


ORIGINAL ARTICLE

Open Access



Development of Three Sets of High-Throughput Genotyped Rice Chromosome Segment Substitution Lines and QTL Mapping for Eleven Traits

Bin Zhang^{1,2†}, Lianguang Shang^{2†}, Banpu Ruan^{1†}, Anpeng Zhang¹, Shenglong Yang¹, Hongzhen Jiang¹, Chaolei Liu¹, Kai Hong², Hai Lin², Zhenyu Gao¹, Jiang Hu¹, Dali Zeng¹, Longbiao Guo¹ and Qian Qian^{1,2*} 

Abstract

Background: Detecting and mapping chromosomal regions that are related to quantitative phenotypic variation in chromosome segment substitution lines (CSSLs) provides an effective means to characterize the genetic basis of complex agronomic trait. CSSLs are also powerful tools for studying the effects of quantitative trait loci (QTLs) pyramiding and interaction on phenotypic variation.

Results: Here, we developed three sets of CSSLs consisting of 81, 55, and 61 lines, which were derived from PA64s × 9311, Nipponbare × 9311 and PA64s × Nipponbare crosses, respectively. All of the 197 CSSLs were subjected to high-throughput genotyping by whole-genome resequencing to obtain accurate physical maps for the 3 sets of CSSLs. The 3 sets of CSSLs were used to analyze variation for 11 major agronomic traits in Hangzhou and Shenzhen and led to the detection of 71 QTLs with phenotypic effect that ranged from 7.6% to 44.8%. Eight QTLs were commonly detected under two environments for the same phenotype, and there were also 8 QTL clusters that were found. Combined with GWAS on grain length and expression profiles on young panicle tissues, *qGL1* detected in CSSLs was fine mapped within a 119 kb region on chromosome 1 and *LOC_Os01g53140* and *LOC_Os01g53250* were the two most likely candidate genes.

Conclusions: Our results indicate that developing CSSLs genotyped by whole-genome resequencing are powerful tools for basic genetic research and provide a platform for the rational design of rice breeding. Meanwhile, the conjoint analysis of different CSSLs, natural population and expression profiles can facilitate QTL fine mapping.

Keywords: Chromosome segment substitution lines (CSSLs), High-throughput resequencing, QTL mapping, Grain length, GWAS

Background

Most important agronomic traits of rice (*Oryza sativa* L.), such as heading date, yield, plant height and grain size, are controlled by quantitative trait loci (QTLs). Although multiple QTLs have been reported over the past few decades, few of them have been isolated or

fine-mapped (Hu et al. 2015; Huang et al. 2009a; Ishimaru et al. 2013; Jiao et al. 2010; Konishi et al. 2006; Song et al. 2007; Wang et al. 2015; Xue et al. 2008; Yano et al. 2000). As the phenotypic manifestations of these quantitative traits show continuous variation and are always influenced by environmental factors and genetic backgrounds, only suitable genetic populations can provide a strong foundation for genetic and genomic research. To date, many types of genetic populations have been developed for QTL research and utilization. Primary genetic mapping populations, such as the F₂, F_{2:3} and BC₁ populations, were often used for analyzing

* Correspondence: qianqian188@hotmail.com

[†]Bin Zhang, Lianguang Shang and Banpu Ruan contributed equally to this work.

¹State Key Laboratory of Rice Biology, China National Rice Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou 310006, China

²Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China

the genetic basis of QTL during early studies (Ahn et al. 1993; Li et al. 1995; Redona and Mackill 1996). However, these types of genetic populations can only detect a few major QTLs with large genetic effects due to complex and unstable genetic backgrounds (Yamamoto et al. 2000). To further explore the genetic basis of complex traits, some permanent mapping populations, such as recombinant inbred lines (RILs) and doubled haploid (DH) have been developed. QTLs with minor effects can also be detected by using these permanent populations, but they are still inadequate for fine mapping and cloning of QTLs (Yano 2001). As a permanent population, chromosome segment substitution lines (CSSLs) can improve the precise detection of QTL regulation of complex traits and make it much easier for QTL fine mapping and cloning (Mei et al. 2006; Takai et al. 2007). Each line of CSSLs has the same genetic background as its recurrent parent, except for a few substituted segments from a donor parent. CSSLs are ideal for mapping genetic population for the elimination of most genetic background noise and dissection of QTLs into the single Mendelian factor (Xu et al. 2010). Although development of CSSLs is laborious and time-consuming, more and more CSSLs have been developed for their significant advantages (Bessho-Uehara et al. 2017; Furuta et al. 2014; Hao et al. 2006; Kubo et al. 2002; Qiao et al. 2016; Shim et al. 2010; Yoshimura 1997).

Marker-assisted selection (MAS) is a powerful technique of selecting target segments during CSSLs construction. With the publication of rice whole genome sequences, many SSR (simple sequence repeat) markers, Indels (insertions and deletions) markers and SNPs (single nucleotide polymorphisms) markers have been exploited for MAS. The density of molecular markers that are distributed in the whole genome determine the accuracy of CSSLs. Low-density molecular makers may easily miss small introgression segments and double-crossovers (Xu et al. 2010). In fact, less than 200 molecular markers were used to develop CSSLs because of laborious work (Bian et al. 2010; Shim et al. 2010; Xu et al. 2007; Zhu et al. 2009). Next-generation sequencing technology provides a high-throughput genotyping method that can construct high-resolution physical maps and calculate the length of substituted segments in the construction of CSSLs. Recently, several RILs and CSSLs have been exactly genotyped by whole-genome resequencing, and high-quality linkage maps were constructed for QTL dissection (Gao et al. 2013; Huang et al. 2009b; Xu et al. 2010; Zhang et al. 2011; Zhu et al. 2015b). Among the CSSLs that were reported previously, few were developed by using the male sterile line and were subjected to high-throughput genotyping by whole-genome resequencing to obtain accurate physical maps. It will be necessary to develop novel CSSLs with high-resolution physical maps.

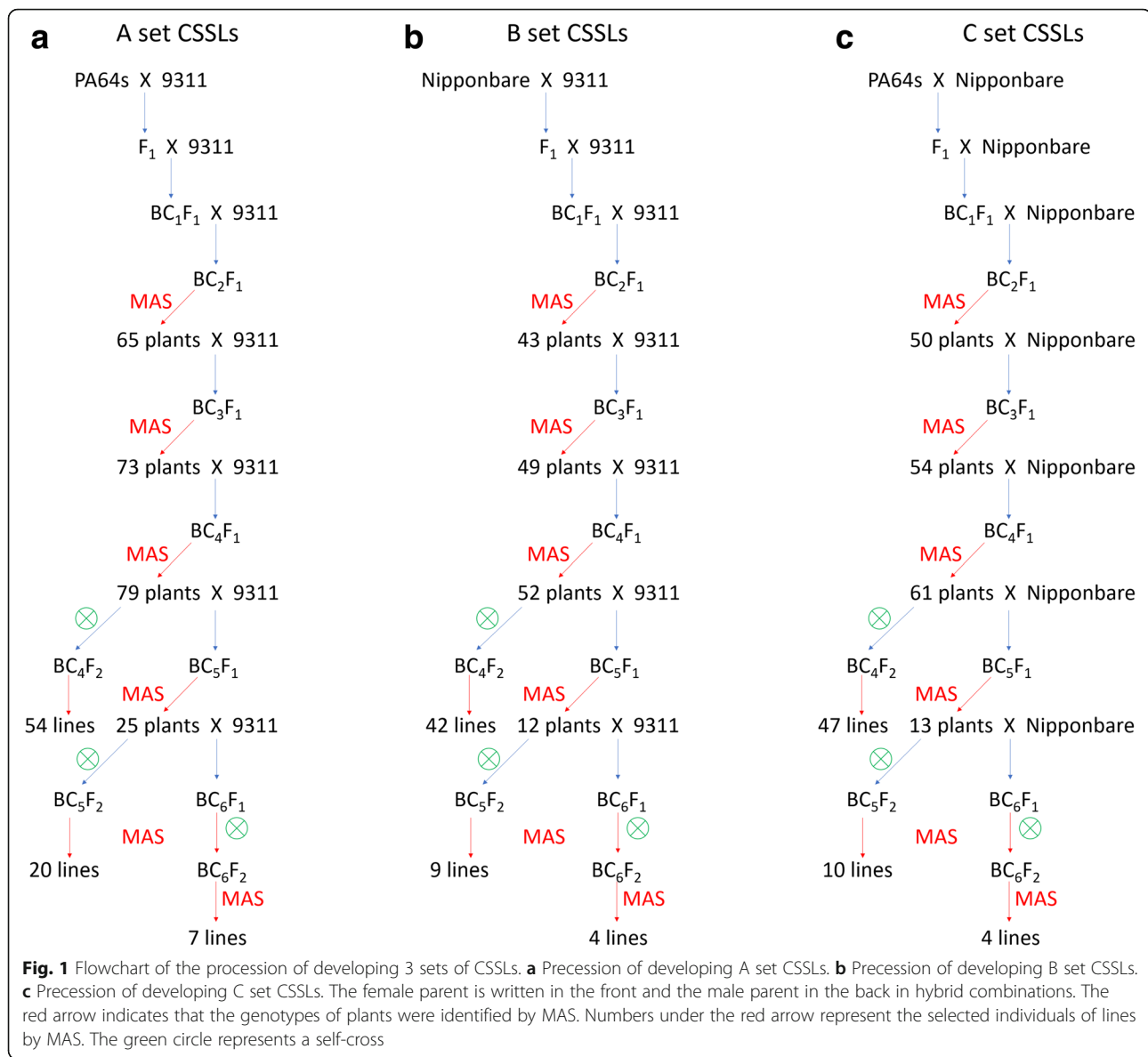
Previously, our lab dissected yield-associated loci using an RIL derived from a cross between PA64s and 9311 (Gao et al. 2013). Since then, to facilitate QTL mapping and cloning, 3 sets of CSSLs were developed and were derived from three representative rice varieties: the first genome-sequenced *japonica* cultivar Nipponbare, an elite restore *indica* cultivar 9311, and a male sterile line PA64s with a mixed genetic background of *indica* and *japonica*. One set of CSSLs was developed from a cross between PA64s as the donor and Nipponbare as the recipient. The variety 9311 was used as recipient parent for the other two sets of CSSLs with Nipponbare and PA64s as donor parents, respectively. To obtain precise physical maps for QTL detection, the three sets of CSSLs were genotyped by high-throughput whole-genome resequencing. QTLs analysis for 11 major agronomic traits of the 3 sets of CSSLs was performed using the bin maps converted from the precise physical maps.

Results

Development of the CSSLs

A total of 311 polymorphic molecular markers were used to assist selection during CSSLs development (Additional file 1). Among them, 136 markers denoted as A, AB or AC were used for MAS in the population derived from PA64s × 9311 crosses, and 162 markers denoted as B, AB or BC were used for CSSL analysis on the population derived from a cross between 9311 and Nipponbare. Then, 138 markers denoted as C, AC or BC were used for the population derived from PA64s × Nipponbare crosses.

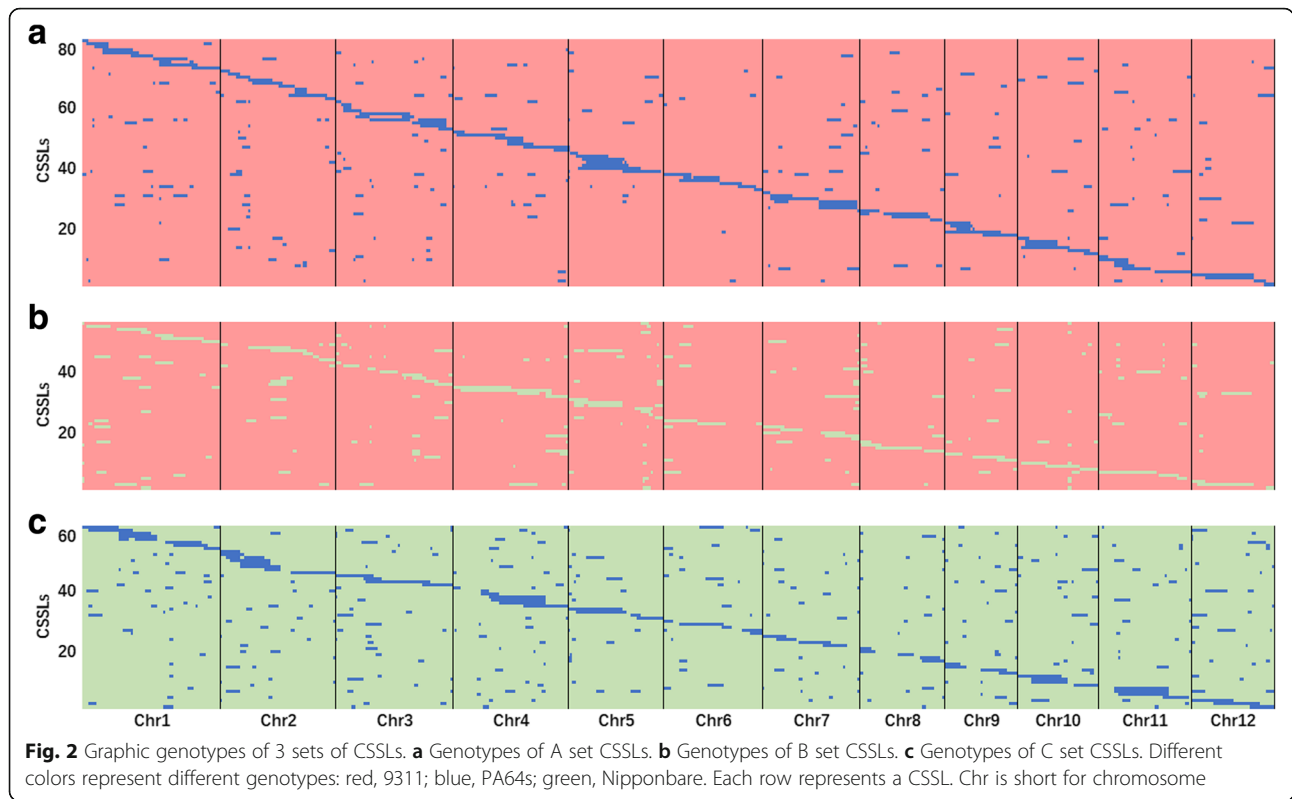
The development procedure for the 3 sets of CSSLs is shown in Fig. 1. Hybrid F₁ plants, namely, LYP9, were generated from a cross between maternal PA64s and paternal 9311. Then, F₁ plants were backcrossed twice with 9311 to produce BC₂F₁ generation. The first MAS was performed for BC₂F₁ generation with 136 polymorphic markers distributed on all 12 chromosomes, and 65 plants were selected from 93 individuals, in which a few of heterozygous substituted segments came from PA64s, while the major genomic regions were homozygous for 9311 alleles. Subsequently, the selected 65 plants from BC₂F₁ generation were backcrossed twice again with 9311 to generate BC₄F₁. Twelve individuals from each BC₃F₁ and BC₄F₁ line were genotyped using the markers on the target-substituted regions. In total, 54 plants from BC₄F₁ that had one to three heterozygous substituted segments were self-crossed to produce BC₄F₂. Then, 12 individuals from each BC₄F₂ were genotyped, and the homozygous substituted individuals on target regions were self-crossed to obtain 54 CSSLs. The remaining BC₄F₁ plants that had a few substituted segments that were continuously backcrossed with 9311 to generate



BC₅F₁ and BC₆F₁. Then, 20 BC₅F₂ plants and 7 BC₆F₂ plants that were selected using MAS were self-crossed to obtain 27 CSSLs. Finally, a total of 81 CSSLs with a few substituted segments were developed, which were named A set CSSLs for easy description (Fig. 1a). Similarly, the other two sets of CSSLs, including 55 lines and 61 lines, were obtained with the same technique using 162 and 138 polymorphic markers, respectively (Fig. 1b and c). Meanwhile, 55 CSSLs with major genomic regions from 9311 had a few substituted segments from Nipponbare, which were named B set CSSLs. Then, 64 CSSLs were named C set CSSLs and contained a few substituted segments from PA64s, in which the major genomic regions were homozygous for Nipponbare.

Characteristics of the CSSLs

To obtain accurate physical maps for the 3 sets of CSSLs, all of the CSSLs were subjected to high-throughput genotyping by whole-genome resequencing. Physical maps for the 3 sets of CSSLs were constructed using the physical locations of filtered SNPs (Fig. 2). The A set CSSLs containing 81 lines were found to carry 310 substituted segments, and there were at least 230 substituted segments in B set CSSLs and 367 substituted segments in C set CSSLs (Table 1). The average number of substituted segments was approximately 26, 19 and 31 per chromosome and ranged from 13 on chromosome 6 to 44 on chromosome 3, 10 on chromosome 6 to 30 on chromosome 3 and 5, and 24 on chromosome 5 to 38 on chromosome 4 for A, B and C sets of CSSLs,



respectively. Statistical data for segment lengths showed that more than 61% segments were shorter than 2.00 Mb, while approximately 11% segments were longer than 7.00 Mb (Fig. 3). The total length of substituted segments was approximately 1152.61 Mb in A set CSSLs and was approximately 868.10 and 1177.96 Mb in B and C sets CSSLs, respectively (Table 1). The average length

of substituted segments varied little, with an average of 3.72, 3.77 and 3.21 Mb in A, B and C sets of CSSLs, respectively. The total coverage length of substituted segments was 363.14 Mb in A set CSSLs, while that length was 325.70 and 331.40 Mb in B and C sets of CSSLs, respectively. The coverage rate of substituted segments was 97.3% in the A set CSSLs, which varied from 88.1%

Table 1 Distribution of substituted segments on 12 chromosomes for 3 sets of CSSLs

Chromosome	Number of segments			Segments length (Mb)			Average length (Mb)			Coverage length (Mb)			Coverage rate		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	38	29	37	129.81	123.45	117.72	3.42	4.26	3.18	42.00	39.45	40.72	97.1%	91.2%	94.1%
2	38	18	37	108.45	77.55	129.26	2.85	4.31	3.49	35.94	28.37	33.42	100.0%	78.9%	93.0%
3	44	30	37	131.20	81.61	109.86	2.98	2.72	2.97	36.16	32.64	36.41	99.3%	89.7%	100.0%
4	28	20	38	100.89	98.40	136.39	3.60	4.92	3.59	35.50	35.50	28.03	100.0%	100.0%	78.9%
5	24	30	24	111.55	95.62	84.14	4.65	3.19	3.51	29.96	24.86	27.41	100.0%	83.0%	91.5%
6	13	10	29	57.40	47.83	91.20	4.42	4.78	3.14	29.97	21.05	28.06	95.9%	67.3%	89.8%
7	25	23	30	108.86	84.12	75.46	4.35	3.66	2.52	27.83	23.50	26.60	93.8%	79.2%	89.6%
8	24	13	25	90.06	55.53	62.97	3.75	4.27	2.52	25.05	27.09	18.28	88.1%	95.2%	64.3%
9	15	14	30	72.87	56.25	71.59	4.86	4.02	2.39	23.01	21.09	17.90	100.0%	91.7%	77.8%
10	19	15	26	75.40	40.02	83.52	3.97	2.67	3.21	23.20	21.46	23.20	100.0%	92.5%	100.0%
11	26	17	26	84.54	57.41	114.19	3.25	3.38	4.39	27.13	27.13	23.34	93.5%	93.5%	80.4%
12	16	11	28	80.58	49.69	99.38	5.04	4.52	3.55	27.19	23.50	27.53	98.8%	85.4%	100.0%
Total	310	230	367	1152.61	868.10	1177.96	3.72	3.77	3.21	363.14	325.70	331.40	97.3%	87.3%	88.8%

A, B and C represent A set CSSLs, B set CSSLs and C set CSSLs, which derived from a cross between PA64s and 9311, Nipponbare and 9311, and PA64s and Nipponbare, respectively

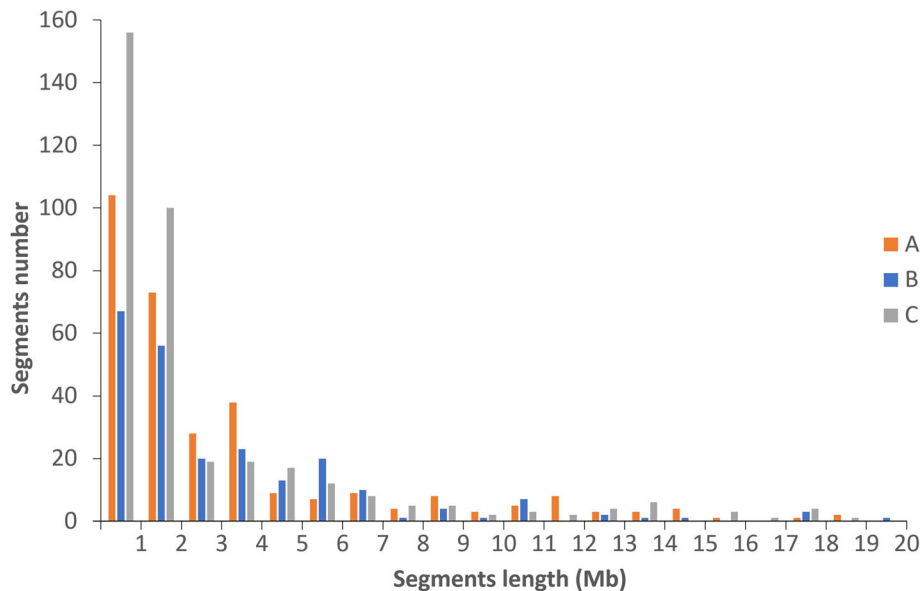


Fig. 3 Distribution of the lengths of the substituted segments in the 3 sets of CSSLs. A, B and C represent A set CSSLs, B set CSSLs and C set CSSLs, respectively

on chromosome 8 to 100.0% on chromosomes 2, 4, 5, 9 and 10. The coverage rate of substituted segments was nearly the same in B and C sets of CSSLs; it was 87.3% and 88.8%, respectively, and ranged from 67.3% to 100.0% on each chromosome.

Phenotypic Variation

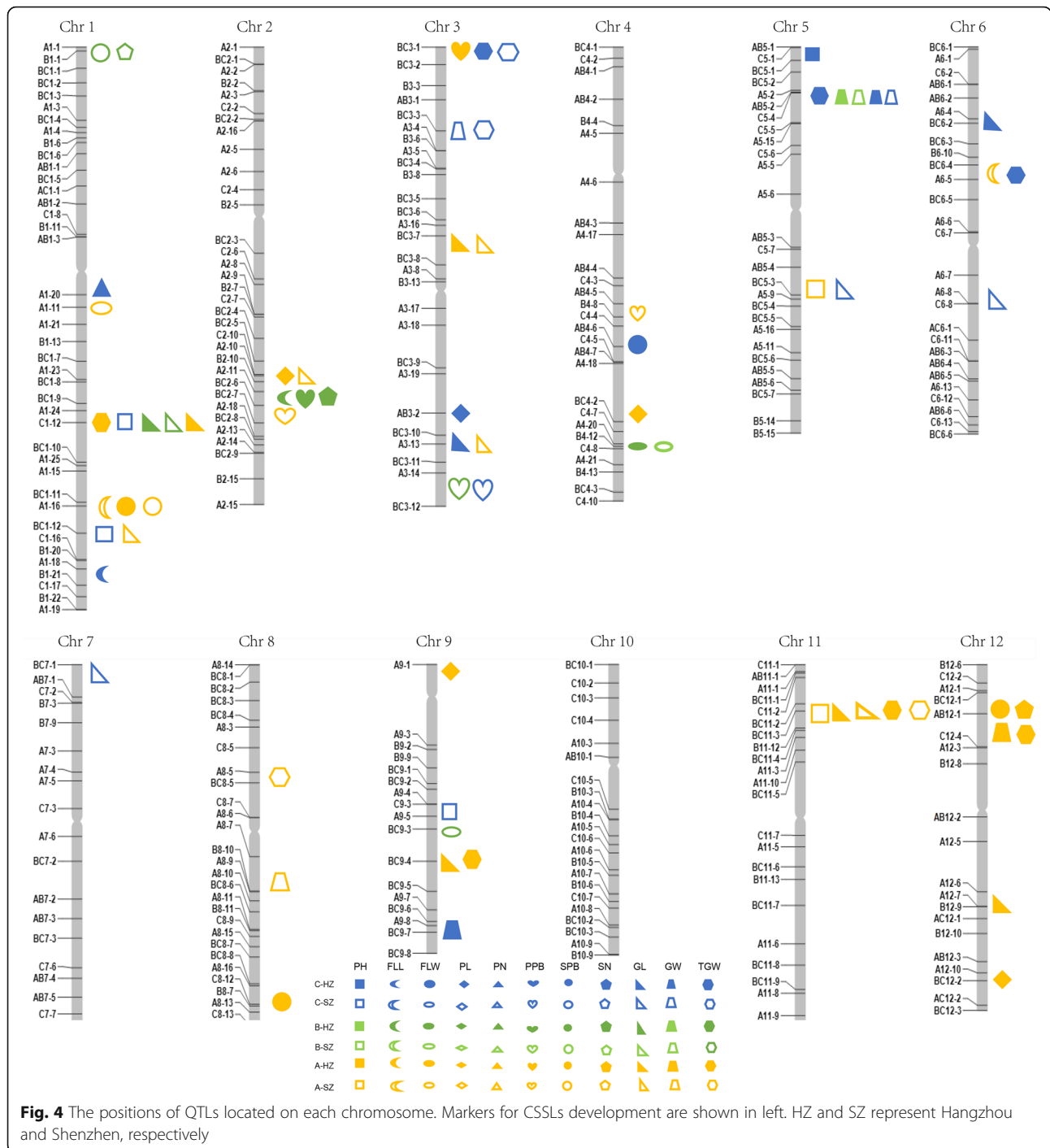
The phenotypic variations of 11 traits were observed between the 3 sets of CSSLs and their parents for QTL analysis. The t-test revealed that the differences between PA64s and Nipponbare were extremely significant concerning the panicle number (PN), primary panicle branch number (PPB), secondary panicle branch number (SPB), spikelet number per panicle (SN), grain length (GL), grain width (GW) and 1000-grain weight (TGW) in Shenzhen, as well as all traits but plant height (PH) and panicle length (PL) in Hangzhou. Both 9311 and Nipponbare showed significant differences in all traits but the flag leaf length (FLL) in Hangzhou and Shenzhen (Additional file 2: Table S1). The correlation of trait variation is shown in Additional file 2: Tables S2 to S7. The results show that extremely significant positive correlations were detected between the grain number and the secondary panicle branch number with the highest correlation coefficient, 0.8883 on average. In both Hangzhou and Shenzhen, 1000-grain weight showed an extremely significant positive correlation with grain length and grain width in the 3 sets of CSSLs, whereas it was negatively correlated with the spikelet number per panicle. Extremely significant positive correlations were also detected between plant height and panicle length in

the 3 sets of CSSLs, except the B set CSSLs in Shenzhen. The phenotypic values of the 11 traits in the 3 sets of CSSLs were all found to be continuous and exhibited normal or skewed distribution patterns across both environments, indicating quantitative inheritance of the 11 characters and satisfying the demands for QTL analysis (Additional file 3: Figure S1).

QTL Analysis for 11 Traits Based on CSSLs

To conduct the QTL analysis, resequencing-based physical maps of the 3 sets of CSSLs were converted into bin maps. A total of 71 QTLs for 11 traits in 3 sets of CSSLs were detected in both Hangzhou and Shenzhen and were distributed on 11 chromosomes, except for chromosome 10 (Fig. 4 and Table 2). Among the 34 QTLs that were detected in Shenzhen and the 37 QTLs in Hangzhou, 35 QTLs were detected in A set CSSLs under both environments, as well as 13 and 23 QTLs in the B and C sets of CSSLs, respectively. Eight QTLs were commonly detected in both Shenzhen and Hangzhou for the same phenotype, and 4 QTLs had overlapping interval locations in different populations. There were also 8 QTL clusters that were detected, including FLL and PPB, SPB and SN, GL and TGW, and GW and TGW, which coincided with significant correlations (Table 2, Additional file 2: Tables S2 to S7).

Compared with the previously identified QTLs by physical location, 22 of the 71 QTLs were reported in the previous studies, and the remaining 49 QTLs were unique. Among the 22 reported QTLs, the interval regions of the four QTLs overlapped with the locations



of four known genes (Additional file 4: Figure S2). For the QTL with the largest effect on plant height in C set CSSLs, *qPH1-2* can explain 26.3% of the phenotypical variation and was located in the region from 38.0 to 38.5 Mb on chromosome 1, which contained *sd1*, the rice green revolution gene that is responsible for rice semi-dwarfness (Sasaki et al. 2002). In the B set CSSLs, the Nipponbare alleles increased the flag leaf width at

qFLW4 under two environments, which overlapped with the location of *Nal1/LSCHLA* that controlled leaf width and yield (Fujita et al. 2013; Qi et al. 2008; Zhang et al. 2014). For the secondary panicle branch number, *qSPB1-2* was mapped to the region that ranged from 35.1 to 36.2 Mb on chromosome 1, while *LAX1*, a gene controlling panicle branch development, was previously cloned (Komatsu et al. 2003). Finally, the major QTL

Table 2 QTLs for 11 traits detected in 3 sets of CSSLs in Shenzhen (SZ) and Hangzhou (HZ)

Trait	Site	Populations	QTL	Chromosome	Location (Mb)	LOD	PVE (%) ^a	Add ^b	Known QTL/ gene
PH	SZ	C	<i>qPH1-1</i>	chr1	30.0–30.5	5.0	22.8	-5.32	
PH	SZ	C	<i>qPH1-2</i>	chr1	38.0–38.4	6.3	26.3	5.88	<i>sd1</i>
PH	HZ	C	<i>qPH5-1</i>	chr5	0.8–1.3	3.7	19.1	7.53	
PH	SZ	A	<i>qPH5-2</i>	chr5	18.6–19.1	3.8	18.0	-4.42	
PH	SZ	C	<i>qPH9</i>	chr9	12.2–13.7	2.9	10.7	-2.93	
PH	SZ	A	<i>qPH11</i>	chr11	3.0–4.1	2.5	11.4	-4.88	
FLL	SZ	A	<i>qFLL1-1</i>	chr1	34.9–35.6	3.5	15.9	1.75	
FLL	HZ	C	<i>qFLL1-2</i>	chr1	42.0–43.2	2.5	16.8	-6.72	<i>qfll1</i> (Yan et al. 1999)
FLL	HZ	B	<i>qFLL2</i>	chr2	27.9–29.3	3.4	24.8	8.75	<i>qFLL5</i> (Wang et al. 2011)
FLL	SZ	A	<i>qFLL6</i>	chr6	9.6–10.5	3.3	14.8	3.05	
FLW	SZ	A	<i>qFLW1</i>	chr1	19.4–20.4	3.4	17.9	-0.11	
FLW	HZ	B	<i>qFLW4</i>	chr4	31.0–31.7	3.3	18.7	-0.21	<i>LSCHL4</i>
FLW	SZ	B	<i>qFLW4</i>	chr4	31.0–31.7	4.4	19.2	-0.25	<i>LSCHL4</i>
FLW	SZ	B	<i>qFLW9</i>	chr9	14.5–14.6	3.3	24.3	0.36	
PL	HZ	A	<i>qPL2</i>	chr2	26.5–26.5	4.3	12.7	-0.77	
PL	HZ	C	<i>qPL3</i>	chr3	30.0–30.4	4.4	27.5	1.20	<i>qPL-3</i> (Shen et al. 2009)
PL	HZ	A	<i>qPL4</i>	chr4	28.3–28.7	7.5	24.4	-0.72	<i>qPL-4-2</i> (Guo and Hong 2010)
PL	HZ	A	<i>qPL9</i>	chr9	0.0–5.7	4.4	13.2	0.60	
PL	HZ	A	<i>qPL12</i>	chr12	25.8–26.2	3.2	9.2	-1.02	
PN	HZ	C	<i>qPN1</i>	chr1	18.8–19.2	3.1	20.2	0.75	
PPB	HZ	B	<i>qPPB2-1</i>	chr2	27.9–29.3	7.1	44.8	2.67	
PPB	SZ	A	<i>qPPB2-2</i>	chr2	28.4–28.7	3.0	13.5	0.56	<i>qPPB3</i> (Gao et al. 2013)
PPB	HZ	A	<i>qPPB3-1</i>	chr3	0.9–1.4	3.4	17.2	0.56	
PPB	SZ	C	<i>qPPB3-2</i>	chr3	35.0–36.4	3.3	20.7	-0.75	
PPB	SZ	B	<i>qPPB3-3</i>	chr3	33.4–35.6	3.4	24.5	-0.88	
PPB	SZ	A	<i>qPPB4</i>	chr4	19.6–20.7	2.6	11.6	0.52	
SPB	SZ	B	<i>qSPB1-1</i>	chr1	0.0–0.5	4.0	28.7	-7.95	
SPB	HZ	A	<i>qSPB1-2</i>	chr1	35.1–36.2	3.5	15.4	5.64	<i>LAX1</i>
SPB	SZ	A	<i>qSPB1-2</i>	chr1	35.1–36.2	3.7	16.3	5.90	<i>LAX1</i>
SPB	HZ	C	<i>qSPB4</i>	chr4	22.8–23.4	4.0	30.5	3.57	
SPB	HZ	A	<i>qSPB8</i>	chr8	26.2–27.9	3.0	13.4	3.01	
SPB	HZ	A	<i>qSPB12</i>	chr12	5.7–6.9	2.6	11.0	2.09	
SN	SZ	B	<i>qSN1</i>	chr1	0.0–0.5	2.8	21.0	-25.21	
SN	HZ	B	<i>qSN2</i>	chr2	27.9–29.3	2.6	19.8	47.00	<i>qGW-2b</i> (Wang et al. 2011)
SN	HZ	A	<i>qSN12</i>	chr12	5.7–6.9	2.7	14.1	6.61	
GL	HZ	A	<i>qGL1-1</i>	chr1	28.5–30.0	3.6	27.3	0.30	
GL	SZ	B	<i>qGL1-2</i>	chr1	28.1–30.0	3.5	20.1	0.26	
GL	HZ	B	<i>qGL1-2</i>	chr1	28.1–30.0	3.3	19.6	0.22	
GL	SZ	A	<i>qGL1-3</i>	chr1	37.7–38.0	3.6	10.4	0.21	
GL	SZ	A	<i>qGL2</i>	chr2	26.5–27.4	3.3	9.3	-0.28	
GL	HZ	A	<i>qGL3-1</i>	chr3	15.9–18.2	3.5	13.6	0.14	
GL	SZ	A	<i>qGL3-1</i>	chr3	15.9–18.2	4.0	16.3	0.26	
GL	HZ	C	<i>qGL3-2</i>	chr3	30.0–31.6	3.5	17.7	0.16	
GL	SZ	A	<i>qGL3-3</i>	chr3	31.1–32.5	4.3	17.8	-0.32	

Table 2 QTLs for 11 traits detected in 3 sets of CSSLs in Shenzhen (SZ) and Hangzhou (HZ) (Continued)

Trait	Site	Populations	QTL	Chromosome	Location (Mb)	LOD	PVE (%) ^a	Add ^b	Known QTL/ gene
GL	SZ	C	<i>qGL5</i>	chr5	18.4–18.9	4.8	17.5	0.31	
GL	HZ	C	<i>qGL6-1</i>	chr6	6.6–7.1	3.4	17.0	0.16	
GL	SZ	C	<i>qGL6-2</i>	chr6	21.1–22.1	4.5	7.6	0.13	<i>qGL6.2</i> (Bian et al. 2010)
GL	SZ	C	<i>qGL7</i>	chr7	0.5–2.1	4.9	8.3	0.16	<i>qGL7</i> (Bian et al. 2010)
GL	HZ	A	<i>qGL9</i>	chr9	15.4–16.1	3.3	8.8	-0.21	<i>qGL9.1</i> (Bian et al. 2010)
GL	SZ	A	<i>qGL11</i>	chr11	2.9–5.8	3.1	8.7	-0.21	
GL	HZ	A	<i>qGL11</i>	chr11	2.9–5.8	4.1	29.2	-0.26	
GL	HZ	A	<i>qGL12</i>	chr12	19.8–20.6	5.2	14.8	-0.30	
GW	SZ	C	<i>qGW3</i>	chr3	7.9–8.6	3.0	11.5	0.06	<i>qGW3</i> (Wang et al. 2011)
GW	HZ	B	<i>qGW5</i>	chr5	4.2–5.5	3.5	25.4	-0.12	<i>GW5</i>
GW	SZ	B	<i>qGW5</i>	chr5	4.2–5.5	4.8	33.3	-0.14	<i>GW5</i>
GW	HZ	C	<i>qGW5</i>	chr5	5.3–5.5	4.8	44.8	0.14	<i>GW5</i>
GW	SZ	C	<i>qGW5</i>	chr5	5.3–5.5	6.3	27.6	0.11	<i>GW5</i>
GW	SZ	A	<i>qGW8</i>	chr8	17.0–17.7	4.3	21.1	-0.06	
GW	HZ	C	<i>qGW9</i>	chr9	21.4–21.9	2.6	15.7	-0.13	
GW	HZ	A	<i>qGW12</i>	chr12	4.5–5.5	3.0	12.1	-0.04	
TGW	HZ	A	<i>qTGW1</i>	chr1	28.5–30.0	4.0	12.0	1.40	
TGW	HZ	C	<i>qTGW3-1</i>	chr3	0.0–1.12	3.6	14.4	-0.98	
TGW	SZ	C	<i>qTGW3-1</i>	chr3	0.0–1.12	4.6	23.7	-1.34	
TGW	SZ	C	<i>qTGW3-2</i>	chr3	7.9–8.6	3.2	15.8	1.06	
TGW	HZ	C	<i>qTGW5</i>	chr5	5.3–5.5	4.9	20.9	1.37	<i>GW5</i>
TGW	HZ	C	<i>qTGW6</i>	chr6	11.7–19.4	6.3	28.5	1.46	
TGW	SZ	A	<i>qTGW8</i>	chr8	8.6–8.7	4.4	19.2	-1.21	
TGW	HZ	A	<i>qTGW9</i>	chr9	15.4–16.1	6.7	22.0	-2.35	
TGW	SZ	A	<i>qTGW11</i>	chr11	2.9–5.8	3.8	16.0	-1.73	<i>qTGW11.1</i> (Bian et al. 2010)
TGW	HZ	A	<i>qTGW11</i>	chr11	2.9–5.8	5.5	17.3	-1.41	<i>qTGW11.1</i>
TGW	HZ	A	<i>qTGW12</i>	chr12	5.7–6.9	3.2	9.3	-1.07	

^aThe percentage of phenotypic variation explained by the detected QTL

^bAdditive effects; The positive value indicates that alleles from recurrent parent increase the effect

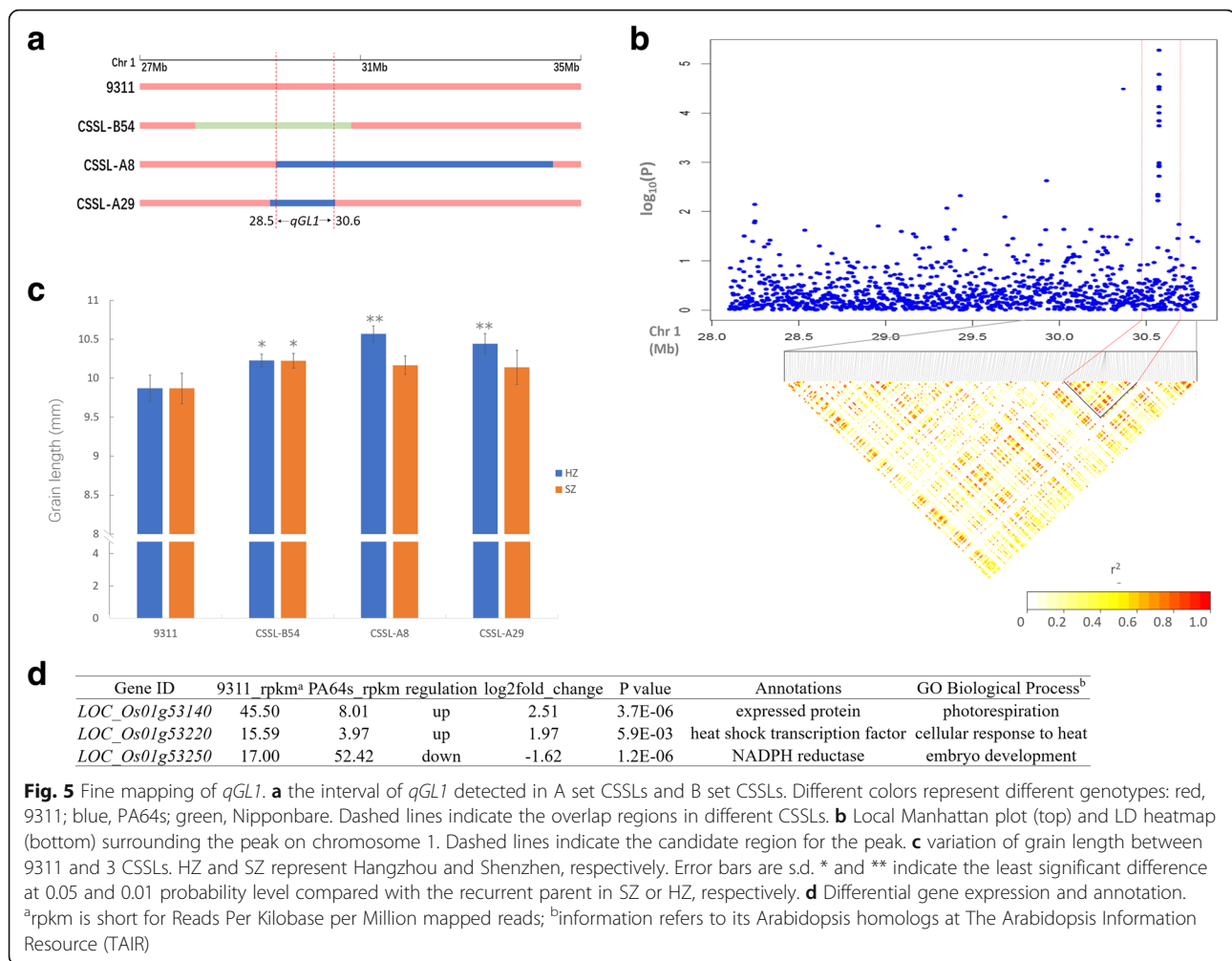
qGW5 with Nipponbare alleles that increased the grain width was detected in B and C set of CSSLs under two environments, while *GW5/GSE5* was located in the region by previous studies (Duan et al. 2017; Liu et al. 2017).

Validation and Fine-Mapping of *qGL1*

qGL1 had a stable effect in B set CSSLs under both environment and was also detected in A set CSSLs in Hangzhou (Fig. 4 and Table 2). As CSSL-A8, CSSL-A29 and CSSL-B54 had overlapped regions on *qGL1* location, the interval of *qGL1* could be narrowed to the region of 2.1 Mb which ranged from 28.503 Mb to 30.611 Mb on chromosome 1 (Fig. 5a). To identify genes associated with grain length and further reduce the interval of *qGL1*, we performed GWAS by analyzing a natural rice population, consisting of 317 rice varieties that were collected from the 3000 Rice Genomes Project (Wang et al. 2018). Using the efficient Mixed-Model, 8 SNPs at *qGL1*

location on chromosome 1 were identified to be significantly associated with grain length. Then, we used pairwise linkage disequilibrium (LD) correlations to estimate a candidate region from 30.492 Mb to 30.652 Mb for *qGL1* (Fig. 5b). Hence, both our GWAS and CSSLs results confirmed that the interval of *qGL1* could be narrowed to a 119 kb region ranged from 30.492 Mb to 30.611 Mb.

There were 18 annotated genes within this 119 kb interval based on Rice Genome Annotation Project Website (<http://rice.plantbiology.msu.edu/>) (Additional file 2: Table S8). Among the 18 genes, 3 genes expressed different significantly between 9311 and PA64s in young panicle tissues based on RNA-seq data. The expression of *LOC_Os01g53140* and *LOC_Os01g53220* in 9311 were higher than that in PA64s, while the expression of *LOC_Os01g53250* was significantly higher in PA64s (Fig. 5d). Moreover, sequence variations of those genes between 9311 and Nipponbare were identified and



expressions at RNA level were also analyzed in young panicles between 9311 and CSSL-B54 (Additional file 6: Figure S4). Both 9311 and Nipponbare had amino acid variations of these three candidate genes, which were caused by SNPs on exons. Nipponbare and CSSL-B54 showed lower expression level on *LOC_Os01g53140* and *LOC_Os01g53220* compared with 9311, which coincided with the RNA-seq data. Similarly, the expression of *LOC_Os01g53250* was higher in Nipponbare and CSSL-B54 than that in 9311. *LOC_Os01g53220* has previously been designated as *OsHsfC1b*, which is a heat shock transcription factor and plays a role in ABA-mediated salt stress tolerance in rice (Mittal et al. 2009; Schmidt et al. 2012). *LOC_Os01g53250* product is a NADPH reductase whose homologous gene in *Arabidopsis* is *ATR3*. *ATR3* is NADPH-dependent diflavin oxidoreductase 1, which is involved in embryo development ending in seed dormancy and oxidation-reduction process (Varadarajan et al. 2010). The product of *LOC_Os01g53140* is an expressed protein. Its homologous protein in *Arabidopsis* is a neuronal acetylcholine receptor subunit alpha-5 which is subunit of

mitochondrial respiratory chain complex I and is involved in photorespiration (Jennifer et al. 2010). Collectively, on the basis of our results, we suggest that *qGL1* is located in a 119 kb region on chromosome 1 and *LOC_Os01g53140* and *LOC_Os01g53250* are the two most likely candidate genes for *qGL1*.

Discussion

CSSLs are the ideal population for detecting the precise QTLs that control complex traits by eliminating genetic background noise (Nadeau et al. 2000; Qiao et al. 2016; Tian et al. 2006; Ye et al. 2010; Zhou et al. 2017). On the other hand, because CSSLs contain a few substituted chromosome segments from donor parent, they can be used as NILs themselves or used to develop NILs with additional cross with the recurrent parent, which are fast and powerful tools for QTL validation and isolation (Ebitani et al. 2005; Mulsanti et al. 2018; Zhu et al. 2017). Under a certain condition, CSSLs are limited for QTL detection. Because several CSSLs carry some long-substituted chromosome segments that can range

between 10 and 20 Mb, they can hardly detect the QTLs that are close to each other with opposite effects. Here, we developed 3 sets of CSSLs using repeated elite parents, which may be much easier for QTL detection and mutual authentication.

Because there are different genetic backgrounds among lines or individuals, the epistatic effects between QTLs for many quantitative traits estimated by using conventional experimental materials are confounded with other genetic factors (Zhu et al. 2015a). Due to the single segment substitution, CSSLs may be ideal materials for verifying and researching the interaction between two QTLs in the additive effect and the epistatic effect by a cross between two related CSSLs. In rice, a number of QTLs and their epistasis on traits, such as plant height, heading date, and yield components, have been well analyzed by using CSSLs or secondary F_2 populations that were derived from CSSLs (Chen et al. 2014; Kubo et al. 2008; Wang et al. 2017; Yang et al. 2018; Zhu et al. 2015a). On the other hand, CSSLs can also be used for QTL pyramiding in plant breeding. *Gn1a* can increase the grain number and improve rice yield simultaneously through inhibiting *OsCKX2* gene expression in the inflorescence meristem, while the green revolution gene *sd1* reduces bioactive gibberellin (GA) abundance, which results in semi-dwarfism to resistant to lodging. NIL-*sd1* + *Gn1a* generated by crossing NIL-*Gn1a* and NIL-*sd1* obtains an increased grain number but decreased plant height compared with Koshihikari, which provides a new strategy for improving crop yield through QTL pyramiding (Ashikari et al. 2005). The rational design rice which pyramided the good appearance genes *GS3* and *qSW5* from 9311 and the superior quality genes *Wx* and *ALK* from Nipponbare showed significantly improved grain appearance and quality compared with Teqing (Zeng et al. 2017). *GW5/GSE5* is a major QTL that regulates grain width and weight predominantly by influencing cell proliferation in spikelet hulls (Duan et al. 2017; Liu et al. 2017). *NALI/LSCHL4*, which regulates the development of the vascular bundle conferred leaf width, enhances the grain productivity of *indica* cultivars by increasing the number of secondary panicle branches and the spikelet number per panicle with declines in the seed-setting rate and in the 1000-grain weight (Fujita et al. 2013; Qi et al. 2008; Zhang et al. 2014). In this study, the confidence intervals of *qFLW4* and *qGW5* that were detected in the B set CSSLs overlapped with the locations of *Nali* and *GW5*, respectively (Table 2). The grain width performance of CSSL-B35 and CSSL-B36, which carried Nipponbare alleles segments at *qGW5*, showed significant differences compared with 9311 in across two environments, as well as the flag leaf width of CSSL-B10 at *qFLW4* (Additional file 4: Figure S2 D and

E). The increasing effects of Nipponbare alleles at both sites suggested a possibility of improving the 9311 yield by pyramiding the good appearance QTL *qFLW4* and *qGW5* from a cross between CSSL-B10 and CSSL-B35 or CSSL-B36.

PCR-based molecular markers are useful tools for selection during CSSLs development. However, they can hardly detect the small substituted segments or determine the length of substituted segments because the number of polymorphic markers used for selection is limited. Resequencing-based high-throughput genotyping technology can overcome these shortcomings by improving the resolution of the physical map accurately and reducing laborious and time-consuming work compared with the conventional molecular marker-based genotyping method (Watanabe et al. 2018; Xu et al. 2016; Zhang et al. 2011; Zhu et al. 2015b). As illustrated in Additional file 5: Figure S3, double-crossovers in CSSL-A35, CSSL-A44 and other CSSLs were detected based on throughput resequencing technology, while these double-crossovers can hardly detect by PCR-based molecular markers for the limited number of molecular markers. In this study, we developed 3 sets of CSSLs by molecular markers for selection and constructed high-density physical maps of the 3 sets of CSSLs by the whole-genome resequencing approach for high-throughput genotyping. A total of 197 lines were selected by 311 molecular markers for the development of the 3 sets of CSSLs. The coverage rates of substituted segments were 97.3%, 87.3% and 88.8% in the A, B and C sets of CSSLs respectively (Table 1). The number and length of substituted segments in 3 sets of CSSLs were also counted precisely using high-throughput genotyping data (Fig. 3), which provided a platform for genetic research and molecular rice breeding.

Due to significant phenotypic variations and the strong potential for heterosis between *indica* and *japonica* subspecies, many populations included some CSSLs that were derived from crosses between *indica* variety and *japonica* variety have been constructed for QTL mapping and gene cloning (Song et al. 2007; Wei et al. 2010; Zhou et al. 2016). To facilitate QTL detection and isolation, we developed 3 sets of CSSLs from three types of rice varieties: a typical *japonica* cultivar Nipponbare, an elite *indica* cultivar restore line 9311, and a middle type cultivar PA64s with a mixed genetic background of *indica* and *japonica*. To evaluate the potential advantages of the 3 sets of CSSLs for QTL detection, the phenotypic variations of 11 agronomic traits were observed between the 3 sets of CSSLs and their parents in Hangzhou and Shenzhen. Among the 71 QTLs detected in the 3 sets of CSSLs for 11 agronomic traits, 22 QTLs, including 4 known genes, were located in the vicinity of QTLs that affect these traits detected in the

other populations, which suggested less epistatic effects of these QTLs in varied genetic backgrounds (Fig. 4 and Table 2). Because of the pleiotropic effects or the tight linkage of genes, multiple QTLs affecting related traits were mapped to the same region to form QTL clusters, such as *qSPB12*, *qTGW12* and *qSN12* on chromosome 12 in A set CSSLs and *qFLL2*, *qPPB2-1* and *qSN2* on chromosome 2 in B set CSSLs, which also coincided with significant correlations, thus indicating that the relationship among these quantitative traits was intricate.

Although CSSLs are powerful tools for QTL detection, the intervals of QTLs are often too large, which are determined by the length of substituted segments. GWAS based on large-scale resequencing provide a powerful platform for finding genetic variants that can be directly used for crop improvement (Yano et al. 2016). In this study, the interval of *qGL1* was located to the region of 2.1 Mb by two CSSLs. In parallel, we also found 8 associated SNPs at *qGL1* location through GWAS and estimated a candidate region from 30.492 Mb to 30.652 Mb for *qGL1* by LD correlations. As the confidence interval resulted from GWAS is overlapped with the CSSLs location, the interval of *qGL1* was narrowed to the region of 119 kb which ranged from 30.492 Mb to 30.611 Mb on chromosome 1. The expression profiles of the diverse varieties of the collection in a particular tissue can more accurately aid in the determination of candidate genes, and annotation of candidate gene in the local region can also provide evidences for gene function (Xuehui et al. 2012). In this study, 3 genes expressed different significantly in the candidate region in young panicle tissue between 9311 and CSSL-B54. Among the 3 differentially expressed genes, *LOC_Os01g53140* and *LOC_Os01g53250* are the two most likely candidate genes for *qGL1*, as they may be involved in embryo development and energy metabolism, respectively, according to homologous alignment and gene function annotation.

Conclusions

In this study, three sets of CSSLs were developed containing 81, 55 and 64 lines, respectively, which provided a platform for basic genetic research and offered ideal materials for molecular rice breeding. High-throughput whole-genome resequencing strategy made the physical maps for the 3 sets of CSSLs more accurate for precise information on location and length of substituted segments, which facilitated novel QTLs discovery. Combined with GWAS on grain length and expression profiles on young panicle tissues, *qGL1* detected in CSSLs was fine mapped within a 119 kb region on chromosome 1 offering a new gene resource for improving grain shape.

Methods

Plant Materials

Nipponbare was the first map-based sequenced *japonica* cultivar. Restorer line 9311 and photothermo-sensitive male sterile line PA64s are the parents of pioneer super hybrid rice LYP9, which has been widely cultivated for commercial production in China. MAS was used for the development of three sets of CSSLs during crossing and backcrossing. Three parents and the populations derived from the three parents were grown annually at the experimental farm of China National Rice Research Institute in Hangzhou (119°54' E, 30°04' N) and Lingshui (110°00' E, 18°31' N).

DNA Extraction and Molecular Analysis

CTAB method was used for the extraction of genomic DNA from fresh young leaves of each individual (Rogers and Bendich 1988). The DNA amplification was performed by PCR (Polymerase Chain Reaction) with the protocol as follows: predenaturation at 94 °C for 4 min; 38 cycles of denaturation at 94 °C for 30 s, annealing at 55–60 °C for 30 s and extension at 72 °C for 30 s, with a final extension at 72 °C for 5 min. The reactions were carried out in 96-well PCR plates in 10 µl volumes containing 50–100 ng template DNA, 0.2 µmol/L of each primer and 5 µl of 2× Taq Master Mix (Vazyme Inc.). Electrophoresis of the amplification products were carried out on 4% agarose gels and photographed using GelDoc XR Gel Documentation System (Bio-Rad Inc.).

Most polymorphic molecular markers used for MAS were SSR markers that were selected from the Gramene Markers Database (McCouch et al. 2002; Project IRGS 2005). The other markers used were InDels markers developed from a BLASTN alignment between the genome sequence of three parents. Primers were designed on Primer3 web (<http://primer3.ut.ee/>).

High-Throughput Genotyping by Whole-Genome Resequencing

At least 1 µg of genomic DNA from each sample of CSSLs was used for resequencing with HiSeq PE150 strategy on Illumina HiSeq Xten. All DNA samples had backups for verification. Approximately 10× coverage sequence reads were generated for each CSSL. The *O. sativa* cv. Nipponbare IRGSP 1.0 reference genome was used for alignment (Kawahara et al. 2013). SNP calling was carried out with SAMtools (Li et al. 2009). A hidden Markov model was used to impute genotypes of recombinational chromosome fragments based on the observed genotypes of SNPs, following the previous method (Xie et al. 2010). The recombinational fragments less than 300 kb were identified as missing data in our study, and a crossover was defined between two adjacent blocks with different genotypes. An interval between

two adjacent crossovers in the whole CSSL population was regarded as a bin (Zhou et al. 2015).

Phenotypic Assessment for CSSLs

The 3 sets of CSSLs and their parents were grown for phenotypic assessment at the experimental farm in Hangzhou (119°54' E, 30°04' N) and Shenzhen (114°03' E, 22°32' N) in 2017. Twenty-five-day-old seedlings of each line were transplanted into an eight-row plot with six plants per row and spacing of 20 × 20 cm. The field management followed normal agricultural practice. The measurements of the plant height, panicle number, flag leaf length and flag leaf width were performed directly in the field at 25 days after heading. For all those traits, five plants of each line were sampled from the middle of each plot, and the main culm of each plant was chosen for trait measurement. At maturity, the main panicles of the individuals were harvested from 5 plants in the middle of each plot to measure the panicle and grain-related traits in the laboratory. For the five main panicles of each line, panicle length was measured with rulers; the primary panicle branch number, secondary panicle branch number and spikelet number per panicle were counted manually. Fully filled grains per panicle were used to calculate the grain length, grain width and 1000-grain weight using the WinSEEDLE Analysis System (Regent Instruments Inc.). Basic statistical analysis and correlation analysis were conducted with SAS 8.0 (SAS Institute). Five replications were performed for each trait.

QTL Analysis for 11 Major Agronomic Traits Based on CSSLs

Bin maps were converted from the physical maps of 3 sets of CSSLs for QTL analysis. Most donor segments between or among the different lines of each CSSL have a little overlap. To perform QTL analysis easily and precisely, the overlapping chromosome segments of the CSSLs were utilized to delineate smaller segment sizes that were described as bins (Huang et al. 2009b). The QTL analysis was performed from these bins that served as genetic markers using QTL IciMapping V4.0 (www.isbreeding.net/software/). The likelihood ratio test based on stepwise regression for the additive QTL (RSTEP-LR-T-ADD) method was employed for power analysis. The mapping parameters of probability in stepwise regression was set at 0.001, and the multicollinearity control by condition number was 1000 as the default. The LOD threshold for each dataset was set based on a permutation test (1000 permutations, $P = 0.05$). When the QTL LOD score was larger than 2.5, the QTL was designated having a major effect.

Genome-Wide Association Study and Linkage Disequilibrium

A total of 317 varieties are selected from 3010 diverse accessions of Asian cultivated rice and the SNPs information was acquired in public database (Wang et al. 2018). GWAS for the grain length were carried out using Efficient Mixed-Model Association eXpedited (EMMAX) software package (Hyun Min et al. 2010). The LD between SNPs in the 317 varieties was evaluated using squared Pearson's correlation coefficient (r^2) as calculated with the $-r^2$ command in the software PLINK. The LD heatmap surrounding peaks in the GWAS was constructed using LDheatmap R package (Shin et al. 2006).

RNA-Seq Analysis

We extracted total RNA from the 2–5 cm long young panicle tissues of rice plants with an RNA extraction kit (Trizol reagent, Invitrogen). The libraries were sequenced on a HiSeq. PE150 paired-end reads were generated. The sequencing reads were aligned to the Nipponbare (*Oryza sativa* L. spp. *japonica*) reference genome (<http://rice.plantbiology.msu.edu>, TIGR Version 7). A gene was considered expressed if the reads per kilobase of transcript model per million mapped reads was > 0.4. Differential gene expression between sample groups was analyzed using Cufflinks software for each gene. Genes were considered differentially expressed if the $|\log_2(\text{fold change})| \geq 1$ and FDR (false discovery rate) < 0.05.

Quantitative Real-Time RT-PCR Analysis

Total RNA was isolated from the 2 cm long young panicle tissues in Shenzhen with an RNA extraction kit (Trizol reagent, Invitrogen). cDNA was synthesized using a Rever Tra Ace[®] qPCR-RT kit (TOYOBA). Real time PCR amplification mixtures (10 μ l) contained 50 ng template cDNA, 2 × SYBR Green PCR Master Mix (Invitrogen), and 200 nM forward and reverse primers. Reactions were run on a StepOnePlus[™] Real-Time PCR System (Thermo Fisher Scientific). The relative expression level of each transcript was obtained by comparing to the expression of the *OsActin1* gene. Primers for candidate genes and *OsActin1* are listed in Additional file 2: Table S9.

Additional files

Additional file 1: Primers used for MAS during CSSLs development. (XLSX 35 kb)

Additional file 2: Table S1. Eleven traits of Nipponbare, PA64s and 9311 observed in Shengzhen (SZ) and Hangzhou (HZ). Mean \pm SD. ($n = 5$). * and ** indicate the least significant difference at 0.05 and 0.01 probability level compared with Nipponbare in SZ or HZ, respectively.

Table S2. Correlation analysis of 11 traits for A set CSSLs in Hangzhou (HZ). P values are shown below and bolds mean $P < 0.01$. **Table S3.** Correlation analysis of 11 traits for A set CSSLs in Shenzhen (SZ). P values

are shown below and bolds mean $P < 0.01$. **Table S4.** Correlation analysis of 11 traits for B set CSSLs in Hangzhou (HZ). P values are shown below and bolds mean $P < 0.01$. **Table S5.** Correlation analysis of 11 traits for B set CSSLs in Shenzhen (SZ). P values are shown below and bolds mean $P < 0.01$. **Table S6.** Correlation analysis of 11 traits for C set CSSLs in Hangzhou (HZ). P values are shown below and bolds mean $P < 0.01$. **Table S7.** Correlation analysis of 11 traits for C set CSSLs in Shenzhen (SZ). P values are shown below and bolds mean $P < 0.01$. **Table S8.** Candidate genes in the 119 kb interval for *qGL1*. Bold indicates differential expressed genes based on RNA-seq data. **Table S9.** Primers used for sequence identification. (DOCX 45 kb)

Additional file 3: Figure S1. Variation of the phenotypic traits in 3 sets of CSSLs. (TIF 1322 kb)

Additional file 4: Figure S2. Sequence difference and phenotypic variations on four reported genes. **A.** Difference sites between the related two parents. The difference sites are shown in red. Amino acids variations are listed in parenthesis. Empty blocks and blue blocks represent non-coding regions and exons, respectively. **B-E.** Phenotypic variation between CSSLs individuals and their recurrent parents. **B.** CSSL-A9 and CSSL-A53 harbor *Lax1* of PA64s allele at *qSPB1-2* locus. **C.** CSSL-C19 harbors *Sd1* of PA64s allele at *qPH1-2* locus. **D.** CSSL-B35 and CSSL-B36 harbor *GW5* of Nipponbare allele at *qGW5* locus. **E.** CSSL-B10 harbors *Nal1* of Nipponbare allele at *qFLW4* locus. Error bars are s.d. * and ** indicate the least significant difference at 0.05 and 0.01 probability level compared with the recurrent parent in SZ or HZ, respectively. NPB represents Nipponbare. (TIF 211 kb)

Additional file 5: Figure S3. Double-crossovers in CSSLs. Different colors represent different genotypes: red, 9311; blue, PA64s; green, Nipponbare. Double-crossovers were shown in black panes. (TIF 1447 kb)

Additional file 6: Figure S4. Quantitative real-time RT-PCR analysis and Sequence variations of predicted genes. **A-C.** RNA relative expression level of 9311, Nip and CSSL-B54 in young panicle for *LOC_Os01g53140*, *LOC_Os01g53140* and *LOC_Os01g53250*, respectively. **D-F.** Sequence variations between Nip and 9311 for *LOC_Os01g53140*, *LOC_Os01g53140* and *LOC_Os01g53250*, respectively. Values represent means \pm SD of three independent assays. ** indicate the least significant difference at 0.01 probability level compared with 9311. NPB represents Nipponbare. SNPs on the exons are shown in bold and Amino acids variations are listed in parenthesis. (TIF 3099 kb)

Abbreviations

CSSLs: Chromosome segment substitution lines; DH: Doubled haploid; FLL: Flag leaf length; FLW: Flag leaf width; GA: Gibberellin acid; GL: Grain length; GW: Grain width; GWAS: Genome-wide association study; Indels: Insertions and deletions; LD: Linkage disequilibrium; MAS: Marker-assisted selection; PCR: Polymerase chain reaction; PH: Plant height; PL: Panicle length; PN: Panicle number; PPB: Primary panicle branch number; QTLs: Quantitative trait loci; RILs: Recombinant inbred lines; SN: Spikelet number per panicle; SNPs: Single nucleotide polymorphisms; SPB: Secondary panicle branch number; SSR: Simple sequence repeat; TGW: 1000-grain weight

Acknowledgements

Not applicable.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 91735304, 31521064 and 31501384), Shenzhen Science and Technology Program (No. KQTD2016113010482651 and JSGG20160608160725473) and National Postdoctoral Program for Innovative Talents (No. BX201600151).

Availability of Data and Materials

The CSSLs materials are available from the corresponding author on reasonable request.

Authors' Contributions

QQ designed the research; BZ, LS, BR, AZ, SY, HJ, CL, KH, HL, ZG, JH, DZ and LG performed research; BZ, LS and B. R. analyzed data; BZ wrote the paper. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 March 2019 Accepted: 22 April 2019

Published online: 10 May 2019

References

- Ahn SN, Bollich CN, McClung AM, Tanksley SD (1993) RFLP analysis of genomic regions associated with cooked-kernel elongation in rice. *Theor Appl Genet* 87:27–32
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A et al (2005) Cytokinin oxidase regulates rice grain production. *Science* 309:741–745
- Bessho-Uehara K, Furuta T, Masuda K, Yamada S, Angeles-Shim RB, Ashikari M et al (2017) Construction of rice chromosome segment substitution lines harboring *Oryza barthii* genome and evaluation of yield-related traits. *Breed Sci* 67:408–415
- Bian JM, Jiang L, Liu LL, Wei XJ, Xiao YH, Zhang LJ et al (2010) Construction of a new set of rice chromosome segment substitution lines and identification of grain weight and related traits QTLs. *Breed Sci* 60:305–313
- Chen JB, Li XY, Cheng C, Wang YH, Qin M, Zhu HT et al (2014) Characterization of epistatic interaction of QTLs LH8 and EH3 controlling heading date in rice. *Sci Rep* 4:4263
- Duan PG, Xu JS, Zeng DL, Zhang BL, Geng MF, Zhang GZ et al (2017) Natural variation in the promoter of *GSE5* contributes to grain size diversity in Rice. *Mol Plant* 10:685–694
- Ebitani T, Takeuchi Y, Nonoue Y, Yamamoto T, Takeuchi K, Yano M (2005) Construction and evaluation of chromosome segment substitution lines carrying overlapping chromosome segments of indica Rice cultivar 'Kasalath' in a genetic background of japonica elite cultivar 'Koshihikari'. *Breed Sci* 55: 65–73
- Fujita D, Trijatmiko KR, Tagle AG, Sapasap MV, Koide Y, Sasaki K et al (2013) NAL1 allele from a rice landrace greatly increases yield in modern indica cultivars. *Proc Natl Acad Sci U S A* 110:20431–20436
- Furuta T, Uehara K, Angeles-Shim RB, Shim J, Ashikari M, Takashi T (2014) Development and evaluation of chromosome segment substitution lines (CSSLs) carrying chromosome segments derived from *Oryza rufipogon* in the genetic background of *Oryza sativa* L. *Breed Sci* 63:468–475
- Gao ZY, Zhao SC, He WM, Guo LB, Peng YL, Wang JJ et al (2013) Dissecting yield-associated loci in super hybrid rice by resequencing recombinant inbred lines and improving parental genome sequences. *Proc Natl Acad Sci U S A* 110:14492–14497
- Guo Y, Hong DL (2010) Novel pleiotropic loci controlling panicle architecture across environments in japonica rice (*Oryza sativa* L.). *J Genet Genomics* 37: 533–544
- Hao W, Jin J, Sun SY, Zhu MZ, Lin HX (2006) Construction of chromosome segment substitution lines carrying overlapping chromosome segments of the whole wild rice genome and identification of quantitative trait loci for rice quality. *J Plant Physiol Mol Biol* 32:354–362
- Hu J, Wang Y, Fang Y, Zeng L, Xu J, Yu H et al (2015) A rare allele of *GS2* enhances grain size and grain yield in Rice. *Mol Plant* 8:1455–1465
- Huang X, Qian Q, Liu Z, Sun H, He S, Luo D et al (2009a) Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet* 41:494–497
- Huang XH, Feng Q, Qian Q, Zhao Q, Wang L, Wang AH et al (2009b) High-throughput genotyping by whole-genome resequencing. *Genome Res* 19: 1068–1076

- Hyun Min K, Jae Hoon S, Service SK, Zaitlen NA, Sit-Yee K, Freimer NB et al (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354
- Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N, Onodera H et al (2013) Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat Genet* 45:707–711
- Jennifer K, Stephanie S, Manfred N, Lothar JN, Hans-Peter B (2010) Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22:797–810
- Jiao YQ, Wang YH, Xue DW, Wang J, Yan MX, Liu GF et al (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat Genet* 42: 541–544
- Kawahara Y, Bastide MDL, Hamilton JP, Kanamori H, Mccombie WR, Shu O et al (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4
- Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H et al (2003) LAX and SPA: major regulators of shoot branching in Rice. *Proc Natl Acad Sci U S A* 100:11765–11770
- Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T et al (2006) An SNP caused loss of seed shattering during Rice domestication. *Science* 312:1392–1396
- Kubo T, Aida Y, Nakamura K, Tsunematsu H, Doi K, Yoshimura A (2002) Reciprocal chromosome segment substitution series derived from japonica and Indica cross of Rice (*Oryza sativa* L.). *Breed Sci* 52:319–325
- Kubo T, Yamagata Y, Eguchi M, Yoshimura A (2008) A novel epistatic interaction at two loci causing hybrid male sterility in an inter-subspecific cross of rice (*Oryza sativa* L.). *Genes Genet Syst* 83:443–453
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
- Li ZK, Pinson SRM, Stansel JW, Park WD (1995) Identification of quantitative trait loci (QTLs) for heading date and plant height in cultivated rice (*Oryza sativa* L.). *Theor Appl Genet* 91:374–381
- Liu JF, Chen J, Zheng XM, Wu FQ, Lin QB, Heng YH et al (2017) GW5 acts in the brassinosteroid signalling pathway to regulate grain width and weight in rice. *Nat Plants* 3:17043
- McCouch S, Teytelman L, Xu Y, Lobos K, Clare K, Walton M et al (2002) Development and mapping of 2240 new SSR markers for rice. *DNA Res* 9: 199–207
- Mei HW, Xu JL, Li ZK, Yu XQ, Guo LB, Wang YP et al (2006) QTLs influencing panicle size detected in two reciprocal introgressive line (IL) populations in rice (*Oryza sativa* L.). *Theor Appl Genet* 112:648–656
- Mittal D, Chakrabarti S, Sarkar A, Singh A, Grover A (2009) Heat shock factor gene family in rice: genomic organization and transcript expression profiling in response to high temperature, low temperature and oxidative stresses. *Plant Physiol Biochem* 47:785–795
- Mulsanti IW, Yamamoto T, Ueda T, Samadi AF, Kamahora E, Rumanti IA et al (2018) Finding the superior allele of japonica-type for increasing stem lodging resistance in indica rice varieties using chromosome segment substitution lines. *Rice* 11:25
- Nadeau JH, Singer JB, Matin A, Lander ES (2000) Analysing complex genetic traits with chromosome substitution strains. *Nat Genet* 24:221–225
- Project IRGS (2005) The map-based sequence of the rice genome. *Nature* 436: 793–800
- Qi J, Qian Q, Bu Q, Li S, Chen Q, Sun J et al (2008) Mutation of the rice narrow leaf1 gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiol* 147:1947–1959
- Qiao W, Lan Q, Cheng Z, Long S, Jing L, Yan S et al (2016) Development and characterization of chromosome segment substitution lines derived from *Oryza rufipogon* in the genetic background of *O. sativa* spp. indica cultivar 9311. *BMC Genomics* 17:580
- Redona ED, Mackill DJ (1996) Mapping quantitative trait loci for seedling vigor in rice using RFLPs. *Theor Appl Genet* 92:395–402
- Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues. *Plant Mol Biol Manual* A6:1–10
- Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D et al (2002) A mutant gibberellin-synthesis gene in rice. *Nature* 416:701–702
- Schmidt R, Schippers JH, Welker A, Mieulet D, Guiderdoni E, Mueller-Roeber B (2012) Transcription factor Osh5fC1b regulates salt tolerance and development in *Oryza sativa* ssp. japonica. *AoB Plants* 2012:pls011
- Shen XH, Cao LY, Chen SG, Zhan XD, Wu WM, Cheng SH (2009) Dissection of QTLs for panicle traits in recombinant inbred lines derived from super hybrid Rice, Xieyou 9308. *Chin J Rice Sci* 23:354–362
- Shim RA, Angeles ER, Ashikari M, Takashi T (2010) Development and evaluation of *Oryza glaberrima* Steud. Chromosome segment substitution lines (CSSLs) in the background of *O. sativa* L. cv. Koshihikari. *Breed Sci* 60:61619
- Shin J-H, Blay S, McNeney B, Graham J (2006) LDheatmap: an R function for graphical display of pairwise linkage disequilibrium between single nucleotide polymorphisms. *J Stat Softw* 16:code snippet 3
- Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet* 39:623–630
- Takai T, Nonoue Y, Yamamoto SI, Yamanouchi U, Matsubara K, Liang ZW et al (2007) Development of chromosome segment substitution lines derived from backcross between indica donor Rice cultivar 'Nona Bokra' and japonica recipient cultivar 'Koshihikari'. *Breed Sci* 57:257–261
- Tian F, Li DJ, Fu Q, Zhu ZF, Fu YC, Wang XK et al (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. *Theor Appl Genet* 112:570–580
- Varadarajan JG, Jocelyne S-J-DC, Picbgu B, Chaboutcb ME, Gomord V, Coolbaugh RC et al (2010) ATR3 encodes a diflavin reductase essential for *Arabidopsis* embryo development. *New Phytol* 187:67–82
- Wang L, Wang AH, Huang XH, Zhao Q, Dong GJ, Qian Q et al (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 W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z et al (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557:43–48
- Wang X, Jin L, Zhu H, Wang S, Zhang G, Liu G (2017) QTL epistatic analysis for yield components with single-segment substitution lines in rice. *Plant Breed* 137:346–354
- Wang Y, Xiong G, Hu J, Jiang L, Yu H, Xu J et al (2015) Copy number variation at the GL7 locus contributes to grain size diversity in rice. *Nat Genet* 47:944–948
- Watanabe S, Shimizu T, Machita K, Tsubokura Y, Xia Z, Yamada T et al (2018) Development of a high-density linkage map and chromosome segment substitution lines for Japanese soybean cultivar Enrei. *DNA Res* 25:123–136
- Wei XJ, Xu JF, Guo HN, Jiang L, Chen SH, Yu CY et al (2010) DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol* 153:1747–1758
- Xie W, Feng Q, Yu H, Huang X, Zhao Q, Xing Y et al (2010) Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc Natl Acad Sci U S A* 107:10578–10583
- Xu HS, Sun Y, Zhou HJ, Yu SB (2007) Development and characterization of contiguous segment substitution lines with background of an elite restorer line. *Acta Agron Sin* 33:979–986
- Xu JJ, Zhao Q, Du PN, Xu CW, Wang BH, Feng Q et al (2010) Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *BMC Genomics* 11:656
- Xu Q, Yang S, Yu T, Xu X, Yan Y, Qi X et al (2016) Whole-genome resequencing of a cucumber chromosome segment substitution line and its recurrent parent to identify candidate genes governing powdery mildew resistance. *PLoS One* 11:e0164469
- Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L et al (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat Genet* 40:761–767
- Xuehui H, Yan Z, Xinghua W, Canyang L, Ahong W, Qiang Z et al (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 44:32
- Yamamoto T, Lin HX, Sasaki T, Yano M (2000) Identification of heading date quantitative trait locus Hd6 and characterization of its epistatic interactions with Hd2 in rice using advanced backcross progeny. *Genetics* 154:885–891
- Yan J, Zhu J, He C, Benmoussa M, Wu P (1999) Molecular marker-assisted dissection of genotype (times) environment interaction for plant type traits in rice (*Oryza sativa* L.). *Crop Sci* 39:538–544
- Yang Z, Jin L, Zhu H, Wang S, Zhang G, Liu G (2018) Analysis of epistasis among QTLs on heading date based on single segment substitution lines in Rice. *Sci Rep* 8:2071
- Yano K, Yamamoto E, Aya K, Takeuchi H, Lo PC, Hu L et al (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet* 48:927
- Yano M (2001) Genetic and molecular dissection of naturally occurring variation. *Curr Opin Plant Biol* 4:130–135

- Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T et al (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. *Plant Cell* 12: 2473–2483
- Ye G, Liang S, Wan J (2010) QTL mapping of protein content in rice using single chromosome segment substitution lines. *Theor Appl Genet* 121:741–750
- Yoshimura KDNIA (1997) The construction of chromosome substitution lines of African rice (*Oryza glaberrima* Steud.) in the background of japonica rice (*O. sativa* L.). *Rice Genet News* 14:39–41
- Zeng DL, Tian ZX, Rao YC, Dong GJ, Yang YL, Huang LC et al (2017) Rational design of high-yield and superior-quality rice. *Nat Plants* 3:17031
- Zhang GH, Li SY, Wang L, Ye WJ, Zeng DL, Rao YC et al (2014) LSCHL4 from japonica cultivar, which is allelic to NAL1, increases yield of Indica super Rice 93-11. *Mol Plant* 7:1350–1364
- Zhang H, Zhao Q, Sun ZZ, Liu QQ (2011) Development and high-throughput genotyping of substitution lines carrying the chromosome segments of indica 9311 in the background of japonica Nipponbare. *J Genet Genomics* 38:603–611
- Zhou Q, Miao H, Li S, Zhang SP, Wang Y, Weng YQ et al (2015) A sequencing-based linkage map of cucumber. *Mol Plant* 8:961–963
- Zhou Y, Dong GC, Tao YJ, Chen C, Yang B, Wu Y et al (2016) Mapping quantitative trait loci associated with root traits using sequencing-based genotyping chromosome segment substitution lines derived from 9311 and Nipponbare in Rice (*Oryza sativa* L.). *PLoS One* 11:e0151796
- Zhou Y, Xie Y, Cai J, Liu C, Zhu H, Jiang R et al (2017) Substitution mapping of QTLs controlling seed dormancy using single segment substitution lines derived from multiple cultivated rice donors in seven cropping seasons. *Theor Appl Genet* 130:1191–1205
- Zhu H, Liu Z, Fu X, Dai Z, Wang S, Zhang G et al (2015a) Detection and characterization of epistasis between QTLs on plant height in rice using single segment substitution lines. *Breed Sci* 65:192–200
- Zhu J, Niu Y, Tao Y, Wang J, Jian J, Tai S et al (2015b) Construction of high-throughput genotyped chromosome segment substitution lines in rice (*Oryza sativa* L.) and QTL mapping for heading date. *Plant Breed* 134:156–163
- Zhu S, Huang R, Wai HP, Xiong H, Shen X, He H et al (2017) Mapping quantitative trait loci for heat tolerance at the booting stage using chromosomal segment substitution lines in rice. *Physiol Mol Biol Plants* 23:1–9
- Zhu W, Lin J, Yang D, Zhao L, Zhang Y, Zhu Z et al (2009) Development of chromosome segment substitution lines derived from backcross between two sequenced Rice cultivars, Indica recipient 93-11 and japonica donor Nipponbare. *Plant Mol Biol Report* 27:126–131

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
