

Exploring the genetic characteristics of 93-11 and Nipponbare recombination inbred lines based on the GoldenGate SNP assay

Renbo Yu^{1,2,3†}, Wei Yan^{2,3‡}, Manzhong Liang^{1†}, Xiaojun Dai¹, Haodong Chen², Yunong Sun², Xingwang Deng², Xiangding Chen^{1*}, Hang He^{2‡} & Liangbi Chen^{1¶}

¹College of Life Sciences, Hunan Normal University, Changsha 410081, China;

²State Key Laboratory of Protein and Plant Gene Research, Peking-Tsinghua Center for Life Sciences, School of Advanced Agriculture Sciences and School of Life Sciences, Peking University, Beijing 100871, China;

³College of Life Sciences, Capital Normal University, Beijing 100048, China

Received March 2, 2016; accepted April 28, 2016; ; published online June 14, 2016

Understanding genetic characteristics in rice populations will facilitate exploring evolutionary mechanisms and gene cloning. Numerous molecular markers have been utilized in linkage map construction and quantitative trait locus (QTL) mappings. However, segregation-distorted markers were rarely considered, which prevented understanding genetic characteristics in many populations. In this study, we designed a 384-marker GoldenGate SNP array to genotype 283 recombination inbred lines (RILs) derived from 93-11 and Nipponbare *Oryza sativa* crosses. Using 294 markers that were highly polymorphic between parents, a linkage map with a total genetic distance of 1,583.2 cM was constructed, including 231 segregation-distorted markers. This linkage map was consistent with maps generated by other methods in previous studies. In total, 85 significant quantitative trait loci (QTLs) with phenotypic variation explained (PVE) values $\geq 5\%$ were identified. Among them, 34 QTLs were overlapped with reported genes/QTLs relevant to corresponding traits, and 17 QTLs were overlapped with reported sterility-related genes/QTLs. Our study provides evidence that segregation-distorted markers can be used in linkage map construction and QTL mapping. Moreover, genetic information resulting from this study will help us to understand recombination events and segregation distortion. Furthermore, this study will facilitate gene cloning and understanding mechanism of inter-subspecies hybrid sterility and correlations with important agronomic traits in rice.

GoldenGate SNP assay, linkage map, segregation-distorted markers, QTLs

Citation: Yu, R., Yan, W., Liang, M., Dai, X., Chen, H., Deng, X., Chen, X., He, H. and Chen, L. (2016). Exploring the genetic characteristics of 93-11 and Nipponbare recombination inbred lines based on the GoldenGate SNP assay. *Sci China Life Sci* 59, 700–708. doi: 10.1007/s11427-016-5082-x

INTRODUCTION

As one of the crops with the largest planted area and yield worldwide, rice (*Oryza sativa* L.) plays an important role in food security. During the domestication of rice, cultivated rice has experienced several large genetic differentiation events involved in adaptation to various environments under

both natural and artificial selection pressures. Rice consists of two major subspecies, *indica* and *japonica*, which were respectively designated as “Hsien” and “Keng” in China approximately 2,000 years ago (Ouyang et al., 2009). Abundant genetic diversity between subspecies, or ecotypes derived from different geographical environments, or ecotypes with different characteristics have led to numerous rice varieties worldwide (Wang et al., 2009). The *indica* and *japonica* subspecies vary in many important aspects during their growth and development process, including leaf color, apiculus hair length, and grain shape (Ouyang et al., 2009).

†Contributed equally to this work

*Corresponding author (email: xdchen@hunnu.edu.cn)

‡Corresponding author (email: hehang@pku.edu.cn)

¶Corresponding author (email: chenliangbi@126.com)

Molecular markers developed from traditional markers like restriction fragment length polymorphisms (RFLPs) and simple sequence repeats (SSRs) to single nucleotide polymorphisms (SNPs) have been widely used to study the evolution of rice varieties (McNally et al., 2009), as well as in molecular breeding of rice (Chen et al., 2011; Chen et al., 2014) and identification of agronomic trait-related loci (Wang et al., 2011). Recently, with the development of sequencing technologies and genotyping platforms (Chen et al., 2013), numerous SNP markers have been discovered among rice varieties and applied in studies using array-based platforms (Chen et al., 2014) and next-generation sequencing methods (Huang et al., 2010; Huang et al., 2012) because of their high density, cost effectiveness, and ability to yield more genetic information than using other traditional molecular markers (Singh et al., 2013).

During the rice breeding, recombinant inbred lines (RILs) are widely used to identify QTLs and phenotype-related genes (Bao et al., 2008; Jiang et al., 2011). Segregation distortion that deviates from Mendelian ratio is a ubiquitous phenomenon in major crop species, which may be induced by gametic or zygotic selection for genetic and physiological factors, leading to changes in inherited genotype frequencies from parents and phenotypes in progenies (Xu and Hu, 2009). The killer-protector system at the *S5* locus regulates fertility in *indica-japonica* hybrids (Yang et al., 2012). Chromosomes with only a few segregation-distorted loci can cause large chromosome segments or entire chromosome to be deviated from Mendelian segregation (Xu, 2008). Moreover, segregation-distorted markers are presumed to be related to QTLs controlling certain traits in selected populations (Cui et al., 2015). Using segregation-distorted markers leads to higher genetic variance and better identification of QTL than using non-distorted markers (Hackett and Broadfoot, 2003; Zhang et al., 2010). However, segregation-distorted markers are usually discarded when building linkage maps and mapping QTLs, resulting in low marker density and failure in identifying biologically interesting segments (Xu and Hu, 2009). Several efficient and powerful programs have been proposed to allow researchers to overcome these problems and include segregation-distorted markers in analysis (Cui et al., 2015; Lorieux, 2012; Shah et al., 2014; Xu, 2008).

To estimate the influences of segregation-distorted markers on the building of linkage map and QTL identification, we designed a GoldenGate SNP array containing 384 experimentally validated SNP markers to genotype 283 *O. sativa* RILs derived from 93-11 and Nipponbare crosses. Segregation-distorted markers comprised 78.57% of the 294 markers selected for the linkage map. Based on the comparative linkage map with a total genetic distance of 1,583.2 cM, 85 significant QTLs with PVE values $\geq 5\%$ were identified, of which, 34 QTLs were identical to known genes/QTLs for relevant traits, and 17 QTLs were overlapped with known sterility-related genes/QTLs. These results suggest that seg-

regation-distorted markers are useful for linkage map construction and QTL mapping, and the segregation-distorted markers might be induced by inter-subspecies hybrid sterility between *indica* and *japonica* rice. RILs with genotypes inherited from 93-11 and Nipponbare subspecies within QTLs harboring segregation-distorted markers showed significantly different phenotypes, indicating these segregation-distorted markers might be responsible for these QTLs.

RESULTS

Genotyping 93-11/Nipponbare RIL populations

In this study, an Illumina BeadXpress system consisting of 384 experimentally verified polymorphic SNP markers was utilized to genotype 298 RIL populations derived from 93-11 and Nipponbare. The genotyping results were analyzed with the Illumina GenomeStudio software and manually adjusted based on the genotypes of the four control samples (two parents and two reciprocal F1 samples) on each plate. Theoretically, the parents should be clustered into two different homozygous alleles (AA/BB), whereas the reciprocal crosses should be clustered into heterozygous allele (AB). After adjustments, 15 RILs with call rates < 0.9 were discarded (Figure S1A). Of all the 384 SNPs, 373 SNPs (97.13%) were polymorphic between 93-11 and Nipponbare. The 11 non-polymorphic SNPs, as well as 50 SNPs with MAFs < 0.05 or missing rates > 0.3 were discarded. The remaining 323 high quality SNPs were evenly distributed along each chromosome. Minor allele frequencies (MAFs) of the most of the SNPs (295 of 323) were > 0.2 (Figure S1B).

The remaining 283 high-call-rate RILs were clustered into two major groups containing 115 RILs and 168 RILs, respectively (Figure S2). Among them, 29 RILs (10.25%) inherited more than 50% of their genome from parent Nipponbare, whereas 148 RILs (52.30%) inherited more than 50% of their genome from parent 93-11 (Figure 1), suggesting a higher heritability of genome segments from 93-11 than Nipponbare among these RILs. The heterozygous ratios of the RILs ranged from 0 to 46.75% (9.37% on average), and only 32 RILs harbored more than 20% heterozygous SNPs (Figure S3).

Recombination block analysis

Genomic segments in progenies tend to be inherited from a single parent together, inducing the form of recombination blocks. It was reported that *indica-japonica* RILs harbor more recombination blocks than *indica-indica* RILs (Jia et al., 2012). In the RIL populations derived from 93-11 (*indica*) and Nipponbare (*japonica*) in this study, 15,997 recombination blocks were identified among all chromosomes, with the fewest blocks (729) identified on chromosome 10 (Figure 2A). The average physical distance per

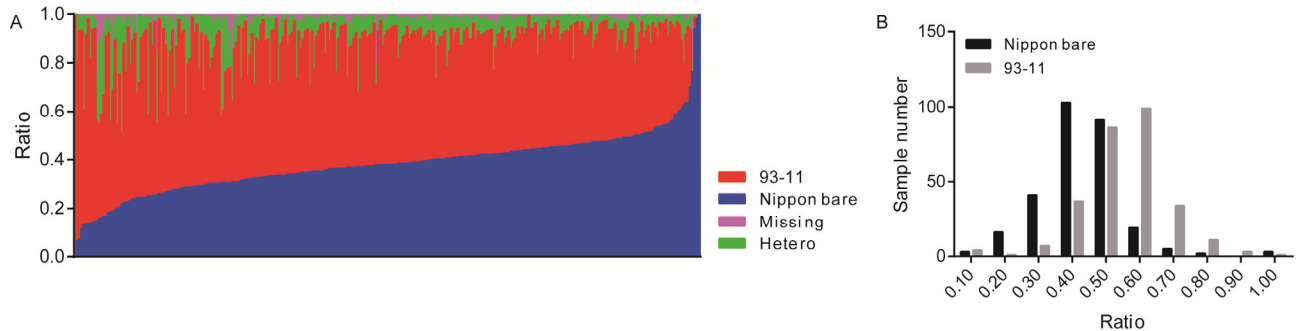


Figure 1 Inheritance ratio from Nipponbare and 93-11 for each RIL. A, Bar plot of the inheritance ratio from each parent for each RIL. Samples are sorted by the inheritance ratio from Nipponbare. B, Statistic of inheritance ratios from each parent among the RILs.

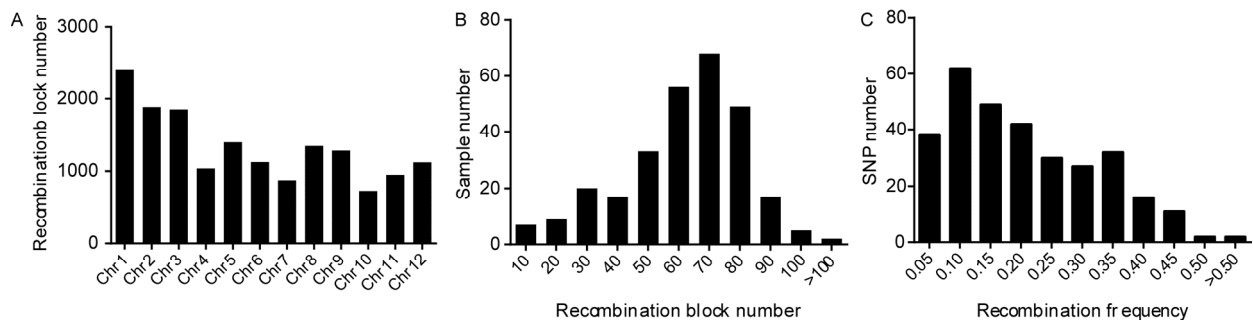


Figure 2 Statistic of recombination blocks on chromosomes (A) and RILs (B), as well as the SNP recombination frequency for SNPs (C).

recombination block on all chromosomes ranged from 17.9 kb to 34.0 kb. Among the 283 RILs, 197 RILs (69.61%) harbor more recombination blocks than average (56.5) (Figure 2B). Moreover, of all 323 high quality SNPs, we found 90 recombinant SNPs (27.86%) present in more than 25% RILs and 15 SNPs (4.64%) present in more than 40% RILs (Figure 2C, Figure S4), which can be regarded as recombination hotspots.

Segregation distortion analysis

Markers not consistent with a typical segregation ratio (1:1) among RILs are defined as segregation-distorted markers (Shah et al., 2014; Xu and Hu, 2009). To identify segregation-distorted SNPs, the chi-square (χ^2) test was utilized to estimate the ratio of 93-11 genotype to Nipponbare genotype for the retained 323 SNPs as described in Methods. In total, 259 SNPs were identified to be segregation-distorted ($P < 0.01$), accounting for 80.19% of all SNPs (Figure S5).

Reflinur et al. have applied segregation-distorted markers to build ten genetic linkage maps based on reciprocal F2 and BC1F1 mapping populations derived from Dasanbyeo (*indica*) and Ilpumbyeo (*japonica*), showing that the segregation distortion tends to be transmitted through male and female gametes in different regions which may be related to reproductive barriers (Reflinur et al., 2014). Several continuous segregation distortion regions were identified in our RIL populations, especially the regions located on chromosome 3, 6, and 7 respectively, harboring more than 50% of

the SNPs inherited 1.5 times more from 93-11 than from Nipponbare, suggesting that these regions may be transmitted through male gametes (Figure 3). On the other hand, much fewer SNPs inherited 1.5 times more from Nipponbare than from 93-11, indicating that 93-11 harbors a higher transmission frequency to progeny as the results from Reflinur et al. After comparing the regions inherited more from one parent with sterility-related genes collected in Q-taro (Yonemaru et al., 2010), 6 reported cross-incompatibility related genes were identified in these regions, especially the region located on chromosome 6 containing three hybrid sterility related genes (Table 1), suggesting that these segregation distortions might be induced by cross-incompatibility.

Linkage map construction with segregation-distorted markers

To add segregation-distorted markers to linkage map, 323 efficient SNPs among the 283 RILs were analyzed with MapDisto 1.7.7, a software capable for analyzing genetic markers that deviate from the expected Mendelian segregation ratio (Lorieux, 2012). After removing SNPs with missing rates ≥ 0.1 and genetic distance ≥ 20 cM between adjacent SNPs, 294 SNPs, including 231 distorted SNPs, were retained on the linkage map and then visualized with MapChart 2.2 (Voorrips, 2002) (Figure S6). The total genetic distance of the linkage map was 1,583.2 cM (Table S1), with an average genetic distance of 5.6 cM between adja-

Table 1 Hybrid sterility-related genes in segregation-distorted regions^{a)}

Region	Type	Gene	Description	Reference
Chr1: 7, 712, 944-8, 683, 495	Nipp	Os01g0258600	<i>DPL1 (DOPPELGANGER1)</i>	Mizuta et al., 2010
	93-11	Os04g0321500	<i>S-28 (Hybrid sterility-28)</i>	Yamagata et al., 2010
		Os06g0184100	<i>DPL2 (DOPPELGANGER2)</i>	Mizuta et al., 2010
Chr6: 4, 069, 976-7, 850, 016	93-11	Os06g0212900	<i>ORF3 (Open Reading Frame 3)</i>	Yang et al., 2012
		Os06g0213100	<i>S5 (Hybrid sterility-5)</i>	Yanagihara et al., 1995
		Os08g0404200	<i>S-27 (Hybrid sterility-27)</i>	Yamagata et al., 2010

a) The type here represents the parent genotype that is inherited more from.

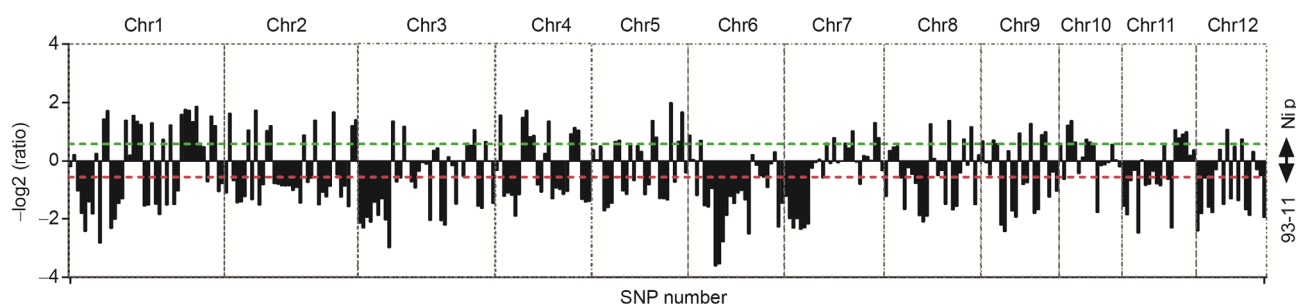


Figure 3 Distribution of segregation distortions along chromosomes. The y-axis represents the $-\log_2$ (ratio) between the sample numbers inherited from 93-11 and Nipponbare. Markers with a positive ratio above the green line represent SNPs that harbor the Nipponbare genotype 1.5 times more than the 93-11 genotype, while markers with a negative ratio below the red line represent the opposite condition. Each bar represents one marker.

cent markers. Chromosome 10 and chromosome 3 harbored the shortest (64.7 cM) and the longest genetic distance (195.6 cM), respectively. The average genetic distance of all chromosomes was 131.5 cM. The average physical distance per cM (the recombination rate) was estimated to identify elevated and suppressed recombination regions. There were 61 regions with recombination rates >10 cM Mb⁻¹, which were located mostly at the telomere regions, especially at the end of chromosome 1 (Chr1: 42,469,719–42,679,635) (Figure 4).

To evaluate the effect of using segregation-distorted markers in the construction of linkage map, we compared our results with a study based on 150 RILs also generated from 93-11 and Nipponbare (Wang et al., 2011). Our linkage map was consistent with map that constructed using 2,330 recombination blocks ($R^2=0.6294$, $P<0.01$) (Figure S7A). Moreover, we also compared the linkage map constructed with segregation-distorted SNPs included for these 150 RILs. By aligning the low-depth re-sequencing data of RILs to the Nipponbare reference genome (MSU6.1), only SNPs with depth between 5 and 100, missing ratios <0.5 were retained, resulting in 1,566 SNPs in total. Ultimately, 734 efficient SNP markers were retained for linkage map construction. The linkages between adjacent markers were estimated using the method described in the Methods section. The linkage distances of each chromosome on the linkage map were consistent with that constructed using the 294 SNP markers in 283 RILs in this study (Figure S7B and C), suggesting that segregation-distorted markers are suitable for linkage map construction.

Mapping of agronomic traits related QTLs

Linkage maps constructed with segregation-distorted markers have been applied in identifying candidate genes or regions relevant to important agronomic traits (Cui et al., 2015; Xu and Hu, 2009; Zhang et al., 2010). To assess the efficiency of QTL mapping with segregation-distorted markers in our analysis, 14 agronomic traits (12 quantitative traits and 2 qualitative traits) were measured for 152 randomly selected RILs. Some of the traits are closely related, such as arista length (AL) and arista color (AC) (Pearson correlation, $r=0.76$), flag leaf length (FLL) and the ratio of the length of the flag leaf to its width (FLL/FLW) ($r=0.69$), and FLW and FLL/FLW ($r=-0.56$) (Figure 5A), suggesting possible shared regions relevant to these related traits.

Based on the linkage map, the LODs between genetic markers and each phenotype were calculated, resulting in 85 significant QTLs that explained more than 5% of phenotypic variation ($PVE\geq 5\%$). Among them, 18 QTLs explained more than 10% of phenotypic variation, including 4 QTLs for AL, 4 QTLs for AC, 3 QTLs for the ratio of grain length to grain width (L/W), 1 QTL for FLL/FLW, 2 QTLs for FLL, and 4 QTLs for FLW (Figure 5B, Table 2). Moreover, 3 shared QTLs with $PVE\geq 10\%$ were identified for AL and AC, corresponding to the strong correlations between these two traits.

To estimate the accuracy of the QTLs identified with the 294 SNPs, QTLs with $PVE\geq 5\%$ were compared to known genes/QTLs collected from Q-taro (Yonemaru et al., 2010), a rice annotation database containing 1,067 reported genes and 471 reported QTLs. In total, 34 (40.00%) out of 85

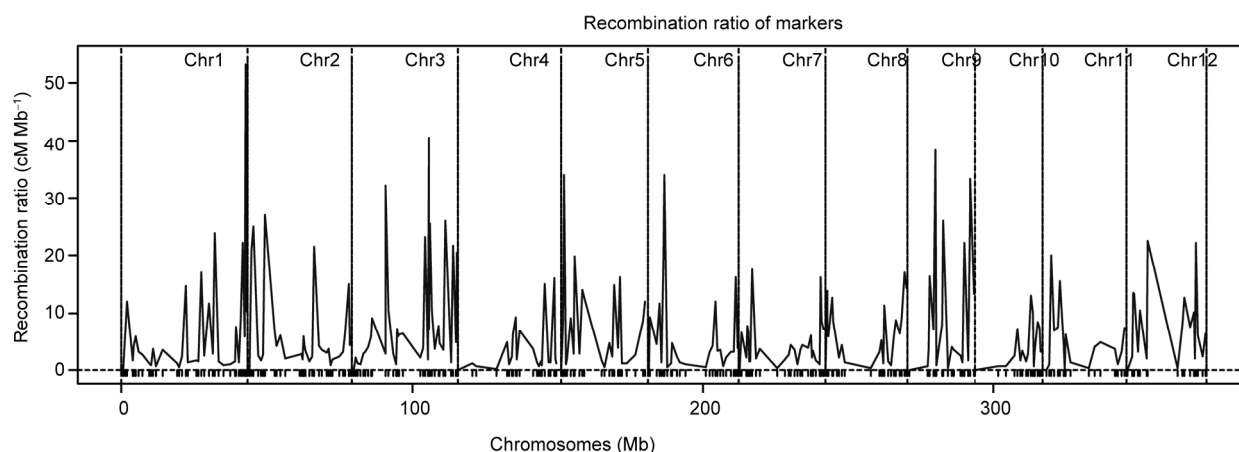


Figure 4 Distribution of recombination rates along chromosomes. The rates are calculated by comparing genetic distance (cM) to physical distance (Mb).

Table 2 List of significant QTLs identified with $PVE \geq 10\%$ ^{a)}

Phenotype	Chr	Peak	Left marker	Right marker	LOD peak	PVE (%)
AL★	Chr4	28,118,063	21,453,456	34,046,417	27.63 ^{*****}	16.88
AL	Chr5	779,216	49,574	1,177,862	17.14 ^{****}	11.27
AL★	Chr7	21,996,206	7,224,894	29,109,953	24.40 ^{*****}	16.23
AL	Chr11	2,950,372	2,527,444	4,075,130	16.97 ^{****}	11.32
AC	Chr2	13,157,393	9,936,823	18,105,224	16.79 ^{****}	11.06
AC★	Chr4	28,531,997	21,453,456	31,246,807	20.10 ^{****}	13.22
AC	Chr5	779,216	390,625	1,177,862	20.52 ^{****}	13.20
AC★	Chr7	21,996,206	7,224,894	26,366,205	17.99 ^{****}	12.49
L/W★	Chr3	16,390,459	12,530,105	27,465,984	17.95 ^{****}	11.97
L/W	Chr5	2,174,722	779,216	3,482,028	16.07 ^{****}	10.57
L/W★	Chr5	4,806,343	4,300,801	7,995,146	29.50 ^{*****}	17.61
FLL/FLW	Chr1	39,489,421	31,396,854	41,726,011	24.98 ^{*****}	15.52
FLL★	Chr1	39,489,421	38,956,081	41,726,011	17.68 ^{****}	11.51
FLL★	Chr2	32,909,780	25,517,693	34,256,548	15.50 ^{***}	10.23
FLW	Chr1	1,157,315	567,295	1,957,319	18.67 ^{****}	11.99
FLW★	Chr2	19,094,255	9,936,823	22,472,089	15.76 ^{***}	10.89
FLW	Chr10	18,473,228	55,729	19,448,657	19.73 ^{****}	12.59
FLW	Chr10	20,429,908	20,084,329	20,765,211	15.11 ^{***}	10.27

a) *, $P < 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.005$; ****, $P \leq 0.001$; *****, $P \leq 0.0005$. Lines marked with a star (★) are QTLs overlapping with known functional genes/QTLs relevant to corresponding traits.

QTLs identified in this study were overlapped with or within 50 kb around previously reported genes/QTLs responsible for corresponding traits, including 9 QTLs with $PVE \geq 10\%$ (Table 2). In addition, 17 QTLs containing segregation-distorted markers were found to be identical to reported sterility-related genes (Table S2), providing possible explanations for the segregation distortions derived from *indica-japonica* crosses. We also noticed that segregation-distorted markers contributed to the phenotypes of RILs at certain QTLs. For the QTLs (Chr1: 37, 373, 294–41, 195, 804) overlapped with the known semidwarf-related gene *SD1*, they were found to be significantly correlated with plant height (PH), panicle length (PL), FLL, FLW, and

FLL/FLW so that we can estimate the correlations between segregation distortions and multiple phenotypes simultaneously (Figure 6A and C). Phenotypes were quite different between RILs with genotypes inherited from 93-11 and those from Nipponbare (Figure 6B and D), suggesting that segregation-distorted markers might be responsible for these QTLs in this region. Similar patterns were found in other QTLs (Figure S8), suggesting the close correlations between segregation-distorted markers and varied phenotypes.

DISCUSSION

With the development of high-throughput sequencing tech-

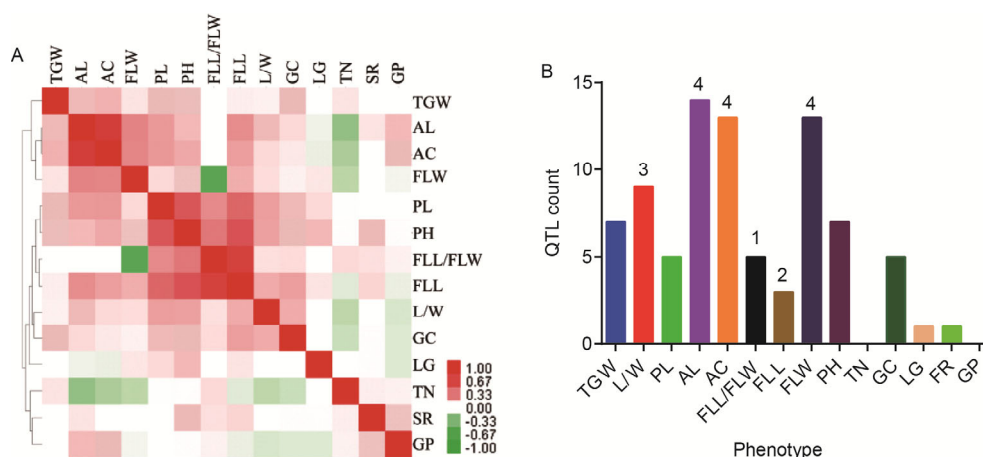


Figure 5 Mapping of QTLs related to 14 agronomic traits. A, Correlations between each trait. Pearson correlations were calculated for each pair. Positive and negative correlations are marked in red and green, respectively. B, Statistic of QTLs identified for the 14 traits with $PVE \geq 5\%$. Numbers above the bars are the numbers of QTLs with $PVE \geq 10\%$ for corresponding trait. TGW, one-thousand-grain weight; AL, arista length; GC, grain color; TN, tiller number; SR, seed setting rate; GP, growth period.

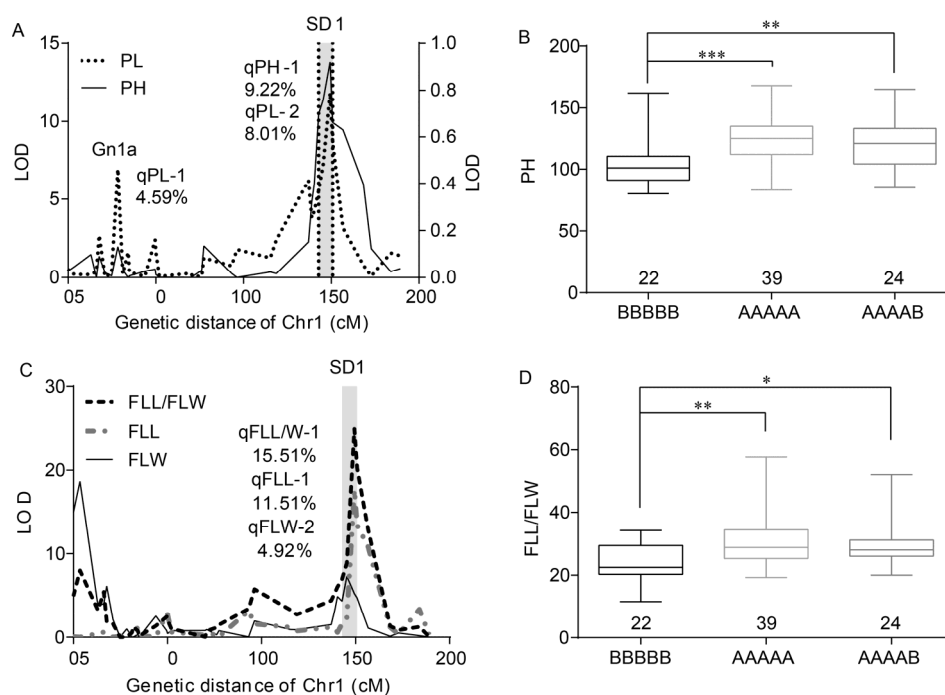


Figure 6 Significant QTLs overlapped with *SD1*. A, LOD curves of qPH-1 ($PVE=9.22\%$) and qPL-2 ($PVE=8.01\%$). B, Box-plot of plant height for RILs with corresponding genotypes (A, 93-11; B, Nipponbare). C, LOD curves of qFLL/W-1 (15.51%), qFLL-1 (11.51%), and qFLW-2 (4.92%). D, Box-plot of FLL/FLW of RILs with corresponding genotypes. This region contains five SNPs. Numbers above the x-axis represent the number of samples with corresponding genotypes. One-way ANOVA is applied to calculate the phenotype values with different genotypes.

nologies, numerous SNPs have been discovered and utilized to obtain the genotypes of various rice varieties and identify candidate regions relevant to important agronomic traits (McNally et al., 2009). In this study, a highly efficient GoldenGate SNP assay consisting of 384 SNP markers was used to genotype 298 RILs derived from 93-11 and Nipponbare. After removing low-call-rate samples and low-MAF or high missing rate markers based on control samples (parents and reciprocal F1 samples), 323 efficient SNPs

were reserved for 283 RILs during our analysis. Interestingly, 148 RILs inherited more from 93-11, while only 29 RILs inherited more from Nipponbare (Figure 1). Through recombination block analysis, 90 recombinant markers present in more than 25% RILs and 15 markers (4.64%) present in more than 40% samples were defined as recombination hotspots (Figure S4). These results indicated the SNP assay we designed was suitable for genetic analysis of RILs derived from 93-11 (*indica*) and Nipponbare (*japonica*).

Segregation distortion deviating from Mendelian ratios has been widely detected in major crop species. Here, 259 SNPs were identified as significantly deviating from the ratio of 1:1 ($P < 0.01$), comprising 80.19% of all markers. It has been reported that segregation distortion transmitted through male and female gametes in different regions may be related to reproductive barriers (Reflinur et al., 2014). We have identified continuous segregation-distorted regions, especially regions on chromosomes 3, 6, and 7, in which half of the markers were inherited 1.5 times more from 93-11 than from Nipponbare (Figure 3), suggesting that these regions are transmitted through male gametes. Moreover, regions located on chromosomes 1, 4, 6, and 8 were found to be identical to 6 reported hybrid sterility genes (Table 1), indicating that segregation distortions in these regions may be induced by cross-incompatibility between *indica* and *japonica*.

Segregation-distorted markers were usually discarded to avoid genetic distance bias when constructing linkage maps. However, due to the wide spread of segregation-distorted markers, lack of intensive investigation of these markers may lead to failed identification of certain important segments. Here, all polymorphic SNP markers were estimated when constructing the linkage map, resulting in a map with a total genetic distance of 1,583.2 cM that was built based on 294 markers (231 segregation-distorted markers included) (Figure S6). By comparison with the RILs also derived from 93-11 and Nipponbare published by Wang et al., the linkage map constructed using segregation-distorted markers was consistent with that using recombination blocks or using re-sequencing high-quality SNPs (Figure S7), indicating that segregation-distorted markers are also suitable for building reliable linkage maps.

Based on the linkage map, we identified 85 significant QTLs with $PVE \geq 5\%$ for 14 agronomic traits (Figure 5B, Table 2). Of them, 34 QTLs were overlapped with reported functional genes or QTLs relevant to corresponding traits. Moreover, 17 QTLs harboring segregation-distorted markers were identified to be overlapped with reported sterility-related genes (Table S2), and the phenotypes of RILs inherited primarily from 93-11 were significantly different from those inherited primarily from Nipponbare (Figure 6). In summary, our results provided a valuable method based on GoldenGate SNP array to study cross-incompatibility induced segregation distortion and identify significant agronomic traits related QTLs in populations derived from *indica* and *japonica*.

MATERIALS AND METHODS

Construction of the RIL population and DNA extraction

The 298 RILs derived from the crosses between *indica* 93-11 (♂) and *japonica* 'Nipponbare' (♀) subspecies consisting of 259 F5 and 39 F6 lines were obtained by sin-

gle-seed descendent method. Total DNA was extracted for each RIL with Qiagen DNA plant extraction kits (69106).

Collection of phenotypic data

Phenotypic data, including TGW, L/W, PL, AL, AC, GC, PH, FLL, FLW, FLL/FLW, TN, GP, SR, and lodging rate (LR), were collected for 152 randomly selected RILs. All samples were grown at Changsha in Hunan province (28.12° N, 112.59° E) from May to October in 2012.

Selection of polymorphic SNPs in re-sequencing data

Whole-genome polymorphic SNPs were identified by comparing re-sequenced 93-11 and Nipponbare to the reference genome of Nipponbare released by MSU (v6.1) (<http://rice.plantbiology.msu.edu/>). Only SNPs with flanking 100-bp regions that were uniquely mapped were considered, resulting in 97,366 SNPs that evenly distributed along the whole genome. Next, 1,517 SNPs (approximately 200 kb per marker) were extracted and experimentally validated by PCR experiments. Then, 60 bp flanking sequences around all experimentally confirmed SNPs were extracted and submitted to the Illumina Infinium ISelect system to estimate their suitability for synthesis on the GoldenGate assay. SNPs with ADT scores larger than 0.6 may have relatively high synthetic successful rates. Ultimately, 384 SNPs evenly distributed across the genome were selected for Oligo Pool Assay (OPA) synthesis.

SNP genotyping by GoldenGate assay

The Illumina BeadXpress GoldenGate assay, which allows genotyping of 96 samples with 384 selected markers in a single plate, was utilized for SNP genotyping in this study. A total of 250 ng of high-quality genomic DNA was prepared for each sample. To obtain reliable genotypes, four samples consist of two parents and two reciprocal F1 samples, were genotyped on each plate as control samples. All genotyping data were analyzed using the GenomeStudio genotyping module with a no-call threshold of 0.25 to obtain reliable genotypes. Incorrectly genotyped samples were manually adjusted based on the genotyping clusters of control samples. Only those SNPs with missing rates < 0.3 and minor allele frequency (MAF) ≥ 0.05 were regarded as efficient SNPs. Samples with call rates < 0.9 were discarded during following analysis.

Identification of segregation-distorted markers

SNP markers in RILs derived from 93-11 and Nipponbare were assessed by the chi-square (χ^2) test to determine whether the segregation ratio was consistent with the normal Mendelian ratio. The χ^2 value was calculated using the following formula:

$$\chi^2 = \sum_{i=1}^2 (|O_i - T_i| - 0.5)^2 / T_i$$

where O_i and T_i represent the real and theoretical number of

samples with the same genotype as parent *i* among all RILs, respectively. The Bonferroni method was used to calculate the *P*-value for each test. Only SNPs with *P*<0.01 were regarded as segregation-distorted markers.

Linkage map construction and QTL mapping

SNP markers selected for the construction of genetic linkage map were assessed using MapDisto 1.7.7 (Lorieux, 2012). SNPs with missing ratio ≥ 0.1 or with a distance larger than 20 cM between adjacent markers were discarded. The linkage map was then visualized with MapChart 2.2 (Voorrips, 2002). Significant QTLs related to the 14 traits described above were identified based on the linkage map. The logarithm of odds (LOD) threshold for identifying significant QTLs was 3.0. Only regions with *P*<0.01 and PVE $\geq 5\%$ were considered as significant QTLs.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National High Technology Research and Development Program of China (2012AA10A304, 2014AA10A602), the National Basic Research Program of China (2013CBA01402) and the National Natural Science Foundation of China (U1031001).*

- Bao, J., Jin, L., Xiao, P., Shen, S., Sun, M., and Corke, H. (2008). Starch physicochemical properties and their associations with microsatellite alleles of starch-synthesizing genes in a rice RIL population. *J Arg Food Chem* 56, 1589–1594.
- Chen, H., He, H., Zhou, F., Yu, H., and Deng, X.W. (2013). Development of genomics-based genotyping platforms and their applications in rice breeding. *Curr Opin Plant Biol* 16, 247–254.
- Chen, H., He, H., Zou, Y., Chen, W., Yu, R., Liu, X., Yang, Y., Gao, Y.M., Xu, J.L., Fan, L.M., Li, Y., Li, Z.K., and Deng, X.W. (2011). Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). *Theor Appl Genet* 123, 869–879.
- Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., Yu, R., Yao, Y., Zhang, W., He, Y., Tang, X., Zhou, F., Deng, X.W., and Zhang, Q. (2014). A high-density SNP genotyping array for rice biology and molecular breeding. *Mol Plant* 7, 541–553.
- Cui, Y., Zhang, F., Xu, J., Li, Z., and Xu, S. (2015). Mapping quantitative trait loci in selected breeding populations: A segregation distortion approach. *Heredity* 115, 538–546.
- Hackett, C.A., and Broadfoot, L.B. (2003). Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90, 33–38.
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., Zhu, C., Lu, T., Zhang, Z., Li, M., Fan, D., Guo, Y., Wang, A., Wang, L., Deng, L., Li, W., Lu, Y., Weng, Q., Liu, K., Huang, T., Zhou, T., Jing, Y., Li, W., Lin, Z., Buckler, E.S., Qian, Q., Zhang, Q.F., Li, J., and Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42, 961–967.
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., Fan, D., Lu, Y., Weng, Q., Liu, K., Zhou, T., Jing, Y., Si, L., Dong, G., Huang, T., Lu, T., Feng, Q., Qian, Q., Li, J., and Han, B. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 44, 32–39.
- Jia, Y., Jia, M.H., Wang, X., and Liu, G. (2012). Indica and japonica crosses resulting in linkage block and recombination suppression on rice chromosome 12. *PLoS One* 7, e43066.
- Jiang, W., Lee, J., Jin, Y.M., Qiao, Y., Piao, R., Jang, S.M., Woo, M.O., Kwon, S.W., Liu, X., Pan, H.Y., Du, X., and Koh, H.J. (2011). Identification of QTLs for seed germination capability after various storage periods using two RIL populations in rice. *Mol Cells* 31, 385–392.
- Lorieux, M. (2012). MapDisto: fast and efficient computation of genetic linkage maps. *Mol Breeding* 30, 1231–1235.
- McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J., Zeller, G., Clark, R.M., Hoen, D.R., Bureau, T.E., Stokowski, R., Ballinger, D.G., Frazer, K.A., Cox, D.R., Padhukasahasram, B., Bustamante, C.D., Weigel, D., Mackill, D.J., Bruskiewich, R.M., Ratsch, G., Buell, C.R., Leung, H., and Leach, J.E. (2009). Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci USA* 106, 12273–12278.
- Ouyang, Y., Chen, J., Ding, J., and Zhang, Q. (2009). Advances in the understanding of inter-subspecific hybrid sterility and wide-compatibility in rice. *Chines Sci Bull* 54, 2332–2341.
- Reflinur, Kim, B., Jang, S.M., Chu, S.H., Bordiya, Y., Akter, M.B., Lee, J., Chin, J.H., and Koh, H.J. (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. *Rice* 7, 3.
- Shah, R., Cavanagh, C.R., and Huang, B.E. (2014). Computationally efficient map construction in the presence of segregation distortion. *Theor Appl Genet* 127, 2585–2597.
- Singh, N., Choudhury, D.R., Singh, A.K., Kumar, S., Srinivasan, K., Tyagi, R.K., Singh, N.K., and Singh, R. (2013). Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS One* 8, e84136.
- Voorrips, R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93, 77–78.
- Wang, L., Wang, A., Huang, X., Zhao, Q., Dong, G., Qian, Q., Sang, T., and Han, B. (2011). Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombinant inbred lines. *Theor Appl Genet* 122, 327–340.
- Wang, S., Tan, Y., Tan, X., Zhang, Z., Wen, J., and Kou, S. (2009). Segregation distortion detected in six rice F₂ populations generated from reciprocal hybrids at three altitudes. *Genet Res* 91, 345–353.
- Xu, S. (2008). Quantitative trait locus mapping can benefit from segregation distortion. *Genetics* 180, 2201–2208.
- Xu, S., and Hu, Z. (2009). Mapping quantitative trait Loci using distorted markers. *Int J Genomics* 2009, 410825.
- Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., Qiu, S., Huang, J., Jiang, Y., Jiang, L., Ding, J., Wang, J., Xu, C., Li, X., and Zhang, Q. (2012). A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337, 1336–1340.
- Yonemaru, J.-i., Yamamoto, T., Fukuoka, S., Uga, Y., Hori, K., and Yano, M. (2010). Q-TARO: QTL Annotation Rice Online Database. *Rice* 3, 194–203.
- Zhang, L., Wang, S., Li, H., Deng, Q., Zheng, A., Li, S., Li, P., Li, Z., and Wang, J. (2010). Effects of missing marker and segregation distortion on QTL mapping in F₂ populations. *Theor Appl Genet* 121, 1071–1082.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

SUPPORTING INFORMATION

Figure S1 Samples and SNPs after genotyping. A, Plot of 10% GC score versus call rate. Samples with a higher call rate (≥ 0.9) tended to have a higher 10% GC score (black). Samples with a lower call rate (< 0.9) diverged from other samples are discarded (blue) during analysis. B, MAF distribution of the remaining 323 SNP markers. Numbers above the bars represent SNP numbers with corresponding MAF.

Figure S2 Clustering of 283 RILs based on 323 SNP markers. Homozygous alleles identical to 93-11 or Nipponbare are labeled in red and green, respectively. Heterozygous alleles are marked in black. Alleles marked in grey represent SNPs without a signal or low-confidence. Each row represents one sample. Each column represents one SNP marker.

Figure S3 Plot of heterozygous ratio for each RIL. Samples are sorted by heterozygous ratio. The red line shows the average heterozygous ratio for all RILs.

Figure S4 Distribution of recombination frequencies of markers along chromosomes. The black vertical lines under the x -axis represent the SNP markers along the chromosomes.

Figure S5 Distribution of segregation-distorted markers along the chromosomes. The black lines under the x -axis represent the locations of 323 markers along the chromosomes. The y -axis represents the $-\log_{10}$ (P -value) of the chi-square test.

Figure S6 Linkage map constructed with 294 SNP markers for 283 RILs derived from 93-11 and Nipponbare.

Figure S7 Comparisons of linkage maps constructed based on the GoldenGate assay with 384 SNPs and 2,330 recombination blocks (A), as well as efficient re-sequencing SNPs (B–C) for RILs derived from 93-11 and Nipponbare reported by Wang et al. The linear regression lines and R^2 are presented for each comparison.

Figure S8 Examples of phenotypes differentiations between RILs with genotypes inherited from 93-11 and Nipponbare within QTLs harboring segregation-distorted markers. (A–B) LOD curve of qTGW-1 and phenotypes of samples with corresponding genotypes on chromosome 2 (A, 93-11; B, Nipponbare). (C–D) LOD curve of qTGW-4 and phenotypes of samples with corresponding genotypes on chromosome 11. Numbers above the x -axis represent the number of samples with corresponding genotypes. One-way ANOVA is applied to calculate the phenotype values with different genotypes.

The supporting information is available online at life.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.