Baños		29.7 b		14.3	4.8	5.2	−3.2
2013WS San Mateo		29.3 b		34.0	13.2	29.2	14.2
2013WS Los Baños		26.1 c		20.4	10.5	20.3	7.3
2013WS Muñoz		22.9 d		35.2	12.9	9.2	−2.1
Summarized by parental group involved in hybrids							
B1		29.0 a		26.6	11.7	17.9	6.8
B2		29.1 a		25.5	10.5	17.5	6.6
B3		28.8 ab		25.0	10.2	16.6	5.5
B4		28.4 b		25.1	8.9	15.1	4.2
B5		27.6 c		26.6	9.8	11.1	0.7
B6		27.3 c		29.1	9.6	12.9	1.8

Different letters mean significant difference at 5 % level

lines (Fig. 1). The genetic structure of the samples was consisted of six subclusters, each with three representative varieties except B5 (Fig. 2). The samples covered 88.9 % of the allelic variation from the 736-line panel and were considered to be fairly representative of the population panel because of the high allelic coverage and similar cluster structure. Among the six subclusters, B3 had the least allelic variation with an average GD of 0.0926 and B1 had the highest GD value (0.2085) followed by B5 (0.2068) (Table 2). The average GD (0.2855) of inter-group parents was significantly ($P < 0.001$) higher than the average GD (0.1643) of intra-group parents.

Analyses of variance

Difference in parental yields was significant across environments with an average yield of 23.9 g plant^{-1}, ranging from 18.1 (IR78365B) to 28.4 g plant^{-1} (TOX3416-170-2-1-1) (Table 1). There was no specific yielding pattern among parents, i.e., parents classified in each group yielded randomly, indicating non-specific relationship between GD

and parental yields. The effects of environment, genotype and GEI all contributed significantly ($P < 0.001$) to the variation of parental yielding which accounted for 38.9, 14.8 and 29.7 % of the total sums of squares, respectively (Table 3).

Hybrid yields differed significantly ($P < 0.001$) for environment, genotype and GEI with an average yield of 28.7 g plant^{-1} that ranged from 20.7 to 34.6 g plant^{-1} (Table 3; Fig. 3). The effects of environment, genotype and GEI contributed 44.9, 11.6 and 26.8 % of the total sums of squares to the hybrid yield variation, respectively. The proportions of GEI in the total variation of parent and hybrid yields were relatively large as expected due to the highly different testing environments involved in different seasons and geographical locations. On average, inter-group hybrids yielded significantly ($P < 0.05$) higher than intra-group hybrids (Table 2). The lowest (22.9 g plant^{-1}) and the highest (35.6 g plant^{-1}) yields of hybrids across environments were observed at the same location (Muñoz), but in different cropping seasons (2013WS and 2013DS). The hybridization patterns with the highest yield, SDHI and SDHH were B1 × B2, B2 × B5 and B1 × B5 (Table 2).

Combining ability, heterosis and correlation of genetic distance with yield heterosis

IR02A127 (P2 in B1) had the largest GCA effect (1.58) on hybrid yield followed by IR79203-105-1-1-3 (P5 in B2) (Table 1). Parents of IR78365B (P17) and IR80157B (P18) had the lowest GCA effects on hybrid yield, both being maintainer lines of hybrid parents classified in B6. The average SCA effect (0.68) of inter-groups hybrids was significantly ($P < 0.001$) higher than the average SCA (−0.49) of intra-group hybrids (Table 2). The hybridization patterns of B3 × B6, B5 × B6, B2 × B5 and B1 × B5 had high SCA values (>1.00). Hybrids of B3 × B6 and B5 × B6 also had the highest MPH values but ranked intermediate for other heterosis measurements and all other hybrids involved with B6 parents performed poorly, indicating a poor yielding level of B6 parents.

All source effects of variations were significant ($P < 0.0001$) for all heterosis measurements with GEI as the major contributor that accounted for 38.8, 41.0, 36.7 and 33.4 % of MPH, BPH, SDHI and SDHH variance, respectively (Table 3). The average MPH was 24.1 % with a range of −5.9 to 58.1 % and the average BPH was 9.4 % with a range of −15.3 to 36.8 %. The hybrids showed averages of 15.5 and 4.7 % yield advantages over inbred and hybrid checks, respectively. Forty-three hybrids (31.6 %) showed more than 20 % of SDHI and 5 % of SDHH. Among the parents involved in those 43 heterotic hybrids, the main contributors were from the groups of B1 (20.9 %), B2 (25.6 %) and B3 (18.6 %), and the rest of other three

Table 3 Analysis of variance, including degrees of freedom (DF), mean squares (MS), and percent contribution to total sums of squares (SS %) across five environments for parent and hybrid yield, hybrid yield mid-parent heterosis (MPH), better parent heterosis (BPH), standard heterosis over inbred CK (SDHI) and hybrid CK (SDHH)

Source	Hybrid													
	Parent yield			Yield			MPH		BPH		SDHI		SDHH	
	DF	MS	SS (%)	DF	MS	SS (%)	MS	SS (%)	MS	SS (%)	MS	SS (%)	MS	SS (%)
Env	4	815.4***	38.9	4	6,024***	44.9	26,415***	8.2	4,423***	2.0	25156***	15.3	14,517***	9.7
Genotype	16	77.7***	14.8	135	46***	11.6	1,813***	19.0	1,223***	18.5	806***	16.5	653***	14.8
GEI	64	38.9***	29.7	540	27***	26.8	926***	38.8	678***	41.0	447***	36.7	369***	33.4
Rep (Env)	5	8.8	0.5	5	25	0.2	3,465***	1.3	3,724***	2.1	10,014***	7.6	24,303***	20.4
Error	80	16.9	16.1	672	13	16.4	625	32.6	484	36.4	233	23.8	193	21.8
Total	169			1,356										

*** Significant at <0.0001 probability level

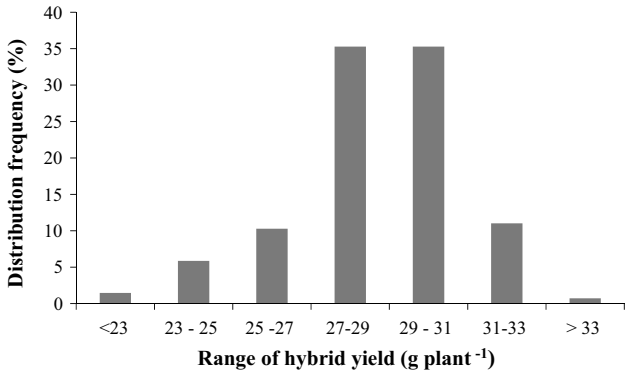

Fig. 3 Range of hybrid yields across environments

groups have less proportions of parents contributing to the heterotic hybrids with a sum of 34.9 %. The highest yielding hybrids (B1 × B2) with the highest yield heterosis over checks produced an average yield of 30.5 g plant^{-1} with 28.2, 14.7, 22.5 and 11.3 % of MPH, BPH, SDHI and SDHH, respectively (Table 2). The lowest yielding and yield heterosis of hybrids were from the crossing patterns of B3 × B5, B5 × B5 and B6 × B6.

All measurements of yield and yield heterosis of inter-group hybrids were significantly ($P < 0.001$) higher than those measured in intra-group hybrids. Among the testing environments, the highest heterosis measurements were observed in 2013WS at San Mateo in Isabela, suggesting that the Isabela area is a better environment than Los Baños and Muñoz for hybrid rice cultivation during wet season (Table 2).

Except for a few small and significant correlations detected in some individual locations and/or heterosis parameters measured, the overall correlations between GD and hybrid, and yield heterosis were insignificant (Table 4).

Discussion

Population structure

The parents selected for the study were from those improved *indica* germplasm currently used in the tropics. Although most of them are from IRRI and the CIAT, they are being widely used in various breeding programs and regions for different purposes either as hybridization parents and elite breeding lines, or as cultivated for rice production. These varieties and breeding lines can be considered to represent the yield levels and genetic divergence of current tropical *indica* germplasm. The samples selected for the present study were fairly representatives of the original divergent population as shown with the similar population structure, distribution of genetic distances and the maximum coverage of allelic variation.

Yield and yield heterosis

It is noted that almost a third of hybrid combinations yielded more than 20 % of SDHI and 5 % of SDHH, showing the great potential and opportunity in using the current tropical *indica* germplasm to significantly increase hybrid rice heterosis. GEI was a major source of variance for parent and hybrid yields, and yield heterosis as shown in the present and previous studies (Xie et al. 2014), indicating the importance of developing hybrids specifically adapted to different environments (locations and seasons) for a maximum utilization of heterosis. Generally, the yields of parents and hybrid were higher in dry season than in wet season in the tropics; however, it does not necessarily mean that hybrids cultivated in dry season have a better yield advantage, measured genetically (MPH and BPH) or agronomically (SDHI and SDHH), over hybrids grown in wet season. For instance, hybrid rice grown during the wet season in Los Baños had a lower yield than in the dry season, but it was able to give a better yield advantage over their parents and the standard variety checks. On the contrary, the hybrids grown at Muñoz were observed to have a significantly better performance in terms of yielding, SDHI and SDHH during the dry season than the wet season, but with lower MPH and BPH (Table 2). These observations suggest that the environment (location and season) should be given an important consideration in cropping decision when selecting a hybrid or inbred variety for production. Given a significant effect of G × E, and as the major source of the variance for parent and hybrid yields, and also with frequent occurrence of extreme weather in the tropics, selection of parents and hybrids with a wide adaptability should be the major factor being considered for minimizing the effect of G × E in developing hybrids for a great area. However, for a national or regional breeding program which has a specific objective of developing products for a particular environment, a large effect of G × E could be applied to develop region- or season-specific hybrids to maximize the heterosis.

The hybrid yields in the present study correlated significantly with mid-parent yield ($r = 0.5629$, $P < 0.001$), and better parent yield ($r = 0.6544$, $P < 0.001$), implying the importance of selecting high yielding, high GCA and genetically distant parents to produce high-yielding hybrids.

Heterotic groups based on markers

It is important to understand the extent of genetic variability and classify the germplasm into heterotic groups in

Table 4 Correlation coefficient (*r*) between SNP marker-based genetic distance and hybrid yield and yield heterosis

Trait	All hybrids ($n = 136$)	Inter-hybrids ($n = 120$)	Intra-hybrids ($n = 16$)	Marker-based groups ($n = 30$)
2013WS San Mateo				
Yield	0.0875	0.0227	−0.1890	0.2417
SCA	0.0683	−0.0126	−0.2525	0.2591
MPH	−0.0402	−0.1429	−0.2094	0.1834
BPH	0.0189	−0.0821	−0.2174	0.3177
2013DS Los Baños				
Yield	−0.1357	−0.2458**	0.0013	−0.0692
SCA	0.0141	−0.0635	−0.0311	0.1185
MPH	0.0510	0.0015	−0.0609	0.2470
BPH	−0.0223	−0.0680	−0.1677	0.0265
2013WS Los Baños				
Yield	−0.0139	−0.2164*	−0.1767	0.1472
SCA	0.1292	−0.0747	−0.0613	0.4452*
MPH	0.0756	−0.0681	−0.1945	0.2263
BPH	0.0585	−0.0724	−0.3667	0.2779
2013DS Muñoz				
Yield	−0.1494	−0.1786	0.2271	−0.1827
SCA	−0.1073	−0.0882	0.0010	−0.2730
MPH	0.0328	0.0614	0.0928	0.0740
BPH	−0.0659	−0.0765	0.0731	−0.0581
2013WS Muñoz				
Yield	0.1509	0.0494	0.1604	0.3360
SCA	0.2015*	0.0931	0.2780	0.4009
MPH	0.0475	−0.0775	0.3840	0.0936
BPH	0.0472	−0.0637	0.3295	0.0873
Five environments combined				
Yield	−0.0068	−0.1752	−0.0063	0.1619
SCA	0.1064	−0.0396	−0.0289	0.3172
MPH	0.0467	−0.0933	0.1717	0.2768
BPH	0.0265	−0.1191	0.0497	0.2571

*, ** Significant at the $P < 0.05$ and <0.01 probability levels, respectively

hybrid crop breeding (Reif et al. 2005). The genetic diversity estimates based on molecular markers are helpful in assigning germplasm into heterotic groups and in assessing pedigree relationship among germplasm (Melchinger and Gumber 1998), but with very limited value in predicting hybrid yield performance as evidenced in hybrid maize (Marsan et al. 1998; Balestre et al. 2008; Legesse et al. 2008) and hybrid rice (Zhang et al. 1995; Zhao et al. 1999; Xu et al. 2002; Singh et al. 2011). With insufficient knowledge of association of functional molecular markers and yield, the information derived from molecular markers currently is limited to the use of assigning parents into germplasm group or heterotic groups, and to provide a general guideline of avoiding heterotic groups from mixture during parent breeding. Further detailed information on functional markers related to yield heterosis may provide specific direction or guideline for combining parents specifically to increase breeding efficiency. Experience of hybrid rice breeding generally showed that the chance of developing heterotic hybrids is much higher when parents are from inter-groups than from intra-group, wherein the groups could mean geographic regions, ecotypes and sub-species. However, it does not necessarily mean that all hybrids from parents of inter-group always give a high yield or heterosis, but only parents from some specific groups, such as those parents in the groups of B1, B2 and B5 in the present study. The preferred hybridization patterns derived from the present study are B1 × B2, B2 × B5 and B1 × B5, because of their high averages of SDHI (>20 %) and hybrid yields.

Heterotic groups in tropical *indica* hybrid rice

It is estimated that more than 90 % of the hybrid rice parents used in Asian countries outside China have included IRRI hybrid rice germplasm. We defined the heterotic groups of IRRI-bred hybrid rice parents in a previous study (Xie et al. 2014) to provide a reference for parent selection considering the wide dissemination of IRRI germplasm in the region. The magnitude of yield heterosis of rice hybrids in the tropics has been low, compared to that of hybrid rice in China and Americas. Limited genetic divergence among hybrid rice parents could be one of the reasons for the constraints of hybrid rice yield and heterosis in the tropics (Xie et al. 2012). To broaden hybrid rice parent gene pools, introduction of germplasm from other sources and integrating them into heterotic groups are essential steps to further enhance heterotic performance of hybrids in the tropics. In our previous study, hybridization patterns of G3 × G5 and G3 × G6 were our preferred choices for developing tropical hybrid rice using IRRI-bred hybrid rice parents because of their high SDHI (>15 %). In the present study, more germplasms, including those from some Asian countries, CIAT and Africa, were evaluated for hybrid performance, and three preferred hybridization patterns (B1 × B2, B2 × B5 and B1 × B5) were included for tropical hybrid rice. A comparison of germplasm clusters and hybridization groups from both studies showed that the G3 group is closely associated with the B1 group (Fig. 4). Most of the lines in the groups of G3 and B1 are restorer lines of "WA" (wild abortive) sterile cytoplasm (Wang et al. 2013), and the highly heterotic response of G3 lines in the previous study is validated with the heterotic performance of the B1 lines, results of which are shown in the present study. In the B2 group, most (97.4 %) of the lines are IRRI inbreds and showed restoring ability to "WA" sterile cytoplasm. The B2 group is highly related to the G2 group which is also a group of restorer lines. Groups of G2, G3, B1, B2 and B3 can be considered as groups of male parents for tropical three-line hybrid rice, but the lines in G3, B1 and B2 are better on hybrid performance than the lines in G2 and B3; therefore, selection for male parents should be focused on the lines in the groups of G3, B1, and B2.

It is noted that Zhengsha 97B, an elite and dominant female parent for the Chinese hybrid rice during 1970–2000s, is included in the B5 group (Wang et al. 2013). The heterotic performance of B2 × B5 hybrids in the present study agrees with the heterotic pattern of Chinese three-line *indica* hybrid rice (early season *indica* in the middle and lower regions of the Yangtze River Valley combined with restorer lines from IRRI). Some of the new IRRI-bred female parents (IR93558B–IR93563B) are included in the B5 group and primary results from testcrosses in IRRI hybrid rice breeding showed the promising heterotic performance of using these new female parents, which indirectly

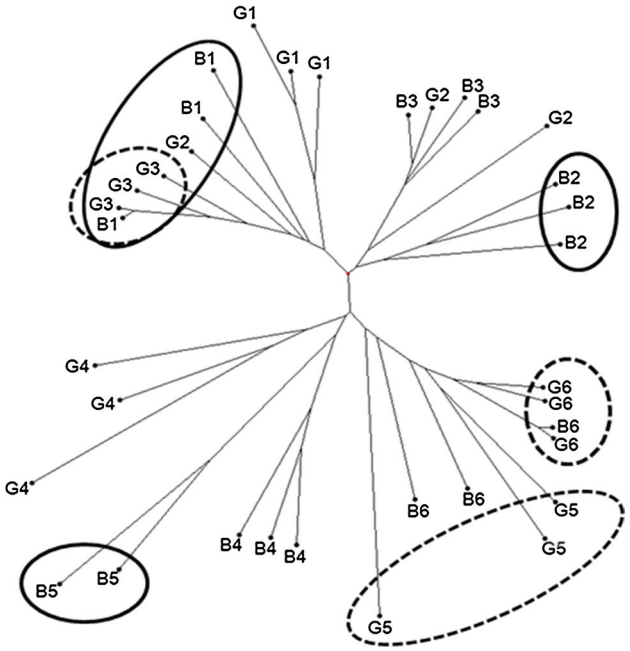

Fig. 4 Cluster of the 35 tropical *indica* parents from two heterotic studies based on genetic distance calculated from 384 SNP markers. *Letters G* and *B* indicate groups from the previous (Xie et al. 2014) and present studies, respectively, and *numbers* refer to the marker-based groups in the two studies

proved the heterotic patterns of B1 × B5 and B2 × B5. The lines in the B6 group are all IRRI-bred female parents developed prior to 2005 and they are closely related to the lines in the G5 and G6 groups. The hybrids from G3 × G5 and G3 × G6 showed good heterotic performance (Xie et al. 2014), but the hybrid performance of B1 × B6, which is the similar hybridization patterns of G3 × G5 and G3 × G6, is ranked relatively low in the present study. It could be due to the low GCA and yield of the B6 lines selected in the present study. It is also interesting to note that most of the CIAT lines are included in the groups of B4 and B5 and these two groups are closely linked. Yields of the hybrids derived from these two groups are ranked as either high (B1 × B5) or relatively high (B1 × B4), confirming the hybridization patterns of B1 × B5, and alternatively of B1 × B4. Considering the existence of maintainers and effective restorers, and the high frequency of effective restorer lines being from South, Southeast Asia and South China, the B4 and B5 groups should be merged as a female parent group to combine with restorer lines from the groups of G3, B1 and B2 and to match the present heterotic pattern of "WA" hybrid rice in China.

Based on the results from the two studies, it can be concluded that four heterotic groups are existing among the current tropical rice germplasm, i.e., for the male parent, it can be classified as the heterotic group (HG) HG1 (B1

and G3) and HG2 (B2), and for the female parent, they are HG3 (G5, G6 and B6), and HG4 (B5 and B4). The hybridization patterns of HG1 × HG3, HG1 × HG4 and HG2 × HG4 should be given a high priority in developing tropical hybrid rice.

Author contribution statement Fangming Xie and Kai Wang designed the experiments. Kai Wang and Wenceslao Larazo produced the hybrids and performed the field experiments. Kai Wang and Madonna Angelita dela Paz did the genetic diversity analyses. Kai Wang, Fulin Qiu and Fangming Xie analyzed the data. Kai Wang and Fangming Xie wrote the paper.

Acknowledgments The authors thank Philippine Rice Research Institute's Thelma F. Padolina and Rustum C. Braceros in Muñoz, Nueva Ecija, and Democrito B. Rebong II in San Mateo, Isabela for providing field facilities and field management of the field trials in the two locations. This research was partially supported by the China Scholarship Council (Grant No. 2011325040).

Conflict of interest This research has no conflict of interest.

References

Akinwalea RO, Badu-Aprakub B, Fakoredea MAB, Vroh-Bib I (2014) Heterotic grouping of tropical early-maturing maize inbred lines based on combining ability in Striga-infested and Striga-free environments and the use of SSR markers for genotyping. Field Crop Res 156:48–62

Balestre M, Machado JC, Lima JL, Souza JC, Nóbrega Filho L (2008) Genetic distance estimates among single cross hybrids and correlation with specific combining ability and yield in corn double cross hybrids. Genet Mol Res 7(1):65–73

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233–257

Choukan R, Hossainzadeh A, Ghannadha MR, Warburton ML, Talei AR, Mohammadi SA (2006) Use of SSR data to determine relationships and potential heterotic groupings within medium to late maturing Iranian maize inbred lines. Field Crop Res 95(2):212–222

Du X, Sun Y, Li X, Zhou J, Li X (2011) Genetic divergence among inbred lines in *Cucurbita moschata* from China. Sci Hortic 127(3):207–213

Fischer S, Melchinger AE, Korzun V, Wilde P, Schmiedchen B, Mohring J, Piepho HP, Dhillon BS, Wurschum T, Reif JC (2010a) Molecular marker assisted broadening of the Central European heterotic groups in rye with Eastern European germplasm. Theor Appl Genet 120:291–299

Fischer S, Maurer HP, Würschum T, Möhring J, Piepho HP, Schön CC, Thiemt EM, Dhillon BS, Weissmann EA, Melchinger AE, Reif JC (2010b) Development of heterotic groups in *Triticale*. Crop Sci 50:584–590

Frascaroli E, Schrag TA, Melchinger AE (2013) Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. Theor Appl Genet 126(1):133–141

Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing system. Aust J Biol Sci 9:463–493

He ZZ, Xie FM, Chen LY, Dela Paz MA (2012) Genetic diversity of tropical hybrid rice germplasm measured by molecular markers. Rice Sci 19:193–201

Legesse BW, Myburg AA, Pixley KV, Twumasi-Afriyie S, Botha AM (2008) Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. Euphytica 162(3):313–323

Li JY, Wang J, Zeigler RS (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. GigaScience 3(1):1–3

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Lu Y, Yan J, Guimaraes CT, Taba S, Hao Z, Gao S, Chen S, Li J, Zhang S, Vivek BS, Magorokosho C, Mugo S, Makumbi D, Parentoni SN, Shah T, Rong T, Crouch JH, Xu Y (2009) Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theor Appl Genet 120(1):93–115

Marsan PA, Castiglioni P, Fusari F, Kuiper M, Motto M (1998) Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor Appl Genet 96(2):219–227

Melchinger AE, Gumber RK (1998) Overview of heterosis and heterotic groups in agronomic crops. In: Lamkey KR, Staub JE (eds) Concepts and breeding of heterosis in crop plants. CSSA, Madison, pp 29–44

Menkir A, Badu-Apraku B, The C, Adepoju A (2003) Evaluation of heterotic patterns of IITA's lowland white maize inbred lines. Maydica 48:161–170

Menkir A, Melake-Berhan A, The C, Ingelbrecht I, Adepoju A (2004) Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. Theor Appl Genet 108:1582–1590

Menz MA, Klein RR, Unruh NC, Rooney WL, Klein PE, Mullet JE (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci 44:1236–1244

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. http://darwin.cirad.fr/Darwin

R Development Core Team (2011) R: a language and environment for statistical computing. R Development Core Team, Vienna. http://www.R-project.org/

Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, Vasal SK, Beck D, Bohn M, Frisch M (2003) Use of SSRs for establishing heterotic groups in subtropical maize. Theor Appl Genet 107:947–957

Reif JC, Hallauer AR, Melchinger AE (2005) Heterosis and heterotic patterns in maize. Maydica 50:215

Reif JC, Zhao Y, Würschum T, Gowda M, Hahn V (2013) Genomic prediction of sunflower hybrid performance. Plant Breed 132:107–114

Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM, Elshire RJ, Acharya CB, Mitchell SE, Flint-Garcia SA, McMullen MD, Holland JB, Buckler ES, Gardner CA (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol 14:R55

SAS Institute Inc (2012) SAS OnlineDoc® 9.3. SAS Institute Inc, Cary

Semagn K, Magorokosho C, Vivek BS, Makumbi D, Beyene Y, Mugo S, Prasanna BM, Warburton ML (2012) Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. BMC Genom 13(1):113

Singh VK, Upadhyay P, Sinha P, Mall AK, Ellur RK, Singh A, Jaiswal SK, Biradar S, Ramakrishna S, Sundaram RM (2011) Prediction of hybrid performance based on the genetic distance of parental

lines in two-line rice (*Oryza sativa* L.) hybrids. J Crop Sci Biotechnol 14:1–10

Sun CQ, Jiang TB, Chen L, Wu CM, Li ZC, Wang XK (2000) Studies on the relation between heterosis and genetic differentiation in hybrid rice (*Oryza sativa* L.). Acta Agron Sin 26:641–649

Suwarno WB, Pixley KV, Palacios-Rojas N, Kaeppler SM, Babu R (2014) Formation of heterotic groups and understanding genetic effects in a provitamin A biofortified maize breeding program. Crop Sci 54:14–24

Tams SH, Bauer E, Oettler G, Melchinger AE (2004) Genetic diversity in European winter triticale determined with SSR markers and coancestry coefficient. Theor Appl Genet 108(7):1385–1391

Wang S, Lu Z (2006) Genetic diversity among parental lines of *Indica* hybrid rice (*Oryza sativa* L.) in China based on coefficient of parentage. Plant Breed 125:606–612

Wang K, Qiu FL, Dela Paz MA, Zhuang JY, Xie FM (2013) Genetic diversity and structure of improved *indica* rice germplasm. Plant Genet Resour Charact Util 1–7

Xie C, Warburton M, Li M, Li X, Xiao M, Hao Z, Zhao Q, Zhang S (2008) Retracted article: an analysis of population structure and linkage disequilibrium using multilocus data in 187 maize inbred lines. Mol Breed 21(4):407–418

Xie FM, Guo LB, Ren GJ, Hu PS, Wang F, Xu JL, Li XQ, Qiu FL, Dela Paz MA (2012) Genetic diversity and structure of indica rice varieties from two heterotic pools of southern China and IRRI. Plant Genet Resour Charact Util 10(3):186–193

Xie FM, He ZZ, Esguerra MQ, Qiu FL, Ramanathan V (2014) Determination of heterotic groups for tropical Indica hybrid rice germplasm. Theor Appl Genet 127:407–417

Xu W, Virmani SS, Hernandez JE, Sebastian LS, Redoña ED, Li Z (2002) Genetic diversity in the parental lines and heterosis of the tropical rice hybrids. Euphytica 127:139–148

Zhang Q, Gao YJ, Maroof MS, Yang SH, Li JX (1995) Molecular divergence and hybrid performance in rice. Mol breed 1:133–142

Zhao MF, Li XH, Yang JB, Xu CG, Hu RY, Liu DJ, Zhang Q (1999) Relationship between molecular marker heterozygosity and hybrid performance in intra-and inter-subspecific crosses of rice. Plant Breed 118:139–144