Theor Appl Genet (2013) 126:1809–1823
DOI 10.1007/s00122-013-2094-9

ORIGINAL PAPER

Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian × Chinese crosses

M. Eugenia Rossi · James H. Orf · Li-Jun Liu · Zhimin Dong · Istvan Rajcan

Received: 27 August 2012 / Accepted: 26 March 2013 / Published online: 18 April 2013

Abstract One of the goals of plant breeding is to increase yield with improved quality characters. Plant introductions (PI) are a rich source of favorable alleles that could improve different characters in modern soybean [*Glycine max* (L.) Merril] including yield. The objectives of this study were to identify yield QTL underlying the genetic basis for differential adaptation of soybeans to the Canadian, United States or Chinese mega-environments (ME) and to evaluate the relationship and colocalization between yield and agronomic traits QTL. Two crosses between high-yielding Canadian cultivars and elite Chinese cultivars, OAC Millennium × Heinong 38 and Pioneer 9071 × #8902, were used to develop two recombinant inbred line (RIL) populations. Both populations were evaluated at different locations in Ontario, Canada; Minnesota, United States (US), Heilongjiang and Jilin, China, in 2009 and 2010. Significant variation for yield was observed among the RILs of both populations across the three hypothetical ME. Two yield QTL (linked to the interval Satt364–Satt591 and Satt277) and one yield QTL (linked to marker Sat_341) were identified by single-factor ANOVA and interval mapping across all ME in populations 1 and 2, respectively. The most frequent top ten high-yielding lines across all ME carried most of the high-yielding alleles of the QTL that were identified in two and three ME. Both parents contributed favorable alleles, which suggests that not only the adapted parent but also the PI parents are potential sources of beneficial alleles in reciprocal environments. Other QTL were detected also at two and one ME. Most of the yield QTL were co-localized with a QTL associated with an agronomic trait in one, two, or three ME in just one or in both populations. Results suggested that most of the variation observed in seed yield can be explained by the variation of different agronomic traits such a maturity, lodging and height. Novel alleles coming from PI can favorably contribute, directly or indirectly, to seed yield and the utilization of QTL detected across one, two or three ME would facilitate the new allele introgression into breeding populations in both North America and China.

Communicated by D. Lightfoot.

M. E. Rossi · 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

J. H. Orf
Department of Agronomy and Plant Genetics, University of Minnesota, 411 Borlaug Hall, 1991 Buford Circle, St. Paul, MN 55108, USA

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

Z. Dong
Jilin Academy of Agricultural Sciences, Changchun, Jilin, People's Republic of China

Z. Dong
Soybean Research Institute of the Northeast Agricultural University, Harbin, Heilongjiang, People's Republic of China

Introduction

From the adaptation perspective, plants with higher yield quantity and quality are better adapted to certain environments (Cooper and Byth 1996). As a result, one of the most important traits to be improved by plant breeders is seed yield with increased resistance to biotic and abiotic

stresses. The variation observed in yield could be due to the variation in different agronomic traits that are correlated with yield.

To combat the potential problem of limited genetic gains in soybean breeding programs, the introgression of new sources of germplasm appears to be a solution. This may increase genetic variability allowing better genetic gains from selection (Ude et al. 2003). Several studies demonstrated that plant introductions (PI) or exotic germplasm, such as modern Asian cultivars, which are genetically different from North American germplasm, represent an important reservoir of favorable alleles that can be used in breeding programs (Orf et al. 1999b; Li et al. 2001a, b; Concibido et al. 2003).

The use of molecular markers has facilitated the identification and localization of quantitative traits loci from exotic germplasm. Reports by Thompson and Nelson (1998), Orf et al. (1999a), Concibido et al. (2003), Ude et al. (2003), Kabelka et al. (2004), Smalley et al. (2004), and Wang et al. (2004) demonstrated that exotic genes can be successfully introduced from wild species for yield enhancement through genetic mapping and molecular markers in a number of crop species.

As an example of the use of exotic germplasm and molecular tools to improve yield and yield components, Orf et al. (1999a, b) compared two recombinant inbred line (RIL) population derived from 'Minsoy' and 'Archer', and 'Noir 1' and Archer. Archer is an elite cultivar from North America (Cianzio et al. 1991) while Minsoy and Noir 1 are PIs (Mansur et al. 1996; Orf et al. 1999a, b). QTL were identified for plant height, lodging, days to flowering and maturity, seed yield, seed weight, and seed oil and protein content. Many of the QTL identified in this study have the beneficial alleles provided by the exotic parents.

In another study, Kabelka et al. (2004) found yield QTL that were associated with significant changes in days to maturity, in plant height and in seed protein concentration in a population derived from a PI and an adapted cultivar. Using a population derived from Canadian × Chinese elite lines, Palomeque et al. (2009a, b) found four yield QTL (tagged with markers Satt100, Satt277, Satt162 and Sat_126) that were detected across the Canadian and Chinese mega-environments. Those QTL were co-localized with QTL associated with different agronomic traits such as 100 seed weight, plant height, days to maturity and oil content. However, none of these seed yield QTL were confirmed in a second population derived from a cross between a different Canadian and Chinese elite soybean line (Pioneer 9071 × #8902) (Palomeque et al. 2010).

Quantitative traits, such as yield, are strongly influenced by environmental factors. As a result, when adaptation is studied, the performance of different genotypes is evaluated across different environments. Mega-environments have been defined by the International Maize and Wheat Improvement Center (CIMMYT) as "a broad, not necessarily contiguous geographic area, occurring in more than one country and frequently transcontinental, defined by similar biotic and abiotic stresses, cropping system requirements, consumer preferences and volume of production" (Braun et al. 1996).

Therefore, the objectives of this study were: (1) to identify QTL for high yield in two populations derived from crosses between Canadian and Chinese elite soybean lines (OAC Millennium × Heinong 38 and Pioneer 9071 × #8902) tested at three hypothetical mega-environments: Ontario, Canada; Minnesota, United States and northeast China, (2) to evaluate the relationship between yield and yield components and seed quality parameters and (3) to determine if any seed yield QTL present in two or more mega-environments is co-localized with any agronomic trait QTL across the same three mega-environments.

Materials and methods

Field experiments

Two populations developed from crosses between Canadian and Chinese soybean lines were used in this experiment. Population 1 consisted of 92 $F_{4:7}$ RILs derived from the cross OAC Millennium × Heinong #38. Population 2 was comprised of 131 lines at $F_{4:7}$ stage originated from the cross between Pioneer 9071 and #8902. OAC Millennium and Pioneer 9071 are high-yielding cultivars well adapted to northern US and Canada (MG = 0 and 0.7, respectively), which were developed by the University of Guelph and Pioneer Hi-Bred, a DuPont Company, respectively. Heinong #38 and #8902 are well-established elite cultivars developed and widely used in the Heilongjiang province of the People's Republic of China. Both parents of each cross as well as commercial varieties from either environment were used as controls in field trials. Because of their elite nature, neither parent in both populations was considered a completed non-adapted PI.

Both populations were grown in three hypothetical mega-environments: Ontario, Canada; Minnesota, United States (US), and northeast China at different locations (Table 1). The parental line Pioneer 9071 was not evaluated in Woodstock in 2009. In Canada and United States, the experimental design was an RCBD Nearest Neighbor with two replications while in China the experiment was arranged as a rectangular lattice design with two replications. The plot size was 1.5 × 5.5 m in the Canadian sites and 1.5 × 2.4 m in the US sites, with four rows and 35 cm

between-row spacing in all sites. For China, the plot size was 2.8 × 3 m with four elevated ridges per plot planted at 70 cm between ridges in Harbin, while in Changchun plots were 2.4 × 3 m with four rows separated by 60 cm between rows. Number of seeds per plot was adjusted based on germination rate. Seed density was the same at all locations within each mega-environment being 18 seeds per meter in Canada and the US, whereas in China it was 28 seeds per meter. Similar plant stands were achieved at all locations. The number and diversity of the environments used in this study along with different resources used by different collaborators added complexity that we could not completely control. Plots were machine harvested in the Canadian and US environments while they were hand-harvested and machine-threshed in the Chinese environments.

Phenotypic scoring

Seed yield (kg/ha) was recorded per plot at each location and year and adjusted to 13 % moisture. Other yield-related traits were recorded per plot at each location and year: days to maturity, as the number of days from planting to the stage when 95 % of the pods show maturity color (R8) (Fehr et al. 1971); lodging score (1 for no lodging to 5 for completely prostrate); plant height (in centimeters) and 100 seed weight (in grams). Each measurement was performed following the protocol detailed by Palomeque (2007). Percentages of oil and protein were analyzed per plot using a portable grain analyzer ZX50 (Zeltex Inc., Hagerstown, MD, USA).

DNA extraction and PCR reaction

Genotypic data from population 1, OAC Millennium × Heinong #38, generated by Palomeque et al. (2009a) was used in this experiment. For population 2, Pioneer 9071 × #8902, tissue samples were collected from parents and RILs in Elora, ON, Canada during 2009, with the aid of a leaf puncher. Leaf disks were collected in pre-labeled 2.0 ml centrifuge tubes, transported on ice and freeze dried with a Labonco FreeZone® freeze dry system (Savant Modulyo, Kansas City, MO, USA) to remove moisture and inactivate endogenous nucleases and then stored at −80 °C.

Genomic DNA was obtained from 10 to 15 stored leaf disks per line previously ground by FastPrep FP 120 (Savant Instruments Inc., Holbrook, NY, USA). GenElute™ Plant Genomic DNA Kit (SIGMA®, Saint Louis, MO, USA) was used for DNA extraction. A Nanodrop ND1000 Spectrophotometer (Nanodrop Technologies Inc., Wilmington, DE, USA) was utilized to check the concentration and quality of each extraction.

For PCR reactions, the extracted DNA was diluted in a 1/100 proportion and stored as template DNA in a 96-well polypropylene plate at 4 °C. Reagents were combined in the following amounts in a 1.5 ml centrifuge tube: 1.5 µl 10× PCR buffer, 1 µl 3 mM dNTPs, 1.5 µl 50 mM $MgCl_2^{Sigma}$, 2 µl primer forward and 2 µl primer reverse, 0.4 µl 2.5 U/µl Taq^{Sigma}, 3 µl 10 ng/µl gDNA and 3.6 µl of deionized and distilled water. In a 96-well PCR plate, 12 µl of master mix (combination of reagents previously described) and 3 µl of template DNA were placed in each

Table 1 Location of field experiments in three hypothetically different mega-environments during 2009 and 2010 using two Canadian by Chinese soybean populations

Year	Province/state country	Populations	
		OAC Millennium × Heinong 38 Locations[a]	Pioneer 9071 × #8902 Locations[a]
2009	Ontario, Canada	Elora	Elora
		St. Pauls	Woodstock
	Minnesota, USA	Lamberton	Lamberton
		Rosemount	Rosemount
2010	Ontario, Canada	Elora	Elora
		St. Pauls	St. Pauls
			Woodstock
	Minnesota, USA	Rosemount	Lamberton
			Rosemount
	Heilongjiang, China	Harbin	Harbin
	Jilin, China	Changchun	Changchun

[a] Elora (43.641044N; 80.405674W), Woodstock (43.18065N; 80.77995W) and St. Pauls (43.16816N; 81.01497W), Lamberton 2009 (44°13′56.63N; 95°18′09.73W), Lamberton 2010 (44°14′22.43N; 95°18′18.43W), Rosemount 2009 (44°43′41.76N; 93°06′06.53W), Rosemount 2010 (44°43′04.30N; 93°06′25.18W), Harbin (45°45′N; 126°36′E), Changchun (43°19′29.97N; 93°06′06.53E)

Table 2 Least square means and standard error for seed yield of OAC Millennium (Canadian parent) and Heinong 38 (Chinese parent), the checks, the recombinant inbred lines (RIL) derived from them and the highest-yielding RIL

Macro-environment		Seed yield mean ± SE (kg/ha)				
Location	Year	Parent		Checks	RIL population	
		OAC Millennium	Heinong 38		Mean	Maximum
Canada						
Elora	2009	2,351 ± 26.6	2,131 ± 75.3	2,741 ± 90.8	2,281 ± 102.7	3,684 ± 218.8
	2010	3,841 ± 116.8	3,929 ± 252.8	4,329 ± 221.4	3,387 ± 203.5	4,539 ± 321.0
St. Pauls	2009	2,524 ± 290.9	2,569 ± 114.4	2,755 ± 254.8	2,386 ± 259.5	4,062 ± 114.9
	2010	4,368 ± 203.9	4,268 ± 232.4	4,789 ± 150.5	3,847 ± 134.1	5,120 ± 115.9
United States						
Lamberton	2009	2,460 ± 78.3	2,655 ± 18.8	2,612 ± 155.7	2,281 ± 172.3	3,542 ± 190.3
Rosemount	2009	2,208 ± 53.1	2,245 ± 387.4	2,308 ± 143.6	1,883 ± 159.2	3,230 ± 315.2
	2010	2,177 ± 57.1	3,067 ± 88.9	2,994 ± 176.4	2,487 ± 133.5	3,499 ± 269.7
China						
Harbin	2010	1,917 ± 389.9	1,625 ± 82.1	1,483 ± 224.9	1,441 ± 171.1	1,882 ± 96.9
Changchun	2010	2,764 ± 78.4	2,694 ± 70.9	2,372 ± 277.4	2,453 ± 173.1	3,206 ± 72.8

The experiments were grown across three different mega-environments: Ontario, Canada; Minnesota, United States; and Heilongjiang and Jilin, China during 2009 and 2010

Table 3 Least square means and standard errors for seed yield of Pioneer 9071 (Canadian parent) and #8902 (Chinese parent), the checks, the recombinant inbred lines (RIL) derived from them and the highest-yielding RIL

Macro-environment		Seed yield mean ± SE (kg/ha)				
Location	Year	Parent		Checks	RIL population	
		Pioneer 9071	#8902		Mean	Maximum
Canada						
Elora	2009	2,596 ± 203.8	2,859 ± 30.3	2,837 ± 99.1	2,009 ± 124.3	3,349 ± 127.9
	2010	4,003 ± 12.5	2,593 ± 242.1	3,807 ± 125.9	3,148 ± 137.9	3,950 ± 226.2
Woodstock	2009	–	2,018 ± 42.8	1,965 ± 104.3	2,017 ± 165.5	3,318 ± 192.4
	2010	4,662 ± 98.9	3,556 ± 77.9	4,391 ± 378.6	4,100 ± 152.1	5,893 ± 683.4
St. Pauls	2010	3,999 ± 508.9	4,154 ± 99.9	4,155 ± 209.7	4,181 ± 308.6	5,031 ± 74.7
United States						
Lamberton	2009	2,408 ± 117.5	2,603 ± 262.1	2,347 ± 180.3	2,272 ± 178.2	3,272 ± 289.2
	2010	3,327 ± 237.6	3,004 ± 140.2	2,877 ± 151.9	3,097 ± 128.3	4,201 ± 62.4
Rosemount	2009	1,666 ± 140.2	1,816 ± 181.1	1,725 ± 55.6	1,814 ± 117.4	2,755 ± 128.6
	2010	2,597 ± 27.5	1,911 ± 221.6	2,674 ± 88.4	2,470 ± 138.4	3,374 ± 348.3
China						
Harbin	2010	2,066 ± 117.2	2,125 ± 69.6	1,602 ± 168.7	1,870 ± 198.4	2,435 ± 69.6
Changchun	2010	2,537 ± 165.2	1,737 ± 86.9	2,460 ± 158.6	2,221 ± 157.2	3,271 ± 84.5

The experiments were grown across three different mega-environments: Ontario, Canada; Minnesota, United States and Heilongjiang and Jilin, China during 2009 and 2010

wall and spun in a centrifuge (Hermle Z180M, Labnet, Edison, NJ, USA). Finally, 12 µl of mineral oil was added on each plate to prevent evaporation.

PCR reactions were performed in the 96-well RoboCycler® (Stratagene, La Jolla, CA, USA) with an amplification program consisting of 2 min at 95 °C, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 47 °C for 1.5 min, and extension at 68 °C for 1.5 min. A final extension step of 72 °C for 5 min was set and the finished reaction mixture was held at 4 °C. Amplification products were separated by electrophoresis on a 4.5 % Metaphor® high-resolution agarose gel (Bio Whittaker Molecular Applications, Rockland, ME, USA) using a Sunrise TM 96 Horizontal Electrophoresis Apparatus

(Gibco BRL, Life technologies, Carlsbad, CA, USA) with 115–130 MA/V of current supplied by an EC105 Electrophoresis Power Supply (ThermoEC, Holbrook, NY, USA). DNA bands were visualized under UV light after staining with ethidium bromide.

The University of Guelph SSR library available consisted of 450 SSR primer pairs selected from the integrated soybean genetic map (Cregan et al. 1999), which covered approximately equidistantly the 20 soybean genetic linkage groups. Both parental lines of the Pioneer 9071 × #8902 population (population 2) were screened with the available SSR markers. One hundred-and-two of the available 450 markers that were polymorphic between the parents were used to genotype the entire RIL population. The polymorphic markers were used to screen the entire RIL population.

Linkage mapping and QTL analysis

The linkage map from OAC Millennium × Heinong 38 population created by Palomeque et al. (2009a) was used in this study for QTL analysis. In that map 26 linkage groups were obtained covering 17 chromosomes in total (Palomeque et al. 2009a). A new map was obtained for population 2, using JoinMap 4® program (van Ooijen 2004a). Segregation distortion was calculated to determine the departure from the expected 1:1 radio. Since segregation distortion can be considered as a normal phenomenon in wide crosses such as the ones used in this study, only the markers that showed extreme segregation distortion (10:1 or greater ration between alleles in the RILs) were not used in the analysis. The parameter that was used to group markers into LGs was recombination frequency. The mapping algorithm was regression and the mapping function selected was Kosambi's (van Ooijen 2004a). Nineteen linkage groups were generated for population 2. The mapping order and distances among markers were similar to the Soybean Composite Map 2003 (Soybase 2011). Marker distance ranged between 1.2 (Ch. 17) and 38.1 (Ch. 7) cM.

Yield QTL were detected using a single-factor ANOVA and interval mapping. The single factor ANOVA analysis was performed using a macro program from SAS ver. 9.1.3 (SAS Institute Inc. 2003) obtained from Dr. Elizabeth Lee of the University of Guelph (Guelph, Ontario, Canada). QTL were also determined using interval mapping analysis (IM) of MapQTL5® (van Ooijen 2004b) with a LOD score of 2.3 and 2.4, for population 1 and 2, respectively. LOD score was calculated by performing a permutation test using MapQTL5® (van Ooijen 2004b) with a set of 10,000 interactions.

Only the SSRs markers that were significantly associated with yield in two or more potential mega-environment were used for QTL analysis of yield-related traits of each population. Only single-factor ANOVA analysis was performed to detect agronomic traits QTL. The type I error (α) was set at 0.01 for each analysis.

Statistical analysis

The variation in seed yield and yield-related traits for the Canadian and American data within a test location and year was analyzed using the protocol for nearest-neighbor RCBD of Agrobase Software (Agronomix Software Inc., Winnipeg, MB, Canada). The variation in yield and other traits for the Chinese experiments was partitioned into effects of genotype, replications and incomplete blocks within replications using PROC MIXED procedure from SAS ver. 9.1.3 (SAS Institute Inc. 2003) for a rectangular lattice design (Bowley 1999). Replications and blocks within replications were considered to be random variables. Test of residuals was evaluated using PROC UNIVARIATE and PROC PLOT procedure from SAS ver. 9.1.3 (SAS Institute Inc. 2003). LSMEANS and PDIFF statements were used to calculate differences in RIL's means. Tukey's multiple range comparison test was utilized for mean comparisons. Linear correlations between yield and yield traits were calculated using PROC CORR and PROC PLOT procedures. Seed yield and agronomic traits means were not averaged over years and locations since significant interactions were between genotypes and years and between genotypes and locations (Reinprecht et al. 2006). The type I error (α) was set at 0.01.

Results

Phenotypic evaluation

The mean yield of the population derived from the cross between OAC Millennium × Heinong #38 (population 1) was lower than the adapted parent, the non-adapted parent and the checks across Canadian, US and Chinese environments in 2009 and 2010 (Table 2). Only at Elora in 2009 and Rosemount in 2010, the population yielded higher than the non-adapted and the adapted parent, respectively. However, the mean of the highest yielding line had an increase in seed yield when it was compared with the respective adapted parent across environments, except at one location. In Changchun in 2010, no significant differences were found between the highest yielding line and the adapted parent, Heinong #38. On average, the highest yielding lines yielded 33, 50 and 17 % more than the adapted parent in Canada, US and China, respectively. In addition, the highest yielding lines exceeded in 40 and 30 % the mean yield of the checks across some of the

Canadian and US environments, respectively. No differences in seed yield were found between the highest yielding lines and the checks at Elora and St. Pauls in 2010 and at the Chinese environments, Harbin and Changchun.

Similar results were found for population 2, Pioneer 9071 × #8902, where the mean yield of the population was lower than the mean yield of the adapted parent, the non-adapted parent and the checks across the three mega-environments, Canada, US and China (Table 3). The population mean was observed to be higher than the adapted parent's mean in Woodstock 2010 and Changchun 2010. Several RILs significantly exceeded the yield of the adapted parent except at two locations. At Elora and St. Pauls in 2010, no significant differences in yield were found between the adapted parent (Pioneer 9071) and the highest yielding lines. On average, the highest yielding lines had an increase in seed yield of 40 and 36 % compared with the adapted parent, Pioneer 9071, in Canada and US environments. In China, the highest yielding lines yielded on average 48 % more than the adapted parent, #8902. In addition, the highest-yielding lines have an increase in seed yield of 30–40 % when compared with the yield of the commercial lines across Canada, US and China, except at one location. At Elora in 2009, no significant differences in yield were found between the highest yielding lines and the mean of the checks.

Correlations between seed yield and agronomic traits

Most of the agronomic traits were correlated with seed yield in population 1 (Table 4) and 2 (Table 5). Maturity (R8; Fehr et al. 1971) was positively correlated with seed yield in 7 out of 9, and in 6 out of 11 environments in populations 1 and 2, respectively. In the case of lodging, it was positively correlated with seed yield in four out of nine of the environments for population 1. However, the correlation was negative for most of the environments of population 2 where the interactions between lodging and seed yield were significant. Disease data were not collected across locations and years as no significant diseases were observed in the trials. No significant abiotic stress was observed in these trials either.

Plant height was also positively correlated with seed yield in five out of seven and in six out of nine environments in population 1 and 2, respectively. A negative correlation between this trait and seed yield was observed in one of the Chinese locations in population 2. Correlation between 100 seed weight and seed yield was significant for population 1 in three of the nine environments being positive in one environment but negative in the other two. In the case of population 2, the correlation was positive for all environments where it was significant.

Oil content was significantly associated with seed yield in just two US environments in population 1. This association was inconsistent since it was positive in one but negative in the other one. In population 2, the correlation was negative at one location in the Canadian mega-environment, one location of the US mega-environment and one location in China. On the contrary, protein content was negatively correlated with seed yield at one, and at three locations in populations 1 and 2, respectively.

QTL analysis

OAC Millennium × Heinong 38 RIL population

Two markers, the interval Satt364–Satt591 and Satt277, were significantly associated with seed yield across the three mega-environment Canada, US and China when single-factor ANOVA and interval mapping were performed (Table 6). The interval Satt364–Satt591 was significantly associated with seed yield at two of the Canadian environments, at one US environment and at one Chinese environment. It is located on chromosome 5 (LG A1) based on the Composite Map 2003 (Cregan et al. 1999; Soybase 2011) and explained between 10 and 23 % of the variation observed in seed yield. The adapted parent, OAC Millennium, contributed the high-yielding allele to this QTL (Table 6) These markers were also linked with a QTL associated with oil content that was identified at one of the US locations explaining approximately 11 % of the variation observed for this trait (Table 7).

Satt277 was also significantly associated with seed yield across three Canadian environments, at two US environments and at one Chinese environment. It was located on chromosome 6 (LG C2) and explained between 9.0 and 14 % the yield variation observed in this study. However, in this case the Chinese parent, Heinong #38, contributed the favorable allele. The QTL tagged by Satt100, located in the same chromosome 6 (LG C2), at approximately 7 cM from Satt277, was the one that explained most of the variation observed in seed yield with an R^2 that varied between 13 and 32.3 % depending on the environment. It was identified as a QTL not only in all the Canadian environments but also in two of the US ones (Table 6). QTL Satt277 and Satt100 were also linked to maturity and height QTL across Canada, US and China. These two yield QTL also overlapped other agronomic trait QTL associated with lodging, 100 seed weight, oil and protein content across Canada and US (Table 7).

Seven markers were associated with seed yield in just two mega-environments, Canada and US. All of them also co-localized with other agronomic trait QTL associated with maturity, lodging, plant height, 100 seed weight, oil

Table 4 Correlation coefficients between yield and other agronomic traits found in a population derived from a cross between Heinong 38 and OAC Millennium across different environments in Canada, United States and China during 2009 and 2010

Year	Mega-environment	Location	Maturity	Lodging	Plant height	100 seed weight	Oil	Protein
2009	Canada	Elora	0.44**	NS	0.63**	ND	ND	ND
		St. Pauls	NS	0.34**	0.61**	−0.31**	NS	NS
	United States	Lamberton	0.46**	0.34**	ND	−0.22*	0.23*	−0.21*
		Rosemount	0.76**	0.44**	0.81**	NS	−0.30**	NS
2010	Canada	Elora	0.26*	NS	0.30**	NS	NS	NS
		St. Pauls	0.41**	0.21*	0.65**	NS	NS	NS
	United States	Rosemount	0.27**	NS	ND	NS	NS	NS
	China	Harbin	0.29**	NS	NS	0.30**	NS	NS
		Changchun	NS	NS	NS	NS	NS	NS

NS correlation not significant ($P > 0.05$), *ND* no data, data could not be recorded

*, ** Values significant at $P < 0.05$ and $P < 0.001$, respectively

Table 5 Correlation coefficients between yield and other agronomic traits found in a population derived from a cross between Pioneer 9071 and #8902 across different environments in Canada, United States and China during 2009 and 2010

Year	Mega-environment	Location	Maturity	Lodging	Plant height	100 seed weight	Oil	Protein
2009	Canada	Elora	NS	NS	0.57**	NS	NS	−0.20*
		Woodstock	0.35**	NS	0.60**	0.28**	−0.31**	NS
	United States	Lamberton	0.22*	NS	ND	NS	NS	−0.21*
		Rosemount	0.35**	0.36**	0.58**	0.31**	−0.27**	NS
2010	Canada	Elora	NS	0.23**	NS	0.21*	NS	NS
		St. Pauls	NS	NS	NS	0.35**	NS	NS
		Woodstock	0.19*	0.18*	0.20*	0.34**	NS	−0.18*
	United States	Lamberton	0.29**	NS	0.20*	NS	NS	NS
		Rosemount	NS	−0.44**	ND	NS	NS	NS
	China	Harbin	0.32**	NS	0.21*	0.20*	−0.21*	NS
		Changchun	0.31**	−0.48**	−0.32**	NS	NS	NS

NS correlation not significant ($P > 0.05$), *ND* no data, data could not be recorded

*, ** Values significant at $P < 0.05$ and $P < 0.01$, respectively

and protein content that were detected not only in Canada and US but also in China (Tables 6, 7).

In addition, 11 markers were associated with seed yield in at least one macro-environment within the mega-environment. Six of those yield QTL were identified in Canadian environments, whereas 2 and 3 out of 11 QTL were associated with yield in the US and Chinese environments, respectively (data not shown). Palomeque et al. (2009a) defined those QTL as mega-environment-specific QTL (QTL_{sp}).

In this RIL population, both parents contributed favorable alleles to the progeny. Among the QTL that were significant in two or more mega-environment, four of them carried the high-yielding allele from Heinong #38, which was the exotic parent in the Canadian and United States mega-environments while being the adapted parent in the Chinese one. In particular, the QTL linked to marker Satt273 inherited the high-yielding allele from the Chinese parent during 2009 and from the Canadian parent during 2010 in Canada and United States. This result could be attributed to more favorable environmental conditions in 2010 that resulted in a longer growing season than in 2009. There was more rain, higher temperature and humidity that could affect some physiological mechanisms that were reflected in differential expression of alleles at loci yield.

The most frequent top ten high-yielding lines across all mega-environments carried most of the beneficial alleles of the QTL that were present in two and three mega-environments. Line 26 was the most frequent line being among the top ten high-yielding lines in six of nine environments. It inherited all the favorable alleles of the QTL that were presented in two and three mega-environments (Table 8).

Table 6 Seed yield QTL identified by single-factor ANOVA and interval mapping (IM) in a minimum of one macro-environment in at least two mega-environments and higher yielding allele in a recombinant inbred population derived from the cross between a Canadian and a Chinese elite soybean line, OAC Millennium × Heinong 38, respectively

Marker	Linkage group	Mega-environment																	
		Canada						United States						China					
		ANOVA		IM		Parent allele	Env[a]	ANOVA		IM		Parent allele	Env[a]	ANOVA		IM		Parent allele	Env[a]
		P value	R^2	LOD	R^2			*P* value	R^2	LOD	R^2			*P* value	R^2	LOD	R^2		
Satt364-Satt591	5	< 0.01	13.8	–	–	CA	E09	< 0.01	16.1	2.5	12	CA	R09	< 0.01	9.8	–	–	CA	C10
		< 0.01	23.0	2.4	12.6	CA	P09	–	–	–	–	–	–	–	–	–	–	–	–
Satt139	4	0.01	9.6	–	–	CH	E09	< 0.01	10.7	–	–	CH	R09	–	–	–	–	–	–
		< 0.01	13.6	–	–	CH	P09	–	–	–	–	–	–	–	–	–	–	–	–
		< 0.01	15.9	3.0	13.8	CH	P10	–	–	–	–	–	–	–	–	–	–	–	–
Satt277	6	< 0.01	14.0	–	–	CH	P09	< 0.01	9.0	–	–	CH	L09	< 0.01	11.3	–	–	CH	H10
		< 0.01	9.1	–	–	CH	E10	< 0.01	12.9	–	–	CH	R09	–	–	–	–	–	–
		0.01	7.1	–	–	CH	P10	–	–	–	–	–	–	–	–	–	–	–	–
Satt100	6	< 0.01	13.2	–	–	CH	E09	< 0.01	16.7	–	–	CH	L09	–	–	–	–	–	–
		< 0.01	13.8	–	–	CH	P09	< 0.01	32.3	–	–	CH	R09	–	–	–	–	–	–
		< 0.01	17.3	–	–	CH	P10	–	–	–	–	–	–	–	–	–	–	–	–
Satt452-Satt263	15	0.01	6.9	–	–	CA	E09	< 0.01	8.0	–	–	CA	R09	–	–	–	–	–	–
Satt162	20	< 0.01	11.6	–	–	CA	E09	< 0.01	7.8	–	–	CA	R10	–	–	–	–	–	–
		< 0.01	10.1	–	–	CA	E10	–	–	–	–	–	–	–	–	–	–	–	–
		< 0.01	8.6	–	–	CA	P10	–	–	–	–	–	–	–	–	–	–	–	–
Satt273	9	< 0.01	13.2	–	–	CA	E10	< 0.01	11.7	–	–	CH	L09	–	–	–	–	–	–
Sat_126	9	< 0.01	14.3	–	–	CA	E09	0.01	9.2	–	–	CA	R09	–	–	–	–	–	–
		< 0.01	10.7	–	–	CA	P09	–	–	–	–	–	–	–	–	–	–	–	–
		< 0.01	15.2	–	–	CA	P10	–	–	–	–	–	–	–	–	–	–	–	–
Satt259	10	< 0.01	9.3	3.2	11.4	CH	E09	< 0.01	12.2	2.6	12.2	CH	L09	–	–	–	–	–	–
		< 0.01	7.3	–	–	CH	P09	< 0.01	21.6	4.8	21.6	CH	R09	–	–	–	–	–	–
		< 0.01	7.3	–	–	CH	P10	–	–	–	–	–	–	–	–	–	–	–	–

The trials were conducted across different environments in Ontario, Canada; Minnesota, United States and Heilongjiang and Jilin, China during 2009 and 2010

Dash indicates that marker or interval was not significant

Percentage of the phenotypic variation explained by the QTL

[a] Environments: (E09), Elora in 2009; (P09), St. Pauls in 2009; (E10), Elora in 2010 and (P10) St. Pauls in 2010 in Canada; (L09), Lamberton in 2009; (R09), Rosemount in 2009 and (R10), Rosemount in 2010 in United States; (H10), Harbin in 2010 and (C10), Changchun in 2010 in China

Table 7 Maturity, lodging, plant height, 100 seed weight, protein and oil content QTL identified by single-factor ANOVA in a recombinant inbred population derived from the cross of OAC Millennium and Heinong 38 across different mega-environments (ME), Canada, United States and China in 2009 and 2010

Marker		Trait / Mega-environment																	
		Maturity			**Lodging**			**Plant Height**			**100 Seed Weight**			**Oil**			**Protein**		
		Canada	US	China	Canada	US	China	Canada	US	China	Canada	US	China	Canada	US	China	Canada	US	China
Satt364-Satt591	*P value*	–	–	–	–	–	–	–	–	–	–	–	–	–	< 0.01	–	–	–	–
	R^2	–	–	–	–	–	–	–	–	–	–	–	–	–	10.6	–	–	–	–
Satt139	*P value*	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	–	–	–	–	–	–	–	–	–	–	–	–	–
	R^2	12.1	11.1	12	13.2	10.4	–	14.2	14.9	19.5	–	–	–	–	–	–	–	–	–
Satt277	*P value*	< 0.01	< 0.01	–	< 0.01	< 0.01	–	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	–	–	–	–	< 0.01	< 0.01	–
	R^2	20.1	25	–	17.4	16.1	–	16.2	28.3	9.5	21.1	12.1	–	–	–	–	8.7	16.2	–
Satt100	*P value*	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	–	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	–	< 0.01	< 0.01	–	< 0.01	–	–
	R^2	41.6	54.1	9.2	19.3	18.9	–	41.4	57.3	16.1	10.25	19	–	15.8	10.1	–	10.6	–	–
Satt452-Satt263	*P value*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	R^2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Satt162	*P value*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	< 0.01	–
	R^2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	7.5	–
Satt273	*P value*	–	–	–	–	–	–	–	–	–	–	< 0.01	–	–	–	–	–	< 0.01	–
	R^2	–	–	–	–	–	–	–	–	–	–	11	–	–	–	–	–	15.5	–
Satt_126	*P value*	0.01	0.01	–	–	< 0.01	< 0.01	< 0.01	< 0.01	–	–	–	–	–	–	–	–	–	–
	R^2	8.7	6.9	–	–	13.3	13.3	13.6	9.5	–	–	–	–	–	–	–	–	–	–
Satt259	*P value*	–	< 0.01	–	–	–	–	< 0.01	< 0.01	–	–	–	–	–	–	–	–	–	–
	R^2	–	9.6	–	–	–	–	11.2	11.2	–	–	–	–	–	–	–	–	–	–

Dash indicates that marker or interval was not significant

Table 8 Ten most frequent RILs that rank among the top ten yielding lines, derived from the cross OAC Millennium × Heinong 38 and their contributing alleles at nine marker loci linked to seed yield QTL detected in a minimum of one macro-environment of at least two mega-environments

RIL	Frequency of detection across 9 environments	Markers linked to seed yield QTL									Number of favorable alleles
		Satt 364–591 High yield alleles[a]	Satt 139	Satt 277	Satt 100	Satt 452–263	Satt 162	Satt 273	Sat_126	Satt 259	
		CA	CH	CH	CH	CA	CA	CA/CH	CA	CH	
26	6	CA	CH	CH	CH	CA	CA	CA	CA	CH	9
28	5	CA	CH	_[b]	CA	CH	CA	_[b]	CA	CA	4
4	4	CH	CH	CH	CH	CA	CA	CA	CA	CH	8
44	4	CA	CH	CA	CA	CA	CA	CA	CA	CH	7
3	3	CH	CA	CH	CH	CA	CA	CA	CA	CH	7
5	3	CA	CA	CH	CH	CA	CA	_[b]	CA	CH	7
15	3	CA	CH	CH	CH	CA	CH	CH	CA	CH	8
20	3	CA	CH	CA	CA	CA	CH	CH	CH	CH	5
62	3	CA	CH	CH	CH	CA	CH	CH	CH	CH	7
70	3	CA	CA	CH	CH	CH	CH	CH	CA	CH	6

[a] Parent alleles: *CA* Canadian parent OAC Millennium and *CH* Chinese parent Heinong 38

[b] Missing data is indicated by a dash (-)

Pioneer 9071 × #8902 RIL population

In this population, Sat_341 was significantly associated with yield across two mega-environments, Canada and China. A possible significant yield QTL was detected associated with this marker in US when the significance level was set at 0.05. Sat_341 was located on chromosome 10 (LG O) based on the Composite Map 2003 (Cregan et al. 1999; Soybase 2011) and explained between 7.5 and 8.2 % of the variation observed in seed yield. In this case, the high-yielding allele was contributed by the Chinese parent #8902 (Table 9). No agronomic traits that were measured in this study co-localized with this seed yield QTL (Table 10).

Markers Satt005 and Satt600 were mapped very close with an estimated average distance between them of 0.12 cM in the published composite map 2003 (Cregan et al. 1999; Soybase 2011). The QTL that was flanked by those markers was the one that explained most of the variation observed for seed yield in this population, with an R^2 that varied, depending on the location, between 8 and 13.2 %. This yield QTL was identified in Woodstock, Canada, in 2009 and Changchun, China, in 2010. When the significance level was set at 0.05, this QTL was also detected at Lamberton and Rosemount, US, in 2010 explaining approximately 6 % of the variation observed for this trait. The Chinese parent, #8902, contributed the high-yielding allele for this QTL (Table 5). Significant associations were found between these markers and maturity, lodging, plant height and oil content across different environments in this population (Table 10).

In addition, 18 QTL were associated with seed yield in only 1 mega-environment, being significant in at least one macro-environment within the mega-environment. Nine of those QTL were found in the Canadian environments, four in the US environments and only one marker was found significant in China. Both parents contributed favorable alleles to the progeny. In this population, line 132 ranked first among the most frequent top 10 high-yielding lines in 6 of 11 environments and contained all the expected favorable alleles for the QTL that were presented in 2 and 3 mega-environment (Table 11).

Genetic similarities between populations

Both populations were screened with the same set of SSR markers and the markers that were polymorphic in both were used to compare the QTL between them. Satt600 tagged a seed yield QTL in both genetic backgrounds at one location in Canada in the population 1 (data not shown). For population 2, this marker was located close to Satt005 and together explained the greatest amount of the variation observed in seed yield in this population. The

Table 9 Seed yield QTL identified by single-factor ANOVA and interval mapping (IM) in at least one macro-environment in a minimum of two mega-environments and the high-yielding allele in a recombinant inbred population derived from the cross between a Canadian and a Chinese elite soybean line, Pioneer 9071 × #8902, respectively

Marker	Linkage group	Mega-environment															
		Canada						United States						China			
		ANOVA		IM		Parent allele	Env[a]	ANOVA		IM		Parent allele	Env[a]	ANOVA		IM		Parent allele	Env[a]
		P value	R^2	LOD	R^2			*P* value	R^2	LOD	R^2			*P* value	R^2	LOD	R^2		
Satt005–Satt600	2	<0.01	8	–	–	CH	W09	–	–	–	–	–	–	<0.01	13.2	2.8	10	CH	C10
Sat120	13	–	–	–	–	–	–	<0.01	7.3	–	–	CA	L10	0.01	5.3	–	–	CA	C10
Sat_341	10	0.01	7.9	–	–	CH	P10	–	–	–	–	–	–	0.01	8.2	–	–	CH	C10
		0.01	7.4	–	–	CH	W10	–	–	–	–	–	–	–	–	–	–	–	–

The trials were conducted across different environments Canada, United States and China during 2009 and 2010

Table 10 Maturity, lodging, plant height, 100 seed weight, protein and oil content QTL identified by single-factor ANOVA in a recombinant inbred population derived from the cross Pioneer 9071 × #8902 across different megs-environments (ME), Canada, United States and China during 2009 and 2010

Trait	ME	Marker					
		Satt005–600		Sat120		Sat_341	
		P value	R^2	*P* value	R^2	*P* value	R^2
Maturity	Canada	<0.01	7.8	–	–	–	–
	US	<0.01	11.05	–	–	–	–
	China	0.01	6.6	–	–	–	–
Lodging	Canada	<0.01	9.1	–	–	–	–
	US	<0.01	8.4	<0.01	11.2	–	–
	China	–	–	–	–	–	–
Plant height	Canada	<0.01	9	–	–	–	–
	US	<0.01	7.2	–	–	–	–
	China	–	–	–	–	–	–
100 seed weight	Canada	–	–	<0.01	5.2	–	–
	US	–	–	<0.01	6.1	–	–
	China	–	–	–	–	–	–
Oil	Canada	<0.01	7.8	<0.01	7	–	–
	US	<0.01	7.3	<0.01	6.5	–	–
	China	–	–	–	–	–	–
Protein	Canada	–	–	–	–	–	–
	US	–	–	–	–	–	–
	China	–	–	–	–	–	–

Chinese parents, Heinong #38 and #8902, contributed the high-yielding allele for the first and second populations, respectively. However, when this marker was compared with other agronomic traits, the results varied between populations. No agronomic traits were detected as co-localizing with this seed yield QTL in population 1 while significant associations were found between this marker and maturity, lodging, plant height and oil content across different environments in population 2 (Table 10).

Satt263, located on chromosome 15 (LG E), was also detected as seed yield QTL in both populations. It was identified in Canadian and US environments in population 1 while it was detected only in US for population 2. This marker was also linked to other QTL that appeared to be significantly associated with 100 seed weight across all mega-environments in population 2 (data not shown). No agronomic traits were associated with this marker in population 1 (Table 7).

Even though Satt277 and Satt100, located on chromosome 6 (LG C2), tagged two important seed yield QTL only population 1, a seed yield QTL linked to marker Satt365, which mapped in the same region of the same linkage group, was identified in population 2 in Elora in 2010. The distance between Satt277 and Satt100 was

Table 11 Ten-most frequent RILs that rank among the top ten yielding lines, derived from the cross Pioneer 9071 × #8902 and their contributing alleles at 3 marker loci linked to seed yield QTL detected in a minimum of one macro-environment of at least two mega-environments

RIL	Frequency of detection/11 tests	Markers linked to seed yield QTL			Number of favorable alleles
		Satt 005–600 High yield alleles[a]	Sat 120	Sat_341	
		CH	CA	CH	
132	6	CH	CA	CH	3
130	5	CH	CA	_[b]	2
2	4	CA	CA	CA	1
8	4	CH	CA	CH	3
7	3	CA	CA	CA	1
21	3	CH	CA	_[b]	2
53	3	CA	CA	CA	1
71	3	CA	CH	CA	0
108	3	CH	CH	CH	2
112	3	CH	CH	CH	2

Eleven tests were conducted across different environments in Canada, United States and China during 2009 and 2010

[a] Parent alleles: *CA* Canadian parent Pioneer 9071 and *CH* Chinese parent #8902

[b] Missing data is indicated by a dash

estimated at approximately 7 cM in the published composite map 2003 (Cregan et al. 1999; Soybase 2011). Satt365 was located on the same linkage group, between those two markers, with an approximately distance of 4 cM from Satt277 and 2 cM to Satt100, based on the composite map (Cregan et al. 1999; Soybase 2011) which indicated the importance of this region of chromosome 6 (LG C2) for breeding purposes. The Chinese parent contributed the favorable allele for these markers. Maturity, plant height, 100 seed weight, oil and protein were associated with this region in both populations.

Discussion

From an agricultural standpoint, plants are considered better adapted to a particular region than others when they have higher seed yield and seed quality than other individuals in certain environments (Cooper and Byth 1996). For decades, yield improvements have been achieved through conventional breeding using elite lines as parents, which could lead to diminishing gains due to genetic similarities among cultivars (Manjarrez-Sandoval et al. 1997; Sneller 1994; Ude et al. 2003). However, the introgression of genes from PI and exotic germplasm into the current soybean genetic base may increase genetic variability and lead to greater gains from selection (Thompson and Nelson 1998).

With the idea of understanding how relevant high-yield QTL behave under different ME, two Canadian × Chinese soybean populations were tested in this study at different locations in Ontario, Canada; in Heilongjiang and Jilin, provinces of China and in a new mega-environment, Minnesota, US. The objectives were to find additional seed yield QTL across one, two and three mega-environments and in both populations. From the reported seed yield QTL that were detected in two or three mega-environments, we attempted to analyze the correlations between yield and yield components and to determine the co-localization with other agronomic trait QTL.

In both crosses across most of the environments, the RIL population yield means were lower than that of the adapted parents. Similar results were reported by Orf et al. (1999a) and Palomeque et al. (2009a). However, several lines exceeded the yield of the adapted parents. Among them, the most frequent top ten high-yielding lines across all mega-environments in both populations carried most of the beneficial alleles for seed yield QTL that were significant in two or three mega-environments. This suggests that PI alleles could be introduced from different germplasm sources without a yield penalty and used as potential sources to increase genetic variability leading to higher and potentially more sustainable yields.

Although some inconsistencies were found, significant correlations existed between seed yield and maturity, lodging, height, seed weight, protein and oil. In addition, most of the seed yield QTL co-localized with agronomic trait QTL. Results confirmed that, in general, yield variation observed in the progenies of both crosses grown in the Canadian, United States and Chinese mega-environments was the consequence of the variation observed in different agronomic traits. Those agronomic traits were correlated with yield and associated with QTL detected in one, two or three mega-environments. In particular, maturity was considered one of the most important traits that influenced seed yield (Egli et al. 1981). Thus, higher yield could be mainly explained by the variation in traits associated either with late maturity or with optimum flowering, which could increase the seed filling period. Similar results were found for population 1 by Palomeque et al. (2009b) and by Orf et al. (1999a), Burton (1987), and Wehrmann et al. (1987) for another population for these agronomic traits.

Brummer et al. (1997) suggested that QTL needed to be stable across environments to be useful in breeding programs and that breeders should be careful about incorporating environmentally sensitive QTL into their cultivars unless they work in specific environments. We assessed stability by evaluating both populations across different mega-environments. The interval Satt364–Satt591 was significantly associated with seed yield in at least one location of each of the three mega-environments in population 1. It co-localized with oil content in US. This association was previously reported in a bi-parental cross between Essex and Williams (Hyten et al. 2004) and also with maturity in another cross (Panthee et al. 2004). The yield QTL associated with marker Satt277 was also identified in all mega-environments in population 1 and co-localized with other agronomic traits QTL such as maturity, lodging, plant height and protein content. It was previously reported in the same population associated with seed yield, 100 seed weight, pods per plant, pods per node, height, R1, R5 and R8 in the Canadian and Chinese mega-environments by Palomeque et al. (2009a, b). It was also identified as a QTL associated with lodging, seed weight, plant height, yield per plant height (Orf et al. 1999a), seed yield (Orf et al. 1999a, b; Reinprecht et al. 2006; Smalley et al. 2004; Specht et al. 2001), oil and protein concentration (Reinprecht et al. 2006) in other populations. In population 2, derived from the cross between Pioneer 9071 and #8902, the QTL tagged by Sat_341 was the only one that was significantly associated with yield across two and three mega-environments. None of the other agronomic traits that were measured in this study were associated with this marker. This result implied that some physiological process might be involved in the variation observed for yield. There is no other previous report of QTL linked to Sat_341 (Soybase 2011). However, unless breeding programs work for specific environments, environmentally stable QTL are usually more useful for marker-assisted selection.

Seven and two QTL were associated with seed yield in two of three mega-environments in populations 1 and 2, respectively. Four of them were previously reported by Palomeque et al. (2009a) as universal QTL in the population 1 in the Chinese and in the Canadian environments. However, in this study, those QTL were only identified in the Canadian and in the US mega-environments. In particular, Satt259 was detected by the same authors (Palomeque et al. 2009a) as a specific seed yield QTL in the Canadian mega-environment while in this study it was found also in the US. Among the markers that were detected in two of the three mega-environments, the QTL tagged by Satt100 and by Satt005–Satt600 were the ones that explained most of the variation observed in seed yield in the population 1 and 2, respectively. Yield QTL Satt100 also co-localized with other agronomic traits. Similar results were reported for the same population by (Palomeque et al. 2009a, b). In the second RIL group, Satt005–Satt600 was previously associated with partial resistance to *Phytophthora sojae* (Burnham et al. 2003). In both cases, the high-yielding allele came from the PI-derived parent, which demonstrates once again that high-yielding alleles can be successfully introduced without penalty.

In general, different QTL were found in particular populations due to differences in their genetic backgrounds (Brummer et al. 1997). However, QTL are generally more useful for breeders if they are identified across different populations. Fasoula et al. (2004) supported the necessity of validating QTL in different populations to increase the spectrum of application for such markers in breeding programs. The universal and specific QTL that Palomeque et al. (2009a, b) found in the population OAC Millennium × Heinong, were not validated in population 2, Pioneer 9071 and #8902, even though the parental lines in both groups were Canadian and Chinese elite lines. Since in the current study both populations were evaluated with the same markers, we identified two seed yield QTL, Satt600 and Satt263 in both genetic backgrounds. QTL tagged by marker Satt263 was previously associated with nitrogen accumulation at growth stage R5 and R6 (Panthee et al. 2004) and with seed weight (Soybase 2011). In addition, it is important for breeders to give special attention to the region tagged by markers Satt277, Satt365 and Satt100 on chromosome 6 (LG C2), as in both populations this genomic region carried alleles with important yield effects. Furthermore, seed yield and other agronomic traits QTL were previously reported in this region in different genetic backgrounds (Orf et al. 1999a; Smalley et al. 2004; Reinprecht et al. 2006; Palomeque et al. 2009a). QTL tested at different

genetic backgrounds provide reliable estimation of marker limitations and strengths for use in marker-assisted selection.

Our results provide a better understanding of how QTL behave under specific environments relative to the adaptation area of each parent. We attempted to determine the value of these markers/QTL in breeding programs for further studies based on their genetic and environmental stability. Both the environmental and genetic stability of QTL play an important role in breeding as they could affect the success of introgression of exotic alleles into adapted germplasm. In addition, most of these seed yield QTL were also linked with some agronomic trait QTL. Results suggested that most of the variation observed in yield could be explained by the variation observed in different agronomic traits. For all these markers, both parents contributed their respective alleles to the progeny and the ratio varied between populations. The importance of both gene pools, Canadian and Chinese, is denoted in this study. They represent a special source of beneficial alleles that could contribute directly or indirectly to increase variability and lead to further increments in soybean yields both in North America and in China. Further studies should focus on inter-mating different lines from these major sources with the objective of discovering new recombination events that could benefit soybean breeding programs.

References

Agrobase (2000) Agrobase Software. Agronomix Software Inc., Winnipeg

Bowley SR (1999) A hitchhiker's guides to statistics in plant biology. Any Old Subject Books, a Division of Plants, Inc., Guelph

Braun HJ, Rajaram S, 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

Burnham KD, Dorrance AE, VanToai TT, St Martin SK (2003) Quantitative trait loci for partial resistance to *Phytophthora sojae* in soybean. Crop Sci 43:1610–1617

Burton JW (1987) Quantitative genetics: results relevant to soybean breeding. In: Wilcox JR (ed) Soybeans: improvement, production, and uses, 2nd edn. American Society of Agronomy, Inc, Madison, pp 211–247

Cianzio SR, Schultz SP, Fehr WR, Tachibana H (1991) Registration of "Archer" soybean. Crop Sci 31:1707

Concibido VC, La Vallee B, McLaird P, Pineda N, Meyer J, Hummel L, Yang J (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. Theor Appl Genet 106:575–582

Cooper M, Byth DE (1996) Understanding plant adaptation to achieve systematic applied crop improvement—a fundamental challenge. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CABI, Wallingford, pp 5–23

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N (1999) A new integrated genetic linkage map of the soybean. Crop Sci 39:1464–1490

Egli DB, Fraser J, Leggett JE, Poneleit CG (1981) Control of seed growth in soya beans [*Glycine max* (L.) Merrill]. Ann Bot 48:171–176

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, Caviness DT, Burmood DT, Pennington JS (1971) Stage of development descriptions for soybeans, *Glycine max* (L.) Merril. Crop Sci 11:929–931

Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt ME (2004) Seed quality QTL in a prominent soybean population. Theor Appl Genet 109:552–561

Kabelka EA, Diers BW, Fehr WR, Leroy AR, Baianu IC, You T, Neece DJ, Nelson RA (2004) Putative alleles for increased yield from soybean plant introductions. Crop Sci 44:784–791

Li Z, Jakkula L, Hussey RS, Tamulonis JP, Boerma HR (2001a) SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. Theor Appl Genet 103:1167–1173

Li Z, Qiu L, Thompson JA, Welsh MM, Nelson RL, Merr L (2001b) Molecular genetic analysis of US and Chinese soybean ancestral lines. Crop Sci 41:1330–1336

Manjarrez-Sandoval P, Carter TE, Webb DM, Burton JW (1997) RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. Crop Sci 37:698–703

Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci 36:1327–1336

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

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

Palomeque L (2007) Mapping of QTLs associated with adaptation to mega-environments using Canadian × Chinese populations in soybean. Ph.D. thesis, Retrieved from Theses database, Department of Plant Agriculture, University of Guelph

Palomeque L, Liu LJ, 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 L, Liu LJ, 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

Palomeque L, Liu LJ, Li W, Hedges B, Cober ER, Smid MP, Lukens L, Rajcan I (2010) Validation of mega-environment universal and specific QTL associated with seed yield and agronomic traits in soybeans. Theor Appl Genet 120:997–1003

Panthee DR, Pantalone VR, Sams CE, Saxton AM, West DR, Rayford WE (2004) Genomic regions governing soybean seed nitrogen accumulation. J Am Oil Chem Soc 81:77–81

Reinprecht Y, Poysa VW, Yu K, Rajcan I, Ablett GR, Pauls KP (2006) Seed and agronomic QTL in low linolenic acid, lipoxygenase-free soybean (*Glycine max* (L.) Merrill) germplasm. Genome 49:1510–1527

SAS ver (2003) SAS. 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

Sneller CH (1994) Pedigree analysis of elite soybean lines. Crop Sci 34:1515–1522

Soybase (2011) Soybean breeder's toolbox genetic map information. http://soybeanbreederstoolbox.org/. Retrieved July 2011

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

Thompson JA, Nelson RL (1998) Utilization of diverse germplasm for soybean yield improvement. Crop Sci 38:1362–1368

Ude GN, Kenworthy WJ, Costa JM, Cregan PB, Alvernaz J (2003) Genetic diversity of soybean cultivars from China, Japan, North America, and North American ancestral lines determined by amplified fragment length polymorphism. Crop Sci 43:1858–1867

van Ooijen JM (2004a) Software for the calculation of genetic linkage maps in experimental populations. JoinMap® 4 Kyazma BV, Wageningen

van Ooijen JM (2004b) Software for the mapping of quantitative trait loci in experimental populations. MapQTL® 5 Kyazma BV, Wageningen

Wang D, Graef GL, Procopiuk AM, Diers BW (2004) Identification of putative QTL that underlie yield in interspecific soybean backcross populations. Theor Appl Genet 108:458–467

Wehrmann VK, Fehr WR, Cianzio SR, Cavins JF (1987) Transfer of high seed protein to high-yielding soybean cultivars. Crop Sci 27:927–931