

## Research Paper

# Identification and dissection of single seed weight QTLs by analysis of seed yield components in soybean

Kenichiro Fujii<sup>1,2)</sup>, Takashi Sayama<sup>1,2,3)</sup>, Kyoko Takagi<sup>2,4)</sup>, Kazumasa Kosuge<sup>5)</sup>, Katsunori Okano<sup>5)</sup>, Akito Kaga<sup>1,2)</sup> and Masao Ishimoto<sup>\*1,2)</sup>

<sup>1)</sup> Institute of Crop Science, National Agriculture and Food Research Organization (NARO), 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8518, Japan

<sup>2)</sup> National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan

<sup>3)</sup> Present address: Western Region Agricultural Research Center, NARO, 1-3-1 Senyu, Zentsuji, Kagawa 765-8508, Japan

<sup>4)</sup> Present address: Tohoku Agricultural Research Center, NARO, 50 Harajukuminami, Arai, Fukushima, Fukushima 960-2156, Japan

<sup>5)</sup> Plant Biotechnology Institute, Ibaraki Agriculture Center, 3165-1 Ago, Kasama, Ibaraki 319-0206, Japan

Single seed weight (SSW), or seed size, is a seed yield components (SYC) in soybean, and it is suggested that the genetic factors regulating SSW are involved in the control of other SYCs. The quantitative trait loci (QTLs) for SSW and their effects on the other SYCs were investigated using a recombinant inbred line population derived from typical small- and large-seeded cultivars that were cultivated in two different environments. QTL analysis detected four environmentally stable QTLs for SSW, two of which coincided with the defined loci, *qSw17-1* and *Ln*. The effects of the other loci, *qSw12-1* and *qSw13-1*, were confirmed by analyzing residual heterozygous line progenies derived from the recombinant population. These four QTL regions were also involved in the control of an additional SYC, namely the large-seeded allele at each locus that reduced either the number of pods per plant or the number of ovules per pod. These results suggest the presence of at least two different regulatory mechanisms for SSW. Isolation of genes responsible for these QTLs provides an important tool in the understanding and utilization of SSW diversity for soybean breeding.

**Key Words:** quantitative trait locus, recombinant inbred line, seed weight, seed yield component, soybean, trade-off.

## Introduction

Single seed weight (SSW), or seed size in soybean [*Glycine max* (L.) Merrill] is a heritable trait (Burton 1987) and soybean germplasm exhibit wide variation in SSW, ranging from 30 to 775 mg (Kaga *et al.* 2012). Seed size is a quality criterion for specific soy-based food products (Cui *et al.* 2004, Gandhi 2009). In general, large-seeded cultivars are used for boiled soybean (nimame), green soybean (edamame), soymilk, and soybean curd (tofu), while small-seeded cultivars are suitable for fermented soybean (natto) and sprout production. Thus, soybean cultivars with considerably diversified seed sizes are cultivated in East Asian countries where there is a tradition of soybean food culture (Cui *et al.* 2004).

SSW is also a component of seed yield, together with

plant density and the number of seeds per plant (NS). NS is divided into the number of pods per plant (NP) and the number of seeds per pod (NSP). Furthermore, NSP is subdivided into the number of ovules per pod (NOP) and the rate of seed-set (RSS), and thus the seed yield per plant (SY) is the product of SSW, NP, NOP, and RSS. Although a positive correlation between SSW and SY has been reported in previous studies (Burriss *et al.* 1973, Smith and Camper 1975), mass selection for small seed with reduced seed weight did not affect SY (Cober and Voldeng 2008) and the correlation between SY and SSW was generally low in several segregating populations (Review by Burton 1987). Hence, the difference in SSW between small- and large-seeded cultivars

Abbreviations: LOD, logarithm of odds; NOP, number of ovules per pod; NP, number of pods per plant; NS, number of seeds per plant; NSP, number of seeds per pod; PIC, polymorphism information content; QTL, quantitative trait locus; RHL, residual heterozygous line; RIL, recombinant inbred line; RSS, rate of seed-set; SSR, simple sequence repeat; SSW, single seed weight; SY, seed yield per plant; SYC, seed yield component.

Communicated by Hideyuki Funatsuki

Received August 18, 2017. Accepted October 23, 2017.

First Published Online in J-STAGE on March 24, 2018.

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

could be compensated for by the change in NS through other seed yield components (SYCs). The narrow leaflet phenotype, controlled by the *Ln* locus (Bernard and Weiss 1973, Jeong *et al.* 2012), is associated with both decreased SSW and increased NOP (Dinkins *et al.* 2002, Jeong *et al.* 2011). The induced mutant line for the *Gm-JAG1* (*Glyma20g25000*) gene clearly demonstrated the increase in NSP compensated reductions of SSW (Sayama *et al.* 2017). Therefore, it is likely that SSW and other SYCs are regulated by common loci and share regulatory mechanisms. However, other regulatory genes for SSW have not been identified; thus, the genetic architecture of SSW regulation remains poorly understood.

In the present study, we detected four significant and environmentally stable quantitative trait loci (QTLs) for SSW in a recombinant inbred line (RIL) population derived from a cross between the typical Japanese small- and large-seeded cultivars ‘Natto-shoryu’ and ‘Tachinagaha’ (**Supplemental Fig. 1**), and examined these QTLs using their effects on other SYCs. Subsequently, the effects of two notable QTLs were verified in the progeny of the residual heterozygous lines derived from the RIL population. Our results suggest the presence of at least two different regulatory mechanisms for SSW through trade-offs among other SYCs, such as between SSW and NP, and between SSW and NOP.

## Materials and Methods

### Plant materials

An RIL population consisting of 181 lines was developed by single seed descent from a cross between ‘Natto-shoryu’ (NARO Genebank JP29161) and ‘Tachinagaha’ (NARO Genebank JP67666), and designated as NsT-RILs. DNA samples were taken from leaflets of single F<sub>5</sub> seedlings from each line for linkage map construction and genotyping for the *E3* (Buzzell 1971), a major flowering locus segregating in the NsT-RILs (see below), and the F<sub>6</sub> population was used for trait evaluation for QTL analysis. The residual heterozygous lines (RHLs; Yamanaka *et al.* 2005) derived from the NsT-RILs were used to verify the effects of environmentally stable and notable QTLs for SSW. The RHLs for the QTLs were selected based on the genotypes of the F<sub>5</sub> RIL population, and genotypes for seeds derived from heterozygous F<sub>6</sub> plants were determined using DNA markers and segregating progenies for each QTL.

### Field experiments

The NsT-RILs and their parental cultivars were evaluated by field experiments at two locations, Mito (Plant Biotechnology Institute, Ibaraki Agricultural Center, located at 36°26'N, 140°26'E) and Tsukuba (National Institute of Agrobiological Sciences, located at 36°01'N, 140°06'E) in 2012. A starter fertilizer was applied at 3, 10, and 10 g m<sup>-2</sup> of N, P<sub>2</sub>O<sub>5</sub>, and K<sub>2</sub>O, respectively. Experimental plots for the NsT-RILs and parental cultivars were arranged according to the *E3* genotypes, because flowering time (maturity)

loci affect plant size and *E3* genotypes differ between the two parental cultivars (Tsubokura *et al.* 2014). Therefore, the *E3* ( $n = 84$  including ‘Natto-shoryu’) and *e3* ( $n = 89$  including ‘Tachinagaha’) genotype groups were planted in separate blocks, and the heterozygous genotypes ( $n = 10$ ) were placed between the two homozygous genotype groups. Within the groups, each line or cultivar was planted in single row plots following a randomized block design with one and two replicates at Mito and Tsukuba, respectively. Each row was placed 0.6 and 0.8 m apart, and each plot was composed of five plants that were spaced 0.15 and 0.2 m apart for the Mito and Tsukuba experimental conditions, respectively. Seeds were sown on June 20, 2012, at Mito and July 2, 2012, at Tsukuba. Three plants in each plot, excluding the plants at both plot ends, were individually harvested.

The RHL progeny and the parental cultivars were grown at Tsukuba in 2014. Seeds were sown in nursery beds on June 19, and robust seedlings were transplanted into the field on June 30. Starter fertilizer with the same composition used in the RIL evaluation was applied before transplantation. The plot was composed of seven plants placed 0.2 m apart, in rows spaced 0.8 m apart, arranged according to a completely randomized block design with four replicates. Three plants, excluding the two plants at both ends of the plots, were individually harvested for further analysis.

### Phenotypic measurements

Trait evaluation for SYCs was conducted on the basis of individual plants. SSW, SY, NP, NOP, and RSS were measured in individual plants from the RIL F<sub>6</sub> population, the RHL progenies, and the parental cultivars. Pods were cut off from the stem and classified based on ovule number per pod by visual inspection as previously reported (Jeong *et al.* 2011). After threshing the pods, seeds were counted and weighed. Immediately after weighing, seed water content was measured using a grain moisture tester (TB-2, OGA Electric, Tochigi, Japan). Pest-damaged seeds were removed from SSW evaluation in order to decrease experimental error, and SY was also corrected by multiplying the SSW and the total number of seeds per plant. The SSW and SY values were displayed to reflect 15% moisture content. NOP was calculated by dividing the number of ovules per plant by NP, and RSS by dividing the number of seeds per plant by the number of ovules per plant.

### Linkage map construction

A whole-genome simple sequence repeat (SSR) marker panel (WGSP) developed by Sayama *et al.* (2011) was modified, and applied to the construction of our linkage map. WGSP was originally developed to construct linkage maps by a simple procedure and at a low cost, using a selection of publicly available SSR markers according to their positions on soybean linkage maps and polymorphism information content (PIC) values (Cregan *et al.* 1999, Hisano *et al.* 2007, Hwang *et al.* 2009, Song *et al.* 2004, Xia *et al.* 2007). This original WGSP is comprised of 304 SSR markers and

has been applied in various genetic studies (Hirata *et al.* 2014, Kato *et al.* 2014, Yamaguchi *et al.* 2014), but some SSRs rarely showed polymorphisms. Therefore, 24 low-polymorphic markers were removed from the original WGSP and 72 novel markers were added to cover whole genomic regions with high polymorphisms. Sixty-eight additional public markers were selected according to the published information (Cregan *et al.* 1999, Hisano *et al.* 2007, Hwang *et al.* 2009, Song *et al.* 2004, 2010, Xia *et al.* 2007), along with four other novel markers designed in reference to the genome sequence published by Schmutz *et al.* (2010) because there were no alternative public markers associated with those chromosomal regions. A fluorescent dye (6-FAM, VIC, PET, or NED) was attached to the 5' end of one primer in each primer pair to determine the fragment length using DNA sequencers. The primer pairs of 64 public SSR markers were redesigned from the target sequences to adjust PCR product sizes in order to analyze two different markers with the same fluorescent dye simultaneously. Then, eight SSR markers (i.e., four dyes with two different fragment sizes) were organized into a multiplex analysis unit. The updated WGSP was composed of a total of 352 SSR markers and released as WGSP ver. 2 (**Supplemental Table 1**). The WGSP ver. 2 newly assigned ID numbers to respective SSR markers.

In addition to WGSP ver. 2, two DNA markers were used to recognize the alleles of the *Ln* and *E3* loci. The 'Natto-shoryu' parental cultivar of the NsT-RILs exhibited an ovate leaflet phenotype (*Ln*), while the 'Tachinagaha' parent had narrow leaflets (*ln*). For genotyping single nucleotide polymorphisms (SNPs) between the *Ln* and *ln* alleles (Jeong *et al.* 2012), the tetra-primer amplification refractory mutation system (ARMS)-PCR method was used (Ye *et al.* 2001), and an M13-tailed primer sequence was added to 5' ends of the *Ln* locus-specific primers so that the polymorphism could be detected using a DNA sequencer (Oetting *et al.* 1995). Additionally, preliminary evaluation showed that the RILs segregated at *E3*, which is a major locus controlling flowering time (Buzzell 1971) and thus, influences plant size and growing period. Sequencing analysis showed that 'Natto-shoryu' and 'Tachinagaha' had late-flowering (*E3-Mi*) and early-flowering (*e3-tr*) alleles at the *E3* locus, respectively (Watanabe *et al.* 2009, Tsubokura *et al.* 2014). A fluorescent primer set for distinguishing the different *E3* alleles was designed using the nucleotide sequence length polymorphism between the alleles. The *Ln* and *E3* marker primer sequences are shown in **Supplemental Table 2**.

### Linkage map construction

Total DNA from each genotype was extracted from approximately 10 mg fresh weight of leaflet or cotyledon tissue using an automatic DNA isolation system (BioSprint 96 DNA Plant Kit, Qiagen, Hilden, Germany). Polymerase chain reaction (PCR) was performed in a GeneAmp PCR System 9700 (Thermo Fisher Scientific, Waltham, USA) in conjunction with the Qiagen Multiplex PCR Kit (Qiagen).

The cycling program consisted of 15 min at 95°C for activation, followed by 35 cycles: 30 s at 95°C for denaturation, 1.5 min at 50°C for primer annealing, and 1.5 min at 60°C for extension. At the end of the cycles, the reaction mixtures were held at 60°C for 30 min, and then cooled to 4°C. The lengths of PCR products were determined using a 3730 DNA Analyzer equipped with POP-7 polymer separation matrix and 36-cm capillaries (Thermo Fisher Scientific), and allele calling and binning was performed with GeneMapper software v5.0 (Thermo Fisher Scientific). Linkage analysis was performed with MAPMAKER/EXP 3.0b software (Lander *et al.* 1987), and genetic distance was calculated with Kosambi's map function (Kosambi 1943).

### Data analysis

Statistical analyses for SY and its components, SSW, NP, NOP, and RSS, were performed using R statistical software ver. 3.1.3 (R Core Team 2015). The effects of experimental factors (genotype/allele, experimental condition, or interaction between genotype/allele and experimental condition) influencing the trait values were assessed by analysis of variance (ANOVA) using type-I sum of square (SS); then the effect size ( $\eta^2$ ) of a given factor was presented as the calculated proportion of SS for the factor to the total SS. Hence, the  $\eta^2$  of the genotype factor of each trait in the RIL population, reflected in a broad sense heritability for that trait. The correlation coefficient is a proportion of the genotypic covariance between two traits to the square root of the product of genotypic variances of the two traits. Trait values between two genotypes were compared by Welch's *t*-test.

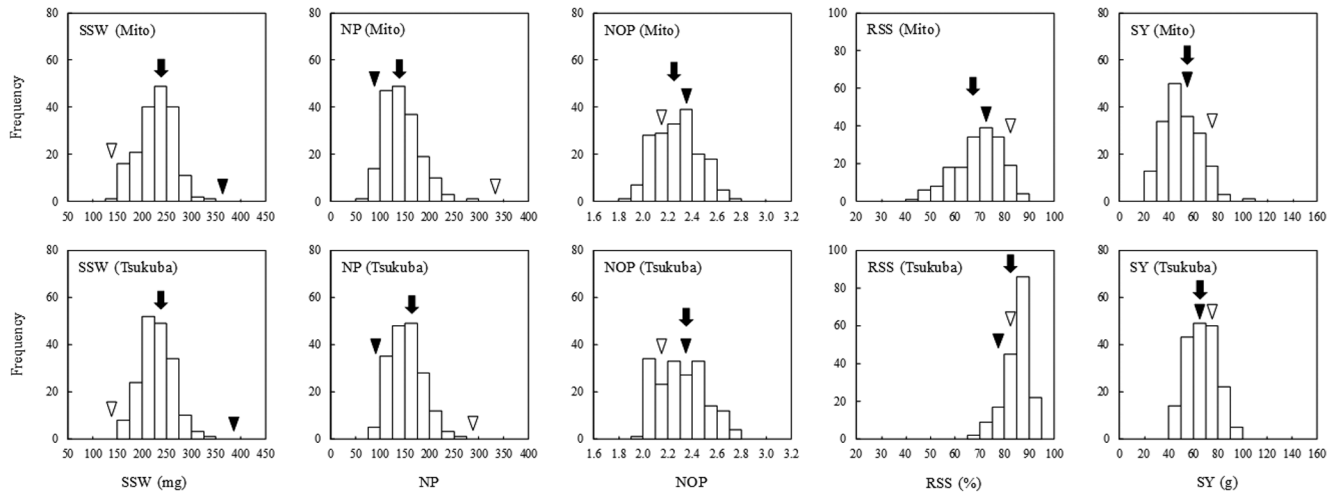
QTL analysis was performed using the Windows QTL Cartographer ver. 2.5 (Wang *et al.* 2013). The composite interval mapping (CIM) method was applied to detect the position and effects of QTLs. The genome was scanned at 1-cM intervals and additive effects and logarithm of odds (LOD) were estimated. A permutation test was conducted 1,000 times for each trait and experiment, and the threshold value of the LOD peak was set at  $P = 0.05$ . The QTL position with the 1- and 2-LOD significant intervals (van Ooijen 1992) on the linkage map was depicted using MapChart software (Voorrips 2002).

## Results

### Phenotypic variation of SYCs in the NsT-RILs

The SSW and other SYCs were evaluated in the NsT-RIL population and their parental cultivars for two different experimental conditions (**Fig. 1**). Other measures of SYCs, including NS and NSP, and additional statistic data are also shown in **Supplemental Table 3**.

SSW values in 'Natto-shoryu' were 128 and 148 mg, and 365 and 400 mg in 'Tachinagaha' at Mito and Tsukuba experimental conditions, respectively. These values indicated significant differences in SSW between the two parental cultivars for each experimental condition. SSW values in the NsT-RILs ranged from 145 to 339 mg and 161 to



**Fig. 1.** Frequency distributions of seed yield components in the recombinant inbred line population derived from a cross between ‘Natto-shoryu’ and ‘Tachinagaha’ (NsT-RILs) at Mito and Tsukuba experimental conditions. Single seed weight (SSW), number of pods per plant (NP), number of ovules per pod (NOP), rate of seed-set (RSS), and seed yield per plant (SY) were evaluated for each individual, and values are averaged within each line. Open and filled triangles represent trait values in ‘Natto-shoryu’ and ‘Tachinagaha’; and arrows represent mean values of the NsT-RILs.

326 mg at Mito and Tsukuba, respectively. These values were intermediate in comparison to the parental SSW measurements, and the effect of genotype on SSW in the NsT-RILs was as high as 0.753. The NP values measured in the RIL population were also intermediate in comparison to the values measured in their parents, where the small-seeded parent produced considerably larger NP than any other genotypes (Fig. 1, Supplemental Table 3). In contrast to SSW and NP, NOP in the RILs displayed transgressive segregation at both environments; the range was from 1.89 to 2.70 at Mito, and 2.00 to 2.79 at Tsukuba. ‘Tachinagaha’ had significantly larger NOP than ‘Natto-shoryu’ and the average difference of NOP between the parents was as small as 0.18. The effect of the genotype on NOP in the NsT-RILs was as high as 0.889. RSS was the only trait that produced no significant difference between the parental cultivars; however, RSS in the RILs and their parental cultivars varied remarkably in individuals, and the experimental condition substantially influenced RSS in the RIL population. SY was significantly larger in ‘Natto-shoryu’ compared to ‘Tachinagaha’, and SY in the RILs also showed transgressive segregation for both experimental conditions. SY varied considerably in individual plants (Supplemental Table 3), and therefore genotype, experimental condition, and their interactive factor had a limited effect on this trait.

We next analyzed the possible correlations between SSW and other SYCs in the NsT-RIL population (Table 1). SSW negatively correlated with NP, NOP, and RSS at both Mito and Tsukuba and positively correlated with SY at Tsukuba, but did not associate with SY at Mito. For the other SYCs, NP and RSS were positively correlated with SY but NP and RSS exhibited opposite correlations between the two experimental conditions. Correlation coefficients between other SYC measurements are also shown in Supplemental Table 4.

**Table 1.** Correlation coefficients between seed yield components in the NsT-RIL population at Mito and Tsukuba experimental conditions

Trait	Experimental condition	SSW	NP	NOP	RSS
NP	Mito	-0.440 ***			
	Tsukuba	-0.211 **			
NOP	Mito	-0.396 ***	0.076 NS		
	Tsukuba	-0.511 ***	-0.231 **		
RSS	Mito	-0.339 ***	0.401 ***	-0.004 NS	
	Tsukuba	-0.362 ***	-0.172 *	0.101 NS	
SY	Mito	-0.071 NS	0.820 ***	0.090 NS	0.610 ***
	Tsukuba	0.157 *	0.838 ***	-0.199 **	-0.066 NS

SSW, single seed weight; NP, number of pods per plant; NOP, number of ovules per pod; RSS, rate of seed-set; SY, seed yield per plant.

\*, \*\*, and \*\*\* represent statistical significance at  $P < 0.05$ , 0.01, and 0.001, respectively; NS = not significant.

### Construction of the genetic linkage map of the NsT-RILs by WGSP ver.2

WGSP ver. 2 was used to construct a linkage map of the NsT-RIL population with a total of 352 SSR markers. Fragment analysis identified polymorphisms between the parental cultivars of the RIL population at 221 loci; 210 of these loci were co-dominantly segregated. Of the 72 new markers that were added in the updated WGSP, 45 were segregated between the parental cultivars and 41 were co-dominant (Supplemental Table 1). These co-dominant SSR markers, as well as the two additional genic markers for the *Ln* and *E3* loci, were used to genotype the NsT-RILs and calculate the genetic distance with Kosambi’s map function (Kosambi 1943) (Supplemental Table 1). The genetic linkage map covered a total of 2,824.3 cM over the 20 soybean chromosomes with an average marker interval of 14.7 cM. The marker order and distance between loci for the constructed linkage maps, including *Ln* and the *E3*, accorded well with

those of the previous linkage maps (Hwang *et al.* 2009, Jeong *et al.* 2012, Sayama *et al.* 2011, Watanabe *et al.* 2009). The *Ln* and *E3* genotypes of each line agreed with phenotypic data from the preliminary evaluation in the F<sub>5</sub> generation (data not shown).

### QTL mapping of SSW and other SYCs

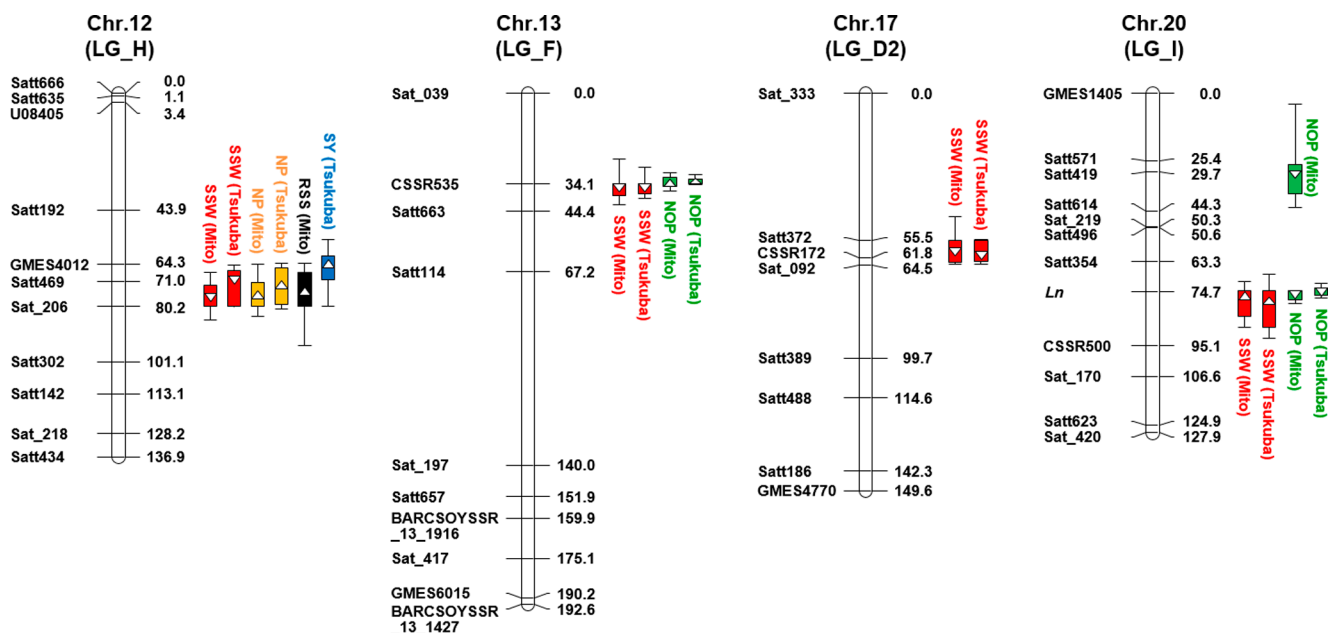
Our QTL analysis detected a total of five QTLs that control SSW (Fig. 2, Supplemental Table 5). Of these, the effects of the QTLs on Chr. 12, 13, 17, and 20 were stable across the two experimental conditions, and the four QTLs explained a total of 34.6% and 43.5% of the phenotypic variations at Mito and Tsukuba, respectively. The small-seeded alleles at the QTLs on Chr. 12, 13, and 17 were derived from ‘Natto-shoryu’, while the small-seeded allele on Chr. 20 was derived from ‘Tachinagaha’. The QTLs on Chr. 13, 17, and 20 mapped to the corresponding regions of *qSw13-1*, *qSw17-1*, and *qSw20-1*, which were identified from an RIL population that also had ‘Tachinagaha’ as a parent (Kato *et al.* 2014). Thus, we also designated these environmentally stable QTLs as *qSw12-1*, *qSw13-1*, *qSw17-1*, and *qSw20-1* based on the previous nomenclature system; the designations represent the initials of trait names with chromosome number and the rank of the additive effect among the QTLs on the same chromosome. The proximal markers of these QTLs were Sat469 or Sat\_206 for *qSw12-1*, CSSR535 for *qSw13-1*, CSSR172 for *qSw17-1*, and *Ln* for *qSw20-1* (Fig. 2).

QTL analyses of NP, NOP, RSS, and SY were also conducted and the detailed results are listed in Supplemental

**Table 5.** Two QTLs for NP were located at the corresponding regions for *qSw12-1* and *E3*, and were stable at both of the experimental conditions. Both the alleles that increased NP were derived from ‘Natto-shoryu’. Two QTLs that controlled NOP were mapped to the same chromosomal regions as the *qSw13-1* and *qSw20-1*. These two QTLs had significant effects on NOP and explained nearly 50% of the phenotypic variation at both experimental conditions. The small-seeded allele at these QTLs was associated with a positive effect on NOP. The QTLs for RSS were closely mapped to the *qSw12-1* and *E3* loci. The QTL close to the *E3* locus was consistently detected in the RIL population grown at the experimental conditions, but the effect on RSS was entirely different between these conditions. For example, the allele from ‘Natto-shoryu’ increased RSS at Mito, but decreased RSS at Tsukuba. The QTLs for SY were mapped to Chrs. 12 and 19, and located near *qSw12-1* and *E3*, respectively. The QTL on Chr. 19 was stable across the experimental conditions. The late flowering allele at *E3* from ‘Natto-shoryu’ increased SY and NP at both conditions. The QTLs for NS and NSP are also shown in Supplemental Table 5.

### Effect of the SSW QTL on other SYCs

The effects of the chromosomal regions regulating SSW on the other SYCs were investigated by ANOVA in the NsT-RILs (Table 2). The genotype of each line was estimated based on the allele genotypes of the flanking markers of the QTLs. Although the flanking markers for *qSw12-1* were different between the experimental conditions, Sat\_206 was



**Fig. 2.** Mapping of QTLs associated with single seed weight (SSW) and other seed yield components (SYCs) in the recombinant inbred line population derived from a cross between ‘Natto-shoryu’ and ‘Tachinagaha’ (NsT-RILs) at Mito and Tsukuba experimental conditions. Environmentally stable QTLs associated with SSW were detected on four chromosomes (Chr. or LG: linkage group); detailed information is noted in Supplemental Table 3. Box and line symbols indicate 1- and 2-LOD significant intervals, respectively. Triangles indicate LOD peak positions. Regular and inverted triangles represent that the ‘Natto-shoryu’ allele has positive and negative effects on the trait, respectively.

**Table 2.** Effects of each QTL allele detected for single seed weight (SSW) on other seed yield components in the NsT-RILs grown at Mito and Tsukuba experimental conditions

QTL	Allele	Experimental condition	SSW (mg)	NP	NOP	RSS (%)	SY (g)
<i>qSw12-1</i> (Sat_206)	‘Natto-shoryu’ ( <i>n</i> = 90)	Mito	216 ± 36	155 ± 39	2.27 ± 0.17	71.6 ± 8.0	53.6 ± 15.8
		Tsukuba	220 ± 30	165 ± 32	2.31 ± 0.19	85.2 ± 4.7	70.2 ± 12.0
	‘Tachinagaha’ ( <i>n</i> = 74)	Mito	244 ± 31	130 ± 28	2.29 ± 0.19	67.8 ± 10.1	48.7 ± 12.2
		Tsukuba	242 ± 29	139 ± 28	2.32 ± 0.20	83.5 ± 5.3	63.7 ± 11.3
	ANOVA	Allele	*** (0.130)	*** (0.132)	NS (0.003)	*** (0.017)	*** (0.034)
	Experimental condition	NS (0.000)	** (0.018)	NS (0.009)	*** (0.492)	*** (0.261)	
	Interaction	NS (0.002)	NS (0.000)	NS (0.000)	NS (0.002)	NS (0.001)	
<i>qSw13-1</i> (CSSR535)	‘Natto-shoryu’ ( <i>n</i> = 89)	Mito	219 ± 39	143 ± 37	2.38 ± 0.14	68.7 ± 9.4	50.3 ± 14.1
		Tsukuba	218 ± 31	151 ± 35	2.42 ± 0.17	85.1 ± 5.2	65.9 ± 12.9
	‘Tachinagaha’ ( <i>n</i> = 78)	Mito	238 ± 32	144 ± 36	2.17 ± 0.14	70.2 ± 9.7	51.7 ± 15.4
		Tsukuba	241 ± 28	157 ± 31	2.21 ± 0.15	84.3 ± 5.1	68.9 ± 11.5
	ANOVA	Allele	*** (0.093)	NS (0.002)	*** (0.321)	NS (0.000)	NS (0.005)
	Experimental condition	NS (0.000)	* (0.020)	* (0.010)	*** (0.499)	*** (0.266)	
	Interaction	NS (0.001)	NS (0.002)	NS (0.000)	NS (0.003)	NS (0.001)	
<i>qSw17-1</i> (CSSR172)	‘Natto-shoryu’ ( <i>n</i> = 96)	Mito	222 ± 36	146 ± 38	2.26 ± 0.18	69.7 ± 9.9	50.1 ± 14.2
		Tsukuba	223 ± 30	156 ± 34	2.29 ± 0.18	85.3 ± 4.7	66.5 ± 12.5
	‘Tachinagaha’ ( <i>n</i> = 70)	Mito	240 ± 35	132 ± 29	2.27 ± 0.17	68.5 ± 9.0	48.6 ± 12.1
		Tsukuba	239 ± 32	148 ± 29	2.31 ± 0.20	84.1 ± 5.4	67.1 ± 11.1
	ANOVA	Allele	*** (0.062)	** (0.024)	NS (0.001)	NS (0.003)	NS (0.000)
	Experimental condition	NS (0.000)	*** (0.032)	NS (0.009)	*** (0.510)	*** (0.317)	
	Interaction	NS (0.000)	NS (0.002)	NS (0.000)	NS (0.000)	NS (0.001)	
<i>qSw20-1</i> ( <i>Ln</i> )	‘Natto-shoryu’ ( <i>n</i> = 79)	Mito	240 ± 34	139 ± 33	2.16 ± 0.13	68.9 ± 9.6	49.0 ± 13.2
		Tsukuba	241 ± 28	159 ± 33	2.17 ± 0.12	85.1 ± 5.0	69.5 ± 12.1
	‘Tachinagaha’ ( <i>n</i> = 91)	Mito	218 ± 35	147 ± 39	2.39 ± 0.14	69.9 ± 9.2	52.8 ± 15.7
		Tsukuba	218 ± 31	147 ± 32	2.45 ± 0.15	84.5 ± 4.9	65.0 ± 11.8
	ANOVA	Allele	*** (0.109)	NS (0.001)	*** (0.465)	NS (0.000)	NS (0.000)
	Experimental condition	NS (0.000)	* (0.018)	* (0.009)	*** (0.511)	*** (0.262)	
	Interaction	NS (0.000)	** (0.020)	NS (0.003)	NS (0.001)	** (0.017)	

The genotype at each environmentally stable QTL for SSW was estimated by the alleles of the flanking markers; Sat\_206 for *qSw12-1*, CSSR535 for *qSw13-1*, CSSR172 for *qSw17-1*, and *Ln* for *qSw20-1*.

SSW, single seed weight; NP, number of pods per plant; NOP, number of ovules per pod; RSS, rate of seed-set; SY, seed yield per plant.

Trait value of each genotype group is shown as mean ± standard deviation.

\*, \*\*, and \*\*\* represent statistical significance at  $P < 0.05$ , 0.01, and 0.001 by ANOVA, respectively; NS = not significant.

Values in parentheses indicate effect size ( $\eta^2$ ) of each factor calculated as a proportion of the sum of squares of the factor to the total sum of squares.

adopted because the estimated effect of the locus on SSW was more evident than that of the other flanking marker (data not shown). Thus, Sat\_206, CSSR535, CSSR172, and *Ln* were employed to estimate *qSw12-1*, *qSw13-1*, *qSw17-1*, and *qSw20-1* allele types, respectively.

The effects of four environmentally stable QTLs on SSW, *qSw12-1*, *qSw13-1*, *qSw17-1*, and *qSw20-1*, were also detected by ANOVA (Table 2). These four QTLs exhibited effect sizes of 0.062–0.130 individually and of 0.394 in total for SSW in the NsT-RILs. The chromosomal regions associated with SSW also contributed to the variation of the other SYCs. The chromosomal regions for *qSw12-1* and *qSw17-1* contained QTLs for NP, and the small-seeded genotypes at both these QTLs had increased NP. The QTL for NP was not detected in the *qSw17-1* region, but the effect of this locus on NP was significant according to an ANOVA including trait data measured at both the experimental conditions. These two QTLs showed an effect size of 0.156 in total for NP in the NsT-RILs. The chromosomal regions for *qSw13-1*

and *qSw20-1* contained QTLs for NOP, and the small-seeded alleles were associated with increased NOP. The effect size of these two QTLs was as large as 0.786 in total for NOP in the NsT-RILs. In addition, the chromosomal region around *qSw12-1* also related to RSS and SY; at this locus, the allele derived from ‘Natto-shoryu’ increased both traits. This analysis indicated the genotype-experimental condition interactions of the *qSw20-1* region on NP and SY, but the effect sizes of these interactions were as small as 0.002 and 0.017, respectively.

### Confirmation of the effects of *qSw12-1* and *qSw13-1* on SYCs using RHL progenies

The effects of the chromosomal regions containing *qSw12-1* and *qSw13-1* on SYCs were verified in the RHL progenies, which were selected from the NsT-RIL population as being heterozygous around each locus based on the genotype of the flanking markers. NsT-RILs 086 and 103 had heterozygous genotypes around the *qSw12-1* regions,

while NsT-RILs 018 and 111 were selected as RHLs for *qSw13-1*. Those RILs did not segregate at the other QTL regions that affected SYCs in the F<sub>5</sub> generation. The ‘Natto-shoryu’ and ‘Tachinagaha’ genotypes in the *qSw12-1* or *qSw13-1* regions were developed from the corresponding RHLs, and then applied for the evaluation of each trait.

For *qSw12-1*, ‘Tachinagaha’ genotypes had significantly larger SSW than the ‘Natto-shoryu’ genotypes in the RHL-086 and RHL-103 genetic backgrounds (Table 3). NP in the ‘Natto-shoryu’ genotype was higher than in the ‘Tachinagaha’ genotype in RHL-103. Similarly, the ‘Natto-shoryu’ genotype had higher NP than the ‘Tachinagaha’ genotype in RHL-086, although this difference was not significant at  $P < 0.05$ . The genotypes in the *qSw12-1* region did not affect NOP and SY in either background. The effects of the chromosomal regions that contain *qSw12-1* on SSW and NP were consistent with the results from the QTL analysis.

SSW and NOP differed among the genotypes in the *qSw13-1* region in both RHL-018 and RHL-111 backgrounds (Table 4). The ‘Natto-shoryu’ genotype decreased SSW and increased NOP. RSS was significantly different between the ‘Natto-shoryu’ and ‘Tachinagaha’ genotypes in RHL-018, but the difference was not observed in RHL-111 backgrounds. The effects of the *qSw13-1* region on SSW and NOP were uniquely identical to the results of the QTL analysis.

## Discussion

The ‘Natto-shoryu’ and ‘Tachinagaha’ cultivars differed substantially in SSW and the RIL population derived from these cultivars, NsT-RILs, exhibited a wide variation in SSW, which ranged from 145 to 339 mg (Fig. 1, Supplemental Table 3). As no obvious correlation between SSW and SY was observed (Table 1), the impact of the difference in SSW on SY would be compensated for by the change in NS (Supplemental Table 4). All the components of NS, namely, NP, NOP, and RSS, negatively correlated with SSW (Table 1). Therefore, the effects of the genetic factors regulating SSW on other SYCs could be dissected using this RIL population.

Composite interval mapping detected four significant and environmentally stable QTLs for SSW from the NsT-RILs that mapped to Chr. 12, 13, 17, and 20; designed as *qSw12-1*, *qSw13-1*, *qSw17-1*, and *qSw20-1*, respectively (Fig. 2, Supplemental Table 5). The four environmentally stable SSW QTLs explained approximately 40% of the phenotypic variance in SSW, and the other part of the phenotypic variance is probably controlled by many other undefined QTLs with smaller effects on SSW. Although the large-seeded allele at *qSw20-1* was derived from the small-seeded parent, the NsT-RILs did not exhibit transgressive segregation in SSW. This suggested that the small- or large-seeded parents accumulated negative or positive effect alleles at most of the SSW QTLs, other than *qSw20-1*; it was extremely difficult to accumulate an adequate number

of positive or negative effect alleles in order to achieve the transgressive segregation because approximately half of the chromosomal segments of each recombinant line was derived from one parental cultivar. The difficulty in achieving extreme allele combination has also been discussed in other reports (Alt *et al.* 2002, Cober and Voldeng 2008, Johnson *et al.* 2001).

The stable QTL on Chr. 17, *qSw17-1*, was mapped to almost the same position as that previously reported by Kato *et al.* (2014), which was identified using two different RIL populations derived from crosses between Japanese large-seeded and US conventional soybean cultivars. The *qSw20-1* is closely linked to the *Ln* locus; whose recessive allele governs narrow leaflet morphology. Previous studies indicated a relationship between small seed and narrow leaflet phenotypes controlled by the *Ln* locus (Dinkins *et al.* 2002, Jeong *et al.* 2011, Sayama *et al.* 2017). ‘Tachinagaha’ is a narrow leaflet cultivar and the small-seeded *ln* allele at *qSw20-1* was certainly derived from this narrow leaflet parent in the NsT-RIL population (Supplemental Table 5). Previously mapped QTLs for SSW were located around the chromosomal regions where *qSw12-1* and the *qSw13-1* were detected (Hoeck *et al.* 2003, Hyten *et al.* 2004, Kato *et al.* 2014, Mian *et al.* 1996, Specht *et al.* 2001), but the presence of these QTLs has not been corroborated to date.

Confirmation of the effect of QTLs in different environmental conditions or genetic backgrounds is an important process because there are several cases of QTLs that are ineffective in subsequent practices, such as marker-assisted selection (Diers *et al.* 1992, Fasoula *et al.* 2004). This is caused by confounding factors such as the genetic structure of tested populations, environmental effects, and differences in marker sets used for genotyping (Beavis 1994). Fasoula *et al.* (2004) verified QTLs for seed protein, seed oil, and SSW in soybean using independent sets of RIL populations derived from the same parental cultivars. As a result, only one third of the QTLs for those traits were verified and the effects of other QTLs were not observed. The effects of the *qSw17-1* and the *qSw20-1* on SSW were confirmed in several genetic studies (Dinkins *et al.* 2002, Hirata *et al.* 2014, Jeong *et al.* 2011, Zhou *et al.* 2015); however, *qSw12-1* and *qSw13-1* have not been corroborated. We have successfully verified the effects of these QTL regions on SSW using segregants of residual heterozygous lines (RHLs; Yamanaka *et al.* 2005) derived from the NsT-RIL population (Tables 3, 4). For the *qSw12-1* region, the ‘Natto-shoryu’ allele decreased SSW, and the effects of the *qSw13-1* allele were in agreement with the QTL mapping results. Therefore, the responsible gene(s) for these traits was expected to be located in the chromosomal region that commonly segregated in the two sets of RHLs.

In the NsT-RILs, SSW was negatively correlated with NP, NOP, and RSS with respect to their genetic variation, indicating that the impact of increased SSW might be largely compensated for by decreases in the other SYCs, which are components of NS (Table 1, Supplemental Table 4).

**Table 3.** Effect of the *qSw12-1* alleles on single seed weight (SSW) and other seed yield components in segregated progeny derived from residual heterozygous lines (RHLs)

RHL	Estimated allele at <i>qSw12-1</i>		Marker allele on Chr. 12				SSW (mg)	NP	NOP	RSS (%)	SY (g)
	Satt192	GIMES4012	Satt469	Sat_206	Satt302	Satt142					
RHL-086	'Natto-shoryu'	N	N	N	N	N	245 ± 14 ***	138 ± 26 NS	2.21 ± 0.04 NS	88.7 ± 2.7 NS	66.5 ± 13.8 NS
	'Tachinagaha'	N	N	T	T	T	278 ± 13	118 ± 24	2.21 ± 0.04	89.4 ± 2.6	64.7 ± 12.0
RHL-103	'Natto-shoryu'	T	N	N	N	T	186 ± 8 ***	239 ± 37 *	2.03 ± 0.02 NS	91.4 ± 1.6 NS	82.0 ± 12.2 NS
	'Tachinagaha'	T	T	T	T	T	219 ± 14	207 ± 39	2.03 ± 0.02	90.2 ± 2.1	83.0 ± 11.1

SSW, single seed weight; NP, number of pods per plant; NOP, number of ovules per pod; RSS, rate of seed-set; SY, seed yield per plant.

Data were obtained from 12 plants (3 plants with 4 replications) for each genotype.

The flanking markers of *qSw12-1* are indicated in boldface type.

N and T indicate the 'Natto-shoryu' and 'Tachinagaha' alleles at each marker locus, respectively.

Trait value of each genotype group is shown as mean ± standard deviation.

\* and \*\*\* represent statistical significance at  $P < 0.05$  and  $0.001$  by Welch's t-test, respectively; NS = not significant.

**Table 4.** Effect of the *qSw13-1* alleles on single seed weight (SSW) and other seed yield components in segregated progeny derived from residual heterozygous lines (RHLs)

Line	Estimated allele at <i>qSw13-1</i>		Marker allele on Chr. 13				SSW (mg)	NP	NOP	RSS (%)	SY (g)
	Sat_039	CSSR535	Satt663	Satt114	Satt114	Satt114					
RHL-018	'Natto-shoryu'	N	N	N	N	N	235 ± 23 **	104 ± 24 NS	2.18 ± 0.05 ***	91.6 ± 1.7 **	49.3 ± 14.8 NS
	'Tachinagaha'	N	T	T	T	N	258 ± 25	91 ± 25	2.04 ± 0.08	89.3 ± 2.0	42.6 ± 15.2
RHL-111	'Natto-shoryu'	T	N	N	N	T	232 ± 15 **	113 ± 18 NS	2.46 ± 0.06 ***	87.1 ± 2.5 NS	56.3 ± 11.8 NS
	'Tachinagaha'	T	T	T	T	T	253 ± 15	128 ± 24	2.24 ± 0.12	88.9 ± 2.6	64.5 ± 13.6

SSW, single seed weight; NP, number of pods per plant; NOP, number of ovules per pod; RSS, rate of seed-set; SY, seed yield per plant.

Data were obtained from 12 plants (3 plants with 4 replications) for each genotype.

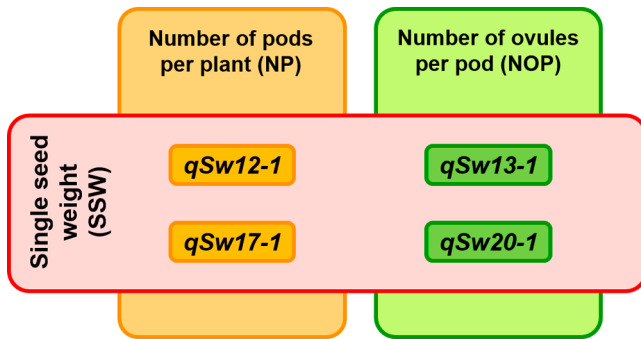
The flanking marker of *qSw13-1* is indicated in boldface type.

N and T indicate the 'Natto-shoryu' and 'Tachinagaha' alleles at each marker locus, respectively.

Trait value of each genotype group is shown as mean ± standard deviation.

\*\* and \*\*\* represent statistical significance at  $P < 0.01$  and  $0.001$  by Welch's t-test, respectively; NS = not significant.





**Fig. 3.** Schematic illustration of the corresponding relationships between the four single seed weight (SSW) QTLs detected in this study and the three seed yield components; SSW, the number of pods per plant (NP), and the number of ovules per pod (NOP). The large seed alleles at *qSw12-1*, *qSw13-1*, and *qSw17-1* were derived from the large-seeded parental cultivar ‘Tachinagaha’, while the large seed allele at *qSw20-1* was derived from the small-seeded parent ‘Natto-shoryu’. These four QTLs had inverse effects on SSW and other seed yield components where *qSw12-1* and *qSw17-1* loci were associated with SSW and NP, and *qSw13-1* and *qSw20-1* loci were involved in the regulation of SSW and NOP.

The QTL analysis and ANOVA suggested genetic control of the potential trade-offs between SSW and other SYCs by the SSW QTL regions, since at each locus the large-seeded allele decreased in either NP or NOP (Fig. 2, Table 2). In addition, the effects of *qSw12-1* and *qSw13-1* were verified by our evaluation of the RHL (Tables 3, 4). The four stable QTLs for SSW detected in the present study could be classified into two different groups according to their genetic effects on SYCs (Fig. 3). The chromosomal regions of *qSw12-1* and *qSw17-1* were involved in the regulation of SSW and NP, whereas the *qSw13-1* and *qSw20-1* regions were associated with the inverse relationship of SSW and NOP. Additionally, the chromosomal region around *qSw12-1* influenced RSS and SY. However, it is not clear whether the co-regulation of these components was mediated by a single genetic factor or by independent factors located closely to each other.

The trade-off between seed/egg size and number is a well-accepted concept in both plant and animal species (Sadras 2007, Smith and Fretwell 1974, Venable 1992). The trade-off takes place under resource-limited conditions, and changes in allocation patterns of the available resources result in an inverse relationship between the size and number of seeds. The present study showed that the large-seeded allele at each QTL for SSW was associated with a decrease in either NP or NOP (Tables 2–4). These findings suggest that genetic factors governing the allocation of resources are capable of pleiotropic properties; if these SYCs are controlled through the allocation of resources, then the regulation of SSW could consist of at least two independent mechanisms that mediate the trade-offs between SSW and NP, as well as between SSW and NOP.

Despite the negative correlation between SSW and RSS

(Table 1), genetic regulation of this trade-off relationship was less evident, in contrast to trade-offs between SSW and NP, and between SSW and NOP; this is possibly because the trade-off between SSW and RSS might be regulated by small-effect genetic factors. Alternatively, there may be no strict genetic control for this trade-off, but potentially heavy SSW might be more vulnerable to decline in RSS. This indicates that the large-seeded genotype carries a greater reproductive sink capacity than the small-seeded genotype but that the resource availability adjusts the actual sink size through RSS. This supposition can also explain the positive correlations between RSS and SY at Mito and between SSW and SY at Tsukuba (Table 1), given that resource availability was limited only in the Mito experimental condition. The positive effect of the *E3* allele on RSS at Mito (Supplemental Table 5) may also reflect increased resource availability in the *E3* genotype, while the inverse relationship at Tsukuba might indicate that RSS is susceptible to the environment to which the plant exposed. More research is needed to better understand the detail of the trade-off relationships between SYCs.

This study shows that the chromosomal regions that regulate SSW also control other SYCs, and thus selection for SSW involves changes in these other SYCs. Deeper understanding of the genetic architecture of SSW regulation would facilitate the manipulation of SSW, and help us understand the genetic basis of seed yield formation. The QTLs detected in the present study are attractive targets for functional analysis, and will contribute to elucidating a more complete picture of SSW regulation.

### Acknowledgements

This study was supported by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan [Genomics-based Technology for Agricultural Improvement (SFC-1001)].

### Literature Cited

- Alt, B.J., W.R. Fehr and G.A. Welke (2002) Selection for large seed and high protein in two- and three-parent soybean populations. *Crop Sci.* 42: 1876–1881.
- Beavis, W.D. (1994) The power and deceit of QTL experiments: Lessons from comparative QTL studies. *In: Proc. of the 49th Annual Corn and Sorghum Industry Research Conference*. ASTA, Washington, DC, pp. 250–266.
- Bernard, R. and M. Weiss (1973) Qualitative genetics. *In: Caldwell, B.E. (ed.) Soybeans: improvement, production, and uses*. American Society of Agronomy, Madison/WI, pp. 117–154.
- 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.
- Burton, J.W. (1987) Quantitative genetics: Results relevant to soybean breeding. *In: Wilcox, J.R. (ed.) Soybeans: improvement, production, and uses*, 2nd edn. American Society of Agronomy, Madison/WI, pp. 211–247.

- Buzzell, R.I. (1971) Inheritance of a soybean flowering response to fluorescent-daylength conditions. *Can. J. Genet. Cytol.* 13: 703–707.
- Cober, E.R. and H.D. Voldeng (2008) Mass selection for small seed size in natto soybean populations and the resulting effect on seed yield. *Crop Sci.* 48: 1337–1340.
- Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. Vantoai, D.G. Lohnes, J. Chung *et al.* (1999) An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39: 1464–1490.
- Cui, Z., A.T. James, S. Miyazaki, R.F. Wilson and T.E. Carter (2004) Breeding of specialty soybeans for traditional and new soyfoods. *In: Liu, K. (ed.) Soybeans as a functional food.* AOCS Press, Champaign, Illinois, pp. 264–322.
- Diers, B.W., P. Keim, W.R. Fehr and R.C. Shoemaker (1992) RFLP analysis of soybean seed protein and oil content. *Theor. Appl. Genet.* 83: 608–612.
- Dinkins, R.D., K.R. Keim, L. Farno and L.H. Edwards (2002) Expression of the narrow leaflet gene for yield and agronomic traits in soybean. *J. Hered.* 93: 346–351.
- Fasoula, V.A., D.K. Harris and H.R. Boerma (2004) Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci.* 44: 1218–1225.
- Gandhi, A.P. (2009) Quality of soybean and its food products. *Int. Food Res. J.* 16: 11–19.
- Hirata, K., R. Masuda, Y. Tsubokura, T. Yasui, T. Yamada, K. Takahashi, T. Nagaya, T. Sayama, M. Ishimoto and M. Hajika (2014) Identification of quantitative trait loci associated with boiled seed hardness in soybean. *Breed. Sci.* 64: 362–370.
- Hisano, H., S. Sato, S. Isobe, S. Sasamoto, T. Wada, A. Matsuno, T. Fujishiro, M. Yamada, S. Nakayama, Y. Nakamura *et al.* (2007) Characterization of the soybean genome using EST-derived microsatellite markers. *DNA Res.* 14: 271–281.
- 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.
- Hwang, T.Y., T. Sayama, M. Takahashi, Y. Takada, Y. Nakamoto, H. Funatsuki, H. Hisano, S. Sasamoto, S. Sato, S. Tabata *et al.* (2009) High-density integrated linkage map based on SSR markers in soybean. *DNA Res.* 16: 213–225.
- Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau-Ellis, T.R. Stefaniak and M.E. Schmidt (2004) Seed quality QTL in a prominent soybean population. *Theor. Appl. Genet.* 109: 552–561.
- Jeong, N., J.K. Moon, H.S. Kim, C.G. Kim and S.C. Jeong (2011) Fine genetic mapping of the genomic region controlling leaflet shape and number of seeds per pod in the soybean. *Theor. Appl. Genet.* 122: 865–874.
- 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.
- Johnson, S.L., W.R. Fehr, G.A. Welke and S.R. Cianzio (2001) Genetic variability for seed size of two- and three-parent soybean populations. *Crop Sci.* 41: 1029–1033.
- Kaga, A., T. Shimizu, S. Watanabe, Y. Tsubokura, Y. Katayose, K. Harada, D.A. Vaughan and N. Tomooka (2012) Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breed. Sci.* 61: 566–592.
- Kato, S., T. Sayama, K. Fujii, 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.
- Kosambi, D.D. (1943) The estimation of map distances from recombination values. *Ann. Eugen.* 12: 172–175.
- Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daley, S.E. Lincoln and L.A. Newburg (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174–181.
- Mian, M.A.R., M.A. Bailey, J.P. Tamulonis, E.R. Shipe, T.E. Carter, Jr., W.A. Parrott, D.A. Ashley, R.S. Hussey and H.R. Boerma (1996) Molecular markers associated with seed weight in two soybean populations. *Theor. Appl. Genet.* 93: 1011–1016.
- Oetting, W.S., H.K. Lee, D.J. Flanders, G.L. Wiesner, T.A. Sellers and R.A. King (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* 30: 450–458.
- R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL, <http://www.R-project.org/>.
- Sadras, V.O. (2007) Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Res.* 100: 125–138.
- Sayama, T., T.-Y. Hwang, K. Komatsu, Y. Takada, M. Takahashi, S. Kato, H. Sasama, A. Higashi, Y. Nakamoto, H. Funatsuki *et al.* (2011) Development and application of a whole-genome simple sequence repeat panel for high-throughput genotyping in soybean. *DNA Res.* 18: 107–115.
- Sayama, T., T. Tanabata, M. Saruta, T. Yamada, T. Anai, A. Kaga and M. Ishimoto (2017) Confirmation of the pleiotropic control of leaflet shape and number of seeds per pod by the *Ln* gene in induced soybean mutants. *Breed. Sci.* 67: 363–369.
- Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng *et al.* (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183.
- Smith, C.C. and S.D. Fretwell (1974) The optimal balance between size and number of offspring. *Am. Nat.* 108: 499–506.
- Smith, T.J. and H.M. 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., G. Jia, Y. Zhu, D. Grant, R.T. Nelson, E.Y. Hwang, D.L. Hyten and P.B. Cregan (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR\_1.0) in soybean. *Crop Sci.* 50: 1950–1960.
- Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, J.H. Orf and K.G. Lark (2001) Soybean response to water: A QTL analysis of drought tolerance. *Crop Sci.* 41: 493–509.
- Tsubokura, Y., S. Watanabe, Z. Xia, H. Kanamori, H. Yamagata, A. Kaga, Y. Katayose, J. Abe, M. Ishimoto and K. Harada (2014) Natural variation in the genes responsible for maturity loci *E1*, *E2*, *E3* and *E4* in soybean. *Ann. Bot.* 113: 429–441.
- van Ooijen, J.W. (1992) Accuracy of mapping quantitative trait loci in autogamous species. *Theor. Appl. Genet.* 84: 803–811.
- Venable, D.L. (1992) Size-number trade-offs and the variation of seed size with plant resource status. *Am. Nat.* 140: 287–304.
- Voorrips, R.E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93: 77–78.
- Wang, S., J. Basten and Z.B. Zeng (2013) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

- Watanabe, S., R. Hideshima, Z. Xia, Y. Tsubokura, S. Sato, Y. Nakamoto, N. Yamanaka, R. Takahashi, M. Ishimoto, T. Anai *et al.* (2009) Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* 182: 1251–1262.
- Xia, Z., Y. Tsubokura, M. Hoshi, M. Hanawa, C. Yano, K. Okamura, T.A. Ahmed, T. Anai, S. Watanabe, M. Hayashi *et al.* (2007) An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F<sub>2</sub> population. *DNA Res.* 14: 257–269.
- Yamaguchi, N., T. Sayama, H. Yamazaki, T. Miyoshi, M. Ishimoto and H. Funatsuki (2014) Quantitative trait loci associated with lodging tolerance in soybean cultivar ‘Toyoharuka’. *Breed. Sci.* 64: 300–308.
- Yamanaka, N., S. Watanabe, K. Toda, M. Hayashi, H. Fuchigami, R. Takahashi and K. Harada (2005) Fine mapping of the *FT1* locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. *Theor. Appl. Genet.* 110: 634–639.
- Ye, S., S. Dhillon, X. Ke, A.R. Collins and I.N. Day (2001) An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* 29: e88.
- Zhou, Z., Y. Jiang, Z. Wang, Z. Gou, J. Lyu, W. Li, Y. Yu, L. Shu, Y. Zhao, Y. Ma *et al.* (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33: 408–414.