

Research Paper

Identification and validation of QTLs for 100-seed weight using chromosome segment substitution lines in soybean

Dequan Liu^{1,2}, Yongliang Yan^{1,3}, Yasunari Fujita^{1,2} and Donghe Xu^{*1}

¹ Japan International Research Center for Agricultural Science (JIRCAS), Ohwashi, Tsukuba, Ibaraki 305-8686, Japan

² Graduate School of Life and Environmental Sciences, University of Tsukuba, Tennodai, Tsukuba, Ibaraki 350-8572, Japan

³ Institute of Crop Germplasm Resources, Xinjiang Academy of Agricultural Sciences, Urumqi, Xinjiang 830000, P. R. China

The 100-seed weight (100SW) is one of the most important traits that control soybean yield. To identify the quantitative trait loci (QTL) of 100SW, 120 BC₃F₅ chromosome segment substitution lines (CSSLs) were cultivated over three years. The CSSLs were developed from a cross between the cultivated soybean variety ‘Jackson’ and the wild soybean accession ‘JWS156-1’, followed by continuous backcrossing using ‘Jackson’ variety as a recurrent parent. A total of nine QTLs (*qSW8.1*, *qSW9.1*, *qSW12.1*, *qSW13.1*, *qSW14.1*, *qSW16.1*, *qSW17.1*, *qSW17.2*, and *qSW20.1*) were detected on eight chromosomes. Of these, *qSW12.1* (LOD = 6.78–12.31) was detected over the three successive years on chromosome 12 as a novel, stable, and major QTL. To validate the effect of *qSW12.1*, a residual heterozygous line (RHL), RHL564, which showed heterozygous at the *qSW12.1* region, was selected from the BC₃F₅ population. Of the two homologous genotypes in the progenies produced by self-pollination of RHL564, a higher seed weight was observed in the ‘Jackson’ genotype plants than that in the ‘JWS156-1’ genotype plants. *qSW12.1* was delimited in an interval of approximately 1,348 kb between the BARCSOYSSR_12_1282 and BARCSOYSSR_12_1347 markers on chromosome 12.

Key Words: 100-seed weight, wild soybean, cultivated soybean, QTL, CSSLs.

Introduction

Soybean [*Glycine max* (L.) Merr.] is a primary crop for protein and oil production. With its increasing industrial and agricultural usages, soybean production has achieved sustainable growth worldwide (USDA-FAS, <http://www.fas.usda.gov/>). Grain yield improvement is always a priority in soybean breeding practice. The seed weight, which is generally expressed as 100-seed weight (100SW), is one of the most important yield components in soybean. Several studies have demonstrated a positive association between seed weight/size and seed yield (Burris *et al.* 1973, Smith and Camper 1975). Seed weight has also been reported to be associated with seed germination and vigor (Edwards and Hartwig 1971). The soybean varieties developed in the tropical and subtropical countries, such as India and Indonesia, tend to have a small seed size compared with varieties from the temperate regions, such as USA, China, and Japan. In addition, seed size is also an important factor for soybean end-use. Although a small seed size is preferred for natto

and soy sprouts, a large seed size is preferred for vegetable soybean and tofu. Therefore, depending on end-use or location of growth, different soybean varieties with different seed sizes have been established.

Over the past decades, many studies have focused on genetic studies of seed weight/size in the soybean, particularly using the DNA marker for quantitative trait locus (QTL) analyses. To date, many QTLs of seed weight/size on all the 20 chromosomes have been reported in SoyBase (<http://www.soybase.org/>). However, owing to the nature of multi-gene effects for controlling seed weight/size, most of the QTLs were minor and not validated. Examples of major seed-weight QTLs include one that is mapped on chromosome 17, *qSw17-1* (Kato *et al.* 2014), and another, *swHCC2-1*, associated with Satt460 on chromosome 6 (Han *et al.* 2012). Both QTLs demonstrated major and stable effects and a high level of phenotypic variation in multiple populations and environments.

Compared with other major crops, such as rice (Duan *et al.* 2015, Hu *et al.* 2015, Wang *et al.* 2015), knowledge about the molecular mechanisms of seed weight control in soybean is limited. To date, only two seed weight/size related genes have been isolated from soybean. The *ln* (narrow leaf) gene has a large effect on the number of seeds per pod and seed size (Jeong *et al.* 2012), and recently, the *PP2C-1*

Communicated by Masao Ishimoto

Received October 19, 2017. Accepted June 3, 2018.

First Published Online in J-STAGE on August 28, 2018.

*Corresponding author (e-mail: xudh@affrc.go.jp)

(protein phosphatase type-2C) allele from wild soybean accession ZYD7 was found to contribute to the increase in seed size (Lu *et al.* 2017). Identification of a major and stable QTL for seed weight is important for facilitating the map-based cloning of seed-weight genes and understanding the molecular mechanisms underlying seed-weight control in soybean.

Most of the previously reported seed-weight QTLs were detected using bi-parental mapping populations, such as recombinant inbred lines. In recent years, QTLs associated with traits of interest have been identified using a more effective molecular breeding study strategy known as advanced backcross QTL (AB-QTL) analysis (Grandillo and Tanksley 2005). Through continuous backcrossing followed by successive self-pollination, different chromosomal segments of the donor parent were introgressed into the genetic background of the recurrent parent to produce chromosome segment substitution lines (CSSLs). In CSSLs, more than 90% of the genome was recovered by the recurrent parent after backcrossing thrice, thus, reducing the genetic background noise and improving the accuracy of gene mapping (Grandillo and Tanksley 2005). Several CSSL mapping populations have been created for soybean and have been used for QTL mapping of agronomy traits (Wang *et al.* 2013, Xin *et al.* 2016). For these CSSL populations, wild soybean (*G. soja* Sieb. & Zucc.), which is believed to be the wild progenitor of the cultivated soybean, was used as donor parent. The wild type has a higher genetic diversity than the cultivated form (Zhou *et al.* 2015), presenting a potential genetic resource pool for improvement of the cultivated soybean. The seed size of the wild soybean is approximately ten times smaller compared with the cultivated soybean. Thus, wild soybean is a particularly good material for genetic analyses of soybean seed weight. In this study, we used a CSSL population derived from crosses between the cultivated soybean variety 'Jackson' and wild soybean accession 'JWS156-1' to identify the seed-weight QTLs.

Materials and Methods

Plant materials

A total of 120 BC₃F₅ CSSLs were used to detect the 100SW QTL. The CSSL population was developed from a cross between the cultivated soybean variety 'Jackson' and the wild soybean accession 'JWS156-1'. 'Jackson' is a soybean variety from the US (PI548657). JWS156-1 was selected from a wild soybean accession B06097, which was provided by the National BioResource Project-LegumeBase (<https://www.legumebase.brc.miyazaki-u.ac.jp/top.jsp>). The 100SW of 'Jackson' was around 22 g and that of 'JWS156-1' was about 2.2 g, which is a ten-fold difference. The wild soybean was the donor parent and the 'Jackson' cultivar was the recurrent parent. Different chromosomal segments of the wild soybean were introgressed into the genetic background of 'Jackson'. These lines were developed by backcrossing of {[(('Jackson' × 'JWS156-1') × 'Jackson') ×

'Jackson'] × 'Jackson'} and five successive generations of self-pollination without any selection (**Supplemental Fig. 1**).

To validate a novel major QTL (*qSW12.1*) detected in the current study, a residual heterozygous line (RHL), RHL564, which was heterozygous at the *qSW12.1* QTL region, was selected from the BC₃F₅ population. Progenies produced by self-pollinating heterozygous plants of RHL564 were used to confirm the effect of *qSW12.1*.

Field trials and phenotypic evaluation of the BC₃F₅ CSSLs

All plant materials were grown in the experimental farm of the Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki, Japan; located at 36.05°N, 140.08°E. The 120 BC₃F₅ CSSLs were grown in 2014, 2015, and 2016. In 2014, a single BC₃F₅ plant for each CSSL was evaluated for seed weight. Seeds harvested from the single plant were used for phenotype evaluation in 2015, and the bulked seeds harvested in 2015 for each CSSL were used for phenotype evaluation in 2016. Soybean seeds were sown on June 24, 2014; June 23, 2015; and June 21 2016. Each CSSL or cultivar was planted in a single 6-m long row plot, which contained 30 plants per row, with a row spacing of 60 cm and at a plant separation of 20 cm. All plots were arranged in a completely randomized block design with one (2014, 2015) or two (2016) replicates. Soybeans were harvested on November 14, 2014; November 12, 2015; and November 16 2016. A single BC₃F₅ plant from each line was harvested in 2014, whereas in 2015 and 2016, BC₃F₅ plants were harvested in bulk from each row. The individual plants at each end of the row were excluded from harvest in the three years.

After harvesting, soybean plants were air-dried in natural conditions for at least four weeks and then threshed. When the seed desiccation was complete, the 100SW trait was measured for each BC₃F₅ CSSL expressed as grams per 100 random healthy seeds.

SSR analysis

Genomic DNA was extracted from young leaf tissues using a modified CTAB method. A total of 235 SSR soybean markers from all 20 chromosomes (Song *et al.* 2004, 2010) were used for genotyping the 120 CSSLs. The 235 SSR markers covered 2,073.79 cM with an average distance of 9.5 cM between two neighbor markers. The SSR reaction mix contained 2 μl (10 pmol) of each primer, 10 μl Quick TaqTMHS DyeMix (Toyobo, Osaka, Japan), and 3 μl (10–50 ng) of genomic DNA template in a total volume of 20 μl. The PCR was performed at 94°C for 30 s, followed by 32 cycles of 30 s at 94°C, 30 s at 55°C, 40 s at 72°C, and a final extension at 72°C for 10 min. After amplification, PCR products were separated on an 8% (w/v) denaturing polyacrylamide gel in 1 × TBE running buffer at 250 V for 3.5 h and stained using ethidium bromide. To detect polymorphisms in fragment size of the PCR products, the gel was scanned using the pharos FX Molecular Imager (Bio-Rad Laboratories, Hercules, USA).

QTL analysis

QTL analysis was conducted with the IciMapping software (Meng *et al.* 2015). The Stepwise regression-based likelihood ratio tests of additive QTL (RSTEP-LRT-ADD) method was used to detect 100SW QTL. A threshold of $LOD \geq 3$ was set to declare existence of a QTL.

Validation of the allelic effect of *qSW12.1*

To further verify the effect of the *qSW12.1* QTL detected in the current study, the residual heterozygous line (RHL), RHL564, was selected from the BC₃F₅ population. Homozygous plants with the ‘Jackson’ and ‘JWS156-1’ genotypes were selected from the progenies produced through the self-pollination of heterozygous plants of RHL564 using the SSR marker Sat_180 for genotyping. Because these two genotypes had a similar genetic background and differed only in the *qSW12.1* region, they can be regarded as near-isogenic lines (NILs). For the homozygous plants with the ‘Jackson’ genotype, we designated the lines as NILs564-C, and the homozygous plants with the ‘JWS156-1’ genotype were designated as NILs564-W. Forty-three NILs564-C plants and 44 NILs564-W plants were evaluated in field condition in 2016. The cultivation conditions were the same as those used for the CSSLs in 2016. A two-tailed student’s t-test was performed to determine the statistical significance between NILs564-C and NILs564-W.

Results

Genotypic characteristics of the BC₃F₅ CSSLs

The graphical genotypes of the 120 CSSLs are represented in **Supplemental Fig. 2**. Of the 235 SSR markers, 224 markers had at least one ‘JWS156-1’ allele in the 120 CSSLs, and 11 markers demonstrated no wild soybean alleles as indicated by the small gaps in the graphical genotypes. The CSSLs were almost recovered by the recurrent parent ‘Jackson’ after backcrossing thrice, and no lines with abnormal growth were observed in the 120 CSSLs. The proportion of the recurrent parent ‘Jackson’ alleles in each CSSL ranged from 80.3% to 99.2% with an average of 92.9 ± 4.0%, which corresponded to the expected value, 93.8%.

The phenotypic variation of the BC₃F₅ CSSLs in the three years

The frequency distribution and statistical analysis of 100SW in the 120 BC₃F₅ CSSLs over the three years are shown in **Fig. 1** and **Table 1**. The 100SW demonstrated a continuous distribution, and the seed size of most CSSLs was intermediate in comparison with that of the two parents. However, in 2014, 2015, and 2016, one, three, and one CSSLs, respectively, demonstrated a higher 100SW than the cultivated soybean ‘Jackson’. ANOVA analyses revealed significant effects ($P < 0.01$) of the different CSSLs and the different years. The 120 CSSLs showed a large variation in terms of 100SW, ranging approximately from 10 g to 22 g, with a two-fold difference. Over the three years, the highest

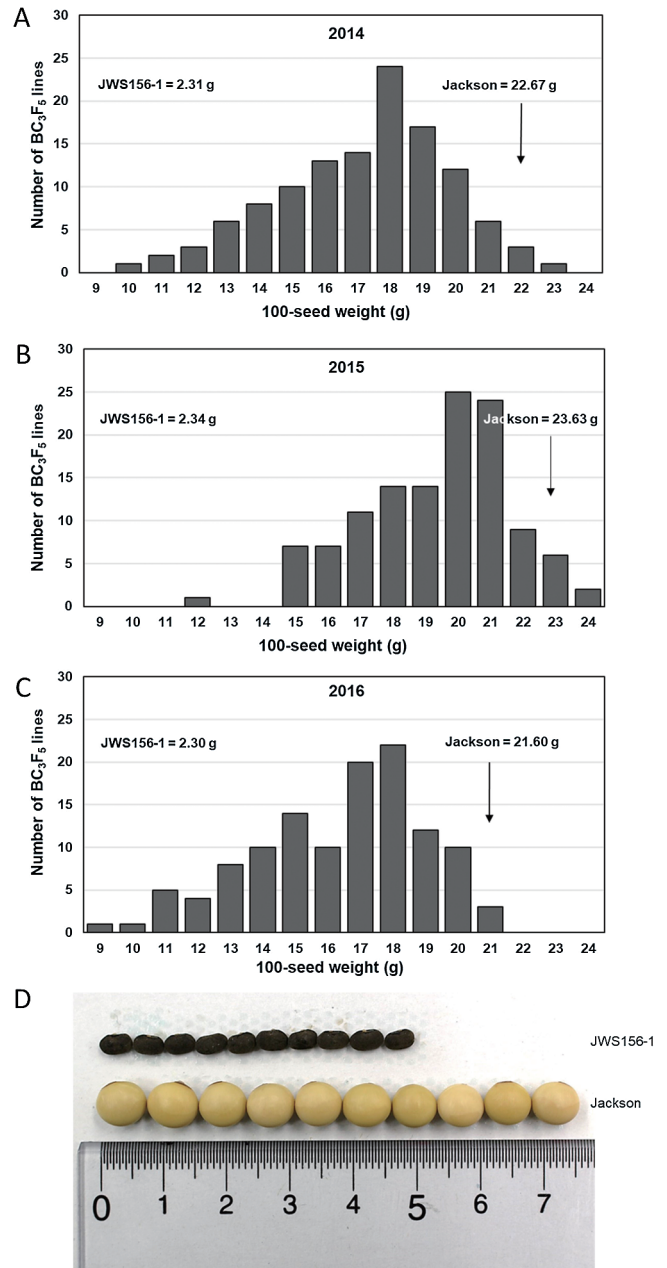


Fig. 1. Frequency distribution of 100-seed weight of the 120 BC₃F₅ chromosome segment substitution lines in the years 2014 (A), 2015 (B), and 2016 (C). Comparison of the seed sizes between the ‘JWS156-1’ line and ‘Jackson’ (D).

average 100SW of all CSSLs was 19.84 g (2015), and the lowest was 16.89 g (2016), whereas the average in 2014 was 17.64 g. These results indicated different growth conditions over the three years. There was a strong positive correlation ($P < 0.01$) between different years (**Fig. 2**), suggesting that seed-weight trait has high heritability in soybean and it is suitable for QTL analysis.

QTLs associated with seed weight

Based on QTL analysis, four, five, and five SSR markers

Table 1. Statistical analysis of 100-seed weight in the BC₃F₅ chromosome segment substitution line (CSSL) population in 2014, 2015, and 2016

Year	Trait	Parents		BC ₃ F ₅ CSSLs								
		Jackson	JWS156-1	SampleSize ^a	Mean	Variance ^b	Std ^c	Skewness ^d	Kurtosis ^e	Minimum	Maximum	Range
2014	100SW (g)	22.67	2.31	120	17.64	6.51	2.55	-0.38	-0.33	10.98	23.04	12.06
2015	100SW (g)	23.63	2.34	120	19.84	5.26	2.29	-0.46	-0.06	12.40	24.46	12.06
2016	100SW (g)	21.60	2.30	120	16.89	6.67	2.58	-0.47	-0.32	9.81	21.88	12.08

^a SampleSize: Size of the mapping population.

^b Variance: Variance of the phenotypic trait.

^c Std: Standard deviation of the phenotypic trait.

^d Skewness is a measure of the phenotypic trait of the probability distribution of a real-valued random variable about its mean.

^e Kurtosis is a measure of the phenotypic trait of the probability distribution of a real-valued random variable.

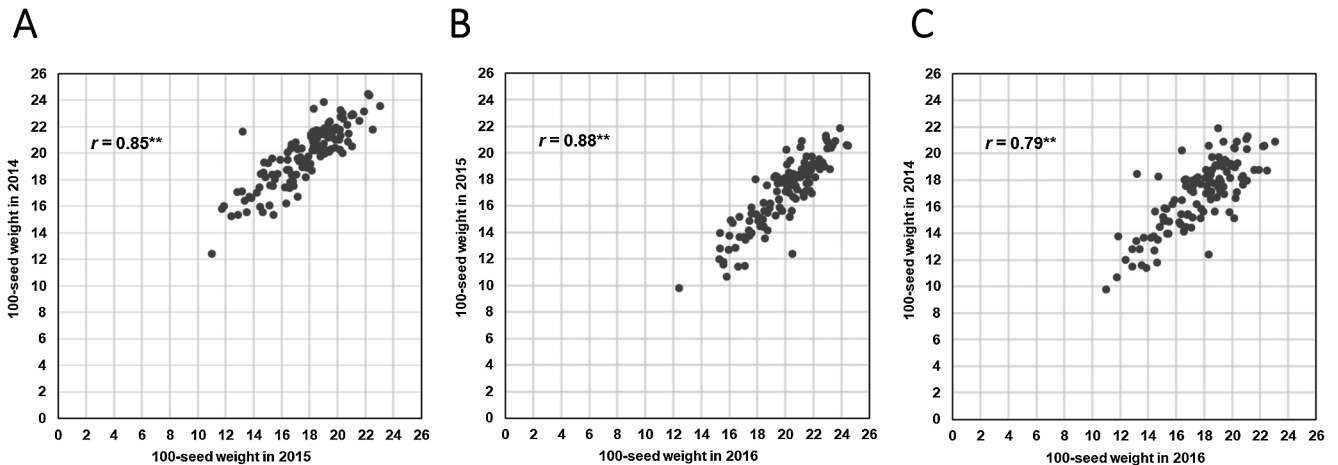


Fig. 2. The correlation between the 100-seed weight of BC₃F₅ chromosome segment substitution lines over different years. ** Indicates significant correlation ($P < 0.01$).

associated with 100SW were detected in 2014, 2015, and 2016, respectively (Table 2). These markers were distributed on chromosomes 8, 9, 12, 13, 14, 16, 17, and 20. BARCSOYSSR_14_0645 ($3.00 < \text{LOD} < 6.63$) located on chromosome 14 and Sat_180 ($6.78 < \text{LOD} < 12.31$) located on chromosome 12 were detected in all three years. Satt582 ($7.16 < \text{LOD} < 8.99$) located on chromosome 17 was detected in both 2014 and 2015. Satt431 ($\text{LOD} = 4.52$, chromosome 16), Satt490 ($\text{LOD} = 3.77$, chromosome 13), Satt354 ($\text{LOD} = 3.83$, chromosome 20), Sat_129 ($\text{LOD} = 7.66$, chromosome 8), Sat_126 ($\text{LOD} = 3.46$, chromosome 9), and Sat_086 ($\text{LOD} = 3.33$, chromosome 17) were detected in only one year. Overall, a total of nine markers were detected over the three years. Sat_180 had the highest LOD scores in 2015 and 2016 and the second highest LOD score in 2014. In all nine loci, the wild soybean allele contributed to decreasing seed size with effects ranging from -1.96 g to -0.87 g.

Based on their chromosome locations, nine 100SW QTLs, corresponding to the nine SSR markers associated with 100SW, were designated as *qSW8.1*, *qSW9.1*, *qSW12.1*, *qSW13.1*, *qSW14.1*, *qSW16.1*, *qSW17.1*, *qSW17.2*, and *qSW20.1* (Table 2). For the three major and stable QTLs (*qSW12.1*, *qSW14.1*, and *qSW17.1*) that were detected in the current study, *qSW14.1* may correspond to the seed-weight QTL reported by Hoeck *et al.* (2003) and Li *et al.* (2007),

whereas *qSW17.1* may correspond to the QTL (*qSw17-1*) detected by Kato *et al.* (2014). In contrast, no seed-weight QTL has been reported for the *qSW12.1* region to date, suggesting that this is a novel major QTL in soybean.

Validating the allelic effect and determining the physical position of *qSW12.1*

To further verify the effect of the *qSW12.1* QTL, RHL564 was selected from the BC₃F₅ population. Homozygous plants with the ‘Jackson’ and ‘JWS156-1’ genotypes were selected from the progenies of RHL564 to produce *qSW12.1* NILs (NILs564-C and NILs564-W). Field evaluation revealed that NILs564-C had 100SW of 18.64 ± 1.03 g, which was significantly ($P < 0.01$) higher than that of the NILs564-W plants (17.71 ± 1.25 g) (Fig. 3). This result confirms that the *qSW12.1* QTL controls seed weight in soybean.

Based on the genotyping of NILs564-C and NILs564-W derived from RHL564, we delimited *qSW12.1* in a 1,348-kb interval between the BARCSOYSSR_12_1282 and BARCSOYSSR_12_1347 markers (Fig. 4). According to the soybean genome sequence of *Glycine max Wm82.a2.v1* at Phytozome 12 (<https://phytozome.jgi.doe.gov>), there are 129 predicted genes in this region.

Table 2. SSR makers associated with 100-seed weights (100SW) in the BC₃F₅ chromosome segment substitution line (CSSL) population over different years

Year	Marker name	QTL	Chr. (LG)	Physical position (bp) ^a / Genetic position (cM) ^b	LOD ^c	PVE (%) ^d	Add ^e	M (QQ) ^f	M (qq) ^g	DHR(%) ^h
2014	Sat_180	<i>qSW12.1</i>	Gm12 (H)	38,388,961/104.37	6.78	15.25	-1.21	12.26	14.69	19.17
	BARCSOYSSR_14_0645	<i>qSW14.1</i>	Gm14 (B2)	13,276,701/-	3.00	6.27	-1.11	12.37	14.59	8.33
	Satt431	<i>qSW16.1</i>	Gm16 (J)	35,718,476/78.57	4.52	9.72	-1.38	12.10	14.86	8.33
	Satt582	<i>qSW17.1</i>	Gm17 (D2)	9,949,907/53.84	8.99	21.16	-1.75	11.72	15.23	11.67
2015	Sat_180	<i>qSW12.1</i>	Gm12 (H)	38,388,961/104.37	10.07	21.06	-1.28	14.01	16.56	19.17
	Satt490	<i>qSW13.1</i>	Gm13 (F)	36,699,447/97.97	3.77	6.94	-0.87	14.41	16.16	14.17
	BARCSOYSSR_14_0645	<i>qSW14.1</i>	Gm14 (B2)	13,276,701/-	5.18	9.80	-1.24	14.05	16.53	8.33
	Satt582	<i>qSW17.1</i>	Gm17 (D2)	9,949,907/53.84	7.16	14.11	-1.28	14.00	16.57	11.67
	Satt354	<i>qSW20.1</i>	Gm20 (I)	34,569,381/46.22	3.83	7.07	-1.24	14.04	16.53	5.83
2016	Sat_129	<i>qSW8.1</i>	Gm8 (A2)	14,730,465/84.08	7.66	13.16	-1.95	7.74	11.63	6.67
	Sat_126	<i>qSW9.1</i>	Gm9 (K)	44,406,619/108.19	3.46	5.46	-1.43	8.25	11.12	5.00
	Sat_180	<i>qSW12.1</i>	Gm12 (H)	38,388,961/104.37	12.31	23.25	-1.64	8.04	11.32	19.17
	BARCSOYSSR_14_0645	<i>qSW14.1</i>	Gm14 (B2)	13,276,701/-	6.63	11.17	-1.62	8.07	11.30	8.33
	Sat_086	<i>qSW17.2</i>	Gm17 (D2)	39,875,727/118.66	3.33	5.25	-1.96	7.72	11.65	2.92

^a Physical positions are based on *Glycine max* Wm82.a2.v1 (Schmutz *et al.* 2010).

^b Genetic positions are based on SoyBase (<https://soybase.org/>).

^c LOD: LOD score calculated from single marker analysis.

^d PVE (%): Phenotypic variation explained by the marker.

^e Add: Estimated additive effect of the marker.

^f M (QQ): Mean value of the QTL genotype of donator parent.

^g M (qq): Mean value of the QTL genotype of recurrent parent.

^h DHR: Genotype ratio of the donor homozygous allele.

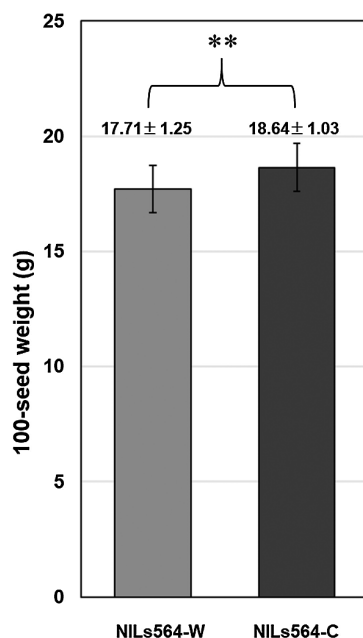


Fig. 3. The allele effect of *qSW12.1* in near-isogenic lines NILs564-C ('Jackson' genotype, $n = 43$) and NILs564-W ('JWS156-1' genotype, $n = 44$). Data are shown as means \pm SD (standard deviation). ** Indicates significant difference between the means ($P < 0.01$).

Discussion

Wild soybean has been found to have the potential to improve cultivated soybean in terms of seed quality, yield-related traits, and tolerance to biotic and abiotic stress

(Concibido *et al.* 2003, Hu *et al.* 2013, Kim *et al.* 2011, Sebolt *et al.* 2000, Tuyen *et al.* 2010). In agreement with previous studies (Liu *et al.* 2007, Wang *et al.* 2016, Xin *et al.* 2016), we demonstrated that alleles from wild soybean contribute to a reduction in seed weight in all nine 100SW QTLs. Soybean breeders are not only interested in increasing seed size, but also require an optimized combination for each yield factor such as seed size, the number of seeds per pod, and the number of pods per plant. The QTLs detected in the current study would be useful to control seed size via genomic breeding by design and positional cloning of the relevant genes.

Here we detected three major and stable QTLs (*qSW12.1*, *qSW17.1*, and *qSW14.1*). These QTLs are good candidates for cloning the genes responsible for 100SW because a major QTL can be steadily detected in different environmental conditions. The QTL *qSW12.1* was identified in three successive years accounting for 15.25%–23.25% of the total phenotypic variation of 100SW as a novel locus associated with this trait. This QTL might differ from a seed-weight QTL previously detected on chromosome 12 (Han *et al.* 2012), occurring at an interval between Satt181 (91.12 cM, genetic position based on SoyBase and so forth) and Satt302 (81.04 cM), and where the proximal marker is approximately 10 cM away from *qSW12.1*.

The *qSW14.1* QTL was detected over three successive years accounting for 6.27%–11.17% of the total phenotypic variation of 100SW. Seed-weight QTL was previously detected in the same region by Hoeck *et al.* (2003) and Li *et al.* (2007), indicating consistency across different environmental conditions and genetic backgrounds.

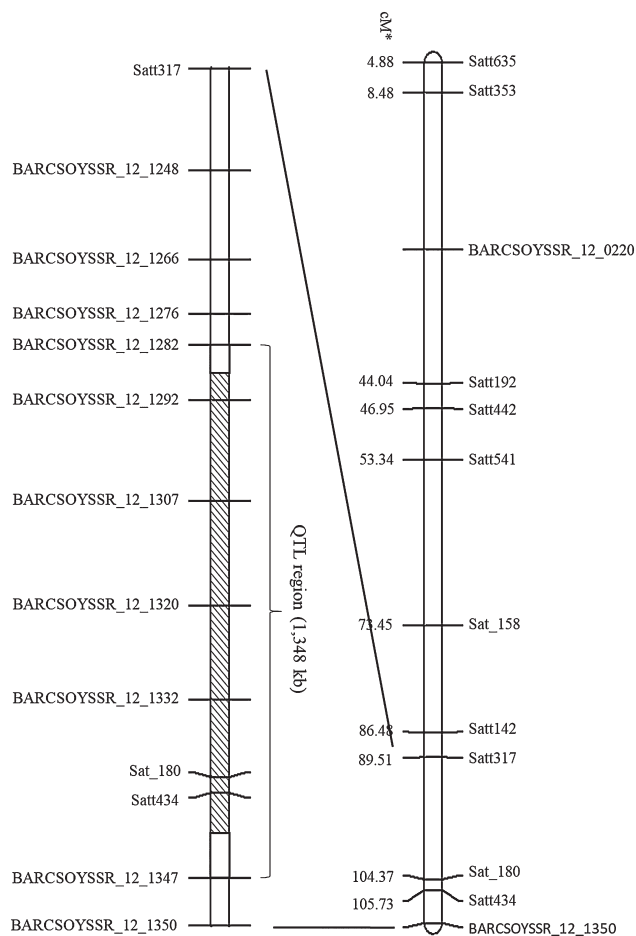


Fig. 4. The novel seed weight QTL (*qSW12.1*) was delimited in a 1,348-kb interval between the BARCSOYSSR_12_1282 and BARCSOYSSR_12_1347 markers on chromosome 12. *marker positions are based on SoyBase (<https://soybase.org/>). The shaded area indicates the different chromosomal region between NILs564-C and NILs564-W.

Although the *qSW17.1* QTL was detected over two years (2014 and 2015) as a major QTL with LOD scores of 8.99 and 7.16, respectively, it was not detected in 2016. As is shown in **Table 1**, the 120 CSSLs and the two parents demonstrated the lowest 100SW in 2016 among the three years. The special growth conditions in 2016 may have affected the effect of *qSW17.1* on seed weight. Other researchers working on different populations have consistently detected QTLs for seed weight in this region (Kato *et al.* 2014, Liu *et al.* 2007, Wang *et al.* 2016, Xin *et al.* 2016). Recently, Lu *et al.* (2017) identified a protein phosphatase type-2C allele (*PP2C-1*) as the causative gene underlying a QTL that contributed a positive effect toward increasing seed weight. This QTL was in the same region of *qSW17.1*. However, the result in Lu *et al.* (2017) differs from our study in which the wild soybean contributed a negative effect, decreasing seed weight. It would be very interesting to determine whether these conflicting results are caused by the different genes or different alleles within the same locus

of the wild soybean genome.

As for the remaining significant QTLs, six were detected in this study (*qSW8.1*, *qSW9.1*, *qSW13.1*, *qSW16.1*, *qSW17.2*, and *qSW20.1*). Of these, *qSW20.1* might correspond to a region associated with seed weight reported by Han *et al.* (2012) and Liu *et al.* (2011). However, these six QTLs require further corroboration because they were detected in one year. Although some CSSLs demonstrated higher 100SW than that of the cultivated parent 'Jackson', no positive effect allele was detected from the wild soybean, suggesting the presence of undetected small effect QTLs or epistasis.

Although, on average, more than 90% of the genome was recovered by the recurrent parent in the BC₃-derived CSSLs, there were several CSSLs with quite low 100SW. For example, CSSLs507 and CSSLs505 had an average of 11.06 ± 1.30 g and 12.74 ± 2.71 g for 100SW over the three years. Genotypic analysis revealed that the CSSLs507 line had the 'JWS156-1' genotypes at four 100SW QTLs (*qSW8.1*, *qSW12.1*, *qSW13.1*, and *qSW17.1*) and the CSSLs505 line had the 'JWS156-1' genotypes at three 100SW QTLs (*qSW9.1*, *qSW12.1*, and *qSW17.1*), both including the two major QTLs, *qSW12.1* and *qSW17.1* (**Supplemental Fig. 1**). Further genetic analysis for 100SW using a segregation population (BC₄) derived from crosses between these CSSLs and 'Jackson' could help us to understand the relationships between the 100SW QTLs.

Analysis of the *qSW12.1* near-isogenic lines, NILs564-C and NILs564-W, not only confirmed the effect of *qSW12.1* but also delimited it in a 1,348-kb interval between the BARCSOYSSR_12_1282 and BARCSOYSSR_12_1347 markers on chromosome 12. There are 129 predicted genes in this region based on the soybean genome sequence of *Glycine max Wm82.a2.v1* at Phytozome 12. However, it is still difficult to identify the candidate gene underlying *qSW12.1* from the 129 predicted genes.

In the current study, we identified a novel major QTL, *qSW12.1*, within a 1,348-kb interval region on chromosome 12. This mapping is sufficiently accurate for DNA marker assisted selection in a breeding program. While the region is currently too large for positional cloning of the causal gene underlying the QTL, our ongoing fine-mapping work using a much larger segregation population will enable us to narrow down the *qSW12.1* QTL into a smaller chromosome region for identifying the candidate gene causal to the seed-weight trait. Compared to genetic studies of seed weight in other plants, such studies are very limited in the soybean. Cloning genes controlling seed weight will not only contribute to a better understanding of the mechanism of seed development but also provide important implications for breeding strategies of soybean.

Acknowledgments

We thank Ms. K. Suzuki, Ms. M. Noda, and Ms. T. Kobayashi for their technical assistance. The work was sup-

ported in part by the scholarship from the China Scholarship Council to D. Q. Liu.

Literature Cited

- Burris, J.S., O.T. Edje and A.H. Wahab (1973) Effects of seed size on seedling performance in soybeans: II. seedling growth and photosynthesis and field performance. *Crop Sci.* 13: 207–210.
- Concibido, V.C., B. LaVallee, P. McIaird, N. Pineda, J. Meyer, L. Hummel, J. Yang, K. Wu and X. Delannay (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor. Appl. Genet.* 106: 575–582.
- Duan, P., S. Ni, J.M. Wang, B.L. Zhang, R. Xu, Y.X. Wang, H.Q. Chen, X.D. Zhu and Y.H. Li (2015) Regulation of *OsGRF4* by *OsmiR396* controls grain size and yield in rice. *Nat. Plants* 2: 15203.
- Edwards, C.J. and E.E. Hartwig (1971) Effect of seed size upon rate of germination in soybeans. *Agron. J.* 63: 429–450.
- Grandillo, S. and S.D. Tanksley (2005) Advanced backcross QTL analysis: results and perspectives. In: Tuberosa, R., R.L. Phillips and M. Gale (eds.) Proceedings of the international congress “In the wake of the double helix: from the green revolution to the gene revolution”. Bologna, Italy, pp. 115–132.
- Han, Y.P., D.M. Li, D. Zhu, H.Y. Li, X.P. Li, W.L. Teng and W.B. Li (2012) QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. *Theor. Appl. Genet.* 125: 671–683.
- Hoeck, J.A., W.R. Fehr, R.C. Shoemaker, G.A. Welke, S.L. Johnson and S.R. Cianzio (2003) Molecular marker analysis of seed size in soybean. *Crop Sci.* 43: 68–74.
- Hu, J., Y.X. Wang, Y. Fang, L. Zeng, J. Xu, H. Yu, Z. Shi, J. Pan, D. Zhang, S. Kang *et al.* (2015) A rare allele of *GS2* enhances grain size and grain yield in rice. *Mol. Plant* 8: 1455–1465.
- Hu, Z.B., D. Zhang, G.Z. Zhang, G.Z. Kan, D.L. Hong and D.Y. Yu (2013) Association mapping of yield-related traits and SSR markers in wild soybean (*Glycine soja* Sieb. and Zucc.). *Breed. Sci.* 63: 441–449.
- Jeong, N., S.J. Suh, M.H. Kim, S. Lee, J.K. Moon, H.S. Kim and S.C. Jeong (2012) *Ln* is a key regulator of leaflet shape and number of seeds per pod in soybean. *Plant Cell* 24: 4807–4818.
- Kato, S., T. Sayama, K. Fuji, S. Yumoto, Y. Kono, T. Y. Hwang, A. Kikuchi, Y. Takada, Y. Tanaka, T. Shiraiwa *et al.* (2014) A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor. Appl. Genet.* 127: 1365–1374.
- Kim, M., D.L. Hyten, T.L. Niblack and B.W. Diers (2011) Stacking resistance alleles from wild and domestic soybean sources improves soybean cyst nematode resistance. *Crop Sci.* 51: 934–943.
- Li, D.D., T.W. Pfeiffer and P.L. Cornelius (2007) Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Sci.* 48: 571–581.
- Liu, B.H., T. Fujita, Z.H. Yan, S. Sakamoto, D.H. Xu and J. Abe (2007) QTL mapping of domestication-related traits in soybean (*Glycine max*). *Ann. Bot.* 100: 1027–1038.
- Liu, W.X., M.Y. Kim, K.J. Van, Y.H. Lee, H.L. Li, X.H. Liu and S.H. Lee (2011) QTL identification of yield-related traits and their association with flowering and maturity in soybean. *J. Crop Sci. Biotechnol.* 14: 65–70.
- Lu, X., Q. Xiong, T. Cheng, Q.T. Li, X.L. Liu, Y.D. Bi, W. Li, W.K. Zhang, B. Ma, Y.C. Lai *et al.* (2017) A *PP2C-1* allele underlying a quantitative trait locus enhances soybean 100-seed weight. *Mol. Plant* 10: 670–684.
- Meng, L., H.H. Li, L.Y. Zhang and J.K. Wang (2015) QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3: 269–283.
- Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Son, J.J. Thelen, J. Cheng *et al.* (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183.
- Sebolt, A.M., R.C. Shoemaker and B.W. Diers (2000) Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Sci.* 40: 1438–1444.
- Smith, T. and H. Camper (1975) Effects of seed size on soybean performance. *Agron. J.* 67: 681–684.
- Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht and P.B. Cregan (2004) A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet.* 109: 122–128.
- Song, Q.J., G.F. Jia, Y.L. Zhu, D. Grant and R.T. Nelson (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. *Crop Sci.* 50: 1950–1960.
- Tuyen, D.D., S.K. Lal and D.H. Xu (2010) Identification of a major QTL allele from wild soybean (*Glycine soja* Sieb. & Zucc.) for increasing alkaline salt tolerance in soybean. *Theor. Appl. Genet.* 121: 229–236.
- Wang, S.K., S. Li, Q. Liu, K. Wu, J.Q. Zhang, S.S. Wang, Y. Wang, X.B. Chen, Y. Zhang, C.X. Gao *et al.* (2015) The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 47: 949–954.
- Wang, W., Q. He, H. Yang, S. Xiang, T. Zhao and J. Gai (2013) Development of a chromosome segment substitution line population with wild soybean (*Glycine soja* Sieb. & Zucc.) as donor parent. *Euphytica* 189: 293–307.
- Wang, W.B., X. Li, S. Chen, S.Y. Song and T.J. Zhao (2016) Using presence/absence variation markers to identify the QTL/allele system that confers the small seed trait in wild soybean (*Glycine soja* Sieb. & Zucc.). *Euphytica* 208: 101–111.
- Xin, D.W., Z.M. Qi, H.W. Jiang, Z.B. Hu, R.S. Zhu, J.H. Hu, H.Y. Han, G.H. Hu, C.Y. Liu and Q.S. Chen (2016) QTL location and epistatic effect analysis of 100-seed weight using wild soybean (*Glycine soja* Sieb. & Zucc.) chromosome segment substitution lines. *PLoS ONE* 11: e0149380.
- Zhou, Z.K., Y. Jiang, Z. Wang, Z.H. Gou, J. Lyu, W.Y. Li, Y.J. Yu, L.P. Shu, Y.J. Zhao, Y.M. Ma *et al.* (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33: 408–414.