

Genetic-basis analysis of heterotic loci in Dongxiang common wild rice (*Oryza rufipogon* Griff.)

XIAO-JIN LUO, XIAO-YUN XIN AND JIN-SHUI YANG*

State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai 200433, P. R. China

(Received 9 January 2012; revised 15 February 2012; accepted 22 February 2012)

Summary

Heterosis is widely used in genetic crop improvement; however, the genetic basis of heterosis is incompletely understood. The use of whole-genome segregating populations poses a problem for establishing the genetic basis of heterosis, in that interactions often mask the effects of individual loci. However, introgression line (IL) populations permit the partitioning of heterosis into defined genomic regions, eliminating a major part of the genome-wide epistasis. In our previous study, based on mid-parental heterosis (H_{MP}) value with single-point analysis, 42 heterotic loci (HLs) associated with six yield-related traits were detected in wild and cultivated rice using a set of 265 ILs of Dongxiang common wild rice (*Oryza rufipogon* Griff.). In this study, the genetic effects of HLs were determined as the combined effects of both additive and dominant gene actions, estimated from the performance values of testcross F_1 s and the dominance effects estimated from the H_{MP} values of testcross F_1 s. We characterized the gene action type at each HL. Thirty-eight of the 42 HLs were over-dominant, and in the absence of epistasis, four HLs were dominant. Therefore, we favour that over-dominance is a major genetic basis of 'wild-cultivar' crosses at the single functional Mendelian locus level.

1. Introduction

Heterosis, or hybrid vigour, has been widely used in genetic crop improvement, and the genetic basis of heterosis has also been investigated. Dominance (Davenport, 1908) and over-dominance (Shull, 1908) are two hypotheses that were proposed a century ago to explain the genetic basis of heterosis. Recent advances in genome research involving a number of molecular-marker techniques and the availability of high-density molecular linkage maps, together with developments in analytical methods (Zeng, 1994), facilitated the analysis of the genetic basis of quantitative traits. Many quantitative trait locus (QTL) mapping studies have provided insight into the genetic basis of heterosis (Xiao *et al.*, 1995; Li *et al.*, 2001, 2008; Hua *et al.*, 2003; Melchinger *et al.*, 2008; Luo *et al.*, 2009a), but reached different conclusions. One known problem in establishing the genetic basis of heterosis is the use of whole-genome

segregating populations, where interactions often mask the effects of individual loci (Semel *et al.*, 2006).

Introgression lines (ILs) are the results of using marker-assisted selection (MAS) to introgress small chromosomal segments from the donor into the recurrent parent by consecutive backcrossing and selfing (Eshed & Zamir, 1994). Any phenotypic difference between such an IL and its recurrent parent should be due to QTLs located on the introgressed segments from the donor. Consequently, ILs are a more precise estimate of the genetic effects of introgression under a relatively uniform and elite lineage background (Tanksley & Nelson, 1996). They are, therefore, well suited for use in the genetic analysis of heterosis.

Recently, we reported QTL analysis of panicle-related traits and identified heterotic loci (HLs) associated with six yield-related traits in a set of 265 ILs (Luo *et al.*, 2009b). The lines were generated from a cross between Guichao 2, a high-yield commercial *Indica* cultivar (*Oryza sativa* L.), as the recurrent parent, and a common wild rice accession collected from Dongxiang County, Jiangxi Province, China, as the donor parent. In our previous study, the mid-parental heterosis (H_{MP}) values were calculated

* Corresponding author: State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai 200433, P. R. China. E-mail: jsyang@fudan.edu.cn

as $H_{MP} = F_1 - (IL + \text{Guichao } 2)/2$. The H_{MP} values were used to identify loci affecting heterosis in the six yield-related traits. An HL is determined when a locus demonstrates a significant difference between the heterozygote and the mean of the two corresponding homozygotes; that is, the HL is a QTL for heterosis (Luo *et al.*, 2011).

In this study, based on the set of 265 ILs and 265 testcross F_1 s (derived from the ILs and the recurrent parent, Guichao 2), QTLs and HLs associated with yield-related traits in testcross F_1 s were analysed. The genetic effects and main features of the HLs were discussed.

2. Materials and methods

(i) Experimental population and field trials

The IL population comprised 265 lines carrying variant introgressed segments of Dongxiang common wild rice (*Oryza rufipogon*) collected from Dongxiang county, Jiangxi Province, in the background of an Indica (*O. sativa* L. ssp. *indica*) cultivar, Guichao 2. The F_1 population comprised 265 testcrosses derived from crosses between the 265 ILs and the recurrent parent Guichao 2. These 265 lines represented 81.5% of the *O. rufipogon* genome. The detailed characteristics of the ILs were presented in Luo *et al.* (2009b, 2011). The F_1 testcross individuals and the corresponding parental ILs were evaluated in the summer of 2004 at the Experiment Station of the China Agricultural University (ESCAU), Beijing (39°N, 116°E).

Field trials of the 265 ILs, the 265 testcross F_1 s and the recurrent parent Guichao 2 were conducted at Beijing–ESCAU in summer 2004. The detailed field trials were presented in Luo *et al.* (2011). Ten plants from each line were harvested at maturity in both the IL and F_1 populations, and the following traits were scored: the number of spikelets per panicle (SP), the number of filled grains per panicle (GP), per cent seed set (SSP), 1000-grain weight (GW), the number of panicles per plant (PP) and grain yield per plant (YP).

(ii) Data analysis

The simple sequence repeat (SSR) markers analysed in this study were taken from previous publications (McCouch *et al.*, 2002). The detailed characteristics of the SSR markers in the ILs were presented in Luo *et al.* (2009b). A total of 160 polymorphic SSR markers were used to genotype the 265 ILs and the recurrent parent Guichao 2, following Tian *et al.* (2006). The F_1 testcross genotypes were deduced based on the genotype of their corresponding parental ILs. The direct trait measurement values from the six yield-related traits obtained from the ILs were used to identify the associated QTLs. The testcross F_1 trait

measurements were used to identify loci affecting testcross F_1 performance.

Based on IL structure, QTLs can be mapped on introgressed chromosome segments. One representative marker for each specific introgressed segment was defined as a QTL (Luo *et al.*, 2011). The association between the phenotype and 160 SSR marker data was investigated by single-point analysis using the software package Map Manager QTXb17 (Manly *et al.*, 2001). The statistical a priori threshold for main effect loci was $P < 0.01$ (the probability that loci had no effect on the trait).

The genetic effects in the testcross F_1 s were defined as follows: $d = H_{MP} = [F_1 - (IL + \text{Guichao } 2)/2]$ (Luo *et al.*, 2011); the trait mean values in the testcross F_1 s were $F_1 = (a + d)$, where a is the additive effects from the performance values of testcross F_1 ; IL is the mean value for the same measured trait in the corresponding IL parent, and the homozygous IL genotype value was $2a$; here Guichao 2 genotype value was assumed as zero for a simple case; subsequently HL effects were inferred by comparisons between the genetic effects on F_1 performance and H_{MP} . HLs with the ratio between dominant and additive effects $d/a \leq 1$ were considered complete or partial dominant loci, and expected to generate an estimate of F_1 performance ($a + d$) equal to or higher than twice the H_{MP} (d). HLs with $d/a > 1$, that is, $2d$ ($2 \times H_{MP}$) $> a + d$ (F_1), or only detectable for H_{MP} , were determined as over-dominant loci (Melchinger *et al.*, 1998).

3. Results

(i) The relationships among the mean trait values of ILs, H_{MP} , and F_1 performance

The correlation coefficients between testcross F_1 mean values, H_{MP} values and parental IL mean values for yield-related traits are shown in Table 1. In general, a positive but lower correlation between IL trait values and the F_1 s was observed; the average R^2 (determination coefficients) was 0.235, which suggested that additive gene action made a small contribution to F_1 performance. A general negative correlation trend was evident between IL and H_{MP} trait values, clearly suggesting that additive and dominant gene action operated independently in the testcross population. In the 265 testcross F_1 s, a highly positive correlation between F_1 performance and H_{MP} values was observed for all traits, with an average $R^2 = 0.580$ (range of 0.428 for GW to 0.775 for YP), indicating that H_{MP} largely influenced F_1 performance.

(ii) Genetic effects of QTLs in F_1 testcross population

IL phenotypic data from six yield-related traits were used to identify the associated QTLs. Fifty-four QTLs

Table 1. Phenotypic correlation (R) and determination coefficients (R^2) for six yield-related traits between IL and testcross F_1 performance values and H_{MP} values

Trait ^a	Between the performance values of IL and testcross F_1		Between the H_{MP} values and performance values of testcross F_1		Between the performance values of IL and the H_{MP} values of testcross F_1	
	R	R^2	R	R^2	R	R^2
SP	0.530*	0.281	0.728*	0.530	-0.192*	0.037
GP	0.589*	0.347	0.732*	0.535	-0.117	0.014
SSP	0.508*	0.258	0.678*	0.459	-0.239*	0.057
GW	0.509*	0.259	0.654*	0.428	-0.362*	0.131
PP	0.375*	0.140	0.869*	0.755	-0.132	0.017
YP	0.352*	0.124	0.881*	0.775	-0.131	0.017
Mean	0.477*	0.235	0.757*	0.580	-0.196*	0.046

^a Trait abbreviations: the number of spikelets per panicle (SP), the number of filled grains per panicle (GP), per cent seed set (SSP), 1000-grain weight (GW), the number of panicles per plant (PP) and grain yield per plant (YP).

*Significance levels $P < 0.01$.

were detected (partial QTLs detected in ILs are indicated in underlined text in Table 2). Trait phenotypic values from F_1 testcrosses were used to infer the QTLs contributing to F_1 testcross performance. Fifty-one QTLs influencing F_1 testcross performance were detected for the six yield-related traits (partial QTLs detected in the F_1 testcrosses are shown in Table 2). In our previous study, the H_{MP} values in F_1 testcross were used to infer which QTLs contributed to heterosis. A total of 42 H_{MP} QTLs (or HLs) associated with H_{MP} values were detected for the six yield-related traits (Luo *et al.*, 2011, all H_{MP} QTLs are shown in Table 2).

Table 2 indicates the genetic overlap of H_{MP} and QTLs detected in the F_1 testcrosses. Of 42 H_{MP} QTLs, 21 loci were only associated with H_{MP} , and showed over-dominant expression. The other 21 loci simultaneously influenced H_{MP} and F_1 performance. A comparison of the genetic effects of loci detected in both H_{MP} and F_1 testcross performance indicated a $d/a \leq 1$ in *hsp2*, *hgp9a*, *hssp5* and *hgw8*, suggesting dominant loci, and over-dominant effects in the remaining 17 loci ($d/a > 1$). In 42 of the HLs, 38 (90.5%) were over-dominant and four appeared dominant. These results indicated that at the single locus level, HLs were predominantly over-dominant. In IL QTL analysis, nine (21.4%) of the above 42 HLs were resolved at the same statistical threshold, and showed less genetic overlap with the six yield-related trait QTLs. These results are consistent with a lower correlation between IL trait values and the corresponding F_1 s.

4. Discussion

The complex nature of heterosis makes it difficult to partition it into individual components, particularly in F_2 , backcrossed, and recombinant inbred populations;

the epistatic interactions among the many segregating loci throughout the genome makes it difficult to define specific heterotic phenotypes and the individual genomic loci that control them (Li *et al.*, 2001; Semel *et al.*, 2006; Xin *et al.*, 2011). To overcome these limitations, we developed a set of ILs carrying a few chromosome segments from the wild rice species *O. rufipogon* (Luo *et al.*, 2009b). This IL population allowed us to partition heterosis into defined genomic regions, eliminating a major part of the genome-wide epistasis. The heterotic effects were determined as the combined effects of both additive and dominant gene actions, estimated from the performance values of testcross F_1 s and the dominance effects estimated from the H_{MP} values of testcross F_1 s. Based on this strategy, we characterized the gene action type at each HL. Forty-two HLs for six yield-related traits revealed two different genetic effects: dominance or over-dominance. These HL data indicated that over-dominance was the major underlying factor of heterosis. Thirty-eight (90.5%) HLs exhibited over-dominant effects and only four HLs showed dominant effects. Notably, Semel *et al.* (2006) carried out quantitative genetic and phenotypic analyses on an IL population of tomato (*Solanum lycopersicum*) carrying a single chromosomal segment from the distantly related wild species *Solanum pennellii*. That study generated results congruent with the present study; in the absence of epistasis, at a single locus level, over-dominant loci had greater effects on tomato yield and fitness.

The exploitation of favourable genes from wild rice might further improve tolerance to biotic and abiotic stress, yield and other important agronomic traits for rice variety. Luo *et al.* (2011) investigated the HLs derived from wild rice, thought that favourable HLs capable of improving agronomic traits are available, and indicated that the identification of HLs

Table 2. The genetic effects of HLs on six yield-related traits in the testcross F_1 s (the HLs were mapped by Luo et al., 2011)

Trait ^a	HL	Marker ^b	The H_{MP}^c of F_1 s			The performance of F_1 s							
			PV ^d	P ^d	d^d	PV ^d	P ^d	a + d^d	d/a^e				
SP	<i>hsp2</i>	RM71	3	0.0082	5.45	6	0.0002	10.80	1.0				
	<i>hsp4</i>	RM307	3	0.0094	11.67								
	<i>hsp8</i>	RM210	3	0.0096	17.91								
	<i>hsp9</i>	RM342B	7	0.0001	46.75	4	0.0060	40.68	2.3				
	<i>hsp11</i>	RM224	8	0.0009	12.78	6	0.0068	17.99	1.4				
GP	<i>hgp6</i>	RM225	3	0.0058	-28.90	22	0.0000	-56.95	1.0				
	<i>hgp9a</i>	RM342B	15	0.0000	-29.18								
	<i>hgp9b</i>	OSR29	8	0.0000	11.64					6	0.0003	12.99	1.8
	<i>hgp9c</i>	OSR12	5	0.0006	11.71					4	0.0037	11.72	2.0
	<i>hgp10</i>	RM222	4	0.0099	9.71								
SSP	<i>hssp4</i>	RM307	10	0.0000	-11.02	10	0.0000	-12.98	1.7				
	<i>hssp5</i>	RM289	6	0.0002	-8.91	17	0.0000	-17.86	1.0				
	<i>hssp6</i>	RM253	9	0.0000	26.72	48	0.0000	-85.25	1.3				
	<i>hssp9a</i>	RM219	5	0.0006	20.89								
	<i>hssp9b</i>	RM342B	30	0.0000	-55.35								
	<i>hssp9c</i>	OSR29	6	0.0002	5.20								
	<i>hssp11</i>	RM21	3	0.0077	8.25								
GW	<i>hgw1</i>	RM259	4	0.0031	1.55					3	0.0061	1.62	1.9
	<i>hgw2</i>	RM263	4	0.0044	1.34					3	0.0097	1.26	2.1
	<i>hgw6a</i>	RM204	5	0.0014	1.50	4	0.0091	1.20	2.5				
	<i>hgw6b</i>	RM217	9	0.0000	3.23	5	0.0010	-1.19	1.0				
	<i>hgw6c</i>	RM253	16	0.0000	6.08								
	<i>hgw8</i>	RM223	4	0.0093	0.59								
	<i>hgw10</i>	RM222	3	0.0084	-0.89								
	<i>hgw11</i>	RM21	3	0.0097	1.30								
	PP	<i>hpp5</i>	RM233B	4	0.0053					0.85	3	0.0061	1.09
<i>hpp6a</i>		RM133	4	0.0051	-0.88					4	0.0036	2.01	2.4
<i>hpp6b</i>		RM217	4	0.0054	-2.46								
<i>hpp6c</i>		RM253	4	0.0026	-3.75								
<i>hpp7</i>		RM234	5	0.0010	2.37								
<i>hpp10</i>		RM222	4	0.0098	1.02								
<i>hpp11</i>		RM21	4	0.0029	-1.87								
YP		<i>hyp2a</i>	RM236	4	0.0089	1.73	4	0.0023	2.66				
	<i>hyp2b</i>	RM71	4	0.0049	1.91	5	0.0010	2.44	1.6				
	<i>hyp5</i>	RM274	7	0.0001	3.79	5	0.0007	4.24	1.8				
	<i>hyp7</i>	RM234	3	0.0096	6.01	12	0.0000	-22.44	1.4				
	<i>hyp9a</i>	RM219	5	0.0012	8.60								
	<i>hyp9b</i>	RM342B	7	0.0001	-15.57								
	<i>hyp9c</i>	OSR29	6	0.0002	3.42					3	0.0095	2.39	2.9
	<i>hyp10</i>	RM222	4	0.0033	4.22								
	<i>hyp12a</i>	RM235	4	0.0093	2.08					4	0.0027	3.07	1.4
	<i>hyp12b</i>	RM17	4	0.0088	-3.01								

^a See Table 1 for abbreviations.

^b Markers indicated in underlined text are QTLs identified in ILs.

^c H_{MP} is the mid-parental heterosis of testcross F_1 .

^d PV, the phenotypic variance explained by the locus; P, the probability that the marker genotype had no effect on the trait; a + d, the additive and dominance effects from the performance values of testcross F_1 ; d, the dominance effect from the H_{MP} values.

^e d/a , the ratio between dominant and additive effects.

between wild rice and cultivated rice could lead to a new strategy for the application of heterosis in rice breeding. It is generally known that the over-dominant HLs is more advantageous than dominant HLs in heterosis utilization. In this study, 28 (66.7%) of 42 HLs showed significantly positive over-dominant

effects ($P < 0.01$) on yield-related traits, suggesting that these markers are viable candidates for marker-aided improvement of rice yield potential.

Previous study (Li *et al.*, 2001) revealed that back-cross F_1 performance was largely determined by dominant gene action. Mei *et al.* (2005) analysed the

correlation between RILs and backcrossed populations for agricultural traits, and considered backcross F_1 performance was mainly determined by non-additive gene action. Our study employed a similar experimental design and found a highly positive correlation between testcross F_1 and H_{MP} F_1 testcross performance values, and a lower positive correlation between IL performance values and F_1 testcross performance values (Table 1). These results indicated that dominance gene action rather than additive gene action was a substantial contributor to F_1 testcross performance. Furthermore, the testcross described in this study corresponds to previous backcross studies (Li *et al.*, 2001; Mei *et al.*, 2005). The negative correlation between IL performance values and H_{MP} values of the F_1 testcross population clearly indicated that additive and dominant gene action acted independently in the testcross population. QTL and HL analyses demonstrated that nine of the 42 HLs were also detected in the QTL analysis (Table 2), and exhibited less genetic overlap with QTLs, consistent with results reported by Hua *et al.* (2003). Therefore, heterosis and trait performance may be conditioned by different sets of loci.

We thank the anonymous referees for their critical comments on this manuscript. This research was supported by a grant (no. 2007CB109002) from the National Basic Research Program of China (973 program).

References

- Davenport, C. B. (1908). Degeneration, albinism and inbreeding. *Science* **28**, 454–455.
- Eshed, Y. & Zamir, D. (1994). Introgressions from *Lycopersicon pennellii* can improve the soluble-solids yield of tomato hybrids. *Theoretical and Applied Genetics* **88**, 891–897.
- Hua, J. P., Xing, Y. Z., Wu, W. R., Xu, C. G., Sun, X. L., Yu, S. B. & Zhang, Q. F. (2003). Single-locus heterotic effects and dominance by dominance interaction can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences USA* **100**, 2574–2579.
- Li, L. Z., Lu, K. Y., Chen, Z. M., Mu, T. M., Hu, Z. L. & Li, X. Q. (2008). Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* **180**, 1725–1742.
- Li, Z. K., Luo, L. J., Mei, H. W., Shu, Q. Y., Tablen, R., Zhong, D. B., Ying, C. S., Stansel, J. W., Khush, G. S. & Paterson, A. H. (2001). Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**, 1737–1703.
- Luo, X. J., Fu, Y. C., Zhang, P. J., Wu, S., Tian, F., Liu, J. Y., Zhu, Z. F., Yang, J. S. & Sun, C. Q. (2009a). Additive and over-dominant effects resulting from epistatic loci are the primary genetic basis of heterosis in rice. *Journal of Integrative Plant Biology* **51**, 393–408.
- Luo, X. J., Tian, F., Fu, Y. C., Yang, J. S. & Sun, C. Q. (2009b). Mapping QTL influencing panicle-related traits from Chinese common wild rice (*Oryza rufipogon* Griff.) using introgression lines. *Plant Breeding* **128**, 559–567.
- Luo, X. J., Wu, S., Tian, F., Za, X. J., Dong, X. X., Fu, Y. C., Wang, X. K., Yang, J. S. & Sun, C. Q. (2011). Identification heterotic loci associated with yield-related traits derived from Chinese common wild rice (*Oryza rufipogon* Griff.). *Plant Science* **181**, 14–22.
- Manly, K. F., Cudmore, J. R. H. & Meer, J. M. (2001). Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* **12**, 930–932.
- McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D. & Stein, L. (2002). Development of 2,240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* **9**, 199–207.
- Mei, H. W., Li, Z. K., Shu, Q. Y., Guo, L. B., Wang, Y. P., Yu, X. Q., Ying, C. S. & Luo, L. J. (2005). Gene actions of QTLs affecting several agronomic traits resolved in a recombinant inbred rice population and two backcross populations. *Theoretical and Applied Genetics* **110**, 649–659.
- Melchinger, A. E., Utz, H. F. & Schön, C. C. (1998). Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **149**, 383–403.
- Melchinger, A. E., Utz, H. F. & Schön, C. C. (2008). Genetic expectations of quantitative trait loci main and interaction effects obtained with the triple testcross design and their relevance for the analysis of heterosis. *Genetics* **178**, 2265–2274.
- Semel, Y., Nissenbaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., Pleban, T., Lippman, Z., Gur, A. & Zamir, D. (2006). Overdominant quantitative trait loci for yield and fitness in tomato. *Proceedings of the National Academy of Sciences USA* **103**, 12981–12986.
- Shull, G. H. (1908). The composition of a field of maize. *American Breeders Association Report* **4**, 296–301.
- Tanksley, S. D. & Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics* **92**, 191–203.
- Tian, F., Li, D. J., Fu, Q., Zhu, Z. F., Fu, Y. C., Wang, X. K. & Sun, C. Q. (2006). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*O. sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theoretical and Applied Genetics* **112**, 570–580.
- Xiao, J. H., Li, J., Yuan, L. P. & Tanksley, S. D. (1995). Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140**, 745–754.
- Xin, X. Y., Wang, W. X., Xin, X. Y., Wang, W. X., Yang, J. S. & Luo, X. J. (2011). Genetic analysis of heterotic locus detected in a cross between *Indica* and *Japonica* in *O. sativa* L. *Breeding Science* **61**, 380–388.
- Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468.