Theor Appl Genet (2010) 120:997–1003
DOI 10.1007/s00122-009-1227-7

ORIGINAL PAPER

Validation of mega-environment universal and specific QTL associated with seed yield and agronomic traits in soybeans

Laura Palomeque · Li-Jun Liu · Wenbin Li · Bradley R. Hedges · Elroy R. Cober · Mathew P. Smid · Lewis Lukens · Istvan Rajcan

Received: 18 June 2009 / Accepted: 18 November 2009 / Published online: 11 December 2009

Abstract The value of quantitative trait loci (QTL) is dependant on the strength of association with the traits of interest, allelic diversity at the QTL and the effect of the genetic background on the expression of the QTL. A number of recent studies have identified QTL associated with traits of interest that appear to be independent of the environment but dependant on the genetic background in which they are found. Therefore, the objective of this study was to validate universal and/or mega-environment-specific seed yield QTL that have been previously reported in an independent recombinant inbred line (RIL) population derived from the cross between an elite Chinese and Canadian parent. The population was evaluated at two field environments in China and in five environments in Canada in 2005 and 2006. Of the seven markers linked to seed yield QTL reported by our group in a previous study, four were polymorphic between the two parents. No association between seed yield and QTL was observed. The result could imply that seed yield QTL were either not stable in this particular genetic background or harboured different alleles than the ones in the original mapping population. QTL_U Satt162 was associated with several agronomic traits of which lodging was validated. Both the non-adapted and adapted parent contributed favourable alleles to the progeny. Therefore, plant introductions have been validated as a source of favourable alleles that could increase the genetic variability of the soybean germplasm pool and lead to further improvements in seed yield and other agronomic traits.

Communicated by D. Mather.

L. Palomeque · M. P. Smid · L. Lukens · I. Rajcan (✉)
Department of Plant Agriculture, Crop Science Building, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada
e-mail: irajcan@uoguelph.ca

L.-J. Liu
Soybean Research Institute of Heilongjiang, Harbin, Heilongjiang, People's Republic of China

W. Li
Soybean Research Institute, Northeast Agricultural University, Harbin, Heilongjiang, People's Republic of China

B. R. Hedges
Pioneer Hi-bred Ltd., A DuPont Company, Chatham, ON, Canada

E. R. Cober
Agriculture and Agrifood Canada, Eastern Cereal and Oilseed Crop Research Centre, Ottawa, ON, Canada

Present Address:
B. R. Hedges
Dow Agrosciences, Indianapolis, IN, USA

Introduction

Marker assisted selection (MAS) is an important plant breeding tool that could be used in the introduction of favourable alleles (Brummer et al. 1997) from plant introductions (PI), especially when the PI is poorly adapted to target environment (Guzman et al. 2007; Reyna and Sneller 2001; Smalley et al. 2004). With the aid of MAS the breeder can focus on introgressing only alleles for the desired quantitative trait loci (QTL) and avoid introgressing those that could have a negative impact on the trait of interest on the general background. The use of phenotypic selection, instead of MAS, tends to be more time consuming, greatly depends on the environment and requires

large populations and much field testing (Reyna and Sneller 2001).

Favourable alleles could be successfully introgressed by using MAS, if they are independent of the environment and the genetic background. Brummer et al. (1997) identified QTL for protein and oil content that were significant in several environments and therefore, were considered as stable QTL. These authors also found that 13.6% of the reported environmentally stable QTLs appeared in more than one population whereas the remainder were dependant on the genetic background (Brummer et al. 1997). QTL associated with traits of agronomic importance, such as seed yield, are generally more useful for the breeder only if they have been confirmed in several backgrounds (Fasoula et al. 2004). There are few reports of validated seed yield QTL in different environments and even fewer validating the reported QTL across diverse genetic backgrounds (Fasoula et al. 2004; Reyna and Sneller 2001). In a previous study, seven seed yield QTL were found in a population derived from a cross between a Canadian and a Chinese elite soybean line (Palomeque et al. 2009a). Two types of QTL were detected, those that were operating in both mega-environments (Canadian and Chinese), which were called universal QTL (QTL_U) and those found in only one mega-environment referred as specific QTL or QTL_{SP}. The objective of this study was to validate seven seed yield QTL (five QTL_U and two QTL_{SP}) and the corresponding co-located agronomic trait QTL reported by our group earlier (Palomeque et al. 2009a) using a population derived from a different cross between an elite Canadian and elite Chinese parent, which was evaluated in both the Canadian and Chinese mega-environments. A mega-environment as referred to in this study is described as "a broad, not necessarily contiguous area, occurring in more than one country and frequently transcontinental, defined by similar biotic and abiotic stresses, cropping system requirements, consumer preferences, and, for convenience, by volume of production" (Braun et al. 1996).

Materials and methods

Experimental design

A population consisting of 133 $F_{4:7}$ recombinant inbred lines (RIL) were derived from a cross between a Canadian and a Chinese soybean elite line. The Canadian parent, Pioneer 9071, is a high yielding cultivar adapted to the Canadian mega-environment bred by Pioneer Hi-bred Ltd., A DuPont Company. The Chinese parent, line # 8902, was developed in Heilongjiang province, in the P. R. China and is a high yielding elite cultivar which was provided by Dr. Gary Ablett at the University of Guelph. F_1 were produced by pollinating Pioneer 9071 with line # 8902. F_2 seeds were advanced by single seed decent (SSD) for one generation in the green house to produce the F_3 and in Los Andes, Chile to produce F_4 seeds. The F_4 seeds were planted in the field at the Woodstock Research Station in Woodstock, ON, Canada, where F_4 single plants were harvested individually and threshed separately. The $F_{4:5}$ lines were grown in Woodstock, ON, Canada and seed was advanced and increased over the next two generations to form $F_{4:7}$ lines, which were grown in yield trials in Canada and China as described below. Parental lines were included in field trials as well as 10 commercial cultivars from either country.

Field trials were arranged in a square lattice design with two replications in the Canadian and the Chinese mega-environments. Field trials were planted at different locations in two mega-environments (Canada and China) as follows. In 2005, two trials were grown in Canada, one in Woodstock (43° 7′ N 80° 45′ W) and one in Ottawa (45° 24′ N 75° 42′ W), both in Ontario. In 2006, two trials were planted in China in Harbin (45° 45′ N 126° 36′ E), in Heilongjiang province at two different sites, one grown by the Soybean Research Institute of the Heilongjiang Academy of Agricultural Sciences (Prof. Liu's group) and the other by the Soybean Research Institute of the Northeast University of Heilongjiang (Dr. Li's group). The two sites are geographically separated and represent different research stations and fields altogether. In Canada, two trials were grown in Ontario, one in Woodstock (43° 7′ N 80° 45′ W) and one in Ottawa (45° 24′ N 75° 42′ W). The plot size was 1.5 m × 5.5 m with four-rows planted at 35 cm between-row spacing in the Canadian mega-environment. In China, the plot sizes were of 2.8 m × 3 m in one site and 2.8 m × 5 m in the second site with four elevated ridges per plot, planted at 70 cm between ridges. Seed per m row length was the same in all locations within each mega-environment being at 18 seeds per meter in Canada and 28 seeds per meter in China.

DNA extraction and SSR markers

Tissue samples from both parental cultivars and from the lines derived from them were collected in 2005 from the trial planted in Woodstock, Canada. DNA extraction was described elsewhere (Palomeque et al. 2009a). For PCR reactions, the extracted DNA was diluted in a 1/100 proportion (1µl DNA in 99 µl deionized and distilled water) and stored as template DNA in a 96-well polypropylene plate (2 ml capacity/well) and kept at 4°C (Primomo et al. 2005). Reagents were combined in the following amounts in a 1.5 ml centrifuge tube: 10× PCR buffer, 1.5 µl; 50 mM $MgCl_2$, 1.5 µl; 5 U/µl Taq DNA polymerase, 0.4 µl (all three reagents from Sigma

JumpStart™); 3 mM dNTP, 1 µl; primer forward and reverse 2 µl each; deionized and distilled water 3.6 µl. After reagents were combined as a master mix, 12 µl of it, plus 3 µl of template DNA, were placed in each well in a 96-well PCR plate. The plate was sealed after 12 µl of mineral oil was placed per well. PCR reactions were performed in the 96-well RoboCycler® (Stratagene, La Jolla, CA, USA) with an amplification program which was previously described (Palomeque et al. 2009a, b). A Sunrise™ 96 Horizontal Electrophoresis Apparatus (Gibco BRL Life Technologies, Carlsbad, CA, USA) was used to separate the PCR products on a 5% MetaPhor® agarose gel (Bio Whittaker Molecular Applications, Rockland, ME, USA). UV light was used to score DNA bands that were stained with ethidium bromide.

Parental lines were screened with seven markers linked to seed yield QTL detected in a different genetic background (Palomeque et al. 2009a). Four out of the total seven markers evaluated (Satt100, Sat162, Satt194 and Satt277) were found to be polymorphic in the present material and were used to screen the RIL population.

For the analysis of parental relatedness 54 SSR markers were used to screen Pioneer 9071 and line # 8902, which were the parents of the population analysed in this study plus OAC Millennium and Heinong #38 which were the parental lines of the population evaluated elsewhere (Palomeque et al. 2009a).

Analysis of parent relatedness

Each SSR allele was scored as either shared (1) or polymorphic (0) for each of the four parental genotypes. Alleles which were amplified in one parent but not amplified in the second were excluded from the pairwise comparisons. A matrix of distances (D) was constructed, using the percentage of fragments that differed between each genotype, $D = \left[1 - \left(\frac{fragments\_shared}{fragments\_scored}\right)\right]$. The proportion of allelic differences between lines was used to calculate a binary, bifurcating tree using Neighbor Joining. Statistical support for groupings was estimated using 1,000 bootstrap resamples and the consensus tree is reported here. In this analysis, all SSR allelic changes were assumed to occur at equal frequency and each SSR has equal weight in inferring relatedness. All analyses were performed using MEGA (Tamura et al. 2007) and R (R Development Team 2006).

Phenotypic scoring

Entire field plots were machine harvested in the Canadian environments and hand-harvested and machine threshed in the Chinese environments. Seed yield (kg/ha) and 100 seed weight were adjusted to 13% moisture at each location and year. Twelve agronomic traits were recorded as follows: emergence score (on a scale 0–10, where 0 corresponded to no emergence and 10 to 100% emergence); R1 (flowering date: date when the first flower is observed), R5 (beginning seed stage: which corresponds to the date when a seed of 3 mm long is observed in a pod in the principal stem placed in any of the four superior pods where a completely developed leaf is observed); R8 (full maturity: date when 95% of the pods in a plot show maturity colour; Fehr 1971); lodging score (on a scale from 1 to 5, 1 corresponding to no lodging and 5–100% lodging); plant height, and oil and protein content. Oil and protein content was measured on a 200 g seed sample using the GrainSpec Near Infra-red Reflectance machine. From five plants randomly taken per plot, the number of pods per node, number of pods per plant and number of seeds per pod was recorded on a plant basis. A protocol was developed with a complete description on how to take each measurement and date to be followed by both the Canadian and the Chinese research groups involved in this study. The protocol was developed to minimize errors in phenotype estimations.

Statistical analysis

Since the interactions between genotype × year and genotype × location were found to be significant, mean were not averaged over years and locations. The variation of seed yield and traits within a test location were partitioned into the effects of genotype, replications and incomplete blocks within replications using the PROC MIXED procedure from SAS ver 9.1.3 (SAS Institute Inc. 2003) for a square lattice design (Bowley 1999). Blocks within replications were considered as random variables. Tests of residuals were performed with PROC UNIVARIATE and PROC PLOT procedure SAS ver 9.1.3 (SAS Institute Inc. 2003). Transformations were performed if assumptions were not met. The linear correlation between emergence and seed yield was evaluated by performing PROC CORR and PROC PLOT procedure from SAS ver 9.1.3 (SAS Institute Inc. 2003). Emergence was included as a covariate in the seed yield and trait models if a significant correlation was observed. No emergence data were recorded in the Chinese mega-environment in 2006. The linear correlation between traits and seed yield was evaluated by performing PROC CORR and PROC PLOT procedure from SAS ver 9.1.3 (SAS Institute Inc. 2003). Type I error rate (α) was set at 0.05 for the procedures mention above. A single factor ANOVA macro for SAS ver 9.1.3 (SAS Institute Inc. 2003) obtained from Dr. Elizabeth Lee at the University of Guelph (Guelph, ON, Canada) was performed to determine the association between phenotypic traits and markers. The type I error rate (α) was set at 0.01.

Results

Validation of seed yield QTL

Two elite lines, the Canadian cultivar Pioneer 9071 and the Chinese elite breeding line # 8902 were crossed to produce a population of 133 $F_{4:7}$ recombinant inbred lines (RIL), which was used to validate seed yield and agronomic trait QTL found in previous studies (Palomeque et al. 2009a, b). Four polymorphic markers were found between Pioneer 9071 and line # 8902 at the seven markers that identified seed yield QTL in a previous study (Palomeque et al. 2009a). Of the polymorphic markers, three were previously identified as tagging seed yield QTL_U and one tagging seed yield QTL_{SP} (Table 1). The polymorphic markers were distributed across two linkage groups, C1 and C2. In contrast to the earlier study, no significant association between seed yield and polymorphic markers Satt100, Sat162, Satt194 and Satt277 was found using single factor ANOVA (Table 1).

Validation of QTL associated with agronomic traits

One of the four polymorphic markers, Satt162 which was linked to a seed yield QTL_U reported by Palomeque et al. (2009a), tagged QTL associated with several agronomic traits in this study (Table 2). Satt162 was linked to a QTL associated with plant height and lodging at individual locations in the Canadian mega-environment, therefore behaving as a QTL_{SP}. In the Chinese mega-environment, the marker Satt162 was tagging a QTL that was associated with number of seeds per pod, R5 and R8, behaving also as a QTL_{SP}. The amount of variation per trait explained by Satt162 ranged between 7 and 9%. The parent that contributed the favourable allele in the environments where Satt162 was associated with an agronomic trait was the Canadian one, Pioneer 9071, except for lodging where the Chinese line 8902 was the donor of the favourable allele, i.e., lower lodging score (Table 2). Satt162 has been previously reported as being linked to a QTL associated with seed yield and described as a QTL_U (Palomeque et al. 2009a). It was also reported as linked to QTL associated with 100 seed weight, number of pods per node, lodging and oil content (Palomeque et al. 2009b). The authors reported that the QTL at marker Satt162 was associated with oil and seed yield, with the favourable allele coming from the Canadian parent. For the rest of the traits, the Chinese parent was the donor of the favourable allele (data not shown).

Analysis of parent relatedness

Among the six pair wise comparisons between the four parents, the highest proportion of shared markers was between the Canadian parent Pioneer 9071 and the Chinese line 8902 (56/83, Table 3). The lowest marker identity was between the Canadian parental line OAC Millennium and Chinese line 8902 (43/87, Table 3). The four remaining pairwise comparisons had 54–55% marker identity (Table 3). A clustering of OAC Millennium with Heinong #38 and Pioneer 9071 with 8902 was well supported in a bootstrap analysis (89%; Fig. 1).

Agronomic traits and their correlation with seed yield

Seed yield was correlated with several of the 11 agronomic traits analysed (Table 4). Four traits (plant height, R1, R5 and R8) were correlated with seed yield in both mega-environments. Plant height and R8 were correlated with seed yield in 50% of all the tests conducted in either mega-environment. However, R1 and R5 were correlated positively in 67% of the tests. One hundred seed weight was positively correlated with seed yield in 50% of the Canadian mega-environments. No correlation was found

Table 1 SSR markers tagging yield QTL reported for the population Heinong 38 × OAC Millennium (Palomeque et al. 2009a) in the validation population derived from the cross between a Canadian and a Chinese elite soybean line, Pioneer 9071 × Line #8902, respectively

Marker	Linkage group	Type of previously reported QTL	Marker polymorphism	Seed yield QTL (*P* value[a])					
				W05	Ott05	W06	Ott06	H(1) 06	H(2) 06
Satt100	C2	QTL_U	Polymorphic	0.7782	0.5613	0.9922	0.4803	0.4087	0.7453
Satt139	C1	QTL_U	Monomorphic	–	–	–	–	–	–
Satt162	I	QTL_U	Polymorphic	0.0453	0.2371	0.5391	0.1853	0.4917	0.0895
Satt194	C1	QTL_{SP}	Polymorphic	0.2583	0.7904	0.1503	0.7355	0.7032	0.8721
Satt259	O	QTL_{SP}	Monomorphic	–	–	–	–	–	–
Satt277	C2	QTL_U	Polymorphic	0.4084	0.4142	0.7957	0.6153	0.278	0.1465
Sat_126	K	QTL_U	Monomorphic	–	–	–	–	–	–

[a] *P* values over 0.01 corresponds to no significant association between the marker and seed yield

W05, Woodstock 2005; Ott05, Ottawa 2005; W06, Woodstock 2006; Ott06, Ottawa 2006; H(1) 06, Harbin 1 2006; H(2) 06, Harbin 2 2006

Table 2 Association of SSR marker Sat162 previously reported as tagging a seed yield QTL (Palomeque et al. 2009a) with different agronomic traits in a validation population derived from the cross between elite Canadian and Chinese lines, Pioneer 9071 × Line # 8902, respectively

Trait	Marker											
	Satt162											
	Canadian mega-environment								Chinese mega-environment			
	Ottawa 2005				Ottawa 2006				Harbin(1)2006			
	R^{2a}	P	Allelic mean		R^2	P	Allelic mean		R^2	P	Allelic mean	
			Canadian	Chinese			Canadian	Chinese			Canadian	Chinese
Plant height	7.1	0.009	92 ± 0.9	89 ± 0.9								
Lodging					9.1	0.0028	3.6 ± 0.14	3.1 ± 0.15				
No. of seeds/pod									8.3	0.0039	2.7 ± 0.03	2.5 ± 0.03
R5									7.5	0.0093	87.5 ± 0.89	84.1 ± 0.96
R8									7.2	0.0094	121.1 ± 0.8	117.9 ± 0.84

[a] Amount of phenotypic variation in the agronomic trait explained by the QTL

Table 3 Analysis of parental relatedness in Phylip, calculated using the Jukes and Cantor analysis

Line	Heinong # 38	Line # 8902	OAC Millennium	Pioneer 9071
Heinong #38	0.000000	0.695990	0.695990	0.721528
Line # 8902	0.695990	0.000000	0.841402	0.426520
OAC Millennium	0.695990	0.841402	0.000000	0.693194
Pioneer 9071	0.721528	0.426520	0.693194	0.000000

The lower the number, the more related the genotypes are

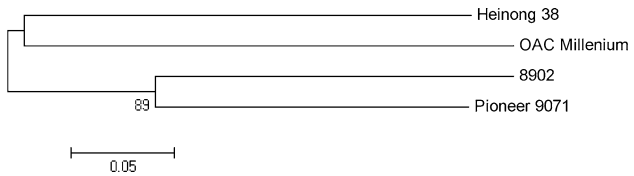

Fig. 1 Tree representation of genetic relatedness. Parental line codes: Pioneer 9071; 8902 = line # 8902; OAC Millennium; Heinong 38

between seed weight and seed yield in any test in China. Number of pods per plant was positively correlated with seed yield in only one location in the Chinese mega-environment. Since no significant difference for number of pods per plant was found between RIL, no linear correlation could be performed between this trait and seed yield in the Canadian mega-environment. Number of pods per node was positively correlated with seed yield in one test in the Canadian mega-environment but negatively correlated in a test in China. Lodging was found to be negatively correlated with seed yield in the Canadian mega-environment as well as oil content in the Chinese mega-environment. In the Canadian mega-environment, oil content was positively correlated with seed yield in only one environment. No significant correlation was found between seed yield and number of seeds per pod in either mega-environment. Protein content was negatively correlated with seed yield in only one environment within the Canadian mega-environment (Table 4).

Discussion

The seed yield QTL linked to Satt100, Satt139, Satt162, Satt194, Satt277, Sat_126 and Satt259, which were identified in the RIL population derived from the cross OAC Millennium × Heinong #38 (Palomeque et al. 2009a), were not validated in the RIL population derived from the cross between Pioneer 9071 and # 8902. The lack of validation in a different genetic background could imply that these seven QTL were not stable even though the parental lines chosen for both populations were elite Canadian and Chinese lines. However, one of the seven QTL evaluated in this study (linked to Satt277) was previously reported as being associated with seed yield in diverse genetic backgrounds and environments by other researchers (Guzman et al. 2007; Orf et al. 1999a, b; Smalley et al. 2004; Specht et al. 2001 and Palomeque et al. 2009a). As mentioned above, epistatic effects could be considered as one of the factors leading to the lack of validation of the QTL effect in the RIL population derived from the cross Pioneer 9071 × line # 8902. Another

Table 4 Correlation analysis between seed yield and 11 agronomic traits found in a RIL population derived from the cross between Canadian and Chinese elite soybean lines, Pioneer 9071 × line # 8902, respectively

Location	R1	R5	R8	100 seed weight	Plant height	Lodging	Pod per plant	Pod per node	Seed per pod	Oil	Protein
W05	$r = 0.25$	$r = 0.18$	$r = 0.38$	$r = 0.59$	$r = 0.41$	ns	No difference[a]	$r = 0.22$	ns	ns	ns
	$P = 0.0047$	$P = 0.0046$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$			$P = 0.013$			
Ott05	$r = 0.41$	$r = 0.42$	$r = 0.58$	$r = 0.64$	$r = 0.30$	$r = -0.19$	Data not recorded	Data not recorde	Data not recorded	ns	ns
	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P = 0.0005$	$P = 0.0308$		d			
W06	ns	$r = 0.36$	$r = 0.37$	Data not	$r = 0.64$	No data	No difference	ns	No difference	No data	No data
		$P < 0.0001$	$P < 0.0001$	Recorded	$P < 0.0001$						
Ott06	$r = -0.21$	$r = -0.34$	ns	ns	ns	$r = -0.36$	Data not recorded	Data not recorded	Data not recorded	$r = 0.27$	$r = -0.20$
	$P = 0.0165$	$P < 0.0001$				$P < 0.0001$				$P = 0.0001$	$P = 0.024$
H(1) 06	$r = 0.18$	ns	$r = 0.36$	ns	$r = 0.44$	No data	$r = 0.33$	$r = -0.22$	ns	ns	ns
	$P = 0.0343$		$P < 0.0001$		$P < 0.0001$		$P = 0.0096$	$P = 0.0096$			
H(2) 06	$r = 0.48$	$r = 0.61$	$r = 0.54$	ns	Data not recorded	Data not recorded	Data not recorded	Data not recorded	Data not recorded	$r = -0.25$	ns
	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$							$P = 0.0035$	

[a] Correlation not calculated since no significant difference for the trait among lines was observed, according to ANOVA

ns, Not significant; no data, data could not be transformed; W05, Woodstock 2005; Ott05, Ottawa 2005; W06, Woodstock 2006; Ott06, Ottawa 2006; H(1) 06, Harbin 1 2006; H(2) 06, Harbin 2 2006

possibility could be that marker loci were fixed in the validation population with parents carrying the same allele or alleles that were different from those in Heinong 38 × OAC Millennium population.

When several agronomic traits were analysed, a QTL_{SP} tagged by Satt162, was observed to be associated with different traits other than seed yield. Satt162 was previously reported as tagging a QTL_{SP} associated with lodging (Palomeque et al. 2009b) and was validated in the current study as a QTL_{SP} associated with the same trait at the same mega-environment. Yield increase could be the result of the RILs with low lodging since both traits, seed yield and lodging, were observed to be negatively correlated. The validation of only mega-environment QTL_{SP} could imply that a breeder could only include QTL_{SP} as a breeding strategy, unless QTL_U were not validated due to the genetic background of the parental lines.

Both the environment and genetic background play an important role in the success of the use and introgression of favourable alleles linked to QTL into adapted germplasm by breeders. QTL for a specific trait are not always stable across environments and/or genetic backgrounds, therefore, their breeding value depends on the strength and stability of trait associations. Brummer et al. (1997) confirmed a QTL associated with protein content that had been previously reported in a different genetic background. The authors suggested that the protein QTL appeared in two different genetic backgrounds due to the fact that both populations were derived from the cross between *Glycine max* × *G. soja* and therefore, the parent contributing the favourable allele could have been *Glycine soja* in both cases (Brummer et al. 1997). When yield QTL were evaluated in different genetic backgrounds a variety of results were obtained. One example is a QTL linked to Satt277, which was associated with yield by several research groups using populations derived from different genetic backgrounds evaluated in distinct environments (Guzman et al. 2007; Orf et al. 1999a, b; Smalley et al. 2004; Specht et al. 2001; Palomeque et al. (2009a). On other hand, Reyna and Sneller (2001) found that three QTL that were previously reported as being associated with yield (Orf et al. 1999b) could not be confirmed across several environments and genetic background. These authors evaluated crosses that included one parent in common (Archer) that was considered as an adapted parent in the study done by Orf et al. (1999a) and as non-adapted in the study done by Reyna and Sneller (2001) since the elite parental lines were southern parents. Reyna and Sneller (2001) suggested that the difference between the research findings could be due to the favourable allele in Orf et al. (1999b) coming from Archer, the elite parental line, which happened to be the non-adapted parent in the alternative population evaluated by Reyna and Sneller (2001).

The authors also suggested that the lack of confirmation of the seed yield QTL could be due to possible population-specific epistatic effects (Reyna and Sneller 2001).

The variability observed between RILs is the result of the allele variation observed between the parental lines, i.e., the RIL population variability depends on how genetically diverse the parents were. In this study, the parents of the population derived from OAC Millennium × Heinong #38 had more variability than the parents of the validation population, derived from the cross Pioneer 9071 × line # 8902. Since the genetic distance between OAC Millennium and Heinong #38 was greater than that observed between line # 8902 and Pioneer 9071, we propose that reduced allele variability between the parents of the validation population could be one of the reasons for the lack of effect of seed yield QTL that were identified in the original population. To test this hypothesis, a new population derived from a cross between OAC Millennium and line # 8902, as the most diverse parent combination, could be used to validate the yield QTL from the OAC Millennium × Heinong #38 population. The latter population would have the advantage of having one parent (OAC Millennium) in common to the original mapping population. It may also be of interest to determine how genetically related were the parental lines of other populations involving exotic × adapted parent crosses (e.g., Orf et al. 1999b), compared to the two populations that we reported. Further studies involving both the OAC Millennium × Heinong #38 and the 'line #8902' × 'Pioneer 9071' RIL populations are being evaluated in additional locations in the Canadian and Chinese mega-environments and a potentially new mega-environment in Minnesota and North Dakota. We anticipate finding additional QTL_U and QTL_{SP} that could be cross-tested and validated between these two populations with the aim of developing new strategies for introgressing novel QTL_U and QTL_{SP} into soybean breeding programs.

Acknowledgments We thank Dr. H. R. Boerma (Univ. of Georgia), S. J. Bowley, K. P. Pauls, G. R. Ablett and D. Falk (Univ. of Guelph) for valuable suggestions on the manuscript. Excellent technical assistance was provided by Wade Montminy, Julia Zilka, Yesenia Salazar, Lin Liao, Chris Grainger, Ron Guillemette, Fernando Pegoraro and technical staff in China. Funding support from the National Science and Research Council of Canada (NSERC), Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and in-kind support from Pioneer Hi-bred International, a Dupont company is gratefully acknowledged

References

Bowley S (1999) A hitchhiker's guide to statistics in plant biology. Any old subject books edn. Plant et al. Inc, Guelph, p 250

Braun H, Rajaram S, van Ginkel M (1996) CIMMYT's approach to breeding for wide adaptation. Euphytica 92:175–183

Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC (1997) Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci 37:370–378

Fasoula VA, Harris DK, Boerma HR (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

Fehr WR (1971) Stage of development descriptions for soybeans, *Glycine max* (L) Merrill. Crop Sci 11:929–933

Guzman PS, Diers BW, Neece DJ, Martin SKS, Leroy AR, Grau CR, Hughes TJ, Nelson RL (2007) QTL associated with yield in three backcross-derived populations of soybean. Crop Sci 47:111–122

Orf JH, Chase K, Adler FR, Mansur LM, Lark KG (1999a) Genetics of soybean agronomic traits: II. Interactions between yield quantitative trait loci in soybean. Crop Sci 39:1652–1657

Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG (1999b) Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. Crop Sci 39:1642–1651

Palomeque LC, Jun LL, Li W, Hedges B, Cober ER, Rajcan I (2009a) QTL in mega-environments: I. Universal and specific seed yield QTL detected in a population derived from a cross of high-yielding adapted × high-yielding exotic soybean lines. Theor Appl Genet 119:417–427

Palomeque LC, Jun LL, Li W, Hedges B, Cober ER, Rajcan I (2009b) 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 × high-yielding exotic soybean lines. Theor Appl Genet 119:429–436

Primomo VS, Poysa V, Ablett GR, Jackson CJ, Gijzen M, Rajcan I (2005) Mapping QTL for individual and total isoflavone content in soybean seeds. Crop Sci 45:2454–2464

Reyna N, Sneller CH (2001) Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. Crop Sci 41:1317–1321

SAS Institute Inc. (2003) SAS ver. 9.3. SAS Institute Inc., Cary

Smalley MD, Fehr WR, Cianzio SR, Han F, Sebastian SA, Streit LG (2004) Quantitative trait loci for soybean seed yield in elite and plant introduction germplasm. Crop Sci 44:436–442

Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG (2001) Soybean response to water: a QTL analysis of drought tolerance. Crop Sci 41:493–509

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

R Development Core Team (2006) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL:http://www.R-project.org