SCIENTIFIC REP⊙RTS

OPEN

Integration of lodging resistance QTL in soybean

Sadal Hwang[1] & Tong Geon Lee[1,2]

Poor lodging resistance could limit increases in soybean yield. Previously, a considerable number of observations of quantitative trait loci (QTL) for lodging resistance have been reported by independent studies. The integration of these QTL into a consensus map will provide further evidence of their usefulness in soybean improvement. To improve informative QTL in soybean, a mapping population from a cross between the Harosoy and Clark cultivars, which inherit major U.S. soybean genetic backgrounds, was used along with previous mapping populations to identify QTL for lodging resistance. Together with 78 QTL for lodging collected from eighteen independent studies, a total of 88 QTL were projected onto the soybean consensus map. A total of 16 significant QTL clusters were observed; fourteen of them were confirmed in either two or more mapping populations or a single population subjected to different environmental conditions. Four QTL (one on chromosome 7 and three on 10) were newly identified in the present study. Further, meta-analysis was used to integrate QTL across different studies, resulting in two significant meta-QTL each on chromosomes 6 and 19. Our results provide deeper knowledge of valuable lodging resistance QTL in soybean, and these QTL could be used to increase lodging resistance.

Lodging is a morphological trait that limits crop yield potential. For example, the effect of lodging decreased yield by approximately 40% in oat, a cereal crop[1]. In dicotyledonous plants, such as soybean, environmental conditions that promote high yield could aggravate lodging by stimulating the height and vegetative growth of the plants[2]. Moreover, lodging could reduce seed yield by up to 10%, suggesting that high-yielding germplasm may be particularly affected by yield loss[3,4]. The physiological mechanisms of yield loss under lodging remain unknown. A reduction in photosynthesis within a canopy may suppress the transport of water and photosynthetic assimilates[5–7]. Stem strength[8], root morphology[9], and node number[10] can influence sensitivity to lodging.

Berry et al.[5] demonstrated two types of lodging: root (bending-type) and stem (breaking-type) lodging. Root lodging occurs at the base of a plant due to a lack of anchorage strength. Stem lodging arises at any point on the stem due to weakened stem bending strength. The former is the most common type of lodging in crop plants, including soybean. The latter usually occurs at the lower portion of a plant when the plant starts grain or seed filling, making the stem dry and brittle.

Improved management practices such as disease control, plant density, planting date, and fertilizer application have been utilized to prevent lodging in soybean. A simulated study has shown that lodging at the growth stages R3, R5, and R6 reduces seed yield by 12–18, 18–32, and 13–15%, respectively[4]. High plant density can also aggravate lodging without increasing yield. Approximately 120,000 to 150,000 seeds/acre sown in 30-inch rows would be ideal to produce maximum yield in the Midwest region of the U.S.[11,12]. Pod or stem disease may easily spread among lodged plants. Early planted short-season soybean could be prone to lodging due to damage by the soybean stem borer[13]. Lodging can also slow down the harvesting process because it is not easy to cut and gather lodged plants into a combine.

The heritability of lodging resistance and its relationship with plant height or seed yield have been evaluated. In a wide range of population studies, a moderate to high heritability of lodging resistance has been reported[14–21]. There is evidence that lodging and height are positively correlated[22–26]. However, lodging and seed yield showed the following discrepancies in the sign of their correlation: not significant[27], positive[23,26,28–31], negative[25,32], and positive or negative[25,33]. These reports suggested that seed yield could be weakly affected by indirect selection on lodging.

Though intensive QTL analysis has been conducted for lodging resistance in soybean[19,22–26,31,33–37], there has not been a comprehensive study that integrates data from a wide range of seasons and populations. A deeper

[1]Gulf Coast Research and Education Center, University of Florida, Wimauma, FL, 33598, USA. [2]Horticultural Sciences Department, University of Florida, Gainesville, FL, 32611, USA. Correspondence and requests for materials should be addressed to T.G.L. (email: tonggeonlee@ufl.edu)

1

Year			1998	1999	1998	1999
Water treatment			Irrigated	Irrigated	Rainfed	Rainfed
Parental	Mean	Harosoy	2.925	2.550	3.225	2.000
		Clark	2.625	2.500	2.125	2.225
	SD[a]	Harosoy	0.245	0.510	0.573	0.513
		Clark	0.393	0.628	0.358	0.734
	t-test[b]		**	ns	**	ns
	Power[c]		82.6	5.90	100.0	20.3
RIL	Mean		2.74[d]	2.80	2.87	2.29
	Minimum		1.00	1.00	1.00	1.00
	Maximum		4.50	4.50	5.00	4.25
	Skewness		0.126	0.364	0.242	0.372
Population	Kurtosis		0.801	−0.257	−0.059	0.328
	Normality[e]	S-W test	0.0061	<0.0001	0.0403	<0.0001
		K-S test	0.0220	<0.0100	0.1500	<0.0100
	H^{2f}			0.679		0.685
	CI of H^2			0.595–0.746		0.602–0.750

Table 1. Population statistics for lodging in the Harosoy x Clark population. [a]The SD represents the standard deviation of each parental line with 20 samples. [b]The significance from a two-tailed t-test was presented as ** and ns, which indicates that P values were less than 0.01 and greater than 0.05, respectively. [c]The statistical power was estimated by a two-tailed t-test. [d]Plant lodging was visually assessed by using a score that ranged from 1 (erect) to 5 (prostrate). [e]The P values are from two types of tests used to test the normality of progeny means for each water-year data set (S-W, Shapiro-Wilk; K-S, Kolmogorov-Smirnov). [f]The 95% of confidence intervals (CIs) of heritability were estimated from the combined two-year data from the irrigated and rainfed field trials.

knowledge of informative QTL has the potential to provide the community of soybean researchers with new tools such as a consensus marker for marker-assisted selection (MAS). Therefore, in the study below, we examine QTL and meta-QTL for lodging resistance across independent populations and a population that we created. We investigate statistically significant QTL-lodging associations and major QTL in the investigated populations.

Results

Development of the Harosoy x Clark genetic map. The genetic map of the Harosoy x Clark (hence-forth referred to as H x C) population was constructed (Supplemental Fig. 1): the total genetic map distance was 3769 cM and the average map distance of the H x C population was 5.21 cM. Since the average map distance of the consensus genetic map was 0.41[38], the map resolution of the H x C population was 12.5 (=5.21 cM/0.41 cM) times lower than that of the soybean consensus map.

Analysis of the field data of H x C population. A parental difference would be a necessary assumption of a single-QTL model, since transgressive segregation requires at least two QTL. The differences between two parental means were tested by a t-test (H_a: Two parental means are not the same) in the irrigated and rainfed trials (Table 1). Parental differences for lodging were statistically significant ($P \leq 0.01$) under irrigated and rainfed conditions in 1998. The statistical power of each t-test in 1998 was greater than 0.8. The parental means in 1999 were not significantly different under either water condition.

The single-QTL model in our study assumes that lodging follows a normal distribution, i.e., that the residuals follow a normal distribution and are independent. A normality test was performed for each water-year data set (Table 1). The average means of lodging ranged from 2.29 to 2.80. The minimum value of lodging was 1, and the maximum value of lodging was close to 5. The Shapiro-Wilk and Kolmogorov-Smirnov tests (H_a: A random variable for lodging does not follow a normal distribution) showed that lodging did not follow a normal distribution. The Kolmogorov-Smirnov test indicated that lodging under rainfed conditions in 1998 followed a normal distribution. The degree of skewness was less than 0.5 and positive. Kurtosis was low under rainfed and irrigated conditions in 1998 and 1999, which indicated no acute peakedness around the distributional mean. Thus, we assumed that our field data for lodging followed a normal distribution. No attempt to improve normality was made in this study.

Based on the analysis of variance (ANOVA) result, the heritability of lodging resistance was estimated over years in each type of water treatment (Table 1 and Supplementary Table 1). While the variance for the Water x recombinant inbred lines (RIL) interaction was not statistically significant, it was ($P < 0.0001$) for RIL and the Year x Water, Year x RIL and Year x Water x RIL interactions (Supplemental Table 1). Since all interactions except for Water x RIL were significant, the RIL means were not averaged across the water and year treatments for subsequent combined-QTL analyses. Instead, the LS mean for each RIL was estimated from each water-year data set. The two heritability values under each water treatment were not substantially different, because the Water x RIL interaction was not statistically significant (Table 1). The heritability of lodging ranged from 0.595 to 0.750 within the upper and lower values of the 95% confidence intervals (CIs).

QTL name	Chromosome number	Year	Water[a]	QTL marker	QTL[b] position	Favorable allele[d]	R[2e]	QTL[f] effect	LOD	Flanking markers and their positions[b] 95% CI[g]
Lg01	7	1998	I	Sat_288	72.83	Clark	0.04	0.13	3.3	BARC-017117-02201 - Satt551 65.88–89.45
Lg02	10	1998	I	BARC-015925-02017	99.69	Harosoy	0.10	0.22	8.4	BARC-050013-09288 - Satt153 94.97–106.32
Lg03	15	1998	I	BARC-057969-15031	77.04	Clark	0.09	0.19	7.0	BARC-053201-11762 - BARC-057969-15031 76.60–77.04
Lg04	18	1998	I	Sat_064	101.82	Clark	0.07	0.17	5.2	Sat_064 - BARC-057845-14952 101.82–103.11
Lg05	14	1998	R	BARC-065009-19043	56.60	Clark	0.07	0.19	4.4	BARC-065009-19043 - Satt474 56.60–63.36
Lg06	10	1999	I	BARC-050013-09288	94.97	Harosoy	0.25	0.40	86.1	Satt592 - Satt581 91.36–95.60
Lg07	13	1999	I	BARC-055613-13490	77.16	Harosoy	0.05	0.23	4.9	BARC-055229–13122 - Satt144 71.89–78.89
Lg08	18	1999	I	Sat_131	32.88	Harosoy	0.07	0.20	5.7	BARC-014395-01348 - Satt324 19.48–35.43
Lg09	10	1999	R	E2	121.41[c]	Harosoy	0.27	0.43	20.3	BARC-063361-18346 - BARC-041935-08142 120.36–122.45
Lg10	15	1999	R	BARC-058675-17461	68.06	Clark	0.07	0.21	3.9	BARC-050109-09389 - BARC-058675-17461 54.94–68.06

Table 2. Lodging resistance QTL in the Harosoy x Clark population. [a]Two different water treatments were applied to irrigated and rainfed field trials. I and R represent irrigated and rainfed field water conditions, respectively. [b]QTL markers and flanking markers were positioned based on the Consensus 4.0 genetic map of soybean. [c]Because $E2$ was not positioned on the Consensus 4.0 genetic map of soybean, the nearest marker, BARC-024447-04891, was considered as the QTL marker of $E2$. [d]Based on the maximum likelihood-estimated QTL positions, alleles with a low plant lodging score were defined as favorable alleles. [e]The amount of phenotypic variation explained by a QTL marker was estimated as a R^2 value. [f]Additive effects were estimated as half the difference between the average effects of two parental alleles at the maximum likelihood-estimated QTL positions. [g]The LOD values with ±1 deviation were used to estimate the 95% confidence invervals of the maximum likelihood-estimated QTL positions.

QTL mapping of H x C population. The associations between lodging resistance and molecular markers were evaluated in the H x C population (Table 2). As we described earlier, although the parental means in 1999 were not different, we conducted QTL analysis for all field trials. A total of ten putative QTL were identified on chromosomes 7, 10, 13, 14, 15, and 18. Three of the ten QTL were identified on chromosome 10. Two QTL, Lg02 and Lg06, which were positioned at 99.69 and 94.97 cM, had overlapping CIs. The phenotypic variation (R^2 value) explained by a QTL marker ranged from 0.04 to 0.27. Interestingly, two QTL markers on chromosome 10, BARC-050013-09288 and $E2$, showed the highest R^2 value, suggesting that these QTL could make major contributions to the lodging resistance of the H x C population. Further, the $E2$ marker showed the highest QTL effect, and the allele derived from Harosoy was detected in this locus.

Confirmation of QTL. Next, we integrated QTL from 19 previous studies and ours (Tables 2 and 3, and Supplemental Table 2). A total of 88 QTL (78 QTL from previous studies and 10 QTL from the H x C population) were projected onto the consensus genetic map of soybean based on 95% CIs to confirm QTL (Fig. 1). QTL identified along with overlapping CIs in two or more mapping populations were considered strongly confirmed QTL[39]. The QTL identified from different environmental conditions (such as location, year, and water treatment) and with overlapping CIs in only one population were considered weakly confirmed QTL[39].

There were 16 QTL clusters on chromosomes 2 (one QTL), 3 (one), 4 (two), 6 (one), 8 (one), 9 (one), 10 (one), 13 (two), 14 (one), 15 (one), 18 (two), 19 (two), and 20 (one). Two QTL clusters on chromosomes 4 (telomere proximal) and 8 were identified in a single population in the same environment[22] (data combined across locations). Therefore, the two QTL clusters on chromosomes 4 and 8 were excluded from subsequent analyses.

Fourteen QTL clusters had overlapping CIs and were thus confirmed as lodging QTL. Two of 14 clusters on chromosomes 9 and 10 were confirmed in different locations or water treatments. The rest of the clusters were confirmed in two or more mapping populations. Based on 95% CIs, six QTL of the H x C population were shared by one or more previous studies[18,19,22,34]. Four new QTL were identified on chromosomes 7 (one QTL) and 10 (three) under irrigated and rainfed field trials in both years.

We investigated whether QTL for stem-related traits, root-related traits, and node number were positioned in the CIs of lodging resistance QTL. A considerable number of QTL were found: QTL for stem strength[8] (40.8, 106.8, 77.3, and 62.9 cM on chromosomes 4, 6, 13, and 14, respectively), stem diameter[40] (76.5 cM on 19), root morphology[9] (64.9 cM on 3), root dry weight[41,42] (17.35 and 28.38 cM on 3), root lateral number[43] (12.3 cM on 19), and node number[10,44,45] (105.8, 31.4, and 61.0 cM on 6, 11, and 13, respectively). Additionally, a locus for major growth habit, $Dt1$[46], was positioned at 78.6 cM on chromosome 19 regardless of whether the parental cross combination was $dt1/dt1$[33] or $dt1/Dt1$[35].

Meta-QTL analysis. Of the 14 QTL clusters, only two were used for meta-QTL analysis, as the others showed low likelihood of odd (LOD) values (or P values) in a single-marker analysis (SMA)[47] due to insufficient

Reference	Population				QTL		
	Name	Size	Type	Cross type	Method[a]	Marker type	Number[b]
[82]	Minsoy x Noir 1	69	$F_{2:5}$ RIL	_G. max_ x _G. max_	IM	RFLP	1
[23]	Minsoy x Noir 1	284	F_7-derived RIL	_G. max_ x _G. max_	SMA	RFLP + SSR	4
[22]	Young x PI 416937	120	F_4-derived RIL	_G. max_ x _G. max_	SMA	RFLP	14
[78]	PI 97100 x Coker 237	111	F_2	_G. max_ x _G. max_	SMA + IM	RFLP + Classical	4
[33]	Minsoy x Noir 1	240	F_7-derived RIL	_G. max_ x _G. max_	IM	SSR + Classical	5
[33]	Minsoy x Archer	233	F_7-derived RIL	_G. max_ x _G. max_	IM	SSR + Classical	1
[33]	Noir 1 x Archer	240	F_7-derived RIL	_G. max_ x _G. max_	IM	SSR + Classical	1
[26]	Minsoy x Noir 1	236	$F_{7:11}$ RIL	_G. max_ x _G. max_	CIM	RFLP + SSR + Classical	5
[16]	Essex x Williams	177	$F_{4:6}$ RIL	_G. max_ x _G. max_	SMA	SSR	1
[36]	PI 468916 x IA2008	110	BC_2F_4-derived	_G. max_ x _G. soja_	CIM	SSR	1
[10]	Kefeng No.1 x Nanong 1138-2	184	$F_{7:10}$ RIL	_G. max_ x _G. max_	CIM	RFLP + SSR	3
[31]	BSR 101 x LG82-8379	167	F_5-derived RIL	_G. max_ x _G. soja_	SMA	SSR	1
[25]	RG10 x OX948	169	F_6 RIL	_G. max_ x _G. max_	CIM	SSR	5
[18]	LG96–6607 x Lawrence[3]	94	BC_3F_2-derived	_G. max_ x _G. max_	SMA	SSR	2
[18]	LG92-1143 x Beeson 80[2]	68	BC_2F_2-derived	_G. max_ x _G. max_	SMA	SSR	2
[18]	LG94–1713 x Kenwood[1]	74	BC_1F_2-derived	_G. max_ x _G. max_	SMA	SSR	3
[24]	N87-984-16 x TN93-99	101	F_6-derived RIL	_G. max_ x _G. max_	CIM	SSR	2
[37]	PI 245331 × 7499	147	BC_2F_4-derived	_G. max_ x _G. soja_	SMA	SSR	8
[20]	OAC Millennium x Heinong 38	98	$F_{4:7}$ RIL	_G. max_ x _G. max_	SMA	SSR	2
[83]	Pioneer 9071 x Line 8902	133	$F_{4:7}$ RIL	_G. max_ x _G. max_	SMA	SSR	1
[34]	PI 436684 x PI 548557	116	BC_2F_3-derived	_G. max_ x _G. max_	CIM	SNP	1
[34]	PI 90566-1 x Williams 82	93	BC_2F_3-derived	_G. max_ x _G. max_	CIM	SNP	1
[84]	OAC Millennium x Heinong 38	92	$F_{4:7}$ RIL	_G. max_ x _G. max_	SMA	SSR	4
[84]	Pioneer 9071 × 8902	131	$F_{4:7}$ RIL	_G. max_ x _G. max_	SMA	SSR	2
[19]	PI 567310B x Wyandot	91	F_7-derived RIL	_G. max_ x _G. max_	SMA + CIM	SNP	4

Table 3. Previous QTL mapping studies for lodging. [a]IM, interval mapping; SMA, single marker analysis; Single marker analysis was based on one-way ANOVA or a paired _t_-test between any two parental alleles. [b]QTL markers from mapping study.

design of the populations or QTL redundancy. A total of 12 QTL on chromosomes 6 (five QTL) and 19 (seven) were projected onto the consensus genetic map as a reference map for meta-analysis (Fig. 2). Across all QTL except for one on chromosome 6, the LOD values were greater than 2.5 (LOD $2.5 \approx P < 0.001$, assuming that the LOD or likelihood ratio (LR) asymptotically follows a χ^2 distribution). The QTL on chromosome 6[48] had a P value less than 0.01 (LOD = 1.75).

Five model selection criteria were used to determine how many meta-QTL could be chosen with the global likelihood value (Table 4). Four and five different mapping populations were used for meta-analysis of chromosomes 6 and 19, respectively. The 5 or 7 QTL were used to evaluate 1 to 5 or 1 to 7 meta-QTL in each QTL cluster. For a given number of meta-QTL, there were differences among the values of each model selection criterion. Except for the corrected Akaike information criterion (AICc), all model selection criteria for both QTL clusters showed the lowest value when two meta-QTL were used to estimate the maximum global likelihood value. When the AICc was used, this model selection criterion showed the lowest values for 3 meta-QTL and 1 meta-QTL on chromosomes 6 and 19, respectively. Considering all five model selection criteria values, the likelihood of two meta-QTL appeared to be the highest. Two QTL identified in the Minsoy x Noir 1 population (98.3 cM on chromosome 6[33]; 81.0 cM on 19[23]) were found to be significant clusters that contributed to the expected meta-QTL in a subsequent meta-analysis.

Consequently, four meta-QTL were identified on chromosomes 6 and 19 (Fig. 2; Table 5). Two of them were located 5.73 cM apart on chromosome 6. There were no overlapping CIs between those with R^2 values ranging from 0.15 to 0.21. Two meta-QTL on chromosome 19 showed slightly overlapping CIs with average R^2 values from 0.27 to 0.39. Notably, all meta-QTL on chromosomes 6 and 19 were predicted major QTL, although meta-analysis generally decreased the CIs of meta-QTL.

Discussion

We performed an integrative QTL study of the H x C population and previous mapping populations for lodging resistance, resulting in 14 significant QTL clusters. Meta-analysis of QTL identified in different studies located clusters of QTL on chromosomes 6 and 9.

There are apparent discrepancies in the QTL detected in different mapping studies, likely because of the type and number of markers incorporated, the population size, heritabilities of target QTL, QTL models, and linkage disequilibrium (LD)[49–51]. It has been suggested that meta-analysis could address such issues to achieve more efficient MAS[39]. Several QTL in previous populations were located in mapping intervals that span a wide range of
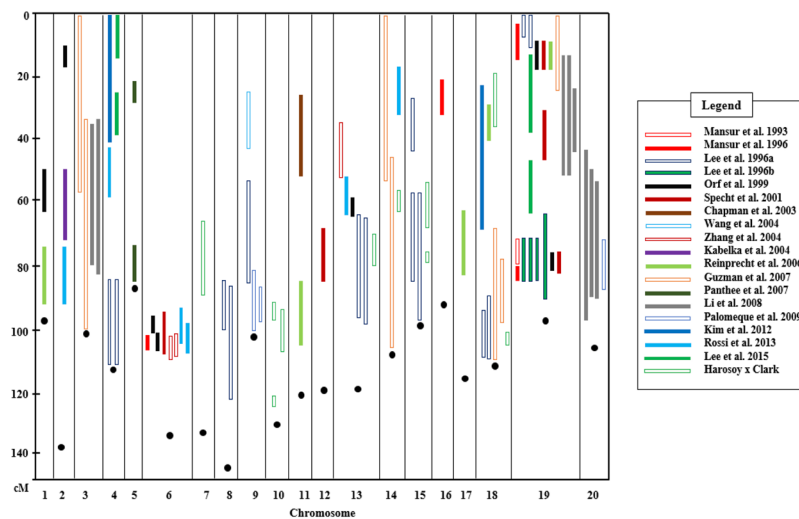
Figure 1. Integration of lodging QTL in soybean. After considering the confidence intervals of all QTL identified in this study and independent researches, significant QTL were projected onto the Consensus 4.0 genetic map of soybean. Black dots indicate the telomere-proximal end of each chromosome based on the consensus map.
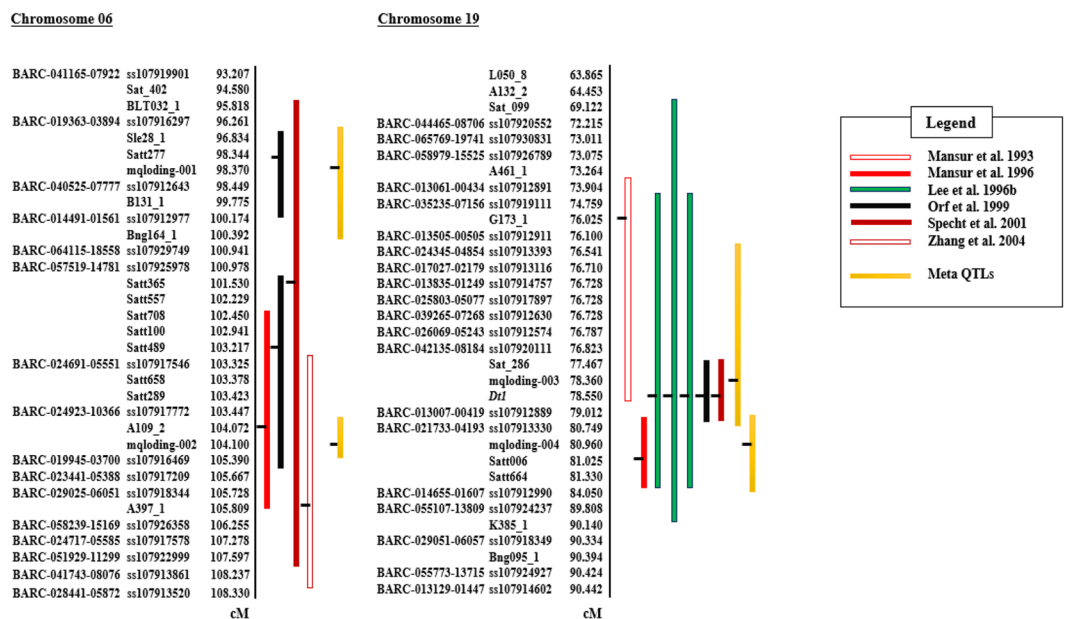

Figure 2. Meta-QTL for soybean lodging resistance. Four meta-QTL (yellow bars) were identified on two different chromosomes, 6 and 19. Positions of maximum likelihood-estimated QTL were determined (black horizontal bars). Colored vertical bars show the 95% confidence intervals. Major markers [either Beltsville Agricultural Research Center (BARC) local ID or SNP identifier in NCBI-dbSNP or both] that overlapped with or flanked QTLs were added to the left side of the chromosome. *cM* centimorgan.

CIs. Presumably, the main causes are population size, population type, and the false-positive level used for QTL confirmation. Considering the polymorphism information content (PIC)[52], a large portion of genome segments in a backcross population may inherit monoallelic information. A small backcross population size could increase map distances between adjacent markers (typically over 20 cM), and the resulting larger map distances could overestimate R^2 values and affect the estimation of CIs for QTL and meta-QTL. In addition, a low false-positive threshold could overestimate the additive effects of QTL and underestimate the CIs of QTL if those QTL originally had small gene effects[49]. Lander and Kruglyak[39] demonstrated that false-positive QTL likely occur under stringent threshold values, such as the family-wise error rate (FWER). In our study, we identified meta-QTL in two QTL clusters in which the QTL mostly had stringent thresholds (LOD > 3.0) and overlapping CIs in multiple mapping populations. Thus, these meta-QTL can be appropriate consensus markers for MAS to improve lodging resistance.

Chromosome	Number of mapping populations[a]	Number of QTL[b]	Number of Meta-QTL[c]	Values of model selection criteria[d] and delta (Δ)[e]									
				AIC	AIC Δ[e]	AICc	AICc Δ	AIC3	AIC3 Δ	BIC	BIC Δ	AWE	AWE Δ
6	5	5	1	53.8	24.2	54.8	21.2	54.8	22.2	53.6	24.6	58.4	19.7
			2	29.6	0.0	41.6	8.0	32.6	0.0	30.0	0.0	38.7	0.0
			3	33.6	4.0	33.6	0.0	38.4	6.0	32.6	3.4	51.8	13.2
			4	.	.	.	.	.	.	.	.	.	.
			5	38.1	8.5	38.1	4.5	44.1	11.5	36.9	7.9	54.9	16.2
19	5	7	1	37.7	6.3	38.5	0.0	38.7	4.3	37.6	6.3	42.6	1.2
			2	31.2	0.0	39.3	0.9	34.4	0.0	31.2	0.0	41.4	0.0
			3	35.3	4.0	95.3	56.7	40.2	6.0	35.0	3.8	54.5	13.0
			4	39.3	8.0	39.3	0.8	46.3	12.0	38.9	7.7	68.2	26.8
			5	.	.	.	.	.	.	.	.	.	.
			6	.	.	.	.	.	.	.	.	.	.
			7	39.6	8.3	39.6	1.2	46.6	12.3	39.3	8.1	65.3	23.9

Table 4. Model selection for meta-QTL analysis of lodging on chromosomes 6 and 19. [a]This number indicates how many populations were used as independent populations for meta-analysis in each QTL cluster. [b]The number of QTL indicates how many QTL in each QTL cluster. [c]The optimal positions and number of meta-QTL were considered based on the number of QTL in each QTL cluster to test the best meta-QTL models. [d]AIC, AICc (or AIC3), BIC, and AWE indicates Akaike information criterion, corrected Akaike information criterion, Bayesian information criterion, and approximate weight of evidence. [e]This could be defined as the difference between model selection criteria values between two meta-QTL models such as the best meta-QTL model and other meta-QTL models.

A positive correlation between lodging and soybean maturity at the R8 growth stage was reported[22,23,25,26,53]. The correlation coefficient (γ) mostly ranged from 0.20 to 0.51. Our data from the H x C population showed a γ of 0.25 ($P < 0.01$). Our data and the independent studies were consistent with results deposited in Germplasm Resources Information Network (GRIN; http://www.ars-grin.gov/cgi-bin/npgs/html/site_holding.pl?SOY). In fact, the *E* series of loci, which are responsible for maturity or photoperiodism, were located in the CIs of QTL and QTL clusters for lodging. Historically, the maturity effect of the *E* series of loci was evaluated by genetic studies with near isogenic lines (NILs) of Harosoy or Clark[54]. *E1*[55,56] and *E2*[55] were positioned at 103.33 and 121.41 cM on chromosomes 6 and 10, respectively. Relative to *E2*, *E1* is particularly well known to have a major effect on flowering time, maturity, and branching[57,58]. *E3*, a type of phytochrome (*GmPhyA3*), was positioned at an interval between 78.26 and 94.5 cM on chromosome 19[59]. This interval included the QTL cluster on chromosome 19 revealed by our meta-analysis. A recent study indicated that allelic variation in *E1*, *E2*, *E3*, and *E4* explained approximately 64% of the phenotypic variation in flowering time among 63 soybean accessions[60]. *E5*[61] and *E7*[62] were closely located to *E2* and *E1*. Although *E4*[63] (31.51 cM on chromosome 20) was not positioned in our study, these results indicate that lodging and maturity are fairly associated with each other.

We found several QTL for root and shoot traits were positioned in the CIs of lodging resistance QTL. Considering the effect of specific gene(s) (for example, *Dt1* and height), the relationships between lodging and other traits suggested that classical markers such as *E* and *Dt1* still had a strong and pleiotropic effect on lodging similar to that of a QTL cluster on chromosome 6.

Although recent QTL studies in plants are increasing our knowledge of the lodging resistance, the molecular mechanisms of lodging resistance remain unknown in most cases. Our study will further narrow down QTL intervals to provide resources for identification of candidate genes: for example, according to the current genome annotation Wm82.a2 (https://phytozome.jgi.doe.gov), there are 163 predicted genes (Glyma.02G293200 to Glyma.02G309400) within the interval to which mqloding-002 is mapped. To identify the sequence of candidate gene(s) from QTL intervals, the map-based cloning or sequence assembly within the intervals from the source of the lodging resistant soybean is needed. Further fine mapping to narrow down such intervals will also greatly facilitate the cloning of candidates.

Putative new lodging resistance QTL on chromosomes 7 and 10 were mapped in the H x C population. The growing number of QTL identified since the early 1990s, and the identification of new QTL demonstrates that there is still potential for discovering new lodging resistance QTL in soybean. Our effort will facilitate the identification of new resistance genes and QTL and will increase the pool of alleles that are important for the control of lodging. Since meta-analysis of QTL identified in different studies (experimental conditions/plant materials) can locate QTL more precisely, the four meta-QTL identified in our analysis will also aid in the development of improved markers to increase soybean breeding efficiency.

Methods

Population development. Two U.S. soybean (*Glycine max*) lines, Harosoy[64] and Clark[65], were used to develop populations. Harosoy and Clark were crossed as females and males, respectively. Each of eight F_1 plants produced approximately 45 F_2 seeds. The plant-to-row method was applied to advance the generation from F_2 to F_6. A total of 25 F_6 plants in each of the 300 F_6-derived rows were bulk threshed by row to generate 300 RILs.

Chromosome	Designation of meta-QTL[a]	Number of meta-QTL[b]	Position of meta-QTL[c]	Mean R²ᵈ	meta-QTL[c]				
					Flaking markers				
					Left	Position	Right	Position	
6	mqloding-001	2	98.370	0.21	Satt277	98.344	BARC-040525-07777	98.449	
	mqloding-002		104.10	0.15	A109_2	104.072	BARC-019945-03700	105.390	
19	mqloding-003	2	78.360	0.27	Sat_286	77.467	*Dt1*	78.550	
	mqloding-004		80.960	0.39	BARC-021733-04193	80.749	Satt006	81.025	
					95% CI of meta-QTL[c]				
					Flaking markers				
					Left	Position	Right	Position	
6	mqloding-001	2	98.370	0.21	BARC-019363-03894	96.261	BARC-064115-18558	100.941	
	mqloding-002		104.100	0.15	BARC-024923-10366	103.447	BARC-019945-03700	105.390	
19	mqloding-003	2	78.360	0.27	BARC-013505-00505	76.100	BARC-021733-04193	80.749	
	mqloding-004		80.960	0.39	BARC-013007-00419	79.012	BARC-014655-01607	84.050	

Table 5. Meta-QTL for lodging resistance on chromosomes 6 and 10. ᵃThe names of meta-QTL were designated for the purpose of using Soybase (https://soybase.org/). ᵇThe number of meta-QTL was based on the values of five model selection criteria. ᶜThe positions of meta-QTL were determined by the maximum joint likelihood values in the search for the best meta-QTL models. All positions were based on the Consensus 4.0 genetic map of soybean. ᵈThe mean R² values were averaged by the R² values of QTL in a QTL cluster. However, in two meta-QTL, mqloding-001 and mqloding-004, because there was no QTL cluster, the R² values were simply based on previous mapping results[23,33]. ᵉThe positions and CIs of meta-QTL were based on the Consensus 4.0 genetic map of soybean.

The number of RILs derived from each F_1 ranged from 34 to 40. Parental lines of Harosoy [Mandarin (Ottawa) (2) x A.K. (Harrow)] or Clark [Lincoln (2) x Richland] were also propagated and harvested each year. These four parental lines are considered soybean ancestors that made significant contributions to the genetic composition of U.S. soybean cultivars.

Genotyping of the population. Leaf tissue was collected in bulk from each of 300 rows of F_6 individuals, eight rows of Harosoy or Clark individuals, and four rows of individuals of each parental line of Harosoy or Clark. DNA was diluted to 20 ng/μl. A modified protocol by Akkaya et al.[66] was used for PCR. Most of the simple sequence repeat (SSR) amplicons were loaded onto a 2.5% agarose gel, and the differences among these were observed after electrophoretic separation at 70 V. Ambiguous SSR amplicons were loaded onto a polyacrylamide gel and separated in a 4300 DNA Analyzer with the LI-COR Saga V. 2 program (LI-COR Bioscience, Lincoln, NE). For single nucleotide polymorphism (SNP) genotyping, the Illumina GoldenGate Assay was performed with the 1536-SNP USLP 1.0 array[38,67]. GenCall software (Illumina, Inc., San Diego, CA) was used to identify allelic variation. Four classical markers including the pubescence color locus (*TT/tt*), hilum color locus (*RR/rr*), hilum color intensity locus (*II/i'i'*), and maturity date locus (*E2E2/e2e2*)[46] were scored in 300 F_6-derived RILs. Both Harosoy and Clark are homozygous (*W1W1*) for purple flower color. Clark (*TTRR*) is dominant over Harosoy (*ttrr*) in pubescence and hilum color. Harosoy (*II*) is dominant over Clark (*i'i'*) in hilum color intensity. Clark (*E2E2*) exhibited late maturity, whereas Harosoy (*e2e2*) displayed early maturity, indicating that late maturity is incompletely dominant over early maturity[46,55]. The genotypic values of only the RILs with definitively extreme maturity dates were scored in the H x C population.

Genetic map construction. A total of 751 markers (4 classical markers, 266 SSRs, and 481 SNPs) were used for linkage analysis in the R/qtl software[68]. We examined linkage groups using the R/qtl function formLinkageGroups() [an initial value of a LOD was 15]. For linkage grouping, LOD value of 3.0 that established known linkage groups was chosen as the significance criterion for multipoint linkage testing. The maximum distance between two flanking markers was 0.372 rf (recombination fraction obtained with the Morgan function). The Kosambi mapping function[69] was used for linkage analysis. To fix gaps in the genetic map, we set the maximum rf value to 0.450, which equaled the Kosambi map distance of 73.61 cM. To estimate linkage distances, a genotyping error of 0.01% was assumed[70]. Linkage distance was estimated by maximum likelihood (ML) with the expectation-maximization (EM) algorithm[71]. The default maximum iteration number was 10,000, and 0.000001 was used as the tolerance value. After checking the genotyping error[68] (Supplemental Fig. 2), we finalized the H x C genetic map with 20 chromosomes and 730 markers (4 classical markers, 260 SSRs, and 466 SNPs) in 285 RILs. Given that our marker data type fits the framework of the Consensus 4.0 genetic map of soybean[38] (all markers used in our study are present in the dataset used for the consensus map), we projected markers of our H x C population onto the consensus map to examine the relative marker positions.

Field experiment. Lodging was visually rated on a scale of 1 (erect) to 5 (prostrate) when plants reached the R8 stage[72]. Field trials were carried out at the west field in East Campus, University of Nebraska-Lincoln Agronomy Research Farm, in 1998 ($F_{6:8}$ plants) and 1999 ($F_{6:9}$). The soil at the test site is a Sharpsburg silty clay loam. A two-replicate balanced incomplete block design (BIBD) was used to control for the effect of maturity.

Sixteen blocks, each containing 20 RILs, were designed based on maturity date. Additionally, each block had one of the parental lines. All entries were planted into 2-row plots with 0.762 m row spacing and a length of 3.05 m, which resulted in a seeding rate of approximately 106,473 seeds/acre. Each plot had two replications under rain-fed and irrigated conditions.

SAS (V. 9.0) was used for randomization, ANOVA, least square mean estimation, independent t-tests, normality tests, and heritability estimation. Two commands, 'PROC UNIVARIATE' and 'PROC GPLOT', were used for the normality test of lodging. Randomization was conducted by the 'PROC PLAN' command. ANOVAs were performed on the phenotypic data collected in each year and on the data across the years of 1998 and 1999. The 'PROC MIXED' model is as follow:

$Y_{ijklm} = a_0 + a_1 \times 1_i + a_2 \times 2_j + a_3 X3_k + a_4 X4_l + a_5 X5_m + a_6 X1_i X2_j + a_7 X2_j X5_m + a_8 X1_i X5_m + a_9 X1_i X2_j X5_m + \varepsilon_{ijklm}$
[Y: Lodging resistance, X1: Water treatment, X2: Year, X3: Replication, X4: Block (Replication), X5: RIL, X1 × 2: Water x Year interaction, X2X5: Year x RIL interaction, X1 × 5: Water x RIL interaction, X1X2 × 5: Water x Year x RIL interactions, a_0: Overall mean, $a_1 \sim a_9$: Coefficient for each factor, ε: Error, which follows the normal distribution (mean: 0, variance: σ^2_ε)].

All classification variables except for the water treatment term were treated as random variables. The command 'PROC MIXED (method = type3)' was used for ANOVA and heritability estimation. Based on the mixed model ANOVA results, least square (LS) means were estimated with the command 'PROC LSMEANS'. The LS means for each water-year data set were used for QTL analysis. The CIs of heritability were estimated by progeny means basis[73]. The correlation coefficients between lodging and other traits were estimated with the command 'PROC CORR'. A 5% false-positive value was used for all statistical tests as the α level criterion.

QTL analysis. An ML approach[74] with the EM algorithm[75] was used to estimate parameters for composite interval mapping (CIM)[76] with a single-QTL model. The null hypothesis (H_o) for CIM was that there was no QTL anywhere in the genome. The alternative hypothesis (H_a) was that there was a QTL in an interval between flanking markers. WinQTL Cartographer V. 2.5 software[77] was used for the QTL analysis, which was applied to the data set of 285 RILs with the 751 combined classical, SSR, and SNP markers. A permutation test (10,000 repeats)[78] was applied to empirically determine the critical LR value for declaring the existence of a QTL. The LR statistic approximately followed a χ^2 distribution. Based on the ANOVA result, LR threshold values were estimated based on each water-year data set. In WinQTL Cartographer V. 2.5, model 6 was chosen to conduct a CIM analysis on the H x C data set. Up to five markers were used as control background markers. For background marker selection, the stepwise selection method was used to mitigate errors arising from using only a forward or only a backward selection approach. To avoid having too many background markers (model overfitting), an alpha value of 0.05 was used for both the forward and backward regression. A window size of 1 cM on either side of the interval flanked by a marker pair was chosen to control background marker effects and to generate an LOD profile across the whole genome with a default window size of 10. To separate multiple close peaks above the LR threshold in WinQTL Cartographer, the parameters for the single-QTL model included a minimum distance of 5 cM between the putative map positions of adjacent putative QTL. The LOD peak values, additive effects, and 95% CIs were estimated based on the ML-estimated QTL positions. An LOD score profile was generated using a 1 cM walking speed.

Data mining and QTL projection. All QTL mapping information was collected from the public database Soybase (http://soybase.org) (Table 3 and Supplemental Table 2). For QTL confirmation and meta-analysis, we collected QTL marker positions, R^2 values, the size of the population, and CIs from previous studies. CIs of previous QTL in independent studies were calculated as follows:

$$95\% \text{ CI} = 530/(N \times R^2) \text{ for a backcross or an } F_2 \text{ population}$$

$$95\% \text{ CI} = 163/(N \times R^2) \text{ for an RIL population}$$

where R^2 was the phenotypic variance explained by a QTL and N was the size of the population. QTL with 95% CIs were projected onto the Consensus 4.0 genetic map of soybean for QTL confirmation[70].

Meta-analysis. For meta-analysis, QTL were projected onto the Consensus 4.0 genetic map of soybean with BioMercator V. 3.0[79] and an algorithm from MetaQTL[80]. QTL clusters, which included the overlapping 95% CIs of two or more QTL, were used to identify the CIs and positions of meta-QTL. The assumption of the meta-analysis was that the variance of the ML-estimated QTL positions followed a normal distribution. The variance of each projected QTL position in a QTL cluster was estimated from the 95% CI[51] as follows:

$$95\% \text{ CI} = 3.92 \times \text{Variance (QTL)}^{1/2}$$

The variance of a meta-QTL position was estimated by the formula:

$$\text{Variance (QTL)} = 1/\sum_{i=1}^{n} \frac{1}{\text{Variance (QTL}i)}$$

where Variance (QTL_i) is the variance of the ith QTL in a QTL cluster with n QTL. In a window size based on the CIs of QTL in a QTL cluster, ML values were obtained with the meta-analysis models. To determine whether there is more than one meta-QTL in each QTL cluster, we evaluated from 1 to n (where n is the number of meta-QTL in each QTL cluster) meta-QTL models using joint-mixture normal functions[81]. We evaluated the

meta-QTL models using five criteria: Akaike information criterion (AIC), corrected Akaike information criterion (AICc and AIC3), Bayesian information criterion (BIC), and approximate weight of evidence (AWE). Using the most appropriate model from these selection criteria, we identified the consensus meta-QTL positions. QTL identified in at least two independent field experiments (either location or year) were used for meta-analysis. If there were multiple QTL in the same field experiment, we chose one with a high LOD or R² value. If a QTL was identified by using means across locations and/or years, it was considered to be from an independent experiment.

References

1. Pendleton, J. W. The effect of lodging on spring oat yields and test weights. *Agron. J.* **46**, 265–267 (1954).
2. Afza, R., Hardarson, G., Zapata, F. & Danso, S. K. A. Effects of delayed soil and foliar N fertilization on yield and N₂ fixation of soybean. *Plant Soil* **97**, 361–368 (1987).
3. Noor, R. B. M. & Caviness, C. E. Influence of lodging on pod distribution and seed yield in soybeans. *Agron. J.* **72**, 904–906 (1980).
4. Woods, S. J. & Swearingin, M. L. Influence of simulated early lodging upon soybean seed yield and its components. *Agron. J.* **69**, 239–242 (1977).
5. Berry, P. M. *et al.* Understanding and reducing lodging in cereals. *Adv. Agron.* **84**, 217–271 (2004).
6. Pinthus, M. J. Lodging in wheat, barley and oats: the phenomenon, its causes and preventative measures. *Adv. Agron.* **25**, 210–263 (1973).
7. Setter, T. L., Laureles, E. V. & Mazaredo, A. M. Lodging reduces yield of rice by self-shading and reductions in canopy photosynthesis. *Field Crops Res.* **49**, 95–106 (1997).
8. Chen, H. *et al.* Quantitative trait loci analysis of stem strength and related traits in soybean. *Euphytica* **179**, 485–497 (2011).
9. Abdel-Haleem, H., Lee, G. & Boerma, R. Identification of QTL for increased fibrous roots in soybean. *Theor. Appl. Genet.* **122**, 935–946 (2010).
10. Zhang, W. *et al.* QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. *Theor. Appl. Genet.* **108**, 1131–1139 (2004).
11. Board, J. Reduced Lodging for Soybean in Low Plant Population is Related to Light Quality. *Crop Sci.* **41**, 379–384 (2001).
12. Cooper, R. L. Influence of early lodging on yield of soybean. *Agron. J.* **63**, 449–450 (1971).
13. Wright, R. J. & Hunt, T. E. Soybean stem borers in Nebraska. University of Nebraska-Lincoln Extension, http://extensionpublications.unl.edu/assets/pdf/g2082.pdf (2011).
14. Brim, C. A. & Burton, J. W. Recurrent selection in soybeans. II. *Selection for increased percent protein in seeds. Crop Sci.* **19**, 494–498 (1979).
15. Byth, D. E. & Caldwell, B. E. Correlated truncation selection for yield in soybeans. *Crop Sci.* **9**, 699–702 (1969).
16. Chapman, A. *et al.* Quantitative trait loci for agronomic and seed quality traits in an F₂ and F₄:₆ soybean population. *Euphytica* **129**, 387–393 (2003).
17. Fehr, W. R. & Weber, C. R. Mass selection by seed size and specific gravity in soybean populations. *Crop Sci.* **8**, 551–554 (1968).
18. Guzman, P. *et al.* QTL Associated with yield in three backcross-derived populations of soybean. *Crop Sci.* **47**, 111–122 (2007).
19. Lee, S., Jun, T., Michel, A. & Mian, M. SNP markers linked to QTL conditioning plant height, lodging, and maturity in soybean. *Euphytica* **203**, 521–532 (2015).
20. Palomeque, L. *et al.* QTL in mega-environments: II. Agronomic trait QTL co-localized with seed yield QTL detected in a population derived from a cross of high-yielding adapted x high-yielding exotic soybean. *Theor. Appl. Genet.* **119**, 429–436 (2009).
21. Smith, R. R., Byth, D. E., Caldwell, B. E. & Weber, C. R. Phenotypic stability in soybean populations. *Crop Sci.* **7**, 590–592 (1967).
22. Lee, S. H. *et al.* Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Sci.* **36**, 728–735 (1996).
23. Mansur, L. M. *et al.* Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* **36**, 1327–1336 (1996).
24. Panthee, D. R., Pantalone, V. R., Saxton, A. M., West, D. R. & Sams, C. E. Quantitative trait loci for agronomic traits in soybean. *Plant Breed.* **126**, 51–57 (2007).
25. Reinprecht, Y. *et al.* Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. *Genome* 49, (1510–1527 2006).
26. Specht, J. E. *et al.* Soybean response to water: A QTL analysis of drought tolerance. *Crop Sci.* **41**, 493–509 (2001).
27. Johnson, H. W. & Bernard, R. L. Soybean genetics and breeding. *In* Norman, A. G. (ed.) The soybean. Academic Press. New York, 1–73 (1963).
28. Anand, S. C. & Torrie, J. H. Heritability of yield and other traits and interrelationships among traits in the F₃ and F4 generation of three soybean crosses. *Crop Sci.* **3**, 508–511 (1963).
29. Kwon, S. H. & Torrie, J. H. Heritability of and interrelationships among traits of two soybean populations. *Crop Sci.* **4**, 196–198 (1964).
30. Simpson, A. K. Jr. & Wilcox, J. R. Genetic and phenotypic associations of agronomic characteristics in four high protein soybean populations. *Crop Sci.* **23**, 1077–1081 (1983).
31. Kabelka, E. A. *et al.* Putative alleles for increased yield from soybean plant introductions. *Crop Sci.* **44**, 784–791 (2004).
32. Byth, D. E., Caldwell, B. E. & Weber, C. R. Specific and non-specific index selection in soybeans, *Glycine max* L. (Merrill). *Crop Sci.* 9, (702–705 1969).
33. Orf, J. H. *et al.* Genetics of soybean agronomic traits: I. comparison of three related recombinant inbred populations. *Crop Sci.* **39**, 1642–1651 (1999).
34. Kim, K. S. *et al.* Identification of positive yield QTL alleles from exotic soybean germplasm in two backcross populations. *Theor. Appl. Genet.* **125**, 1353–1369 (2012).
35. Lee, S. H. *et al.* Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. *Theor. Appl. Genet.* **92**, 516–523 (1996).
36. Wang, D., Graef, G. L., Procopiuk, A. M. & Diers, B. W. Identification of putative QTL that underlie yield in interspecific soybean backcross populations. *Theor. Appl. Genet.* **108**, 458–467 (2004).
37. Li, D. D., Pfeiffer, T. W. & Cornelius, P. L. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Sci.* **48**, 571–581 (2008).
38. Hyten, D. L. *et al.* A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. *Crop Sci.* **50**, 1–9 (2010).
39. Lander, E. & Kruglyak, L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241–247 (1995).
40. Keim, P., Diers, B. W., Olson, T. C. & Shoemaker, R. C. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* **126**, 735–742 (1990).
41. Brensha, W. *et al.* Genetic Analysis of root and shoot traits in the 'Essex' by 'Forrest' recombinant inbred line (RIL) population of soybean [*Glycine max* (L.) Merr.]. *J. Pl. Genom. Sci.* **1**, 1–9 (2012).
42. Zhang, D. *et al.* Detection of quantitative trait loci for phosphorus deficiency tolerance at soybean seedling stage. *Euphytica* **167**, 313–322 (2009).

43. Liang, H. *et al*. Inheritance and QTL mapping of related root traits in soybean at the seedling stage. *Theor. Appl. Genet.* **127**, 2127–2137 (2014).
44. Chen, Q. *et al*. QTL analysis of major agronomic traits in soybean. *Agr. Sci. China* **6**, 399–405 (2007).
45. Gai, J., Wang, Y., Wu, X. & Chen, S. A comparative study on segregation analysis and QTL mapping of quantitative traits in plants- with a case in soybean. *Front. Agr. China* **1**, 1–7 (2007).
46. Palmer, R. G. & Kilen, T. C. Quantitative Genetics. *In* Wilcox, J. R. (ed.), Soybeans: Improvement, Production, and Uses. 2nd Edition, American Society of Agronomy, Madison, Wisconsin, Ch. 5, 135–209 (1987).
47. Kearsey, M. J. & Hyne, V. QTL analysis: a simple marker regression approach. *Theor. Appl. Genet.* **89**, 698–702 (1994).
48. Jiang, C. & Zeng, Z. B. Multiple trait analysis and genetic mapping for quantitative trait loci. *Genetics* **140**, 1111–1127 (1995).
49. Darvasi, A. & Soller, M. A simple method to calculate resolving power and confidence interval of QTL map location. *Behav. Genet.* **27**, 125–132 (1997).
50. Darvasi, A., Vinreb, A., Minke, V., Weller, J. I. & Soller, M. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* **134**, 943–951 (1993).
51. Lande, R. & Thompson, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743–756 (1990).
52. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. Construction of genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314–331 (1980).
53. Mancuso, N. & Caviness, C. E. Association of selected plant traits with lodging of four determinate soybean cultivars. *Crop Sci.* **31**, 911–914 (1991).
54. Zhang, L. X., Kyei-Boahen, S., Zhang, J., Zhang, M. H. & Freeland, T. B. Modifications of optimum adaptation zones for soybean maturity groups in the USA. *Crop Manage.* **6**, 52–53 (2007).
55. Bernard, R. L. Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* **11**, 242–244 (1971).
56. Xia, Z. *et al*. Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1, which regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. USA* **109**, E2155–2164 (2012).
57. Owen, F. V. Inheritance studies in soybeans. II. Glabrousness, color of pubescence, time of maturity, and linkage relations. *Genetics* **12**, 519–529 (1927).
58. Sayama, T. *et al*. Mapping and comparison of quantitative trait loci for soybean branching phenotype in two locations. *Breed. Sci.* **60**, 380–389 (2010).
59. Watanabe, S. *et al*. Map-based cloning of the gene associated with the soybean maturity locus E3. *Genetics* **182**, 1251–1262 (2009).
60. Tsubokura, Y. *et al*. Natural variation in the genes responsible for maturity loci E1, E2, E3 and E4 in soybean. *Ann. Bot.* **113**, 429–441 (2014).
61. Dissanayaka, A. *et al*. Quantitative trait locus mapping of soybean maturity gene *E5*. *Breed. Sci.* **66**, 407–415 (2016).
62. Cober, E. R. & Voldeng, H. D. A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. *Crop Sci.* **41**, 698–701 (2001).
63. Liu, B. *et al*. Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics* **180**, 995–1007 (2008).
64. Weiss, M. G. & Stevenson, T. M. Registration of soybean varieties, V. *Agron. J.* **47**, 541–543 (1955).
65. Johnson, H. W. Registration of soybean varieties, VI. *Agron. J.* **50**, 690–691 (1958).
66. Akkaya, M. S., Shoemaker, R. C., Specht, J. E., Bhagwat, A. A. & Cregan, P. B. Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Sci.* **35**, 1439–1445 (1995).
67. Fan, J. B. *et al*. Highly parallel SNP genotyping. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 69–78 (2003).
68. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890 (2003).
69. Kosambi, D. D. The estimation of map distances from recombination values. *Ann. Eugen.* **12**, 172–175 (1944).
70. Hwang, S. *et al*. Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. *Theor. Appl. Genet.* **128**, 2047–2065 (2015).
71. Lander, E. S. *et al*. Mapmaker an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174–181 (1987).
72. Holshouser, D. Virginia soybeans: Lodging and it's effect on yield. Virginia Cooperative Extension, https://blogs.ext.vt.edu/ag-pest-advisory/lodging-and-its-effect-on-soybean-yield/ (2015).
73. Knapp, S. J., Stroup, W. W. & Ross, W. M. Exact confidence intervals for heritability on a progeny mean basis. *Crop Sci.* **25**, 192–194 (1985).
74. Weller, J. I. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics* **42**, 627–640 (1986).
75. Meng, X. L. & Rubin, D. B. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **80**, 267–278 (1993).
76. Zeng, Z. B. Precise mapping of quantitative trait loci. *Genetics* **136**, 1457–1468 (1994).
77. Wang, S., Basten, C. J. & Zeng, Z. B. Windows QTL Cartographer V. 2.5. Department of Statistics, North Carolina State University, https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm (2011).
78. Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 967–971 (1994).
79. Arcade, A., Labourdette, A., Falque, M. & Mangin, B. Biomercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* **20**, 2324–2326 (2004).
80. Veyrieras, J., Goffinet, B. & Charcosset, A. MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinformatics* **8**, 49–64 (2007).
81. Goffinet, B. & Gerber, S. Quantitative trait loci: a meta-analysis. *Genetics* **155**, 463–473 (2000).
82. Mansur, L. M., Lark, K. G., Kross, H. & Oliveria, A. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). *Theor. Appl. Genet.* **86**, 907–913 (1993).
83. Palomeque, L. *et al*. Validation of mega-environment universal and specific QTL associated with seed yield and agronomic traits in soybeans. *Theor. Appl. Genet.* **120**, 997–1003 (2010).
84. Rossi, M., Orf, J., Liu, L., Dong, Z. & Rajcan, I. Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian x Chinese crosses. *Theor. Appl. Genet.* **126**, 1809–1823 (2013).

Acknowledgements

Author Contributions

S.H. wrote the first draft of the main manuscript text and prepared figures. T.G.L. substantively revised the manuscript text and figures. All authors approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-42965-6.

Competing Interests: The authors declare no competing interests.

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