

SCIENTIFIC REPORTS



OPEN

Exploitation of heterosis loci for yield and yield components in rice using chromosome segment substitution lines

Received: 20 April 2016
Accepted: 21 October 2016
Published: 11 November 2016

Yajun Tao^{1,*}, Jinyan Zhu^{2,*}, Jianjun Xu¹, LiuJun Wang¹, Houwen Gu¹, Ronghua Zhou¹, Zefeng Yang¹, Yong Zhou¹ & Guohua Liang¹

We constructed 128 chromosome segment substitution lines (CSSLs), derived from a cross between *indica* rice (*Oryza sativa* L.) 9311 and *japonica* rice Nipponbare, to investigate the genetic mechanism of heterosis. Three photo-thermo-sensitive-genic male sterile lines (Guangzhan63-4s, 036s, and Lian99s) were selected to cross with each CSSL to produce testcross populations (TCs). Field experiments were carried out in 2009, 2011, and 2015 to evaluate yield and yield-related traits in the CSSLs and TCs. Four traits (plant height, spikelet per panicle, thousand-grain weight, and grain yield per plant) were significantly related between CSSLs and TCs. In the TCs, plant height, panicle length, seed setting rate, thousand-grain weight, and grain yield per plant showed partial dominance, indicating that dominance largely contributes to heterosis of these five traits. While overdominance may be more important for heterosis of panicles per plant and spikelet per panicle. Based on the bin-maps of CSSLs and TCs, we detected 62 quantitative trait loci (QTLs) and 97 heterotic loci (HLs) using multiple linear regression analyses. Some of these loci were clustered together. The identification of QTLs and HLs for yield and yield-related traits provide useful information for hybrid rice breeding, and help to uncover the genetic basis of rice heterosis.

Food shortages are becoming a serious global problem because of the increasing population and decreasing area of available arable land. The world's population is expected to reach 9 billion by 2050¹, and the area of farmland per capita will be just 0.151 ha. Rice is one of the most important staple foods, feeding half of the world's population. Improvements to rice yield will significantly contribute to addressing food security issues. Since the 1970s, the utilization of heterosis in inter-subspecies has provided an avenue for rice breeding^{2,3}. The yield production reached 10.5 t/ha in 'super-hybrid rice breeding' Phase I in 2000, and increased to 12 t/ha and 13.5 t/ha in 2004 and 2012, respectively⁴. In China, hybrid varieties account for approximately 50% of all cultivated rice⁵.

Heterosis, also known as hybrid vigor, is defined as the superior performance of a hybrid relative to its parents, including yield and adaptability to abiotic/biotic environmental conditions⁶. Although this phenomenon is widely acknowledged in animal and plant breeding⁷⁻⁹, the genetic basis of heterosis remains unclear. Three classical hypotheses, including dominance¹⁰⁻¹³, overdominance¹⁴⁻¹⁶ and epistasis^{17,18} have been proposed to explain heterosis. Molecular markers and genetic linkage maps provide a means to identify loci for heterosis. Quantitative trait loci (QTLs) mapping has proved to be an informative approach for dissecting complex traits and heterosis in rice and other crops, including maize, cotton, and oilseed rape, and many QTLs for yield and other important agronomic traits have been detected. In earlier studies, some non-permanent segregating populations such as F₂, F_{2,3}, BC₁F₂ were used to reveal the genetic basis of heterosis. However, the phenotypic data from these populations are not be allowed to be repeated under multiple environmental conditions. To address this issue, some permanent populations, such as immortalized F₂ and recombinant inbred lines (RILs), were developed to investigate heterosis loci.

¹Jiangsu Key Laboratory of Crop Genetics and Physiology/Co-Innovation Center for Modern Production Technology of Grain Crops, Key Laboratory of Plant Functional Genomics of the Ministry of Education, Yangzhou University, Yangzhou 225009, China. ²Institute of Food Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Y.Z. (email: zhoyong@yzu.edu.cn) or G.L. (email: ricegb@yzu.edu.cn)

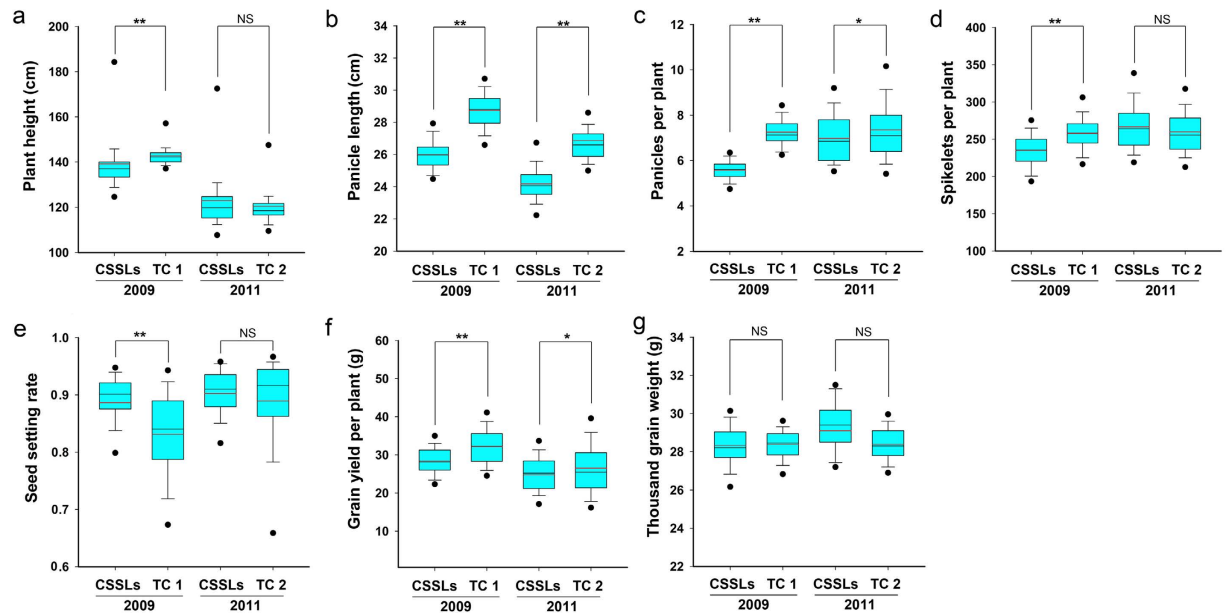


Figure 1. Means and ranges of yield-related traits measured in CSSLs, TC 1 in 2009, and TC 2 in 2011. (* $p \leq 0.05$, ** $p \leq 0.01$, NS not significant).

A more accurate estimate of the effects of QTLs, including those for heterosis, can be obtained using advanced populations such as near-isogenic lines (NILs) and chromosome segment substitution lines (CSSLs). For instance, 21 heterosis loci with significant effects on hybrids were identified in rice using 66 CSSLs and their 66 corresponding F_1 plants with the recurrent parent¹⁹. Fifteen QTLs contributing to heterosis of plant height were identified by comparing performance among 15 CSSLs, their hybrids, and the recurrent parent²⁰. Recently, combining QTL mapping and transcriptome profiling, multiple QTLs for yield-component heterosis were characterized²¹. Among these QTLs, one major QTL *RH8* (*rice heterosis 8*) was cloned. Interestingly, *Ghd7*, which also controlled rice flowering-time, has been found contributed to heterosis in hybrid as well²². These results help to clarify the genetic basis of rice heterosis.

In the application of rice heterosis, the F_1 offspring originate from hybridization between restorer lines/male parents and male sterile lines/female parents, rather than the parents themselves. Therefore, we generated three TC (testcross) populations by mating three photo-thermo-sensitive-genic male sterile (PTGMS) lines with a set of whole-genome-resequencing-genotyped CSSLs. The phenotypes of these populations were estimated and used to investigate heterosis loci for seven yield and yield-related traits through QTLs analyses.

Results

Performance of the populations. To identify QTLs for rice heterosis, three PTGMS lines (Guangzhan63–4s, 036s, and Lian99s) were used as females and were crossed with the 128 CSSLs to generate four test-cross populations (TCs): TC 1 (Guangzhan63-4s \times CSSLs in 2008), TC 2 (Guangzhan63-4s \times CSSLs in 2010), TC 3 (036s \times CSSLs in 2014), and TC 4 (Lian99s \times CSSLs in 2014). We also generated F_1 hybrids of Guangzhan63-4s \times 9311, 036s \times 9311, and Lian99s \times 9311, which were tested along with the two parents.

Seven yield-related agronomic traits of the CSSLs and four TC F_1 hybrids were measured in 2009, 2011, and 2015. The seven quantitative traits investigated were plant height (PH), panicle length (PL), panicle per plant (PPP), spikelet per panicle (SPP), seed setting rate (SSR), grain yield per plant (GYPP), and thousand-grain weight (TGW). The means and ranges of these seven quantitative traits are summarized in Figs 1 and 2. Genetic variations were observed for all traits in all three data sets.

In 2009, except for PH and SSR, the performance of all traits showed approximately normal distributions in the CSSLs. In the TC 1 population, PL, SPP, SSR, GYPP, and TGW data showed normal distributions (Table S1). We compared the phenotypic values between CSSLs and TC 1 hybrids. The means of most traits, except for SSR and TGW, in the TC 1 population were higher than the corresponding values in the CSSLs, suggesting that the heterozygous F_1 was superior to the corresponding parental lines (Fig. 1, Table S4). In 2011, PL, SPP, PPP, and GYPP data showed normal distributions in the CSSLs. The PPP, SPP, GYPP, and TGW data showed normal distributions in the TC 2 population (Table S2). In addition, the mean values of PL, PPP, and GYPP in TC 2 were higher than those in the CSSLs in 2011 (Fig. 1, Table S5). Based on the combination of data from 2009 and 2011, we concluded that Guangzhan 63-4s might have a high general combining ability.

The data collected in 2015 were subjected to Kolmogorov–Smirnov (K–S) tests, which showed that only SPP, GYPP, and TGW had normal distributions in the CSSLs. However, except for PH, PPP, and SSR, the other traits showed normal distributions in the TC 3 and TC 4 populations (Table S3). Only PL and PPP showed better performance in the TC 3 and TC 4 populations than in the CSSLs. The mean values of SPP and GYPP in the CSSLs were higher than those in the TC 3 population, and lower than those in the TC 4 population (Fig. 2, Table S6).

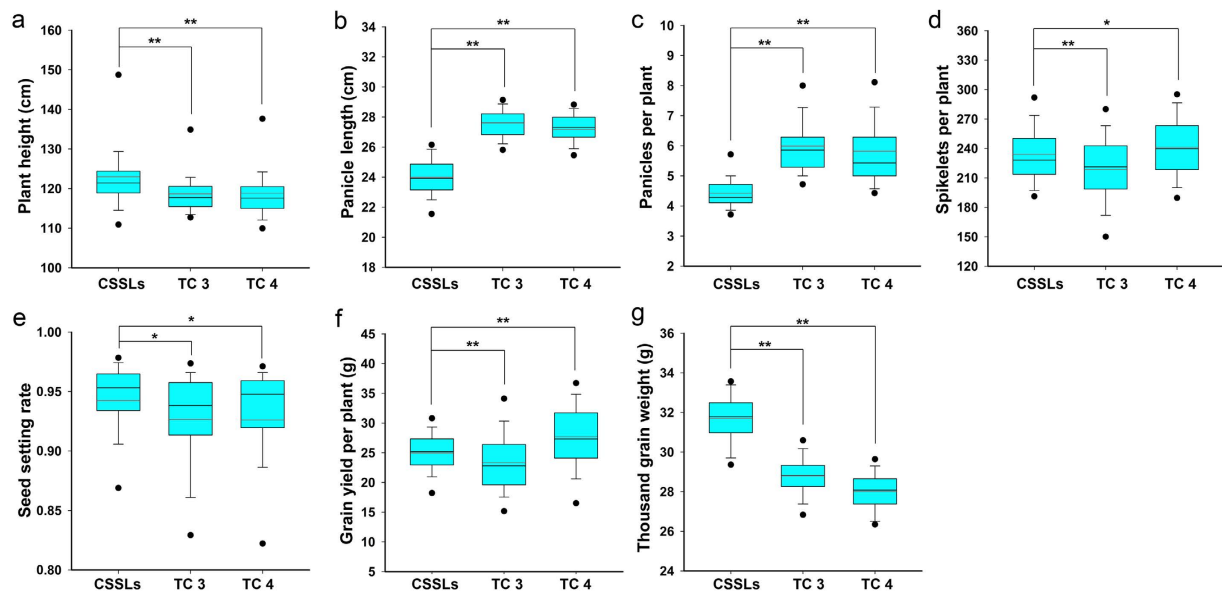


Figure 2. Means and ranges of yield-related traits measured in CSSLs, and TC 3 and TC 4 in 2015. (* $p \leq 0.05$, ** $p \leq 0.01$, NS not significant).

Traits	Population	PH	PL	PPP	SPP	SSR	GYPP
PL	CSSLs	0.564**					
	TC 1	0.401**					
PPP	CSSLs	-0.232	-0.173				
	TC 1	-0.245	-0.101				
SPP	CSSLs	0.469**	0.396**	-0.135			
	TC 1	0.514**	0.737**	-0.152			
SSR	CSSLs	0.518**	0.201	-0.110	0.143		
	TC 1	0.202	0.398**	-0.240	0.391**		
GYPP	CSSLs	0.596**	0.480**	0.184	0.553**	0.495**	
	TC 1	0.521**	0.631**	-0.012	0.665**	0.581**	
TGW	CSSLs	0.204	-0.201	0.047	-0.161	0.033	-0.071
	TC 1	0.077	0.105	-0.146	0.102	0.270	0.185

Table 1. Correlation coefficients between seven traits in CSSLs and TC 1 in 2009. *Correlation significant at the 0.05 level (two-tailed). **Correlation significant at the 0.01 level (two-tailed).

These results suggested that different PTGMS may have different combination abilities and that Guangzhan63-4s and Lian 99s have higher general combination abilities than 036s for yield production.

Relationships between traits in CSSLs and TCs. We conducted correlation analyses to determine the relationships for seven yield-related traits between the CSSLs and TCs and to discover whether each trait was correlated between TCs and their parental CSSLs. The correlation coefficients for the seven phenotypic values in CSSLs and TCs are summarized in Tables 1–3.

In the CSSL populations, significant correlations were detected between the mean values of PH and PL, and between the mean values of PH and GYPP, over the 3 years (Tables 1–3). Some traits were significantly correlated over 2 years, such as PH and SPP (2009, 2015) (Tables 1 and 3), GYPP and PPP (2011, 2015) (Tables 2 and 3), GYPP and SPP (2009, 2015) (Tables 1 and 3), and GYPP and SSR (2009, 2015) (Tables 1 and 3). Table 4 summarizes the correlation coefficients between the phenotypic values of the individual TCs and those of their parental CSSLs for the seven traits. Except for PPP and SSR, the other five traits were significantly correlated between the TCs and CSSLs. Among them, the correlation between PH and TGW was significant in all four TCs–CSSLs, while the correlation between SPP and GYPP was significant in three TCs–CSSLs.

Detection of QTLs and heterotic loci. The estimated broad sense heritability of the seven traits ranged from 21.47% to 79.85% in the TCs (Table 5). Among the seven traits, PH, PL, GYPP, and TGW showed high heritability (>50%), indicating that genetic factors played dominant roles in these four traits. The lowest broad sense heritability was for PPP, suggesting that it was strongly affected by environmental factors. We also estimated narrow sense heritability to detect the additive effects for each trait. Additive effects mainly controlled PH, GYPP, and

Traits	Population	PH	PL	PPP	SPP	SSR	GYPP
PL	CSSLs	0.187 [*]					
	TC 2	0.523 ^{**}					
PPP	CSSLs	-0.172	-0.275 ^{**}				
	TC 2	0.039	0.163				
SPP	CSSLs	-0.041	0.125	-0.067			
	TC 2	0.372 ^{**}	0.284 ^{**}	0.348 ^{**}			
SSR	CSSLs	0.005	0.100	-0.111	-0.353 ^{**}		
	TC 2	0.324 ^{**}	0.374 ^{**}	-0.168	0.263 ^{**}		
GYPP	CSSLs	0.191 [*]	-0.006	0.220 [*]	0.047	0.044	
	TC 2	0.218 [*]	0.240 [*]	0.459 ^{**}	-0.013	0.257 ^{**}	
TGW	CSSLs	-0.037	-0.080	0.116	-0.084	-0.031	0.362 ^{**}
	TC 2	0.014	-0.071	0.196 [*]	0.015	0.052	0.301 ^{**}

Table 2. Correlation coefficients between seven traits in CSSLs and TC 2 in 2011. *Correlation significant at the 0.05 level (two-tailed). **Correlation significant at the 0.01 level (two-tailed).

Traits	Population	PH	PL	PPP	SPP	SSR	GYPP
PL	CSSLs	0.507 ^{**}					
	TC 3	0.373 ^{**}					
	TC 4	0.189 [*]					
PPP	CSSLs	-0.278 ^{**}	0.051				
	TC 3	-0.066	0.208 [*]				
	TC 4	-0.061	0.326 ^{**}				
SPP	CSSLs	0.448 ^{**}	-0.153	-0.214 [*]			
	TC 3	-0.129	0.022	0.302 ^{**}			
	TC 4	0.405 ^{**}	0.059	0.059			
SSR	CSSLs	0.155	0.045	-0.173 [*]	0.002		
	TC 3	0.369 ^{**}	0.179	0.015	-0.151		
	TC 4	0.083	-0.022	-0.075	0.064		
GYPP	CSSLs	0.179 [*]	-0.058	0.259 ^{**}	0.234 ^{**}	0.210 [*]	
	TC 3	-0.060	-0.088	-0.002	-0.066	-0.010	
	TC 4	0.268 ^{**}	0.159	0.222 ^{**}	0.322 ^{**}	0.122	
TGW	CSSLs	0.100	0.058	-0.110	-0.067	0.048	0.248 ^{**}
	TC 3	0.037	0.003	-0.314 ^{**}	-0.303 ^{**}	0.020	-0.130
	TC 4	0.132	0.347 ^{**}	-0.107	-0.149	0.133	0.310

Table 3. Correlation coefficients between seven traits in CSSLs, TC 3 and TC 4 in 2015. *Correlation significant at the 0.05 level (two-tailed). **Correlation significant at the 0.01 level (two-tailed).

	PH	PL	PPP	SPP	SSR	GYPP	TGW
CSSLs – TC 1	0.761 ^{**}	0.464 ^{**}	-0.113	0.421 ^{**}	0.077	0.448 ^{**}	0.513 ^{**}
CSSLs – TC 2	0.836 ^{**}	0.092	-0.119	0.832 ^{**}	0.049	0.288 ^{**}	0.397 ^{**}
CSSLs – TC 3	0.666 ^{**}	0.062	0.022	0.048	0.051	0.252 ^{**}	0.470 ^{**}
CSSLs – TC 4	0.561 ^{**}	0.146	0.023	0.264 ^{**}	0.022	-0.020	0.327 ^{**}

Table 4. Correlation coefficients between seven traits in each CSSL and its corresponding TC. *Correlation significant at the 0.05 level (two-tailed). **Correlation significant at the 0.01 level (two-tailed).

TGW. In general, additive effects are steadily inherited. Therefore, these three traits could be effectively selected for in early generations in hybrid rice breeding.

Among the 3 years, a total of 62 QTLs associated with yield-related traits were detected in the CSSLs, and 97 heterotic loci (HLs) were identified in four TC populations. Most of these QTLs and HLs explained more than 10% of total variation (Tables S7–13).

Plant height. For PH, 20 QTLs and 21 HLs were detected in the CSSLs and TCs, respectively (Table S7). Among the detected QTLs, *qPH1-1*, which was located in bin x36, was identified in all 3 years. This region contains the semi-dwarf gene *sd1*²³. The QTL *qPH2-1*, located in bin x107 on chromosome 2, was detected in 2009 and 2015.

Trait	$h_b^2(\%)$	$h_n^2(\%)$	Additive effect	Dominant effect	Degree of dominance
PH	79.85	63.29	63.29	16.56	0.51
PL	51.22	33.95	33.95	17.27	0.71
SPP	43.49	9.72	9.72	33.77	1.86
PPP	21.47	5.21	5.21	16.26	1.77
SSR	44.91	29.38	29.38	15.53	0.73
GYPP	75.50	60.66	60.66	14.84	0.51
TGW	76.75	61.42	61.42	15.33	0.53

Table 5. Genetic parameters estimates for seven traits.

The QTL *qPH2-2*, located in bin x108 on chromosome 2, was detected in 2009 and 2015. The HL *hPH1-4*, also located in bin x36, was identified in all TCs, and the HL *hPH8-1* in bin x278 was stably identified in TC 1 and TC 2.

Panicle length. Only five QTLs for PL were detected over the 3 years. Among them, three were on chromosome 1 and the other two were on chromosome 8 (Table S8). The QTL *qPL1-1*, located in bin x36, was identified in 2009 and 2015. The interval sizes of the two QTLs on chromosome 8, *qPL8-1* and *qPL8-2*, were just 154,545 bp and 103,885 bp, respectively, illustrating the advantage of CSSLs to narrow down regions of interest for further mapping of yield-related genes. In the TC populations, 16 HLs were identified on chromosomes 1, 2, 3, 5, 6, 8, and 11 (Table S8). One HL, *hPL1-2*, located in bin x36, was identified in all three TCs.

Panicles per plant. Four QTLs and 13 HLs were identified in the CSSLs and TCs, respectively (Table S9). Except for the HLs located in bin x275 in TC 1 and TC 3, there were no repeated bins found in all populations. These results indicated that PPP is a complicated trait that is significantly affected by environmental factors. The QTL with the largest effect was mapped to bin x279 and explained 39.89% of the phenotypic variance in PPP.

Spikelets per plant. Seven QTLs were detected for SPP. The QTL *qSPP1-1*, located in bin x36, was identified in all 3 years. Among the HLs, *hSPP1-1*, also located in bin x36, was identified in three TCs, and *hSPP6-1* and *hSPP6-2* were simultaneously detected in TC 3 and TC 4 (Table S10).

Seed setting rate. We identified 17 QTLs for SSR. Only *qSSR3-4*, located in bin x154; *qSSR3-5*, located in bin x155; and *qSSR6-1*, located in bin x232, were identified in 2 years. We identified 13 HLs for SSR on seven chromosomes and there were no repeatedly detected bins (Table S11). These results indicated that SSR is a complicated quantitative trait that is affected by environmental factors.

Thousand-grain weight. For the CSSLs, *qTGW3-1*, *qTGW6-1*, *qTGW6-2*, *qTGW6-3*, and *qTGW10-1* were detected over the 3 years. Among these QTLs, *qTGW6-1* was detected in 2011 and 2015. In the TCs, 10 HLs were identified on chromosomes 2, 4, 5, 6, and 8 (Table S12).

Grain yield per plant. The GYPP is an important agronomic trait, and QTLs and HLs associated with GYPP are highly significant for rice breeding and improving rice outputs. Four QTLs for GYPP (*qGYPP1-1*, *qGYPP6-1*, *qGYPP6-2*, and *qGYPP8-1*) were identified in all 3 years (Table S13); their interval sizes were 791,655 bp, 872,289 bp, 259,355 bp, and 538,176 bp, respectively. Twelve HLs underlying GYPP were identified in the four TCs. In TC 1, *hGYPP8-1* was identified in bin x278, the same bin harboring *qGYPP8-1* (Table S13).

Some bins harbored QTLs and HLs for several traits. For example, bin x36 harbored QTLs and HLs for PH, PL, SPP, and GYPP, suggesting pleiotropic characteristics or tight linkage among genes controlling these traits.

Discussion

The grain yields of hybrid rice lines are 20~30% higher than those of conventional varieties. However, the yields have plateaued and progress in high-yielding hybrid rice breeding is slowing down. One of the key reasons is the limited genetic variability. *Indica* and *japonica* are the two main subspecies of cultivated rice. Significant molecular and physiological differences have been detected between these two rice groups^{24–26}. One study confirmed that inter-subspecific hybrid rice lines have ~25% greater yield production than that of inter-varietal hybrid rice²⁷. Consequently, it is important to identify QTLs for yield and yield-related traits in *indica* and *japonica* rice, and then pyramid them to advance hybrid rice breeding.

The production of CSSLs is a powerful strategy for precise QTLs mapping in a genomic region and to validate QTLs in earlier or segregating generations. CSSLs give a more accurate estimate of genetic effects in a specific background. Previously, we developed a set of CSSLs derived from the crossing and back-crossing of two sequenced rice cultivars: 9311, an elite *indica* cultivar, as the recipient, and Nipponbare, a *japonica* cultivar, as the donor. The 9311 variety is an elite line with excellent yield, quality, and resistance, and it has been widely used as the male parent of two-line hybrid rice in China. As a result, any improved lines in the CSSLs population due to the introgressed Nipponbare chromosome segments can be used directly to generate hybrid rice.

The genetic basis of rice heterosis is complicated, and many attempts have been made to identify key and valuable loci for hybrid breeding. In this study, to accurately identify useful loci for rice heterosis, three PTGMS

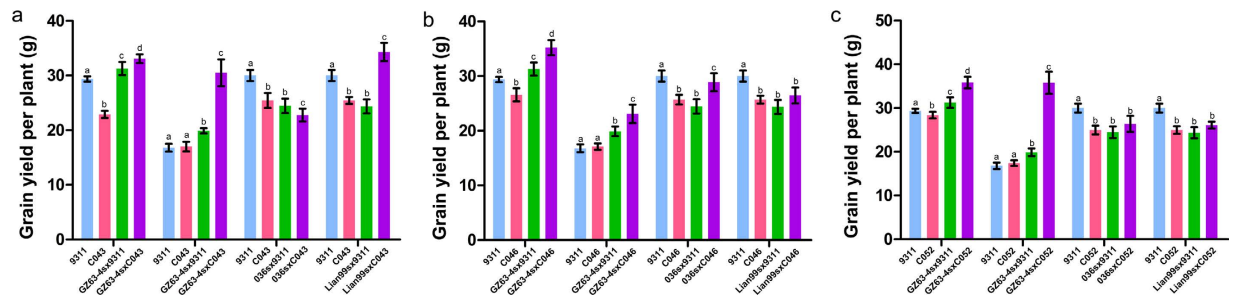


Figure 3. Phenotypic variation of GYPP in 9311, CSSLs, and their F₁ hybrids. (a) Phenotypic variation of GYPP in 9311, C043, and their F₁ hybrids; (b) Phenotypic variation of GYPP in 9311, C046, and their F₁ hybrids; phenotypic variation of GYPP in 9311, C052, and their F₁ hybrids. Different letters following mean values indicate significant difference ($P \leq 0.05$, Tukey's honestly significant difference test. GZ63-4s, Guangzhan 63-4s.

lines were crossed with 128 CSSLs and 9311 to develop four sets of testcross populations. We totally found 62 QTLs and 97 HLs, and additional studies will be needed to fine-map these genes. Analyses of phenotypic data for seven traits across 3 years indicated that four traits (PH, SPP, GYPP, and TGW) were significantly related between CSSLs and TCs. In all four crosses, PH showed the highest correlation, suggesting that plant height of hybrid rice depends largely on male parent. However, it was difficult to predict the performance of PPP and SSR. Rice heterosis is usually detected in terms of yield production. As summarized in Table 4, the performance of the F₁ hybrid was closely linked to that of the parent lines. That is, better parents, better hybrids. However, the grain yields of several combinations (Guangzhan63-4s × C043, Guangzhan63-4s × C046, and Guangzhan63-4s × C052) were significantly higher than that of Guangzhan63-4s × 9311 (Fig. 3). In contrast, the grain yield per plant in C043, C046, and C052 was lower than in 9311. This result highlights the complexity of rice heterosis, and indicates that overdominance may play an important role in specific rice hybrids.

Variance analysis showed that the average broad sense heritability of PH, TGW, and GYPP in four TCs was 79.85%, 76.75%, and 75.50%, respectively. Five traits (PH, PL, SSR, GYPP, and TGW) showed partial dominance, indicating that partial dominance is an important contributor to heterosis for yield-traits in rice (Table 5). Very recently, Huang *et al.*²⁸ used integrated method to uncover the mechanisms of heterosis in different hybrid systems and found that the heterozygous state of these grain-yield QTLs always show positive partial dominance, which supports our results. Zhou *et al.*²⁹ reported that overdominance/pseudo-overdominance is the most important contributor to heterosis of yield and number of grains per panicle. In our research, we also found that SPP showed overdominance, which is consistent with Zhou's results.

The QTLs and HLs analyses were conducted for seven yield-related traits in 2009, 2011, and 2015. Thus, it is time-consuming and laborious to detect stable and useful QTLs and HLs. Encouragingly, we identified some stable QTLs in all 3 years, which suggested the possibility of using a marker-based strategy to identify superior parents of inter-subspecies hybrids. Among them, a large-effect QTL was located in bin x36, the region harboring the rice 'green revolution' gene *sd1*³⁰. The *SD1* gene encodes a defective enzyme in the GA-biosynthetic pathway. The utilization of *sd1* in cereal crops has enhanced resistance to lodging and increased the harvest index, ultimately increasing yield and production. We also found some HLs in more than one TC population, at identical sites or at nearby intervals on the same chromosomes. Interestingly, some detected QTLs were also identified in TC populations. These results showed that there is indeed a significant relationship between TC hybrid and the parental lines, which is consistent with phenotypic correlation analyses. Therefore, mapping of QTLs and HLs is a feasible strategy that should receive more attention. As shown in Fig. 4, we observed that some QTLs and HLs had a clustered distribution (near RM11935 on chromosome 1, near RM17811 on chromosome 5, near RM19713 on chromosome 6, and near RM22271 on chromosome 8). These QTLs and/or HLs 'hot spots' should be analyzed in further studies.

Determining the mechanism of heterosis is a worldwide challenge in biology, mainly because heterosis is a complex biological phenomenon controlled by multiple genes and the heterosis of different traits is affected by combinations of partially non-redundant loci. To resolve the molecular biological basis of heterosis, we should concentrate on particular traits, such as yield-related traits, and identify their major HLs. Using near isogenic lines or CSSLs containing these QTLs as materials could remove the genetic background noise. The development of plant genomics and cloning of heterosis-related QTLs or genes will have important driving effects on studies of the mechanisms of heterosis.

Materials and Methods

Plant materials. The population comprising 128 CSSLs was developed from a cross between the *japonica* cultivar Nipponbare as the donor, and the *indica* cultivar 9311 as the recurrent parent. Each CSSL line was genotyped, and a high-quality physical map of ultrahigh-density single nucleotide polymorphisms (SNPs) based on whole-genome re-sequencing data was constructed³¹. Each CSSL was genotyped on the basis of the physical locations and genotypes of SNPs. A physical map of the 128 CSSLs was constructed and was used for QTL mapping of yield-related traits in rice. Three photo-thermo-sensitive-genic male sterile (PTGMS) lines (Guangzhan63-4s, 036s, and Lian99s) were selected as females and were crossed with the 128 CSSLs to generate the following four TCs: TC 1 (Guangzhan63-4s × CSSLs in 2008), TC 2 (Guangzhan63-4s × CSSLs in 2010), TC 3 (036s × CSSLs

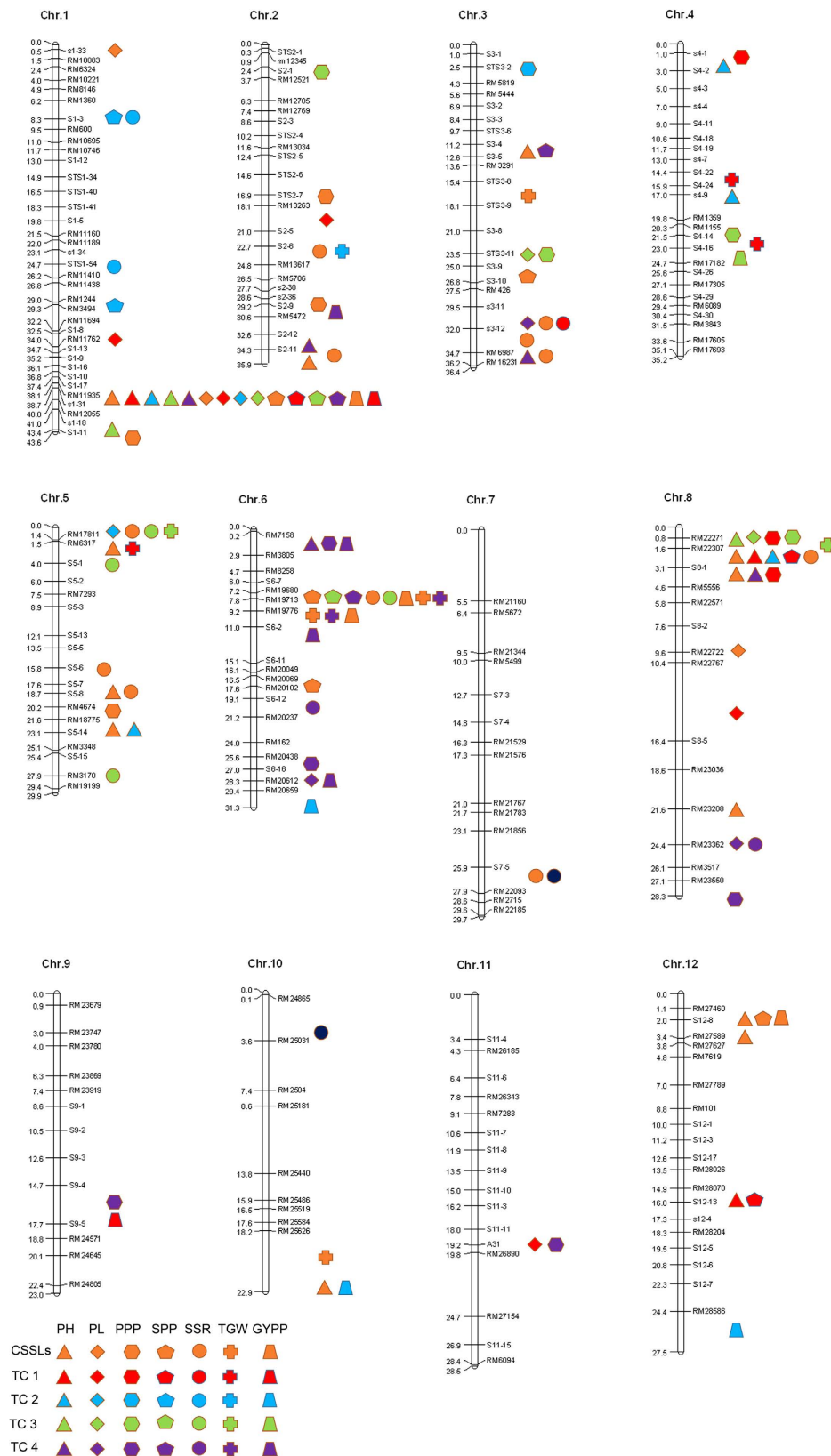


Figure 4. Positions of QTLs detected in CSSLs and four TCs.

in 2014), and TC 4 (Lian99s × CSSLs in 2014). The F₁ hybrids of Guangzhan63-4s × 9311, 036s × 9311, and Lian99s × 9311, along with the two parents, were included in the yield evaluations.

Phenotypic evaluation. The TC populations were generated in the summers of 2008, 2010, and 2014, and the phenotypic evaluations were conducted in the summers of 2009, 2011, and 2015 at the experimental field of Yangzhou University (Jiangsu, China). Four TCs, the CSSLs, the F₁ hybrids (Guangzhan63-4s × 9311, 036s × 9311, and Lian99s × 9311) and the parental lines (9311 and Nipponbare) were planted with two replications. Each plot consisted of four rows, and each row had 10 plants. Materials were sown on May 14th in each year. Thirty-day-old seedlings were transplanted to an experimental field in which the plants were spaced at a distance of (26.4 + 16.5) × 13.2 cm. The recommended agronomic practices for hybrid rice were applied in the experimental plots: the middle five plants in the central row of each plot were used for data collection. The seven quantitative traits investigated were PH, PL, PPP, SPP, SSR, GYPP, and TGW.

Data analysis, QTLs and HLs mapping. The means and standard deviations of the values of the seven phenotypic traits among TCs and CSSLs were analyzed using Microsoft Excel 2013. Statistical analyses, including correlation analyses, *t*-tests, and K-S tests were conducted using SPSS version 13.0 (SPSS, Inc., Chicago, IL, USA). Box-plots of the seven traits among TCs and CSSLs were drawn using SigmaPlot version 10.0 (Systat Software, Inc., CA, USA).

Analyses of QTLs and HLs were conducted as described previously³¹. The following multiple linear model was used:

$$y_i = b_0 + \sum_{k=1}^m b_k x_{ik} + e_i$$

where y_i is the mean value of the i th line of the CSSL population and the heterosis value (HV) of the i th line of the TC population. The definition of HV is $HV = (F_1 - CK)/CK$, where F_1 is the mean value of the i th line of the TC population and CK is the mean value of "PTGMS × 9311". b_0 is the overall mean of the population, m is the total number of bins in the whole genome, and b_k is the main effect associated with bin k . x_{ik} is an indicator variable, denoting $x_{ik} = 1$ for the donor parent bin and $x_{ik} = -1$ for the recurrent parent bin. e_i denotes the residual error following a normal distribution. Contributions of the target bins to phenotypic variation were estimated using a multiple linear regression analysis with the stepwise option in the REG procedure of the SAS software package. Analysis of variance (ANOVA) was conducted for hybrids, where variance components, broad sense heritability, and narrow sense heritability were estimated according to the model described by Mo³² (Table S14).

References

- Gregory, P. J. & George, T. S. Feeding nine billion: the challenge to sustainable crop production. *J. Exp. Bot.* **62**, 5233–5239 (2011).
- Yuan, L. P. In "Advances in hybrid rice technology" *virmani, s.s., e.a. siddiq and k.muralidharan* (eds). International Rice Research Institute, Los Banos, Philippines. p. 27–33 (1998).
- Peng, S., Cassman, K. G., Virmani, S. S., Sheehy, J. & Khush, G. G. Yield potential trends of tropical rice since release of IR8 and the challenge of increasing rice yield potential. *Crop Sci.* **39**, 1552–1559 (1999).
- Li, J. M., Xin, Y. Y. & Yuan, L. P. Hybrid rice technology development: Ensuring China food security. *International food policy research institute* (2009).
- Wen, J. *et al.* Genetic dissection of heterosis using epistatic association mapping in a partial NCII mating design. *Sci. Rep.* **5**, 18376 (2015).
- Fu, D. H. *et al.* What is crop heterosis: new insights into an old topic. *J. Appl. Genet.* **56**, 1–13 (2015).
- Ferreira, V. C., Rosa, G. J. M., Berger, Y. M. & Thomas, D. L. Survival in crossbred lambs: Breed and heterosis effects. *J. Anim. Sci.* **93**, 912–919 (2015).
- Dezetter, C. *et al.* Inbreeding and crossbreeding parameters for production and fertility traits in Holstein, Montbeliarde, and Normande cows. *J. Dairy Sci.* **98**, 4904–4913 (2015).
- Dan, Z. W. *et al.* Balance between a higher degree of heterosis and increased reproductive isolation: a strategic design for breeding inter-subspecific hybrid rice. *PLoS One* **9**, e93122 (2014).
- Bruce, A. B. The Mendelian theory of heredity and the augmentation of vigor. *Science* **32**, 627–628 (1910).
- Keeble, F. & Pellew, C. The mode of inheritance of stature and of time of flowering in peas (*Pisum sativum*). *J. Genet.* **1**, 47–56 (1910).
- Jones, D. F. Dominance of linked factors as a means of accounting for heterosis. *Genetics* **2**, 466–479 (1917).
- Collins, G. N. Teosinte in Mexico. *J. Hered.* **12**, 339–350 (1921).
- Shull, G. H. The composition of a field of corn. *Am. Breed. Assoc. Rep.* **4**, 296–301 (1908).
- Shull, G. H. Hybrid seed corn. *Science* **103**, 547–550 (1946).
- East, E. M. Inbreeding in corn. *Rep. Conn. Agric. Exp. Stn.* 419–428 (1908).
- Minvielle, F. Dominance is not necessary for heterosis: a two-locus model. *Genet. Res.* **49**, 245–247 (1987).
- Schnell, F. W. & Cockerham, C. C. Multiplicative vs. arbitrary gene action in heterosis. *Genetics* **131**, 461–469 (1992).
- Wang, Z. *et al.* Identification of *indica* rice chromosome segments for the improvement of *Japonica* inbreds and hybrids. *Theor. Appl. Genet.* **124**, 1351–1364 (2012).
- Shen, G. J., Zhan, W., Chen, H. X. & Xing, Y. Z. Dominance and epistasis are the main contributors to heterosis for plant height in rice. *Plant Sci.* **215**, 11–18 (2014).
- Li, D. *et al.* Integrated analysis of phenome, genome, and transcriptome of hybrid rice uncovered multiple heterosis-related loci for yield increase. *Proc. Natl. Acad. Sci. USA*, doi: 10.1073/pnas.1610115113 (2016).
- Xue, W. Y. *et al.* Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* **40**, 761–767 (2008).
- Sasaki, A. *et al.* Green revolution: A mutant gibberellin-synthesis gene in rice - New insight into the rice variant that helped to avert famine over thirty years ago. *Nature* **416**, 701–702 (2002).
- Zhang, Q. F., Saghai Maroof, M. A., Lu, T. Y. & Shen, B. Z. Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theor. Appl. Genet.* **83**, 495–499 (1992).
- Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S. & McCouch, S. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**, 1631–1638 (2005).
- Nakamura, Y. *et al.* The fine structure of amylopectin in endosperm from Asian cultivated rice can be largely classified into two classes. *Starch-Starke* **54**, 117–131 (2002).
- Khush, G. S. Breaking the yield frontier of rice. *GeoJournal* **35**, 329–332 (1995).
- Huang, X. *et al.* Genomic architecture of heterosis for yield traits in rice. *Nature* **537**, 629–633 (2016).

29. Zhou, G. *et al.* Genetic composition of yield heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* **109**, 15847–15852 (2012).
30. Spielmeyer, W., Ellis, M. H. & Chandler, P. M. *Semidwarf* (*sd-1*), “green revolution” rice, contains a defective gibberellin 20-oxidase gene. *Proc. Natl. Acad. Sci. USA* **99**, 9043–9048 (2002).
31. Xu, J. J. *et al.* Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *BMC Genomics* **11**, 656 (2010).
32. Mo, H. D. The analysis of combining ability in $p \times q$ mating pattern (in Chinese). *J. Jiangsu Agri. Col.* **3**, 51–57 (1982).

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (31271679), the Key Science and Technology Support Program of Jiangsu Province (BE2015341), the Natural Science Fund for Colleges and Universities in Jiangsu Province (15KJA210005), the Natural Science Foundation of Jiangsu Province (BK20150010), and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Author Contributions

Yajun Tao generated TC population in 2014, conducted phenotypic evaluations in 2015, and wrote the manuscript. Jinyan Zhu generated the CSSL populations and analyzed data. Jianjun Xu generated the CSSL and TC populations in 2008. Liujun Wang generated the TC population in 2010. Houwen Gu and Ronghua Zhou conducted phenotypic evaluations in 2015. Zefeng Yang analyzed phenotypic data. Guohua Liang and Yong Zhou designed the experiment and guided manuscript writing.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tao, Y. *et al.* Exploitation of heterosis loci for yield and yield components in rice using chromosome segment substitution lines. *Sci. Rep.* **6**, 36802; doi: 10.1038/srep36802 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016