CONFIRMING QTL FOR SEED YIELD FROM EXOTIC SOYBEAN GERMPLASM

BY

CHARLES COLE HENDRIX

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Professor Randall L. Nelson

ABSTRACT

The genetic improvement of soybean (*Glycine max* (L.)) cultivars in North America (N.A.) has, for the most part, been accomplished by intermating elite cultivars. This breeding strategy, combined with the limited N.A. genetic base, has resulted in a very narrow gene pool. Plant introductions (PI) have been used to expand the N.A. genetic base with limited success when conventional breeding methods have been used. Quantitative trait loci (QTL) mapping has been used to identify genetic regions within PIs that could contribute both genetic diversity and improved seed yield potential in the N.A. gene pool. Of the putative QTL for seed yield that have been identified, only a small percentage have been tested in confirmation trials and even fewer have been confirmed. The objective of this study is to confirm putative QTL for seed yield derived from Exotic germplasm that were identified in previous QTL mapping studies. One BC_1F_9 confirmation population was developed from the cross Kenwood x LG94-1713 to test the QTL associated with SSR loci Satt405 (linkage group (LG), J chromosome (chr) 16) and two BC_1F_{11} populations were developed to test Satt477 (LG O, chr 10) and Satt557 (LG C2, chr 6). Four F_8 confirmation populations were developed from the cross of BSR 101 x LG82-8379 to test the QTL linked to Satt142 (LG H, chr 12), Satt225 (LG A1, chr 5), Satt363 (LG C2, chr 6), and Satt544 (LG K, chr 9) and two F_9 populations were developed to test Satt168 (LG B2, chr 14) and Satt358 (LG O, chr 10). Unfortunately, no putative QTL for seed yield were confirmed in any of the populations developed from the BSR 101 x LG82-8379 mapping population. A QTL for plant height, maturity, and seed yield associated with Satt557 was confirmed in two populations developed from Kenwood x LG94-1713 with the beneficial allele coming from the LG94-1713 parent but these results were confounded by the tight linkage of Satt557 to the *E1* locus. In both populations maturity was delayed by slightly more than 5 days in the lines homozygous for the LG94-1713 allele. However, there were differences between the two populations for both plant height and seed yield. The allele from LG94-1713 in one population increased plant height by 4.6 cm and seed yield by 0.32 Mg ha^{-1} more than in the second population. We hypothesize that a crossover occurred in one of the populations that separated the putative QTL from Satt557 but not from *E1* and this

QTL is responsible for the increase in seed yield and plant height. Several polymorphic single nucleotide polymorphism (SNP) markers have been identified between the two parents, which are on either side of Satt557. The lines in both populations are being tested with these markers to determine if and where the crossover occurred. If our assumption is correct, this confirmed QTL for seed yield may add additional genetic diversity and higher seed yield potential to the N.A. gene pool.

TABLE OF CONTENTS

INTRODUCTION……1
MATERIALS AND METHODS……31
RESULTS AND DISCUSSION……36
REFERENCES……44
APPENDIX……53

INTRODUCTION

The limited number of major soybean (*Glycine max* (L.) Merr.) ancestors introduced to the North American (N.A.) continent (Gizlice et al., 1994) in combination with over 75 years of intense selective breeding pressure has been shown to be a significant ($P < 0.05$) genetic bottleneck in the history of N.A. soybean breeding (Hyten et al., 2006). The genetic advancement of N.A. soybean cultivars has been historically achieved by crossing elite x elite cultivars. This breeding pattern has led to an average annual seed yield increase of 0.022 Mg ha^{-1} yr^{-1} (www.nass.usda.gov) but, future yield gains may be negatively impacted by the low genetic diversity in modern cultivars (Hyten et al., 2006).

Utilizing diverse exotic germplasm to improve the N.A. genetic diversity for seed yield has been met with mixed results and can be a very slow process when traditional breeding methods have been used. Ininda et al. (1996) found that utilizing soybean plant introduction (PI) parentage in recurrent selection programs showed no increase in genetic gain over elite cultivar parentage. Sneller et al. (1997) tested southern maturity group (MG) PIs against northern and southern elite cultivars and found that the majority of the PIs had unacceptable agronomic performances but there were a few exceptions that could be competitive with the northern elite cultivars. Thompson and Nelson (1998a) were also able to show examples of lines developed from N.A. cultivar x PI crosses that had yield performances greater than both parents but the majority of those lines had maturities later than the N.A. parent. Although some positive results have been found by Sneller et al. (1997) and Thompson and Nelson (1998a) and there have been several recent germplasm releases which were derived from PI parentage (Nelson and Johnson, 2006a; Nelson and Johnson, 2006b; Shannon et al., 2005), traditional breeding methods may still involve several cycles of selection before meaningful testing can begin. Quantitative trait loci (QTL) mapping uses polymorphic molecular markers and phenotypic data to associate genetic regions with the traits that were measured. However, due to environmental effects, putative QTL need to be confirmed in a separate population. If the QTL is

confirmed, then the molecular marker information can be used to transfer the QTL into improved cultivars.

North American Soybean Genetic Variability

In 1994, Gizlice et al. calculated coefficient of parentage estimates for all of the 258 public N.A. soybean cultivars that were released between 1947 and 1988, assuming a 50% genetic contribution from each of the parental genotypes. They were able to account for more than 99% of the N.A. genetic base with 80 ancestors. The percent of contribution of the 80 ancestors ranged from around 0.00001 to 12% with 20 of the 80 ancestors accounting for about 86% of the genetic base. For the southern United States, 46% of the genetic base was accounted for by the cultivars CNS and S-100, and 10 ancestors accounted for more than 83% of the southern genetic base. The northern United States had about 24% of the genetic base accounted for by the two unknown parents of Lincoln.

Gizlice et al. (1994) also use coefficient of parentage estimates of 133 cultivars and six breeding lines (defined as first progeny) from the 80 previously mentioned ancestors to quantify the genetic base in a second analysis. Gizlice et al. (1994) found that 91 first progeny explained 99% of the N.A. genetic base which consisted of 5 breeding lines, 8 old cultivars, and 78 modern cultivars. The remaining 1% of the genetic base was explained by a single breeding line and 47 modern cultivars which contained minimal amounts of unique genetic contributions from the 78 ancestors. Of the aforementioned 91 first progeny, 16 cultivars and one breeding line accounted for nearly 75% of the genes found in modern cultivars released before 1960. The northern and southern genetic bases were fully defined by 70 and 46 first progeny, respectively. This narrowness of the N.A. gene pool is of major concern for breeders, who require genetic variation in order to continue the improvement of seed yield and other agronomic traits.

Since the genetic relationships among the 80 ancestors prior to N.A. introduction were unknown, the Gizlice et al. (1994) results may be underestimating the amount of

relatedness between the N.A. ancestors. Looking at the genetic similarities rather than pedigree analysis, may be a better means of determining the genetic relationship between parents. In order to test this hypothesis, Manjarrez-Sandoval et al. (1997) developed five $F_{6:8}$ populations of varying coefficient of parentage estimates between the parents. The five populations were derived from crosses between Davis x N73-1102 (20 lines), Young x N73-1102 (28 lines), Forrest x Bay (27 lines), N77-179 x Forrest (23 lines), and Essex x Vance (22 lines) with coefficient of parentage estimates of 0.06, 0.15, 0.17, 0.27, and 0.50, respectively. Restriction fragment length polymorphism (RFLP) markers for 42 loci were used to determine the genetic similarity estimates between the two parents of each population which were 57%, 63%, 59%, 60%, and 75% similar, respectively. The five populations were tested in two environments, with two replications in one and three replications in the other. The experimental design was a split-split plot design with the whole plots divided into maturity ranges, subplots to populations, and sub-sub plots to lines nested within each subplot. Data were collected on seed yield and was adjusted for plant maturity so that the genetic variation for yield could be calculated.

The coefficient of parentage estimates were found to be significantly correlated ($P < 0.05$) with the variation observed with the RFLP genetic similarly estimates (RFLP-GS) ($r = 0.91$) and with the genetic variance for seed yield ($r = -0.81$). However, the RFLP-GS were not found to be significantly correlated with the genetic variance for seed yield ($r = -0.58$). Although the correlation between the RFLP-GS and the genetic variance for seed yield was not significant, the results did show a correlation in the negative direction. It is quite possible, that if this analysis would have used a greater number of polymorphic loci and more genome coverage, the correlation between RFLP-GS and the genetic variance for seed yield would have been significant. This negative correlation suggests that increasing the number of polymorphic loci between two parents would result in progeny with increase genetic variance for seed yield and could increase the chances of identifying transgressive segregates (Manjarrez-Sandoval et al., 1997). However, caution should be used when considering these correlation results because the results were based on only five comparisons which could be heavily influenced by a single observation. This is evident when the Essex x Vance population is removed from

the data set and the correlation between coefficient of parentage and RFLP-GS for the remaining four populations drops from $r = 0.91$ to $r = 0.41$.

The conclusions of Manjarrez-Sandoval et al. (1997) were supported by the USDA-ARS breeding program at North Carolina Sate University. They developed the five populations used by Manjarrez-Sandoval et al. (1997) but selected lines for seed yield independently of the 1997 report. A comparison between the numbers of lines derived from each of the five aforementioned populations that were tested in advanced yield trials showed that the greatest number of lines were derived from Davis x N73-1102 and Young x N73-1102 with three and two lines in regional preliminary testing, respectively, and one line each in uniform testing compared with no lines in either test for the other three populations (Manjarrez-Sandoval et al., 1997). In the Manjarrez-Sandoval et al. (1997) study, these two crosses had the lowest coefficients of parentage estimates and higher genetic variation for seed yield. These results indicate that using coefficient of parentage estimates to select parents for seed yield improvement programs may be practical. However, the Forrest x Bay population had a coefficient of parentage estimate similar to the Young x N73-1102 population and a greater genetic variance, but only three lines out of 168 were advanced to local preliminary testing and were then discarded. Coefficient of parentage estimates may be used as a tool to predict variability but may not predict the overall success of a yield improvement population. Although the RFLP-GS were not as consistent as the coefficient of parentage estimates in predicting genetic variation, they were able to identify the two populations with the highest (Davis x N73-1102 (RFLP-GS = 57%) and Young x N73-1102 (RFLP-GS = 59%)) and one of the populations with lowest (Essex x Vance (RFLP-GS = 75%)) genetic variances. This suggests that with refinement, genetic similarity estimates may become an important tool in selecting parents for breeding programs.

Thompson et al. (1998c) also used genetic markers to determine the relationships between genotypes. Their objectives were to evaluate the genetic relationships among and the diversity present within 18 major N.A. ancestors and determine if selected PIs, based on progeny performance, are genetically diverse from the N.A. ancestors.

Thompson et al. (1998c) used 833 random amplified polymorphic DNA (RAPD) markers and four cluster analysis methods to compare 17 soybean PIs to 18 ancestors and first progeny that account for over 85% of the N.A. genetic base (Gizlice et al., 1994). The PIs (mostly from Asia) had been successfully used in developing high yielding experimental lines (unpublished data). The Asian origin is significant because soybean domestication occurred in China (Li et al., 2008b), so germplasm from this geographical region should contain greater genetic diversity. The region(s) where soybean was domesticated in China has not been positively identified. There is historical and geographical evidence that suggests that soybean domestication occurred in northeastern China around the 11th century B.C. (Hymowitz, 1970). RFLP marker data of chloroplast and mitochondrial DNA of wild soybean indicated that the Yangtze River Valley was the center of cytoplasmic diversity (Shimamoto et al., 2000). However, the geographical distributions, genetic diversity, and multivariate variation coefficients of 6,172 accessions helped identify three centers of origin: Northeastern China, the Yellow River Valley, and the Southeast Coasts of China with each center being independent of each other or the domestication flowed from Northeastern China to the Yellow River Valley and then to the Southeast Coast (Dong et al., 2001). The Yellow River Valley was also identified as a center of origin using 1,863 Chinese landraces and 59 SSR loci (Li et al., 2008b). The consensus of these reports that soybean was domesticated in China but there is no agreement about where within China domestication occurred or if there was a single domestication event.

Thompson et al. (1998c) observed that of the 833 RAPD markers used, only 281 (34%) were polymorphic. This is about half the level of RAPD polymorphism observed in other self-pollinated crops. For example, Heun et al. (1994) used 177 RAPD markers to compare 24 accessions of *Arena sterilis* L. and found 155 markers (65%) to be polymorphic, which is similar to Ko et al. (1994) who used 144 RAPD markers to compare 37 varieties of *Oryza sativa* L. and found 96 markers (67%) to be polymorphic among the varieties. Furthermore, the amount of polymorphism observed by Thompson et al. (1998c) is also about one-third the amount found in outcrossing species. This can be seen in the Heun and Helentjaris (1993) report where 140 RAPD markers were used to

compare 21 genotypes of maze and 111 markers (80%) were found to be polymorphic among the genotypes.

Two hierarchal and two non-hierarchal clustering methods were applied to the Thompson et al. (1998c) RAPD data. Across all four methods, the 35 lines were placed in the same clusters 87% of the time. On average, 7 clusters were created with two containing mostly northern cultivars, two containing mostly southern cultivars, and three containing mostly Exotic germplasm. This study indicates that the 17 PIs with promising progeny performance are genetically distinct from both the northern and southern genetic bases so the integration of Exotic germplasm could add useful variation to both genetic pools.

With the advent of simple sequence repeat (SSR) markers, the use of RAPD and RFLP markers by soybean breeders became less common due to the fact that SSR markers are codominant, highly polymorphic in soybean, fairly inexpensive to use in breeding programs, can be combined with polymerase chain reactions (PCR) for faster analysis, primers can be developed for specific loci, and SSR loci are more or less randomly distributed throughout the soybean genome (Akkaya et al., 1992; Cregan et al., 1999). To assess the amount of SSR diversity that may reside among modern N.A. cultivars and PIs, Narvel et al. (2000) conducted a study that analyzed the extent of polymorphism among 40 PIs and 39 N.A. cultivars. A total of 74 SSR markers distributed across all 20 linkage groups (LG) were used to characterize each line. Each LG contained 1 to 7 SSR markers which were evenly distributed, with a few exceptions, within the LGs. The exceptions were LG B1 (chromosome (chr) 11) which contained only one SSR marker and both LG J (chr 16) and LG K (chr 9) had markers that were clustered or were unevenly spaced. The PIs and N.A. cultivars were selected based on high seed yield potential observed during replicated yield trials. Across both genetic pools, 72 of the 74 SSR markers were polymorphic with the most common allele occurring at a frequency of 0.95 or less. Within the PI genetic pool, 70 of the 74 loci were polymorphic with a total of 365 different alleles across all SSR markers. Within the N.A. cultivar genetic pool, 63 of the 74 loci were polymorphic with a total of 259 alleles

across all SSR markers. Of the total number of alleles detected across the two genetic pools, 227 were synonymous, 138 were specific to the PIs, and 32 were specific to the N.A. cultivars. Of the 72 loci that were analyzed, 63 had alleles that discriminated between the PIs and the N.A. cultivars. These results show that SSR markers are capable of distinguishing between elite and exotic genotypes.

A more in-depth investigation into the genetic variation of soybean as a species was conducted by Hyten et al. (2006) to determine how the variation has been altered from domestication through the development of elite cultivars. Samples of DNA were collected from 25 N.A. *G. max* cultivars, 17 of the N.A. *G. max* ancestors described by Gizlice et al. (1994), 52 *G. max* accessions of Asia landraces, and a diverse set of 26 *Glycine soja* (Sieb. and Zucc.) accessions (*G. soja* is the wild progenitor of *G. max*). For each genotype, single nucleotide polymorphism (SNP) alleles were detected among 102 randomly selected genes. Within the 102 genes 496 SNPs were detected of which, 84 were non-synonymous, 59 were synonymous, and 353 were in non-coding regions. The number of polymorphism sites in a genotypic sample corrected for sample size (θ) (Watterson, 1975) for the elite soybean cultivars was relatively low ($\theta = 0.00083$) when compared to *Sorghum bicolor* ($\theta = 0.0023$) (Hamblin et al., 2004) and very low when compared to modern maze inbred lines ($\theta = 0.00627$) (Wright et al., 2005).

The number of polymorphic sites within the N.A. cultivar, N.A. ancestor, Asian landrace, and *G. soja* genetic pools were tested against each other with the Duncan's Multiple Range test in order to determine the significant differences ($P < 0.05$) between the genetic pools. When the N.A. cultivars ($\theta = 0.00083$) were compared to the N.A. ancestors ($\theta = 0.001$) there was no significant difference between the two genetic pools with the N.A. cultivars retaining 83% of the diversity present within the N.A. ancestors. There was also no significant difference found between the N.A. ancestors and the Asian landraces ($\theta = 0.00115$) even though there was a loss of 78% of the 98 low frequency alleles found in the Asian landraces, but the N.A. ancestors did retain 87% of the Asian landrace diversity. However, the cumulative effects of both the N.A. ancestor bottleneck

and the N.A. breeding bottleneck resulted in the N.A. cultivars having significantly less diversity than the Asian landraces. The N.A. cultivars only retained 72% of the diversity found in the Asian landraces and lost close to 5% of the low frequency alleles found in the N.A. ancestor genetic pool. The *G. soja* accessions had the most polymorphism sites per sample ($\theta = 0.00235$) which was significantly higher than all of the other genetic pools with 65% more polymorphism diversity than the N.A. cultivars. Although the *G. soja* accessions had more SNP diversity when compared to *G. max*, the Hyten et al. (2006) data indicates that *G. soja* has unusually low diversity when compared to *Arabidopsis* ($\theta = 0.0071$) (Schmid et al., 2005), wild barley ($\theta = 0.0081$) (Morrell et al., 2005), and teosinte ($\theta = 0.0109$) (Wright et al., 2005). Thus, this low degree of genetic variability in the wild progenitor of the modern soybean may be the main reason why N.A. cultivars have lower genetic variability compared with other cultivated crops.

The Hyten et al. (2006) results are surprising because they indicate that N.A. soybean breeders have not significantly narrowed the genetic variability that was first available from the 17 most important N.A. ancestors. One could argue that only looking at 102 genes across the soybean genome is too few and that many more alleles could have been lost between the N.A. cultivars and the N.A. ancestors, especially for regions affecting agronomic traits such as seed yield. The results also indicated that the low N.A. cultivar diversity is mostly due to the unusually low level of genetic diversity found in the wild progenitor, *G. soja*, and the domestication bottleneck. Many of the rare alleles found in the *G. soja* genetic pool were lost when soybean was domesticated and during the introduction to N.A. This suggests that utilizing Asian landraces and *G. soja* accessions would allow breeders to tap into unused genetic resources that have the potential of revealing unique alleles that could be of benefit to the N.A. gene pool.

The previously mentioned reports have shown, through different measures of genetic contribution and genetic variability, that the N.A. soybean genetic base is quite limited. Gizlice et al. (1994) determined through coefficient of parentage estimates that 80 ancestors defined more than 99% of the N.A. genetic base, but the genetic relationship between these ancestors was unknown. Manjarrez-Sandoval et al. (1997) recognized the

potential problems with relying on pedigree data for determining diverse germplasm and that looking at genetic similarities could prove more useful. The genetic similarity estimate data indicated that greater genetic similarity may lead to less phenotypic variation but a significant correlation between genetic diversity and phenotypic diversity was not able to be found. This may be due to the low number of molecular markers that were used in this study.

Thompson et al. (1998c) used RAPD markers to genetically separate PIs from the N.A. ancestors using cluster analysis and found that some PIs were genetically distinct from the N.A. ancestors. Narvel et al. (2000) was able to show that, even with a low number of SSR loci, it is possible to detect greater amounts of genetic variability in the PI genetic pool than the N.A. genetic pool. Hyten et al. (2006) investigated SNP frequencies found within the wild progenitor (*G. soja*), Asian landrace, N.A. ancestor, and modern N.A cultivar genetic pools so as to determine the loss of genetic variation due to these genetic bottlenecks. There was no significant difference between the variability observed in the N.A. ancestors and the modern N.A. cultivars after over 75 years of intense breeding selection. However, one may speculate that analyzing 102 random genes is not representative of the soybean genome and that many more alleles could have been lost, especially in the regions affecting the agronomic traits that are subjected to the most selection intensity, such as seed yield. These studies suggest that exotic germplasm can be used as a source of unique genetics to both increase seed yield in N.A. cultivars and expand the N.A. genetic diversity.

Utilizing Exotic Soybean Germplasm

In order to determine whether or not having greater genetic diversity would lead to greater genetic gains, Ininda et al. (1996) conducted a study that looked at the rate of gain for seed yield in five populations with different percentages of PI parentage (Vello et al., 1984) after three cycles of recurrent selection. The percentages of PI parentage were 100%, 75%, 50%, 25%, and 0% in populations AP10, AP11, AP12, AP13, and AP14, respectively. Each population contained 200 F_4-derived lines from different intermatings

of 40 PIs and 40 N.A. cultivars. After seed yield data were collected from the 200 F_4-derived lines in each cycle, the highest yielding 20 individuals were selected as parents for the next cycle. To test the rate of yield gain, only the 20 parents for each cycle were analyzed byIninda et al. (1996). The population parents were planted in a randomized complete block design (RCBD) in three environments with two replications. Seed yield data were adjusted for plant maturity. The response to selection was calculated by linear regression and the genetic variance estimates for seed yield were based on 90 random F_4-derived lines from each selection cycle. These 90 random lines were planted in a RCBD in two environments with two replications. Significant differences between populations were determined by ANOVA using the linear response data.

The linear response to selection was significant ($P < 0.01$) for all populations. However, the differences in percent yield gain were not significantly different ($P < 0.05$) among populations AP 10 (2.5%), AP 11 (2.0%), AP 12 (3.1%), and AP 13 (2.8%), only AP 14 (5.4%) had a significantly different response. The AP 14 population also had the highest parental seed yield mean and the highest yielding parent in each cycle, the greatest average seed yield in cycle 3 (3291 kg ha $^{-1}$), and the genetic variance within each cycle was higher than most of the other populations. Having one of the highest genetic variances come from AP 14 was not expected due to the fact that it was derived from the most related pedigree and should have the least amount of genetic variation. This unexpected result may be due to each cycle being limited to the alleles in the 20 highest yielding plants from the previous cycle and that all of the populations may have become more similar.

Although, there was no significant difference in the rate of genetic gain between the four populations with PI parentage, the average seed yield between the four populations in cycle 3 showed a range of 2826 to 3053 kg ha $^{-1}$ with AP 10 and AP 11 being significantly lower yielding than AP 12 and AP 13. This indicates that although the rate of increase may not have been significantly different among AP 10, AP 11, AP 12, and AP 13, the populations with the lower percentages of PI parentage tended to yield more than populations with higher percentages. This could be due to difficulties in

eliminating deleterious traits and/or an increased number of beneficial alleles derived from the N.A. parents. Plant maturity was found to significantly increase across selection cycles for all populations, except AP 13, and plant lodging was found to significantly decrease ($P < 0.01$) across selection cycles for populations AP 10 and AP 13.

TheIninda et al. (1996) study showed that after three cycles of recurrent selection, the population with 0% PI parentage (AP 14) had a significantly higher rate of genetic improvement for seed yield when compared with populations containing 25 to 100% PI parentage. It also showed that there is no significant difference in the rate of genetic improvement for seed yield in populations containing 25 to 100% PI parentage. However, the populations containing 50 to 100% N.A. parentage did have significantly greater yields than the populations containing 0 to 25%. It seems that utilizing populations derived from elite cultivars continues to be the quickest means of developing higher yielding cultivars when conventional breeding methods are used. Still, there was some genetic gain made in AP 10, AP 11, AP 12, and AP 13 so it would be interesting to see what alleles were selected for in these four populations and whether or not the same alleles can be found in each population. If the same alleles were selected for across the four populations, these alleles would have been derived from a PI parent and this parent could be a potential source for adding genetic variability to the N.A. gene pool.

It has been previously indicated (Gizlice et al., 1994) that utilizing diverse northern cultivars and exotic germplasm could potentially contribute additional genetic variation to southern cultivars. Sneller et al. (1997) conducted a study to determine whether utilizing exotic germplasm with more acceptable agronomic traits or northern cultivars would lead to greater genetic variability and increased seed yields in southern cultivars. This study was conducted with 15 southern cultivars, 31 PIs in MGs similar to southern cultivars, 11 populations derived from southern x northern cultivar crosses, and 9 populations derived from southern x southern cultivar crosses. Each population consisted of 70 F_4-derived lines. The populations were planted in a RCBD in five environments with two replications. Fifty three RFLP probes were used to detect 60

polymorphic markers among the PIs, 57 southern cultivars, and the five northern cultivars that were used as parents.

The analysis showed that both the PIs and the southern x northern cultivar populations had seed yields significantly less than that of the southern cultivars and the southern x southern cultivar populations which were not significantly different from each other. Close to 69% of the PIs yielded less than the lowest yielding southern cultivar, no PI yielded higher than the average southern cultivar, around 78% shattered more, and over 33% had greater lodging. Nearly 46% of the southern x northern cultivar populations had lower average seed yields than the lowest yielding southern cultivar. The southern x northern cultivar populations were used to estimate the performance of the northern parents (NK S19-90, Archer, Sturdy, DSR-262, and ASG A2234) in the southern environments. Due to their early maturities, the northern parents were not included in the study. This was done by assuming that the mean of the population was equal to the mid-parent value and the northern parent was estimated as: southern parent yield + [2 x (yield of population - southern parent yield)]. Three of the five northern parents (NK S19-90, Sturdy, and DSR-262) had estimated mean seed yields greater than the average seed yield of the PIs.

To determine the genetic variability between the genetic pools, the RFLP data were used in a cluster analysis which created three clusters: one with the northern cultivars, one with 25 of the PIs, and one with the southern cultivars and six of the PIs. The cluster analysis results indicate that the PIs are generally genetically diverse from both northern and southern cultivars and that the northern cultivars are distinct from southern cultivars. Combining the agronomic data and the cluster analysis results, Sneller et al. (1997) concluded that PIs with agronomic qualities suitable for southern environments are competitive with northern cultivars and that the use of both diverse northern cultivars and PIs can contribute additional genetic diversity and potentially increase seed yields in southern soybean cultivars.

Another study that investigated the potential use of exotic germplasm was conducted by Thompson and Nelson (1998a). Their objective was to assess progeny with 25% to 100% exotic germplasm (based on pedigree) for seed yield potential. Using early generation yield data, 57 F_8, F_9, or F_{10} experimental lines derived from 13 crosses between seven PIs and seven N.A. cultivars were selected. Each line was characterized by the 35 core RAPD primers described by Thompson and Nelson (1998b) and these RAPD data were combined with the 28 N.A. ancestor and PI RAPD data from Thompson et al. (1998c). The experimental lines were planted in a RCBD in seven environments with two replications, and were separated into three tests based on MG. Significant differences between lines were determined by ANOVA and the genetic distances between lines were determined by cluster analysis of the RAPD data.

All of the exotic PI parents and many progeny with 100% PI genetics were among the lowest yielding lines in each MG. The exception was the 100% PI experimental line LG85-3343 which was the sixth highest yielding MG II entry with an average seed yield of 2975 kg ha^{-1}. There were six cases where the experimental lines with < 100% PI genetics yielded significantly ($P < 0.05$) higher than their N.A. parent. Four of these lines may have yielded more due to later maturity, LG91-5270 (2959 kg ha^{-1}) was 5 days later than Beeson 80 (2760 kg ha^{-1}) and LG91-7359 (3066 kg ha^{-1}), LG91-7323 (3094 kg ha^{-1}), and LG91-7350 (3205 kg ha^{-1}) were all about 18 days later than BSR 101 (2834 kg ha^{-1}). However, the other two experimental lines were LG90-2614 (2958 kg ha^{-1}) and LG86-7537 (3202 kg ha^{-1}) and they matured equal to and 4 days earlier than Beeson 80 and Lawrence (2847 kg ha^{-1}), respectively. The cluster analysis results showed that the experimental lines generally clustered with their N.A. parents. This may be the result of selecting progeny with favorable agronomic characteristics which may have been inherited from the N.A. parent and would have resulted in the experimental lines being more genetically similar to their N.A. parent.

Several significantly different low-yielding lines were also included in the cluster analysis so that genetic distances can be compared with their high-yielding full-sibs. It was found, that in some cases the low-yielding lines were more closely related to the

N.A. parents than were the high-yielding lines. For example, the pairwise distance between the high-yielding LG91-5270 (2959 kg ha^{-1}) and Beeson 80 was 0.45, while the distance between the low-yielding LG91-5258 (2609 kg ha^{-1}) and Beeson 80 was 0.35. Also, the distance between the high-yielding LG91-5644 (2459 kg ha^{-1}) and BSR 101 was 0.47, while the distance between the low-yielding LG91-5602 (2459 kg ha^{-1}) and BSR 101 was 0.43. The comparisons between these high and low-yielding full-sibs indicate that the genetic diversity found in the PI parents has been maintained in these high-yielding progeny, even after selection for seed yield, and that Exotic germplasm may offer favorable genetics for seed yield that may not be found in their N.A. counterparts.

Smalley et al. (2004) used the yield data and the AP 10, AP 12, and AP 14 populations from the Ininda et al. (1996) study along with a fourth selection cycle to test the hypothesis that beneficial PI alleles for increased seed yield were conserved through the recurrent selection process. Smalley et al. (2004) genotyped the 40 PI parents, the 39 N.A. parents (one line did not germinate), the 20 highest yielding Cycle 0 lines in AP 10 and AP 14, the 13 highest yielding Cycle 4 lines in AP 12, and the 15 highest yielding Cycle 4 lines in AP 10 and AP 14 with 184 fluorescent SSR markers, of which, only 54 SSR markers were polymorphic. The polymorphic markers showed that there were 16 alleles at 15 loci unique to the PI parents, 9 alleles at 9 loci unique to the N.A. parents, and 41 alleles at 36 loci found in both the PI and N.A. parent genetic pools. The allele frequencies were then calculated for each population and these data were used to determine the allele frequency changes between the parents used for the Cycle 0 populations and the highest yielding lines in the Cycle 4 populations. The 20 highest yielding Cycle 0 lines were used to differentiate between the allele frequency changes that would have occurred due to the selection on alleles associated with QTL for seed yield and the allele frequency changes due to the initial genetic drift which would have influenced the following selection cycles. Any of the SSR marker alleles that had a significant ($P < 0.05$) deviation from the expected frequency were inferred to be associated with alleles that were selected for and were determined to be QTL for seed yield.

There were several alleles with significant frequency changes in all of the populations: 27 alleles at 25 loci in AP 10, 21 alleles at 20 loci in AP 12, and 19 alleles at 18 loci in AP 14. There were a low number of alleles unique to the PI parents with significant frequency changes in cycle 4. This could have been due to the rarity of these PI alleles and their higher probability of being lost in the beginning selection cycles. In total, 43 QTL for seed yield were identified and of these QTL, only two were associated with SSR markers (Satt436 (LG D1a+Q, chr 1) and Satt317 (LG H, chr 12)) that were unique to the PI parents and had a significant allele frequency increases across AP 10 and AP 12. These results show that there were alleles derived from the PI parents, which were associated with putative QTL for seed yield, conserved throughout the selection process. However, before these putative QTL can be used in a breeding program, they should first be confirmed in a more direct manner by testing the effect of having the beneficial alleles being present or absent in different lines in a separate confirmation population.

When using recurrent selection, Ininda et al. (1996) determined that the use of populations developed from elite x elite cultivar parentage is the most effective means of quickly developing higher yielding cultivars, when compared to populations derived from 25 to 100% PI parentage. Sneller et al. (1997) showed that even though southern cultivars have better performance in southern environments than both diverse northern cultivars and PIs with southern maturities, both northern cultivars and PIs have the potential to add beneficial genetic variability to the southern gene pool. The increased seed yield potential that can be found in Exotic germplasm was observed by Thompson and Nelson (1998a). They were able to show evidence that unique alleles found in PI parents were present in the high-yielding progeny and that these unique alleles may have contributed to the higher seed yield potential observed in these progeny. A report by Smalley et al. (2004) was able to identify two genetic regions in the Ininda et al. (1996) populations derived from Exotic germplasm that could be linked to QTL for seed yield. The two putative QTL were identified at the SSR loci Satt436 and Satt317, and are potentially unique to the N.A. gene pool. However, these putative QTL for seed yield should first be confirmed before being used in a breeding program.

QTL mapping detects significant differences between phenotypic means by dividing a mapping population into groups based on the allele class at a given marker locus. If a significant difference is detected, then a QTL affecting the agronomic trait of interest is assigned to that marker locus (Collard et al., 2005). The benefits of QTL mapping are that it does not depend on transgressive segregates (reducing the cycles of selection needed to identify beneficial alleles), both large and small effect QTL can be detected, allows breeders to focus breeding efforts to experimental lines containing beneficial alleles thus reducing the amount of wasted expense, and allows for smaller populations when compared to conventional breeding methods (Tanksley, 1993). The goal is that once these unique beneficial QTL are identified, they can then be backcrossed out of their PI donor parent and into elite N.A. cultivars so as to improve the overall N.A. seed yield and add genetic variability. The following studies are examples of how breeders have applied QTL mapping with the goal of identifying QTL for seed yield.

QTL Mapping for Seed Yield

In 2001, Specht et al. conducted a QTL mapping study for seed yield under drought conditions with a mapping population consisting of 236 $F_{7:11}$ lines derived from a cross between two exotic *G. max* accessions, Minsoy x Noir 1. The mapping population was planted in a RCBD in two environments with two replications. A water gradient was designed to replenish 0 to 100% of the transpirational water loss weekly. The experimental design was a split-split plot with the irrigation as the main plot, the lines divided into 29 maturity blocks, and the individual lines as the sub-subplots. Associations with 665 RFLP, SSR, and morphological markers with plant height, lodging, maturity, and seed yield were identified by interval mapping and composite interval mapping with a significant logarithm of odds (LOD) score of 3.4.

The analysis revealed a QTL for seed yield associated with the genetic region between Satt277 and Satt489 (LG C2, chr 6) that accounted for 7 to 13% of the variability in each of the environments. The beneficial allele was derived from the Minsoy parent and an additive effect of 123 to 189 kg ha^{-1} was observed for seed yield.

The LG C2 QTL was also associated with increased plant height, increased lodging, and later maturity, and accounted for 5 to15%, 5%, and 13 to15% of the variability for each trait, respectively. Specht et al. (2001) attributed the QTL association with seed yield as a result of the segregation of the maturity gene *E_1* which resides 4 cM below Satt277. A second QTL for seed yield was associated with the genetic region between Satt150 and Satt567 (LG M, chr 7) that accounted for 33 to 38% of the variability in each of the environments. The beneficial allele was derived from the Noir 1 parent and an additive effect of 265 to 311 kg ha^{-1} was observed for seed yield. The LG M QTL was also associated with increased plant height and delayed maturity, and accounted for 14 to 15% and 23 to 29% of the variability, respectively. This QTL association with seed yield was also assumed to be the result of another, yet unknown, *E* gene which hastened maturity of the lines homozygous for the Minsoy-derived genetic region.

Four other QTL for seed yield were identified in only one environment and were associated with Sat_074 (LG F, chr 13), Satt314 (LG H, chr 12), the region between G173B and *Dt_1* (LG L, chr 19), and *Rpg_4* (LG N, chr 3). The beneficial alleles for the QTL associated with Satt314 and *Rpg_4* were derived from the Minsoy parent and had additive effects of 84 and 91 kg ha^{-1}, respectively. The QTL for seed yield associated with Satt314 was attributed to the *Ps* gene for semi-sparse pubescence derived from the Noir 1 parent. The *Ps* gene tends to cause plants to be more lodging prone which will affect yield measurements. The Minsoy-derived QTL for seed yield associated with the *Rpg_4* bacterial blight (*Pseudomonas glycinea* Coerper) resistance gene was only significant in an environment where bacterial blight was observed late in the season. The beneficial alleles for the QTL associated with Sat_074 and the region between G173B and *Dt_1* were derived from the Noir 1 parent and had additive effects of 80 and 142 kg ha^{-1}, respectively. The QTL for seed yield associated with the region between G173B and *Dt_1* was attributed to the *Dt_1* gene for growth habit. The determinant Minsoy-derived *dt_1* allele reduced plant heights 9 to 10 cm and the planting density caused the internodes to be shortened, which increased combine harvest loss. Due to five of the six putative QTL for seed yield having been explained by existing knowledge of governing genes near the QTL, inconclusive results in terms of identifying pure QTL for seed yield in drought

conditions were obtained. However, a positive QTL for seed yield was identified as having significance in one environment associated with Sat_074 which could be further investigated.

Another QTL mapping study was conducted by Wang et al. (2004) where four putative QTL for seed yield were identified in the *G. soja* accession PI 468916. A total of 468 BC_2F_4 lines were developed from a cross between IA2008 x PI 468916 which were then separated into five populations: 110 lines in populations 324B and 330A, and 79, 57, and 112 lines in populations 326, 334A, and 338B, respectively. The populations were planted in a RCBD in four environments with two replications, except for 326 and 334A which were only grown in three environments. Each line was genotyped with 302 SSR markers of which, only 52% were polymorphic in at least one population and 26% were polymorphic in at least two populations. Composite interval mapping with a LOD score of 3.0 (equal to an experiment-wise error of $P < 0.05$) was used to determine the marker-trait associations.

There were no lines in any of the populations with seed yields higher than IA2008 (the recurrent parent). Four QTL derived from IA2008 were significantly ($P < 0.001$) associated with seed yield in the regions between Satt134 and *T* (LG C2, chr 6), Satt137 and Satt178 (LG K, chr 9), Satt567 and Satt463 (LG M, chr 7), and close to Satt575 (LG E, chr 15) with additive effects ranging between 65 and 147 kg ha^{-1}. The QTL on LG C2 was significant in 334A across all environments, the QTL on LG E and M were significant in 338B across three and four environments respectively, and the QTL on LG K was significant in 324B and 330A across all environments. Plant height was also significantly associated ($P < 0.01$) with all of the identified QTL. The QTL on LG E, K, and M were associated with an increase in plant height of approximately 2 cm and the QTL on LG C2 decreased plant height by 5 cm. The QTL on LG C2 hastened maturity by almost 8 days whereas the QTL on LG M delayed maturity by almost 3 days. Plant lodging was associated with the QTL on LG K with an average lodging score of 0.4 less than the PI 468916-derived genetic region.

Due to not having identified any QTL for seed yield derived from PI 468916, a lower significance threshold of $P < 0.05$ comparison-wise error was used and four putative QTL for seed yield derived from PI 468916 were associated with Satt350 (LG D1b, chr 2), Satt491 (LG E, chr 15), Satt006 (LG L, chr 19), and Satt257 (LG N, chr 3) with additive effects ranging from 35 to 62 kg ha^{-1}. Wang et al. (2004) attributed the lack of PI 468916-derived QTL for seed yield to not having any selection pressure for seed yield in the BC_1 population. This led to poor plant performance in the following generations and a tendency for pod dehiscence, which is commonly found in *G. soja* accessions. The lack of selection pressure may have led to unreliable data which hindered the QTL mapping process. When compared to *G. soja*, *G. max* accessions in general have less severe deleterious agronomic traits (Li et al., 2008a; Sneller et al., 1997) (ex. pod dehiscence); therefore utilizing *G. max* accessions may yield better QTL mapping results for seed yield.

Two exotic *G. max* accessions were used by Kabelka et al. (2004) to identify several putative QTL for seed yield. The mapping population consisted of 167 F_5 lines derived from a cross between the cultivar BSR 101 and the experimental line LG82-8379. LG82-8379 was derived from a cross between PI 68508 and FC 04007B. The mapping population was divided into three maturity sets with Set 1 having 54 lines (MG II), Set 2 having 55 lines (MG III), and Set 3 having 58 lines (MG IV). The population was planted in 12 environments in a nested split-plot design with the maturity sets as the main plot and the lines as the sub-plots with two replications. Data were collected on plant height, lodging, maturity, seed oil and protein concentration, and seed yield. DNA samples from the BSR 101 parent and seven DNA bulks of 22 random F_5 lines were used to identify 145 polymorphic SSR markers. The 22 F_5 lines were used instead of LG82-8379 because this experimental line is heterogeneous and the exact parental genotypes were unknown. The marker-trait associations were identified by single-trait analysis and composite interval mapping with a significant LOD score of 2.5. The significant difference ($P < 0.05$) between the homozygous marker classes for a given trait was determined by ANOVA.

In total, 15 putative QTL for seed yield were identified, nine of which had the beneficial allele derived from LG82-8379. Many of the QTL were also associated with other agronomic traits. The nine QTL derived from LG82-8379 are of particular interest because they would have been derived from one of two PIs that have been shown to be genetically distinct from the major N.A. ancestors (Thompson et al., 1998c). The QTL for seed yield were associated with SSR markers Satt225 (LG A1, chr 5), Satt168 (LG B2, chr 14), Satt363 (LG C2, chr 6), Satt186 (LG D2, chr 17), Satt394 (LG G, chr 18), Satt142 (LG H, chr 12), Satt544 (LG K, chr 9), Satt308 (LG M, chr 7), and Satt358 (LG O, chr 10). The QTL for seed yield associated with Satt363 and Satt358 were identified across all maturity sets whereas the QTL associated with Satt225, Satt394, and Satt544 were only found in Set 1, Satt168 and Satt186 in Set 2, and Satt142 and Satt308 in Set 3. The LG82-8379 marker class showed a 2.1% yield increase over the BSR 101 marker class at Satt225, 2.4% at Satt168, 2.2% at Satt363, 2.4% at Satt186, 2.4% at Satt394, 5.4% at Satt142, 4.1% at Satt544, 3.2% at Satt308, and 1.7% at Satt358. Seed protein content was also associated with Satt142, Satt168, Satt225, Satt308, Satt358, Satt363, and Satt544 with differences between marker classes ranging between -3 and +3 g kg^{-1}. Plant height was associated with Satt142 and Satt544 with the LG82-8379 marker class being 2 cm shorter and 3 cm taller respectively, and plant maturity was also associated with Satt544 with the LG82-8379 marker class maturing 1 day later. Kabelka et al. (2004) stated that the LG82-8379-derived QTL associated with Satt358 should be of particular interest because of the additive effects of +47 kg ha^{-1} for seed yield and +3 g kg^{-1} for seed protein content. It was stated that the experiment-wise error rate was close to 100%, so it should be assumed that some of these putative QTL are false positives. Also, the number of lines within each maturity set was low and could have resulted in an overestimation of the QTL effects.

A QTL mapping study conducted by Guzman et al. (2007) identified eight putative QTL for seed yield among three different backcross (BC) populations. The three mapping populations were derived from Beeson 80 x LG82-8224, Kenwood x LG94-1713, and Lawrence x LG96-6607, and consisted of 68 BC_2F_5 lines, 74 BC_1F_5 lines, and 94 BC_3F_2 lines, respectively. All three of the experimental lines were derived from

crosses that involved at least one exotic *G. max* accession. In each of the BC populations, the lines crossed to their respective recurrent parent had greater seed yields than the recurrent parent. The parents of each mapping population were genotyped with 602 SSR markers which resulted in 45 polymorphic markers in the Beeson 80 population, 84 in the Kenwood population, and 30 in the Lawrence population. There were very few polymorphic markers because within each population the lines would theoretically have 75 (BC_1) to 94% (BC_3) of the genes of the recurrent parent. The mapping populations were planted in a RCBD in eight environments with two replications and data were collected on plant height, maturity, and seed yield. The marker-trait associations were identified by single-trait analysis and interval mapping in the Beeson 80 and Lawrence populations, and composite interval mapping in the Kenwood population with significant LOD scores of 1.5. The significant difference ($P < 0.05$) between the homozygous marker classes for a given trait was determined by ANOVA and the means were adjusted with nearest neighbor analysis.

The QTL analysis identified two putative QTL for seed yield in the Beeson 80 mapping population and three in both the Kenwood and Lawrence populations where the beneficial alleles were derived from the experimental lines. The putative QTL for seed yield in the Beeson 80 population were associated with Satt215 (LG J, chr 16) and Satt547 (LG J, chr 16) with additive effects of 0.04 and 0.09 Mg ha^{-1}, respectively. The QTL associated with Satt547 also increased plant height with an additive effect of 1.2 cm. The putative QTL for seed yield in the Kenwood population were associated with Satt557 (LG C2, chr 6), Satt405 (LG J, chr 16), and Satt477 (LG O, chr 10) with additive effects of 0.11, 0.05, and 0.07 Mg ha^{-1}, respectively. The QTL associated with Satt557, which is also very close to the maturity gene *E_1*, was associated with increased plant height and delayed maturity with additive effects of 4.1 cm and 3.3 days respectively. The QTL associated with Satt477 also delayed plant maturity with an additive effect of 2 days. The putative QTL for seed yield in the Lawrence population were associated with Satt300 (LG A1, chr 5), Satt474 (LG B2, chr 14), and Satt622 (LG J, chr 16) with additive effects of 0.06, 0.04, and 0.04 Mg ha^{-1}, respectively. Guzman et al. (2007) stated that these putative QTL for seed yield were identified with a significant LOD score of 1.5

which is equal to a comparison-wise probability of 0.05 for a type 1 error which resulted in the experiment-wise error rate being much greater than 0.05.

Palomeque et al. (2009) conducted a QTL mapping study where the primary objective was to determine whether or not evaluation of a mapping population in different mega-environments results only in mega-environments-specific QTL, only universal QTL, or both. The objective of the Palomeque et al. (2009) study is not in direct alignment with the objectives of this current study. However, due to the germplasm selected by Palomeque et al. (2009), their report allows for the opportunity to potentially determine if using a modern PI would contribute a greater number of QTL or larger effect QTL for seed yield. The mapping population consisted of 93 $F_{4:7}$ lines derived from a cross between Heinong 38, a cultivar from Heilongjiang, China, and OAC Millennium, a N.A. cultivar. The experimental design was a rectangular lattice design planted in six environments in Canada with two replications. The parents were characterized with 450 SSR markers of which, only 105 SSR markers were found to be polymorphic. The marker-trait associations were identified using single-trait analysis, interval mapping with a significant LOD score of 2.6 ($P < 0.01$), and multiple QTL mapping where the QTL identified by interval mapping were used as cofactors.

The multiple QTL mapping technique identified one putative QTL for seed yield in two environments associated with the genetic region between Satt139 and Sat_042 (LG C1, chr 4) which was derived from the Heinong 38 parent and accounted for 16 to 20% of the variability. The single-trait analysis method identified putative QTL for seed yield in three environments associated with Satt277 (LG C2, chr 6) and in four environments associated with Satt100 (LG C2, chr 6). Both of these QTL were derived from Heinong 38, are estimated to be 10 cM apart, and accounted for 15 to 18% and 13 to 35% of the variability, respectively. Single-trait analysis also identified a putative QTL for seed yield associated with Satt162 (LG I, chr 20) which was derived from OAC Millennium, was significant in two environments, and accounted for 9 to 12% of the variability. Although most of the putative QTL for seed yield that were identified came from the exotic modern cultivar and each QTL accounted for 15 to 35% of the variability,

the Palomeque et al. (2009) results show no obvious advantages to using exotic modern cultivars over historical PIs when compared to the findings by Guzman et al. (2007), Kabelka et al. (2004), Specht et al. (2001), and Wang et al. (2004). However, the confirmation of the putative QTL identified in the aforementioned studies have yet to be pursued so the relationship between the number of confirmed QTL and the historical age of the exotic donor parent has yet to be determined.

QTL mapping has allowed breeders to identify several putative QTL for seed yield with the potential to both improve the yield of N.A. cultivars and expand the N.A. genetic base. Specht et al. (2001) identified six putative QTL for seed yield in drought conditions from two *G. max* PI parents. However, five of the six QTL (the QTL on LG C2 being the exception) could be explained by the genes dt_1, E_1, *Ps*, R_{pg4}, or an unknown *E* gene that may have been segregating in the mapping population. Wang et al. (2004) identified four *G. soja*-derived putative QTL for seed yield but with a more relaxed significance cut-off so these putative QTL are questionable. Both Kabelka et al. (2004) and Guzman et al. (2007) were able to identify putative QTL for seed yield that were derived from exotic *G. max* germplasm. Kabelka et al. (2004) identified nine putative QTL in a single mapping population and Guzman et al. (2007) was able to identify two putative QTL for seed yield in a single BC mapping population and three putative QTL in two other BC populations. Both Kabelka et al. (2004) and Guzman et al. (2007) have stressed that the confirmation of these putative QTL for seed yield should be pursued.

The goal of the Palomeque et al. (2009) study to determine whether or not more beneficial QTL for seed yield could be derived from an exotic modern cultivar as opposed to a historic PI was unsuccessful. The identification of three putative QTL for seed yield by Palomeque et al. (2009) was achieved using a P value of 0.01. It would be difficult to determine whether or not more QTL for seed yield were identified because of the high P value that was used which may have prevented smaller effect QTL from being detected. Another possible way to compare the different QTL mapping studies would be to test the putative QTL identified in the various reports in confirmation populations.

This would allow one to calculate the success rate of each QTL mapping study in terms of identifying real QTL for seed yield.

In order to confirm a QTL in accordance with the Soybean Genetics Committee (Soybase: http://www.soybase.org/ (verified: July 2010)) the confirmation population should be created from a separate meiotic event with at least one parent in common between the original mapping study and the confirmation study and be tested in different environments than the mapping population. Separate meiotic events can be a new set of F_2 lines derived from the original cross, a new backcross population, or lines derived from a plant in the original mapping population that was heterozygous for the region carrying the QTL. The QTL must also be confirmed at a P value ≤ 0.01. Once the above criteria have been met the confirmed QTL receives a *cq* designation to notify breeders that it has been confirmed under these strict conditions. QTL can be confirmed by other means, for example, selecting homozygous sub-lines from a heterogeneous line found in the original mapping population can be used as the confirmation population. However, the confirmed QTL will not receive the *cq* designation. There have been very few putative QTL tested in a confirmation population which is unfortunate due to the importance of the confirmation process in separating the real QTL from the false positives. The putative QTL for seed yield identified in the aforementioned studies are listed in Table 1. Studies that have attempted to confirm putative QTL for seed yield are summarized below.

Confirmation of Putative QTL for Seed Yield

Orf et al. (1999) used three separate mapping populations derived from different combinations of three parents and used the data from each population to confirm the results found in the other populations. The mapping populations were derived from Minsoy x Noir 1 (MN), Minsoy x Archer (MA), and Noir 1 x Archer (NA). Both Minsoy and Noir 1 are exotic *G. max* accessions (Specht et al., 2001) and Archer is a N.A. cultivar. The three mapping populations consist of 233 F_7 lines in the MA population and 240 F_7 lines in the MN and NA populations. All three mapping populations were planted

in a RCBD in four environments with two replications for the MA and NA populations and three replications for the MN population. Data were collected for the MA and NA populations and the data obtained by Mansur et al. (1996) was used for the MN population. The marker-trait associations using RFLP and SSR markers were determined by interval mapping and a significant LOD score of 3.0.

No significant QTL for seed yield was identified in the MA population. Between the MN and NA populations, four QTL for seed yield were associated with Satt277 (LG C2, chr 6), Satt002 (LG D2, chr 17), Satt144 (LG F, chr 13), and Satt150 (LG M, chr 7). The Archer-derived QTL for seed yield associated with Satt002 and Satt144 were identified in the NA population accounting for 8 and 13% of the variation respectively, but the putative QTL were not confirmed in the other populations. The Noir 1-derived QTL for seed yield associated with Satt150 was identified in the MN population accounting for 19% of the variation, but was also unable to be confirmed in the other populations. The Noir 1-donated QTL for seed yield associated with Satt277 was identified in the NA population and accounted for 11% of the variation and was confirmed in the MN population. A QTL for seed weight was also associated with Satt277 with the Noir 1 allele having the most benefit. Of the four QTL that were identified by Orf et al. (1999), only one was able to be confirmed in a different genetic background. The uniqueness of the Noir 1-derived QTL associated with the Satt277 locus will need to be shown before the significance of this QTL in the expansion of the N.A. gene pool can be determined. A first step could be to identify the allele(s) that are already common in N.A. cultivars at the Satt277 locus. If the Noir 1-derived allele remains unique, then further testing of the linked QTL and its influence in different backgrounds can be pursued.

As stated previously, the northern and southern N.A. soybean cultivar gene pools are genetically distinct (Gizlice et al., 1994; Thompson et al., 1998c) so utilizing northern germplasm could potentially expand the southern genetic base. Orf et al. (1999) identified two Archer-derived putative QTL for seed yield in northern environments in a cross with a northern MG PI, Noir 1. Before the putative QTL for seed yield can be

utilized in southern cultivars, it must be tested in southern environments and genotypes to determine its potential value. Confirmation populations were created by Reyna and Sneller (2001) to test the Archer-derived QTL in southern genetic backgrounds (Pioneer 9641 and Asgrow A5403). The putative QTL were associated with Satt002 (LG D2, chr 17), Satt144 (LG F, chr 13), and Sct_33 (LG F, chr 13). For each marker, four sets of near isogenic lines were developed with each set containing 2 to 10 $F_{7:9}$ homozygous lines derived from the same F_6 plant. The confirmation population testing the QTL for seed yield associated with Satt144 had a total of 8 and 7 lines homozygous for the Archer and Pioneer 9641 marker class, respectively. The confirmation populations testing the QTL associated with Satt002 and Sct_33 consisted of three sets of near isogenic lines derived from Archer x Pioneer 9641 and one set of near isogenic lines derived from Archer x Asgrow A5403. To confirm the QTL associated with the Satt002 marker, there were a total of 8 lines with the Archer allele and 9 lines with the allele from the southern cultivars. The confirmation populations with the Sct_33 marker had a total of 9 lines with the Archer allele and 11 lines with the allele from the southern cultivars. Each confirmation population was tested in a split-plot design planted in four Arkansas environments, with the sets as the whole plots and the lines as the sub-plots, with two replications in two environments and three replications in the other two. The significant differences ($P < 0.05$) between the Archer homozygous marker classes and the southern parent homozygous marker classes were determined using ANOVA.

No significant difference was found between the Archer marker class and the southern parent marker class in any of the confirmation populations for plant height, maturity, or seed yield. The Reyna and Sneller (2001) confirmation study was unable to confirm the putative QTL for seed yield (Orf et al., 1999) derived from Archer in southern environments and genetic backgrounds. It is quite possible that the reason the Archer marker class did not show a significant difference from the southern marker class is that southern cultivars may already posses an allele of equal effect to the Archer allele. Another reason may be that, although the Archer allele was shown to be superior to the alleles found in Minsoy and Noir 1 when tested in northern environments, the Archer

allele may not perform in the same manner in southern environments where the Archer allele may be inferior to alleles in southern cultivars.

In 2003, Concibido et al. conducted a QTL mapping study with the objective of identifying QTL for seed yield derived from a *G. soja* accession (PI 407305). PI 407305 was crossed to a *G. max* line (HS-1) to create a mapping population consisting of 265 BC_2F_1 lines. The mapping population was tested in 10 environments which were considered replications. The linkage map was created by interval mapping using 212 amplified fragment length polymorphism (AFLP) and four morphological markers with a marker linkage LOD score of 2.0 and QTL association of $P < 0.001$. Due to several complications with the AFLP markers, Concibido et al. (2003) developed a F_2 population consisting of 96 lines from the original HS-1 x PI 407305 cross. These 96 lines were analyzed with an additional 600 SSR markers to aid in assigning the AFLP markers to their respective LG. Single-factor ANOVA was conducted to determine the significant difference ($P < 0.0001$) between the homozygous *G. soja* and *G. max* marker classes. A single QTL for seed yield was identified within the 42.71 cM region between the SSR markers Satt168 and Satt560 (LG B2, chr 14) with the *G. soja* marker class yielding 9.3% higher than the *G. max* marker class.

To validate the results of the mapping population, Concibido et al. (2003) created a further inbred population of BC_2F_4 lines developed from a single heterozygous BC_2F_3 plant from the original cross. From this single plant, 25 plants for each homozygous marker class were selected and the seed within each class were bulked. The validation population was planted in a RCBD in two environments with four replications. Single-factor ANOVA was conducted to determine the significant difference ($P < 0.05$) between the homozygous *G. soja* and *G. max* marker classes in the region between Satt560 and Satt168. It was found that the marker classes were significantly different in only one environment and within that environment, the *G. soja* marker class yielded 18% higher than the *G. max* marker class and was on average 6.5 cm taller. In order to determine whether or not the differences in plant height between marker classes significantly affected seed yield, five plants from each of the lines homozygous for the *G. soja* marker

class and heterozygous or homozygous for the *G. max* marker class were selected to test the differences in pod number on the main stem. No significant difference in pod number between marker class was detected so it was determined that plant height was not a factor in the significant difference for seed yield.

To determine the uniqueness of the *G. soja*-derived QTL for seed yield in the region between Satt168 and Satt560, Concibido et al. (2003) characterized 70 lines including wild and commercial soybean varieties with SSR markers spanning the region between Satt168 and Satt560. It was determined, that the *G. soja* haplotype was in fact unique to PI 407305 when compared to this limited sample of genotypes. Concibido et al. (2003) also tested the utility of the QTL for seed yield in elite backgrounds by identifying a BC_2F_4 plant from the mapping population that was homozygous for the *G. soja* marker class and crossing the BC_2F_4 plant to Asgrow lines AG2401, QR4459, QP4459, QR4544, and QP4604. Five BC_1 populations were created with the Asgrow lines as the recurrent parents. A single BC_2 population with AG4501 as the recurrent parent was also created. In each population, about 25 plants were identified as being homozygous within each of the marker classes and seed within each marker class was bulked. Each population was planted in a RCBD with two replications, AG4501 and AG2401 populations were grown in five environments, and the other populations were grown in one environment. Single-factor ANOVA was conducted to determine the significant difference ($P < 0.05$) between the homozygous *G. soja* and *G. max* marker classes.

Significant differences were only found in the AG4501 and QP4459 populations with the *G. soja* marker class yielding 9 and 5% higher than the *G. max* marker class, respectively. In the AG2401 population, the *G. soja* marker class yielded 7% more than the *G. max* marker class but the difference was determined to be non-significant. This is most likely due to the higher coefficient of variation (CV) (30%) found in this population, whereas the AG4501 population had a CV = 6% and the QP4459 population had a CV = 5%. The QR4459, QR5444, and QR4604 populations also showed a non-significant difference between the marker classes but the *G. soja* marker class on average

yielded 2% less in all three populations. The Concibido et al. (2003) results show that the *G. soja*-derived QTL for seed yield in the region between Satt168 and Satt560 could be a major QTL for increased seed yield (AG4501 and QP4459 populations) but is likely to already be present in the N.A. gene pool (AG2401, QR4459, QR5444, and QR4604 populations). Despite not showing a significant effect across all six genetic backgrounds, the *G. soja*-derived QTL for seed yield was confirmed in two different backgrounds which indicate that this QTL is real and may be useful in genetic backgrounds similar to AG4501 and QP4459.

Another QTL mapping and confirmation study using *G. soja* was conducted by Li et al. (2008a). A cross between *G. soja* accession PI 245331 and *G. max* variety 7499 was used to create two populations of BC_2F_4 lines. The first population was used as the QTL discovery population and consisted of 147 lines divided into three maturity sets with 49 lines each. The second population was used as the QTL confirmation population and consisted of 148 lines divided into three maturity sets with 49 lines in the first two sets and 50 in the third set. The lines in the confirmation population were developed from the same BC_2 plants as the discovery population. The experimental design was a split-plot design planted in four environments with the population as the main plot and the maturity sets as the sub-plots with two replications. The parents were characterized with 534 SSR markers of which only 120 were found to be polymorphic. The SSR markers were combined with the seed yield data from the discovery population and the marker-trait associations were made with marker regression, interval mapping, and composite interval mapping. A significant LOD score of 3.6 and 3.8 were used for interval mapping and composite interval mapping respectively, both were equal to a $P < 0.01$ significance.

All three statistical methods identified a single QTL for seed yield that was derived from the *G. soja* parent and associated with the genetic region between Satt050 and Satt511 (LG A1, chr 5). In all three methods, about 13% of the seed yield variation was explained by this QTL. The additive effects were 191 kg ha^{-1} for the marker regression method, 223 kg ha^{-1} for interval mapping, and 235 kg ha^{-1} for composite interval mapping. This QTL was further investigated in the confirmation population

using ANOVA to determine the significant difference ($P < 0.25$) between the *G. soja* and the *G. max* homozygous marker classes. Due to the relaxed significance cutoff for the confirmation population, the QTL for seed yield derived from the *G. soja* parent associated with the region between Satt050 and Satt511 was confirmed with an average seed yield of 6% greater than the *G. max* marker class. The confirmation of this QTL for seed yield is important due to the potential genetic variability that the *G. soja*-derived genetics can introduce into the N.A. gene pool (Hyten et al., 2006). However, the QTL was identified in similar environments and similar genetic backgrounds to the mapping population, so the utility of this QTL in different environments and backgrounds is still unknown. As was shown in the previous study (Concibido et al., 2003), QTL for seed yield may not perform consistently across environments and genetic backgrounds.

The confirmation of putative yield QTL after their identification should be a high priority. Due to the genetic complexity of yield, the performance of yield QTL can be highly dependent on the environment as seen in the studies conducted by Li et al. (2008) and Reyna and Sneller (2001), and the genetic background as seen in the studies conducted by Concibido et al. (2003), Li et al. (2008a), Orf et al. (1999), and Reyna and Sneller (2001). It has also been noted by Guzman et al. (2007) and Kabelka et al. (2004) that when mapping studies are being performed, there is greater concern for type 2 error, at the expense of committing more type 1 error, so there is already an innate high probability that a given putative QTL is a false positive. The putative QTL that have been confirmed in the aforementioned studies are listed in Table 1.

The objectives of this research are to confirm the putative PI-derived QTL for seed yield identified by Guzman et al. (2007) and Kabelka et al. (2004) and to confirm the PI parents from which the putative QTL for seed yield were derived.

MATERIALS AND METHODS

Confirmation Population Development

Populations were developed to confirm three putative yield QTL derived from the Kenwood x LG94-1713 mapping population (Guzman et al., 2007) and six putative yield QTL from the BSR 101 x LG82-8379 mapping population (Kabelka et al., 2004). The SSR loci associated with these QTL are Satt405 (linkage group (LG) J, chromosome (chr) 16), Satt477 (LG O, chr 10), and Satt557 (LG C2, chr 6) from Guzman et al. (2007) and Satt142 (LG H, chr 12), Satt168 (LG B2, chr 14), Satt225 (LG A1, chr 5), Satt358 (LG O, chr 10), Satt363 (LG C2, chr 6), and Satt544 (LG K, chr 9) from Kabelka et al. (2004).

Confirmation populations were created by identifying heterogeneous lines for the markers of interest within the original mapping populations. Lines from the BSR 101 x LG82-8379 mapping population heterogeneous for Satt142, Satt168, Satt225, Satt358, Satt363, and Satt544 and lines from the Kenwood x LG94-1713 mapping population heterogeneous for Satt405 were planted at Urbana, IL in 2006. DNA samples were collected from individual plants and genotyped for the segregating marker. Fourteen plants were genotyped for Satt405, 15 plants for Satt168, 16 plants for Satt544, 29 plants for Satt142 and Satt363, 30 plants for Satt225, and 45 plants for Satt358. No heterozygous plants were found for Satt142, Satt225, Satt363, and Satt544 so homozygous plants for the LG82-8379 or BSR 101 marker classes were selected from these genotyped plants and increased during the winter of 2006-2007 in Chile. These populations are in the F_8 generation and the number of lines homozygous for each marker class is presented in Table 2.

Heterozygous plants segregating for Satt168, Satt358, or Satt405 were identified in 2006 and a single heterozygous plant was selected for each marker of interest. The following December, 120 seeds from each of the heterozygous plants were planted in sand benches in the greenhouse. From the 120 seeds planted, 96 plants were genotyped

for the segregating marker and homozygous plants were selected, transplanted, and grown to maturity in the greenhouse. Seeds harvested from the selected plants were planted at Urbana, IL during the 2007 growing season for increase, populations are in the F_9 generation. The number of lines homozygous for each marker class is presented in Tables 2 and 3.

The two other confirmation populations developed from the Kenwood x LG94-1713 mapping population were created to test the two putative QTL for seed yield associated with the SSR markers Satt477 and Satt557. An inbred line segregating for both markers was selected from the Guzman et al. (2007) mapping population and was planted at Urbana, IL in 2004. One hundred eighty five plants were genotyped for both markers of interest and three plants were selected that were heterozygous for both loci. Seeds from the three selected plants were planted at Urbana, IL in 2005. From these three plant rows, 18, 21, and 27 plants where harvested and planted to plant rows at Urbana, IL in 2006. The plants within two plant rows within the population of 27 plant rows (Population A) and three plant rows within the population of 21 plant rows (Population B) were genotyped with the two markers of interest and a single heterozygous plant segregating for both markers of interest was selected from each population. The population of 18 plant rows was not used. The following December, 120 seeds derived from each of the selected plants were planted in sand benches in the greenhouse. Of the 120 seeds planted, 96 progeny were characterized by the two markers of interest and plants homozygous for at least one of the markers of interest were selected, transplanted, and grown to maturity in the greenhouse. Seeds harvested from the selected plants were then planted to plant rows at Urbana, IL during the 2007 growing season and each plant row became a single line within its confirmation population, populations are in the F_{11} generation. The number of lines homozygous for each marker class is presented in Table 2.

Field Procedures

The confirmation populations testing QTL at Satt225 and Satt363 were planted at Urbana, IL on May 18 in 2007 with two replications and at Bellflower (June 16) and DeKalb, IL (May 20) in 2008 and at Bellflower (May 30), DeKalb (June 7), and Urbana, IL (June 1) in 2009 with four replications. The confirmation populations testing Satt142 and Satt544 were planted at Urbana, IL on May 18, 2007 with two replications and at Arthur (May 21) and Urbana, IL (May 28) in 2008 and at Arthur (May 22), Hume (June 2), and Urbana, IL (June 1) in 2009 with four replications. The confirmation populations testing Satt168 and Satt358 were planted at Bellflower (June 16) and DeKalb, IL (May 20) in 2008 and at Bellflower (May 30), DeKalb (June 7), and Urbana, IL (June 1) in 2009 with two replications. The confirmation population testing Satt405 was planted at Bellflower (June 16) and Pontiac, IL (May 9) in 2008 and at Bellflower (May 30), Pontiac (May 26), and Urbana, IL (June 1) in 2009 with two replications. The two confirmation populations testing Satt477 and Satt557 were planted at Bellflower (June 16) and Pontiac, IL (May 9) in 2008 and at Bellflower (May 30), Pontiac (May 26), and Urbana, IL (June 1) in 2009 with two replications. All populations were planted in a randomized complete block design at every location. Planting was delayed in 2008 and 2009 at most locations due to wet field conditions.

Field plots in 2007 and 2009 at Urbana and 2008 and 2009 at Bellflower for the Satt168, Satt225, Satt358, Satt363, Satt405, Satt477, and Satt557 confirmation populations and for the Urbana 2007 and Hume 2009 locations for the Satt142 and Satt544 confirmation populations were four rows wide with 0.76 m spacing 3.05 m long. Seeds were sown at a rate of 30 seeds m^{-1}. Field plots at the Pontiac 2008 and 2009 locations for Satt477 and Satt557 Population A were four rows wide with 0.76 m spacing and were 3.2 m long. Seed were sown at a rate of 31 seeds m^{-1}. The field plots in the DeKalb 2008 and Pontiac 2008 and 2009 locations for the Satt168, Satt225, Satt358, Satt363, Satt405, and Satt477 and Satt557 Population B confirmation populations and Arthur 2008 and 2009 and Urbana 2008 and 2009 locations for the Satt142 and Satt544 confirmation populations were two rows wide with 0.76 m spacing and were 3.2 m long.

Seed were sown at a rate of 31 seeds m^{-1}. Conventional tillage practices were followed at all locations.

Plant lodging was scored at maturity based on a 1 to 5 scale (1 = all plants are erect, 5 = all plants are prostrate). Plant height (cm) was determined as the height from the soil surface to the top of the plant. Plant height was recorded for all replications at each environment, except for Urbana 2007, Bellflower 2008, and Urbana 2008, where only the first replication was measured. Plant maturity was recorded as the number of days after May 31 when approximately 95% of the pods had reached mature pod color or R8 (Fehr et al., 1971). For four row plots, the middle two rows were harvested mechanically and the seed was uniformly dried to 7% moisture and seed yield (Mg ha^{-1}) was adjusted to 13% moisture. For two row plots, both rows were harvested mechanically and seed moisture was measured on the combine and the yield data was adjusted to 13% moisture.

DNA Extraction and Analysis

Ten plants from PI 68508 and FC 04007B (parents of LG82-8379), PI 297544 and PI 391583 (parents of LG94-1713), BSR 101, Kenwood, and each of the confirmation population lines were grown in the greenhouse. Leaf samples were taken from each plant and bulked within the line. DNA was extracted from the bulked leaf tissue using a modified CTAB procedure (Keim et al., 1988). The DNA samples were characterized with both fluorescently and non-fluorescently labeled SSR primers. The SSR primers developed by P. B. Cregan (USDA-ARS, Beltsville, MD) were obtained from Integrated DNA Technologies (Coralville, IA). The polymerase chain reactions (PCR) were performed according to the procedure by Cregan and Quigley (1997). The non-florescent PCR products were analyzed on 6% nondenaturing polyacrylamide gels according to the methods described by Wang et al. (2003) and stained with ethidium bromide. An ABI Prism 377 Genetic Analyzer (Applied Biosystems, Foster City, CA) was used to analyze the florescent PCR products.

Data Analysis

The data for plant height, lodging, maturity, and seed yield were analyzed separately as a nested randomized complete block design using the PROC MIXED function in SAS (Statistical Analysis System version 9.1, SAS Institute, Cary NC). Each year by location combination was considered a single environment. The analysis of variance between the two homozygous marker classes within each confirmation population was determined across and within environments. The environment and marker class were considered fixed. The replications nested within the environments and lines within marker class variation were considered random. A small number of plots were dropped from the statistical analysis due to water damage or for being extreme outliers when the residuals were analyzed using PROC UNIVARIATE. Significant differences between the two homozygous marker classes were determined at a 0.05 probability of a type 1 error.

RESULTS AND DISCUSSION

There were no significant differences for seed yield between the sets of near isogenic lines for any of the confirmation populations developed from the BSR 101 x LG82-8379 lines, however there were some differences between and within sets of near isogenic lines for other agronomic traits.

The mean of the LG82-8379 allele class associated with Satt142 (linkage group (LG) H, chromosome (chr) 12) showed no significant difference ($P < 0.05$) for any of the traits when compared to the mean of the BSR 101 allele class (Table 3). The allele classes had significant ($P < 0.01$) within class variability for plant maturity with a range of 2.3 days for the LG82-8379 allele class and 3.4 days for the BSR 101 allele class (Table 4 Appendix). There was a 0.07 Mg ha^{-1} average yield difference ($P > 0.2$) across all environments between the two allele classes but the difference was not large enough to be declared significant, so the putative quantitative trait loci (QTL) for seed yield was not confirmed (Table 3). Kabelka et al. (2004) published results that reported the LG82-8379 allele yielding about 0.15 Mg ha^{-1} over the BSR 101 allele and was selected with a LOD score of 2.7 but was only identified within one of the maturity sets. Both of the PI parents had the same allele at Satt142 which was different from the allele in BSR 101.

The lines homozygous for the putative LG82-8379-derived QTL for seed yield associated with Satt168 (LG B2, chr 14) showed no significant difference ($P < 0.05$) from the lines homozygous for the BSR 101-derived allele for any of the traits measured (Table 3). There was no significant ($P < 0.01$) within class variability observed for any of the traits either (Table 5 Appendix). The average yield difference between the allele classes across all environments was only 0.01 Mg ha^{-1} ($P > 0.2$) so the QTL for seed yield was not confirmed. Kabelka et al. (2004) reported that the LG82-8379-derived putative QTL for seed yield produced about 0.07 Mg ha^{-1} over the BSR 101-derived allele. This allele was originally selected based on a LOD score of 6.7 but was only detected in one of the three maturity sets. The PI donor parent for the allele at Satt168 was PI 68508

which is contradictory of the Kabelka et al. (2004) report which identified FC 04007B as the source of that allele donor parent.

The allele from LG82-8379 at the quantitative trait loci (QTL) associated with Satt225 (LG A1, chr 5) produced plants that were significantly ($P < 0.05$) taller (2.1 cm) and had greater lodging (2.9 vs. 2.8) when averaged across all environments (Table 3). With approximately 200 observations per allele class, these differences were statistically significant but were not large enough to be of biological importance. There was significant within allele class variation ($P < 0.01$) across environments for plant lodging and maturity. Within the LG82-8379 allele class there was a 3.2 day range in maturity compared to a 1.3 day difference within the BSR 101 allele class (Table 6 Appendix). Within class differences for lodging were small with a range of 0.2 for the LG82-8379 allele class and 0.4 for the BSR 101 allele class (Table 6 Appendix). The yield difference between allele classes was only 0.04 Mg ha^{-1} so the yield effect of this QTL was not confirmed (Table 3). The lines with the LG82-3879 allele ranged in yield from 2.95 to 3.13 Mg ha^{-1} and the line with the BSR 101 allele yielded between 3.02 to 3.23 Mg ha^{-1} (Table 6 Appendix). Kabelka et al. (2004) reported a yield advantage of 0.06 Mg ha^{-1} for the LG82-8379 allele that was selected with a LOD score of 2.6 but was only detected in one of the three maturity sets. PI 68508 was confirmed to be the source of the Satt225 allele from the exotic germplasm.

The experimental lines in the LG82-8379 allele class for the QTL associated with Satt358 (LG O, chr 10) were significantly 1.2 cm taller ($P < 0.05$), than the experimental lines in the BSR 101 allele class (Table 3) but again this difference is very small. No significant differences were determined for the other traits. Significant ($P < 0.01$) within allele class variation occurred for both plant height and maturity. The LG82-8379 allele class had an 8.2 cm range in height and a 2.3 day range in maturity compared with the BSR 101 allele class which had a 3.6 cm range in height and a 1 day range in maturity (Table 7 Appendix). The average yield between the two allele classes across environments had a 0.01 Mg ha^{-1} difference ($P > 0.2$) which is too small of a difference to be declared significant (Table 3). The results from the Kabelka et al. (2004) report

showed the LG82-8379 allele class had about a 0.05 Mg ha^{-1} yield advantage over the BSR 101 allele class and was detected with a LOD score of 3.0 across all three maturity sets, but this QTL had the smallest effect when compared to the other identified QTL for seed yield. The PI parent that donated the Satt358 allele different from the allele in BSR 101 was confirmed to be FC 04007B.

The lines homozygous for the LG82-8379 QTL allele associated with Satt363 (LG C2, chr 6) were significantly ($P < 0.05$) later maturing (0.6 days) than the BSR 101 allele when averaged across environments but, no other traits showed any significant differences (Table 3). This difference in maturity is too small to be of biological importance. There was also significant within allele class variation ($P < 0.01$) across environments for all four traits recorded. The within class height difference for the both LG82-8379 and BSR 101 alleles was approximately 5 cm (Table 8 Appendix). For lodging, the within class difference for the LG82-8379 and the BSR 101 allele classes were 0.4 and 0.3, respectively (Table 8 Appendix). The LG82-8379 allele class had a maturity range of 1.6 days and the BSR 101 allele class had a maturity range of 0.5 days (Table 8 Appendix). The within allele class yields ranged from 3.06 to 3.3 Mg ha^{-1} for the LG82-8379 allele and 3.04 to 3.26 Mg ha^{-1} for the BSR 101 allele class (Table 8 Appendix). Although all of these differences were statistically significant, the ranges within each of the classes were very similar. The yield difference between allele classes was 0.06 Mg ha^{-1} with a P value of 0.14 (Table 3). Kabelka et al. (2004) reported the same yield difference with the LG82-8379 allele being superior to the BSR 101 allele across all three maturity sets. However, Kabelka et al. (2004) identified this QTL using single-trait analysis which has a higher probability for error than the composite interval mapping method which was used to identify most of the other QTL in their report. Both PI parents had the same allele at Satt363, which was different from the allele in BSR 101.

The LG82-8379-derived QTL associated with Satt544 (LG K, chr 9) caused homozygous lines for this QTL to be, on average, significantly ($P < 0.05$) but only slightly shorter (3.9 cm) than lines homozygous for the BSR 101-derived QTL (Table 3). No significant differences were detected for the other measured traits. There was

significant ($P < 0.01$) within allele class variation for plant maturity with the LG82-8379 allele class having a range of 2.4 days and the BSR 101 allele class having a range of 3.1 days (Table 9 Appendix). There was no yield difference between the two allele classes when averaged over the environments so this QTL for seed yield was not confirmed. The Kabelka et al. (2004) report showed a LG82-8379 allele with a 0.11 Mg ha^{-1} yield advantage over the BSR 101 allele and was detected with a LOD score of 5.9, but this QTL was only identified in one of the three maturity sets. The Satt544 allele different from the allele in BSR 101 was from FC 04007B which is contradictory to the Kabelka et al. (2004) report which identified PI 68508 as the donor parent. The discrepancies between the present study and the Kabelka et al. (2004) report in terms of identifying the PI parents that donated the beneficial alleles could be due to the analysis of the PCR products which were analyzed on different machines and by different researchers for both studies.

Unfortunately, we could not confirm any of the putative QTL for seed yield identified by Kabelka et al. (2004). This is surprising considering the previous reports that also identified QTL for seed yield in the same genetic regions. Li et al. (2008a) identified a QTL in the same region as Satt225, Guzman et al. (2007), Orf et al. (1999), Palomeque et al (2009), and Specht et al. (2001) all identified a QTL in the same region as Satt363, Smalley et al. (2004) and Specht et al. (2001) also identified a QTL around Satt142, and Concibido et al. (2008) and Guzman et al (2007) identified QTL close to Satt168 (Table 1). Our results showed that only two of the six LG82-8379-derived putative QTL for seed yield had numerically higher yields than the BSR 101 alleles across all environments (Satt363 0.06 Mg ha^{-1} and Satt358 0.01 Mg ha^{-1}). However, the LG82-8379 allele for the putative QTL for seed yield associated with Satt363 had a higher yield than the BSR 101 allele in five of the six environments it was tested in, with differences ranging from 0.02 to 0.17 Mg ha^{-1} (Appendix: Tables 15, 17, 18, 22, 24, and 25). It is possible that this QTL could be affecting yield but the effect is very small and the precision of our test was not sufficient to declare it significant. There have been previous reports of non-significant results within confirmation populations, and in most cases, the non-significance was attributed to the different environments used by the

confirmation and mapping studies (Li et al., 2008a; Reyna and Sneller, 2001). Another potential reason for the lack of confirmation is that Kabelka et al. (2004) used smaller mapping populations (54-58 lines) to identify most of the QTL which could have over estimated the effects of the QTL. Kabelka et al. (2004) also stated that there was an experiment-wise error rate of 100% and a high probability of false positives.

There was a significant yield difference between the allele classes for 1 of the 3 putative QTL identified in the Kenwood x LG94-1713 populations. This difference was also associated with a significant delay in maturity; however, we tested this QTL in two populations which had large differences in yield and similar differences in maturity. There were significant differences between allele classes for maturity, height, and lodging associated with other putative QTL as well.

The analysis of the confirmation population testing the QTL for seed yield associated with Satt405 (LG J, chr 16) found that there were significant differences ($P < 0.05$) in plant height, lodging, and maturity but not for yield. The lines with the allele from LG94-1713 averaged 2.3 cm taller, had an average lodging score of +0.05, and matured an average of 0.5 days later than the Kenwood allele class (Table 3). These significant differences are so small they are of little biological importance. Guzman et al. (2007) also associated a putative QTL for plant maturity with Satt405 with the LG94-1713 allele maturing 1.1 days later than Kenwood, but did not report effects for either lodging or height. There was no significant within class variation for any of the traits measured (Table 10 Appendix). Despite the differences in the other agronomic traits, the two allele classes at Satt405 had identical seed yields so this QTL was not confirmed. Guzman et al. (2007) reported the LG94-1713-derived putative QTL for seed yield associated with Satt405 yielded 0.05 Mg ha^{-1} over the Kenwood allele when averaged across years but with a LOD score of only 1.7. The investigation to determine which PI parent of LG94-1713 donated the allele at the Satt405 locus found that both PI parents could have been the donor.

The effects of the QTL for seed yield associated with Satt477 (LG O, chr 10) were tested in two populations. A significant difference ($P < 0.05$) in plant lodging was observed in Population B but not in Population A even though the allele class differences were nearly the same in both populations (0.09 vs. 0.1 respectively) (Table 3). There were no other significant differences between allele classes but there was significant within class variation ($P < 0.01$) for all of the agronomic traits in both Population A and B (Table 11 and 12 Appendix). This variability is likely due to segregation of maturity gene E_1 (Bernard, 1971) which was segregating within the two populations but not linked to Satt477. There was an 8.5 day maturity range among lines within each allele class for Population A and a 7 day range within each allele class for Population B (Table 11 and 12 Appendix). Guzman et al. (2007) reported a putative QTL for plant maturity associated with Satt477 with the LG94-1713 allele maturing 2 days later than the Kenwood allele. Even though this difference was not observed in this research, in Population B the allele from LG94-1713 delayed maturity an average of 1.2 days (Table 3). The putative QTL for seed yield associated with the Satt477 locus produced an increase of 0.1 Mg ha^{-1} in the original study and was detected in both years and in the combined analysis (Guzman et al., 2007). However, in our study a yield loss by the LG94-1713 allele class (-0.03 Mg ha^{-1}) compared to the Kenwood allele class was observed in Population A and a yield gain (0.11 Mg ha^{-1}) was observed in Population B (Table 3). The within class variation for seed yield was very large within both populations. In Population A the LG94-1713 allele class ranged from 3.58 to 3.89 Mg ha^{-1} and the Kenwood allele class ranged from 3.49 to 3.92 Mg ha^{-1} (Table 11 Appendix). In population B the LG94-1713 allele class ranged from 3.11 to 3.72 Mg ha^{-1} and the Kenwood allele class ranged from 3.09 to 3.75 Mg ha^{-1} (Table 12 Appendix). The large differences in maturity also likely affected seed yield in both populations. It was determined that PI 391583 donated the allele at Satt477 found in LG94-1713.

When the confirmation populations testing the QTL for seed yield associated with Satt557 (LG C2, chr 6) were analyzed across environments, significant differences ($P < 0.0001$) for all measured agronomic traits were detected. The lines with the LG94-1713 allele were 6.1 cm taller, had lodging scores 0.3 higher, matured 5.8 days later, and

yielded 0.14 Mg ha^{-1} greater than the lines with the Kenwood allele in Population A and in Population B, the lines with the LG94-1713 allele were 10.7 cm taller, had lodging scores 0.3 higher, matured 5.5 days later, and yielded 0.46 Mg ha^{-1} greater than the lines with the Kenwood allele (Table 3). These results confirm the presence of the PI-derived putative QTL for plant height, maturity, and seed yield identified by Guzman et al. (2007) associated with the Satt557 locus and is supported by the findings of other QTL mapping studies (Orf et al., 1999; Palomeque et al., 2009; Specht et al., 2001) (Table 1). The significant differences observed in both populations for all of the agronomic traits could be explained by the 1 cM distance between Satt557 and the maturity gene *E1* locus (Molnar et al., 2003). The significant differences in plant height, lodging, and seed yield are most likely pleotropic effects. Within the allele classes, significant ($P < 0.0001$) variability occurred for plant height and maturity in both Population A and B. The allele classes had a range of 6.8 cm and 1.7 days for the LG94-1713 allele class and 15.9 cm and 7.2 days for the Kenwood allele class in Population A and 5.3 cm and 3.7 days for the LG94-1713 allele class and 12.2 cm and 6.1 days for the Kenwood allele class in Population B (Table 13 and 14 Appendix). Significant within class variation was also observed for yield in Population A ($P = 0.0009$) and Population B ($P = 0.046$). The yields ranged from 3.62 to 3.92 Mg ha^{-1} in the LG94-1713 allele class and from 3.49 to 3.87 Mg ha^{-1} in the Kenwood allele class for Population A and from 3.53 to 3.75 Mg ha^{-1} in the LG94-1713 allele class and from 2.96 to 3.43 Mg ha^{-1} in the Kenwood allele class in Population B (Table 11 and 12 Appendix). Guzman et al. (2007) identified this QTL for seed yield across years with a LOD score of 5.9 and determined that the LG94-1713 allele caused a 0.11 Mg ha^{-1} yield increase over the Kenwood allele. This QTL was also associated with plant height and maturity and caused a 4.1 cm height increase and a 3.3 day delay in maturity. PI 391583 was the donor parent of the beneficial allele in LG94-1713.

The *E1* maturity gene linked with the LG94-1713-derived allele for Satt557 makes it difficult to separate the effects of the putative QTL with those of the delayed maturity; however differences between Population A and B indicate that not all differences can be explained by the changes at the *E1* locus. In both Populations A and B, there was an

average lodging score difference of 0.3 and less than a 6 day maturity difference between the LG94-1713 allele class and the Kenwood allele class (Table 3). These differences could be attributed to changes at the *E1* locus, although differences of approximately 2 weeks difference in maturity dates were reported for an allelic substitution at the *E1* locus in near isogenic lines of Clark (Bernard, 1971). Even though maturity differences were similar between the allele classes in both populations, there were large differences in seed yield between the allele classes in the two populations. In Population A the LG94-1713 allele class yielded 0.14 Mg ha^{-1} more than the Kenwood allele class but in Population B that yield difference was 0.46 Mg ha^{-1} (Table 3). Plant height differences also showed inequality between the two populations. A 6.1 cm difference was observed in Population A and a 10.7 cm difference was observed in Population B (Table 3). If the *E1* maturity gene was the only segregating gene influencing plant performance, then the differences in plant height and seed yield between LG94-1713 and Kenwood should be the same in both Populations A and B (like with plant lodging and maturity).

Our hypothesis is that both Population A and B have the *E1* maturity gene linked to Satt557 derived from LG94-1713 and this resulted in similar allele class maturity differences (6 days) (Table 3). However, the yield difference observed by Population A (0.14 Mg ha^{-1}) was solely due to the *E1* maturity gene whereas the yield difference observed by Population B (0.46 Mg ha^{-1}) (Table 3) was due to both the *E1* maturity gene and the LG94-1713-derived QTL for seed yield and that the increased height is due to pleiotropism. This could have occurred if there was a crossover between the Satt557 locus and the yield QTL in Population A but not in Population B. The parents, LG94-1713 and Kenwood, have both been characterized with SNP markers and several polymorphic markers were found above Satt557 and below the *E1* locus. Characterization of the lines in both populations will be done to determine if a crossover did occur and if so, where it occurred on the chromosome. This will also aid in more accurately identifying where the QTL for seed yield is on chromosome 6. If we can establish a haplotype difference between these two populations near Satt557, then this QTL may add both genetic variability and higher seed yield potential to the N.A. gene pool.

REFERENCES

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131-1139.

Bernard, R.L. 1971. 2 major genes for time of flowering and maturity in soybeans. Crop Sci. 11:242-&.

Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer and E.C.K. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196.

Concibido, V.C., B. La Vallee, P. Mclaird, N. Pineda, J. Meyer, L. Hummel, J. Yang, K. Wu and X. Delannay. 2003. Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. Theor. Appl. Genet. 106:575-582.

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, L. Chung and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464-1490.

Cregan, P.B., and C. Quigley. 1997. Simple sequence repeat DNA marker analysis. P. 173-185. *In* G. Caetano-Anolles and P.M. Gresshoff (ed.) DNA markers: Protocols, applications and overview. J. Wiley and Sons, New York.

Dong, Y.S., B.C. Zhuang, L.M. Zhao, H. Sun and M.Y. He. 2001. The genetic diversity of annual wild soybeans grown in china. Theor. Appl. Genet. 103:98-103.

Fehr, W.R., C.E. Caviness, D.T. Burmood and J.S. Pennington. 1971. Stage of development descriptions for soybeans, Glycine max (L.) Merrill. Crop Sci. 11:929-933.

Gizlice, Z., T.E. Carter and J.W. Burton. 1994. Genetic base for North-American public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143-1151.

Guzman, P.S., B.W. Diers, D.J. Neece, S.K. St. Martin, A.R. Leroy, C.R. Grau, T.J. Hughes and R.L. Nelson. 2007. QTL associated with yield in three backcross-derived populations of soybean. Crop Sci. 47:111-122.

Hamblin, M.T., S.E. Mitchell, G.M. White, W. Gallego, R. Kukatla, R.A. Wing, A.H. Paterson and S. Kresovich. 2004. Comparative population genetics of the panicoid grasses: Sequence polymorphism, linkage disequilibrium and selection in a diverse sample of *Sorghum bicolor*. Genetics 167:471-483.

Heun, M. and T. Helentjaris. 1993. Inheritance of RAPDs in F1 hybrids of corn. Theor. Appl. Genet. 85:961-968.

Heun, M., J.P. Murphy and T.D. Phillips. 1994. A comparison of RAPD and isozyme analyses for determining the genetic-relationships among *Avena-sterilis* L. accessions. Theor. Appl. Genet. 87:689-696.

Hymowitz, T. 1970. Domestication of soybean. Econ. Bot. 24:408-&.

Hyten, D.L., Q. Song, Y. Zhu, I. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker and P.B. Cregan. 2006. Impacts of genetic bottlenecks on soybean genome diversity. Proc. Natl. Acad. Sci. U. S. A. 103:16666-16671.

Ininda, J., W.R. Fehr, S.R. Cianzio and S.R. Schnebly. 1996. Genetic gain in soybean populations with different percentages of plant introduction parentage. Crop Sci. 36:1470-1472.

Kabelka, E.A., B.W. Diers, W.R. Fehr, A.R. LeRoy, I.C. Baianu, T. You, D.J. Neece and R.L. Nelson. 2004. Putative alleles for increased yield from soybean plant introductions. Crop Sci. 44:784-791.

Keim, P., T.C. Olson, and R.S. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. Soybean Genet. Newsletter 15:150-152.

Ko, H.L., D.C. Cowan, R.J. Henry, G.C. Graham, A.B. Blakeney and L.G. Lewin. 1994. Random amplified polymorphic DNA analysis of Australian rice (*Oryza-sativa* L.) varieties. Euphytica 80:179-189.

Li, D., T.W. Pfeiffer and P.L. Cornelius. 2008a. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. Crop Sci. 48:571-581.

Li, Y., R. Guan, Z. Liu, Y. Ma, L. Wang, L. Li, F. Lin, W. Luan, P. Chen, Z. Yan, Y. Guan, L. Zhu, X. Ning, M.J.M. Smulders, W. Li, R. Piao, Y. Cui, Z. Yu, M. Guan, R. Chang, A. Hou, A. Shi, B. Zhang, S. Zhu and L. Qiu. 2008b. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in china. Theor. Appl. Genet. 117:857-871.

Manjarrez-Sandoval, P., T.E. Carter, D.M. Webb and J.W. Burton. 1997. RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. Crop Sci. 37:698-703.

Mansur, L.M., J.H. Orf, K. Chase, T. Jarvik, P.B. Cregan and K.G. Lark. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci. 36:1327-1336.

Molnar, S.J., S. Rai, M. Charette and E.R. Cober. 2003. Simple sequence repeat (SSR) markers linked to E1, E3, E4, and E7 maturity genes in soybean. Genome 46:1024-1036.

Morrell, P.L., D.M. Toleno, K.E. Lundy and M.T. Clegg. 2005. Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp spontaneum) despite high rates of self-fertilization. Proc. Natl. Acad. Sci. U. S. A. 102:2442-2447.

Narvel, J.M., W.R. Fehr, W.C. Chu, D. Grant and R.C. Shoemaker. 2000. Simple sequence repeat diversity among soybean plant introductions and elite genotypes. Crop Sci. 40:1452-1458.

Nelson, R.L. and E.O.C. Johnson. 2006a. Registration of soybean germplasm lines LG97-7012, LG98-1445, and LG98-1605. Crop Sci. 46:1822-1824.

Nelson, R.L. and E.O.C. Johnson. 2006b. Registration of LG96-1797 soybean germplasm. Crop Sci. 46:1403-1403.

Orf, J.H., K. Chase, T. Jarvik, L.M. Mansur, P.B. Cregan, F.R. Adler and K.G. Lark. 1999. Genetics of soybean agronomic traits: I. comparison of three related recombinant inbred populations. Crop Sci. 39:1642-1651.

Palomeque, L., Liu Li-Jun, W. Li, B. Hedges, E.R. Cober and I. Rajcan. 2009. QTL in mega-environments: I. universal and specific seed yield QTL detected in a population derived from a cross of high-yielding adapted x high-yielding exotic soybean lines. Theor. Appl. Genet. 119:417-427.

Reyna, N. and C.H. Sneller. 2001. Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. Crop Sci. 41:1317-1321.

Schmid, K.J., S. Ramos-Onsins, H. Ringys-Beckstein, B. Weisshaar and T. Mitchell-Olds. 2005. A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. Genetics 169:1601-1615.

Shannon, J.G., R.L. Nelson and J.A. Wrather. 2005. Registration of S99-11509 and S99-11986 improved soybean germplasm with diverse pedigree. Crop Sci. 45:1672-1673.

Shimamoto, Y., J. Abe, Z. Gao, J.Y. Gai and F.S. Thseng. 2000. Characterizing the cytoplasmic diversity and phyletic relationship of Chinese landraces of soybean, *Glycine max*, based on RFLPs of chloroplast and mitochondrial DNA. Genet. Resour. Crop Evol. 47:611-617.

Smalley, M.D., W.R. Fehr, S.R. Cianzio, F. Han, S.A. Sebastian and L.G. Streit. 2004. Quantitative trait loci for soybean seed yield in elite and plant introduction germplasm. Crop Sci. 44:436-442.

Sneller, C.H., J.W. Miles and J.M. Hoyt. 1997. Agronomic performance of soybean plant introductions and their genetic similarity to elite lines. Crop Sci. 37:1595-1600.

Specht, J.E., K. Chase, M. Macrander, G.L. Graef, J. Chung, J.P. Markwell, M. Germann, J.H. Orf and K.G. Lark. 2001. Soybean response to water: A QTL analysis of drought tolerance. Crop Sci. 41:493-509.

Tanksley, S.D. 1993. Mapping polygenes. Ann. Rev. Genet. 27:205-233.

Thompson, J.A. and R.L. Nelson. 1998a. Utilization of diverse germplasm for soybean yield improvement. Crop Sci. 38:1362-1368.

Thompson, J.A. and R.L. Nelson. 1998b. Core set of primers to evaluate genetic diversity in soybean. Crop Sci. 38:1356-1362.

Thompson, J.A., R.L. Nelson and L.O. Vodkin. 1998c. Identification of diverse soybean germplasm using RAPD markers. Crop Sci. 38:1348-1355.

Vello, N.A., W.R. Fehr and J.B. Bahrenfus. 1984. Genetic-variability and agronomic performance of soybean populations developed from plant introductions. Crop Sci. 24:511-514.

Wang, D., G.L. Graef, A.M. Procopiuk and B.W. Diers. 2004. Identification of putative QTL that underlie yield in interspecific soybean backcross populations. Theor. Appl. Genet. 108:458-467.

Wang, D., J. Shi, S.R. Carlson, P.B. Cregan, R.W. Ward and B.W. Diers. 2003. A low-cost, high-throughput polyacrylamide gel electrophoresis system for genotyping with microsatellite DNA markers. Crop Sci. 43:1828-1832.

Watterson, G.A. 1975. Number of segregating sites in genetic models without recombination. Theor. Popul. Biol. 7:256-276.

Wright, S.I., I.V. Bi, S.G. Schroeder, M. Yamasaki, J.F. Doebley, M.D. McMullen and B.S. Gaut. 2005. The effects of artificial selection of the maize genome. Science 308:1310-1314.

Table 1: A listing of the reports that have identified putative QTL for seed yield, the genetic location of the putative QTL, and the exotic source from which the putative QTL were derived.

LG/Chromosome	Report	Genetic Marker	LG Position (cM)	Exotic Source
A1/5	Guzman et al. (2007)	Satt300	30.9	Exotic *G. max*
	Li et al. (2008a)†	Satt050 to Satt511	46.5 to 94.2	*G. soja*: PI 245331
	Kabelka et al. (2004)	Satt225	95.2	PI 68508
B2/14	Kabelka et al. (2004)	Satt168	55.2	FC 04007B
	Concibido et al. (2003)†	Satt168 to Satt560	55.2 to 97.9	*G. soja*: PI 407305
	Guzman et al. (2007)	Satt474	75.3	Exotic *G. max*
C1/4	Palomeque et al. (2009)	Satt139 to Sat_042	74.5 to 82.5	Heinong 38
C2/6	Kabelka et al. (2004)	Satt363	98.1	FC 04007B
	Orf et al. (1999)†	Satt277	107.6	Noir 1
	Palomeque et al. (2009)	Satt277	107.6	Heinong 38
	Specht et al. (2001)	Satt277 to Satt489	107.6 to 113.4	Minsoy
	Guzman et al. (2007)	Satt557	112.2	Exotic *G. max*
	Palomeque et al. (2009)	Satt100	114.0	Heinong 38
D1a+Q/1	Smalley et al. (2004)	Satt436	70.7	Exotic *G. max*
D1b/2	Wang et al. (2004)	Satt350	76.6	*G. soja*: PI 468916
D2/17	Kabelka et al. (2004)	Satt186	105.4	PI 68508
E/15	Wang et al. (2004)	Satt491	43.6	*G. soja*: PI 468916
F/13	Specht et al. (2001)	Sat_074	142.4	Noir 1
G/18	Kabelka et al. (2004)	Satt394	43.4	PI 68508
H/12	Specht et al. (2001)	Satt314	69.1	Minsoy
	Kabelka et al. (2004)	Satt142	86.5	FC 04007B
	Smalley et al. (2004)	Satt317	89.5	Exotic *G. max*
J/16	Guzman et al. (2007)	Satt405	12.4	Exotic *G. max*
		Satt622	42.3	Exotic *G. max*
		Satt215	44.1	Exotic *G. max*
		Satt547	67.8	Exotic *G. max*
K/9	Kabelka et al. (2004)	Satt544	43.3	PI 68508
L/19	Specht et al. (2001)	G173B to Dt_1	86.6 to 89.1	Noir 1
	Wang et al. (2004)	Satt006	92.0	*G. soja*: PI 468916
M/7	Orf et al. (1999)	Satt150	18.6	Noir 1
	Specht et al. (2001)	Satt150 to Satt567	18.6 to 33.5	Noir 1
	Kabelka et al. (2004)	Satt308	130.8	PI 68508
N/3	Specht et al. (2001)	Rpg_4	55.4	Minsoy
	Wang et al. (2004)	Satt257	92.6	*G. soja*: PI 468916
O/10	Kabelka et al. (2004)	Satt358	5.4	FC 04007B
	Guzman et al. (2007)	Satt477	82.1	Exotic *G. max*

† Putative QTL that have been confirmed.

Table 2: The number of lines that are homozygous for each allele class within each of the confirmation populations testing a single putative QTL for seed yield.

Marker†	Lines homozygous for the LG82-8379 allele	Lines homozygous for the BSR 101 allele
Satt142 (H/12)	5	9
Satt168 (B2/14)	15	18
Satt225 (A1/5)	9	10
Satt358 (O/10)	19	18
Satt363 (C2/6)	7	9
Satt544 (K/9)	9	7
	Lines homozygous for the LG94-1713 allele	**Lines homozygous for the Kenwood allele**
Satt405 (J/16)	21	19
Satt477 (O/10) Population A‡	18	24
Satt557 (C2/6) Population A	19	19
Satt477 (O/10) Population B	20	21
Satt557 (C2/6) Population B	20	21

† The SSR marker that is being investigated followed by its linkage group/chromosome number.
‡ Population A and Population B both have Satt477 and Satt557 segregating within the population.

Table 3: The population means across all environment for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Population	Donor Parent	N†	Mat (days)†	Ldg (1-5)†	Plant Height (cm)	Seed Yield (Mg/ha)
Satt142‡	PI Parent	110	116.6	3.3	83.7	3.10
	BSR 101	198	116.6	3.3	84.5	3.17
P Value			>0.2	>0.2	>0.2	>0.2
Satt168	PI 68508	149	112.0	2.8	75.7	2.99
	BSR 101	179	112.1	2.8	74.6	3.00
P Value			>0.2	>0.2	0.0717	>0.2
Satt225	PI 68508	195	114.8	2.9	78.8	3.05
	BSR 101	220	114.6	2.8	76.7	3.09
P Value			>0.2	0.0184	0.0011	>0.2
Satt358	FC 04007B	190	116.9	2.9	77.3	3.42
	BSR 101	180	117.0	2.8	76.1	3.41
P Value			>0.2	>0.2	0.0306	>0.2
Satt363	PI Parent	148	116.5	2.9	79.2	3.18
	BSR 101	194	115.9	2.8	78.4	3.12
P Value			0.0168	0.0696	>0.2	0.1382
Satt544	FC 04007B	195	115.7	4.2	88.0	3.05
	BSR 101	150	116.5	4.2	91.9	3.05
P Value			0.0559	>0.2	0.0065	>0.2
Satt405	PI Parent	204	115.0	2.95	95.6	3.66
	Kenwood	183	114.5	2.90	93.3	3.67
P Value			<0.0001	0.0193	0.0003	>0.2
Population A Satt477	PI 391583	176	113.1	2.0	77.7	3.74
	Kenwood	235	113.7	2.1	80.0	3.77
P Value			>0.2	>0.2	0.1122	>0.2
Population A Satt557	PI 391583	186	115.9	2.2	81.9	3.81
	Kenwood	187	110.1	1.9	75.8	3.67
P Value			<0.0001	<0.0001	<0.0001	<0.0001
Population B Satt477	PI 391583	199	111.6	1.64	76.2	3.53
	Kenwood	210	110.4	1.55	74.5	3.42
P Value			0.1173	0.0198	0.1829	0.0739
Population B Satt557	PI 391583	199	112.9	1.7	79.5	3.64
	Kenwood	208	107.4	1.4	68.8	3.18
P Value			<0.0001	<0.0001	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

APPENDIX

Table 4: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt142 confirmation population.

Satt142†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI Parent	LG07c-1086	22	115.6	3.1	81.2	2.97
	LG07c-1090	22	116.2	3.2	81.3	3.12
	LG07c-1092	22	115.9	3.4	85.2	3.07
	LG07c-1093	22	117.9	3.5	86.7	3.15
	LG07c-1098	22	117.4	3.4	84.2	3.20
BSR 101	LG07c-1084	22	118.7	3.3	86.0	3.05
	LG07c-1085	22	115.6	3.2	85.0	3.29
	LG07c-1087	22	115.8	3.3	84.1	3.12
	LG07c-1088	22	117.0	3.4	83.9	3.16
	LG07c-1089	22	115.5	3.2	88.4	3.24
	LG07c-1091	22	115.3	3.3	82.6	3.24
	LG07c-1094	22	118.1	3.4	83.6	3.11
	LG07c-1095	22	115.8	3.3	82.9	3.21
	LG07c-1099	22	118.0	3.4	84.1	3.09
Within allele class variation (P value)			<0.0001	>0.2	>0.2	0.0184

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 5: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt168 confirmation population.

Satt168†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 68508	LG07-4160	10	112.0	2.9	74.3	2.90
	LG07-4161	10	112.4	2.9	75.6	2.96
	LG07-4162	10	112.4	2.7	75.1	3.02
	LG07-4163	10	111.6	2.7	77.8	2.87
	LG07-4164	10	111.5	3.1	76.2	3.05
	LG07-4165	9	111.9	2.9	75.6	3.06
	LG07-4166	10	113.0	2.8	76.4	3.10
	LG07-4167	10	111.9	2.6	75.8	2.96
	LG07-4168	10	112.3	2.8	72.2	2.94
	LG07-4169	10	112.3	2.9	74.3	2.97
	LG07-4170	10	111.9	3.0	78.7	3.11
	LG07-4171	10	111.5	3.0	76.3	3.06
	LG07-4172	10	111.8	2.8	74.3	3.02
	LG07-4175	10	111.7	2.9	75.0	2.96
	LG07-4176	10	112.0	3.0	77.1	2.93
BSR 101	LG07-4174	10	112.3	2.9	75.6	3.02
	LG07-4180	10	112.5	2.8	72.6	3.06
	LG07-4182	10	111.4	3.0	75.4	3.13
	LG07-4183	10	111.6	2.8	76.1	2.98
	LG07-4184	10	112.5	2.7	72.6	2.93
	LG07-4185	10	112.1	2.8	75.0	2.88
	LG07-4186	10	112.1	2.9	73.3	3.01
	LG07-4187	10	112.4	2.8	76.4	2.99
	LG07-4188	10	111.9	3.0	73.6	2.97
	LG07-4190	10	112.8	3.0	76.0	3.03
	LG07-4192	10	112.3	2.9	74.9	2.97
	LG07-4194	9	112.0	2.7	76.5	2.99
	LG07-4195	10	110.8	2.9	75.2	2.90
	LG07-4197	10	112.2	2.8	76.7	3.03
	LG07-4198	10	112.6	2.8	74.6	3.04
	LG07-4200	10	112.1	2.8	70.7	2.92
	LG07-4201	10	112.3	2.7	73.2	2.93
	LG07-4202	10	112.4	2.9	73.9	3.12
Within allele class variation (P value)			0.1157	0.107	>0.2	0.0167

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 6: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt225 confirmation population.

Satt225†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 68508	LG07c-1048	22	114.9	2.8	78.8	3.13
	LG07c-1053	22	115.0	2.9	77.9	3.00
	LG07c-1055	22	114.6	2.9	76.6	2.97
	LG07c-1058	22	115.5	3.0	79.9	2.95
	LG07c-1059	22	115.4	2.9	79.3	3.12
	LG07c-1062	22	114.5	2.9	78.3	3.05
	LG07c-1063	21	114.8	3.0	80.1	3.08
	LG07c-1064	20	114.3	3.0	81.6	3.06
	LG07c-1066	22	114.2	2.9	76.7	3.06
BSR 101	LG07c-1049	22	113.8	2.9	75.8	3.10
	LG07c-1050	22	115.3	2.9	78.6	3.07
	LG07c-1051	22	114.5	2.9	76.4	3.11
	LG07c-1052	22	115.1	2.9	78.2	3.14
	LG07c-1054	22	114.6	2.5	76.6	3.05
	LG07c-1056	22	115.0	2.8	75.4	3.03
	LG07c-1057	22	112.7	2.5	75.8	3.03
	LG07c-1060	22	114.0	2.9	75.5	3.02
	LG07c-1065	22	115.9	2.8	76.3	3.23
	LG07c-1067	22	114.8	2.8	77.8	3.07
Within allele class variation (P value)			<0.0001	0.0054	>0.2	0.0532

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 7: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt358 confirmation population.

Satt358†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
FC 04007B	LG07-4133	10	117.5	3.0	78.4	3.38
	LG07-4135	10	117.2	3.0	79.2	3.47
	LG07-4136	10	117.1	3.0	77.9	3.44
	LG07-4137	10	117.1	2.8	74.2	3.45
	LG07-4139	10	116.8	2.9	76.2	3.36
	LG07-4140	10	115.2	2.7	74.0	3.35
	LG07-4141	10	117.1	2.9	77.8	3.48
	LG07-4142	10	117.3	3.1	80.0	3.46
	LG07-4143	10	115.4	2.9	74.8	3.36
	LG07-4144	10	117.3	2.9	81.8	3.41
	LG07-4145	10	117.3	2.9	78.6	3.41
	LG07-4146	10	117.1	2.8	76.6	3.45
	LG07-4150	10	117.4	2.7	78.3	3.39
	LG07-4152	10	116.7	2.9	75.9	3.33
	LG07-4153	10	117.1	3.0	76.8	3.36
	LG07-4155	10	117.2	2.8	80.7	3.61
	LG07-4156	10	116.1	2.8	73.6	3.28
	LG07-4157	10	117.0	2.9	76.8	3.40
	LG07-4158	10	116.9	2.9	77.7	3.52
BSR 101	LG07-4107	10	117.3	2.9	77.1	3.55
	LG07-4108	10	117.0	2.8	75.1	3.30
	LG07-4113	10	116.8	2.9	74.7	3.47
	LG07-4114	10	116.4	2.9	74.9	3.35
	LG07-4116	10	117.0	3.0	77.7	3.39
	LG07-4117	10	116.6	2.9	74.2	3.38
	LG07-4119	10	117.1	2.8	76.9	3.37
	LG07-4120	10	116.7	2.7	77.3	3.59
	LG07-4121	10	117.3	2.9	77.8	3.40
	LG07-4123	10	117.2	2.8	76.4	3.56
	LG07-4124	10	116.5	2.8	76.9	3.39
	LG07-4125	10	117.2	2.9	76.2	3.27
	LG07-4126	10	116.5	2.7	74.3	3.44
	LG07-4128	10	117.4	2.9	74.9	3.37
	LG07-4129	10	117.4	2.9	76.8	3.53
	LG07-4130	10	117.4	2.9	76.6	3.38
	LG07-4131	10	117.0	2.9	75.8	3.34
	LG07-4132	10	117.4	3.1	75.6	3.37
Within allele class variation (P value)			<0.0001	>0.2	0.0036	0.044

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 8: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt363 confirmation population.

Satt363†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI Parent	LG07c-1033	22	116.3	2.8	79.2	3.26
	LG07c-1035	21	117.6	3.1	82.3	3.19
	LG07c-1036	22	116.1	2.9	79.2	3.21
	LG07c-1037	21	116.7	3.0	77.2	3.06
	LG07c-1039	21	116.2	2.7	78.5	3.30
	LG07c-1041	21	116.0	2.9	77.2	3.06
	LG07c-1046	20	116.4	3.0	80.9	3.17
BSR 101	LG07c-1031	21	115.9	2.7	78.0	3.26
	LG07c-1032	22	115.6	2.6	78.2	3.06
	LG07c-1034	22	116.0	2.9	77.5	3.04
	LG07c-1038	22	116.0	2.9	78.7	3.07
	LG07c-1040	22	115.8	2.6	76.1	3.04
	LG07c-1042	21	115.7	2.9	81.3	3.19
	LG07c-1043	22	116.0	2.7	78.0	3.11
	LG07c-1044	21	116.1	2.8	81.0	3.21
	LG07c-1045	21	115.7	2.7	77.0	3.13
Within allele class variation (P value)			0.0001	0.0013	0.0054	0.0095

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 9: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt544 confirmation population.

Satt544†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
FC 04007B	LG07c-1068	21	117.3	4.5	88.8	3.03
	LG07c-1069	22	115.5	4.0	87.3	3.09
	LG07c-1070	22	115.5	4.1	86.8	3.09
	LG07c-1071	22	115.2	4.2	88.7	3.05
	LG07c-1072	22	114.9	4.1	89.0	2.99
	LG07c-1075	22	115.5	4.1	89.8	3.01
	LG07c-1077	21	114.9	4.0	87.6	3.14
	LG07c-1078	21	116.0	4.3	85.8	3.03
	LG07c-1083	22	116.5	4.3	88.2	3.03
BSR 101	LG07c-1073	22	116.0	4.1	89.6	3.03
	LG07c-1074	20	118.5	4.3	95.9	3.02
	LG07c-1076	21	116.0	4.5	94.6	3.10
	LG07c-1079	22	116.8	4.2	91.8	2.96
	LG07c-1080	21	116.9	4.2	92.1	3.04
	LG07c-1081	22	115.4	4.0	87.1	3.08
	LG07c-1082	22	116.2	4.2	92.9	3.12
Within allele class variation (P value)			<0.0001	0.065	0.0475	>0.2

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 10: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes in the Satt405 confirmation population.

Satt405†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI Parent	LG07-4016	10	114.3	2.8	91.8	3.64
	LG07-4017	10	115.3	3.1	95.0	3.73
	LG07-4018	10	115.7	3.0	96.2	3.63
	LG07-4019	9	116.2	3.1	95.3	3.73
	LG07-4020	10	114.6	3.0	98.6	3.60
	LG07-4021	10	115.1	3.0	96.0	3.74
	LG07-4022	10	114.7	3.0	95.1	3.62
	LG07-4023	9	116.8	2.9	94.0	3.64
	LG07-4024	10	115.8	3.0	97.3	3.63
	LG07-4025	9	115.3	2.9	95.9	3.64
	LG07-4026	10	114.3	3.0	98.6	3.63
	LG07-4028	10	114.3	3.0	93.3	3.58
	LG07-4030	9	115.2	2.9	93.9	3.66
	LG07-4031	9	114.6	2.8	96.3	3.62
	LG07-4032	10	114.7	3.0	99.1	3.68
	LG07-4033	10	115.2	3.0	97.7	3.65
	LG07-4034	9	114.0	3.0	97.3	3.69
	LG07-4035	10	115.3	2.9	94.2	3.71
	LG07-4039	10	114.9	3.0	94.1	3.62
	LG07-4040	10	114.4	3.0	95.4	3.69
	LG07-4059	10	114.9	2.9	93.3	3.67
Kenwood	LG07-4041	9	115.8	2.9	93.4	3.66
	LG07-4042	10	114.0	2.9	92.1	3.67
	LG07-4043	10	114.5	2.9	93.8	3.65
	LG07-4044	10	114.0	2.9	94.1	3.64
	LG07-4045	9	114.3	2.7	96.0	3.62
	LG07-4046	9	115.7	2.9	93.1	3.78
	LG07-4047	10	114.4	3.1	94.0	3.73
	LG07-4048	10	114.0	2.9	93.1	3.61
	LG07-4049	10	114.5	2.9	96.1	3.73
	LG07-4051	10	114.1	2.9	91.2	3.71
	LG07-4052	10	114.4	2.8	94.1	3.70
	LG07-4053	10	114.8	3.0	92.2	3.71
	LG07-4054	9	113.6	2.9	92.0	3.76
	LG07-4055	9	116.2	3.2	93.9	3.70
	LG07-4056	10	114.2	3.0	92.3	3.56
	LG07-4057	10	113.7	2.8	91.7	3.59
	LG07-4058	10	114.0	2.8	94.7	3.65
	LG07-4060	9	113.9	2.9	92.8	3.54
	LG07-4062	9	116.1	2.9	93.0	3.79
Within allele class variation (P value)			>0.2	>0.2	0.1655	>0.2

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 11: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes for the Satt477 confirmation population A.

Population A Satt477†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 391583	LG07-3875	10	107.8	1.7	73.7	3.62
	LG07-3876	9	109.6	1.9	72.8	3.58
	LG07-3877	10	108.8	2.0	74.2	3.61
	LG07-3878	10	108.9	1.8	73.2	3.67
	LG07-3879	10	116.2	2.4	86.0	3.78
	LG07-3880	10	115.4	2.1	78.1	3.74
	LG07-3881	10	115.4	2.1	80.9	3.85
	LG07-3882	8	115.8	2.1	78.4	3.82
	LG07-3883	10	114.0	2.1	76.7	3.72
	LG07-3884	10	113.7	2.0	76.8	3.87
	LG07-3885	10	113.6	2.0	77.6	3.75
	LG07-3886	10	114.4	2.1	76.8	3.71
	LG07-3887	10	116.0	2.2	80.7	3.89
	LG07-3888	9	113.6	1.9	78.9	3.69
	LG07-3890	10	115.2	2.3	79.4	3.62
	LG07-3891	10	113.7	2.0	79.9	3.87
	LG07-3892	10	115.1	2.1	76.8	3.82
	LG07-3913	10	109.6	2.1	76.6	3.78
Kenwood	LG07-3893	10	109.7	1.9	73.9	3.56
	LG07-3894	10	109.9	1.9	72.0	3.68
	LG07-3895	10	108.8	2.0	77.8	3.85
	LG07-3896	10	116.7	2.3	80.2	3.92
	LG07-3897	10	116.4	2.3	84.1	3.85
	LG07-3898	10	116.5	2.2	82.9	3.88
	LG07-3899	9	115.1	2.1	82.0	3.74
	LG07-3900	10	115.7	2.2	86.0	3.80
	LG07-3901	10	116.8	2.3	82.8	3.87
	LG07-3902	9	108.8	1.7	72.9	3.62
	LG07-3903	9	114.3	2.0	79.0	3.78
	LG07-3904	10	113.9	2.1	82.9	3.85
	LG07-3906	9	115.7	2.0	80.9	3.90
	LG07-3907	10	114.9	2.2	84.3	3.75
	LG07-3908	10	114.8	2.2	79.3	3.76
	LG07-3910	10	114.0	2.1	81.4	3.73
	LG07-3917	10	109.5	1.8	71.7	3.49
	LG07-3918	10	111.0	2.2	87.6	3.87
	LG07-3919	9	108.3	1.7	74.8	3.76
	LG07-3924	10	115.6	2.2	79.2	3.62
	LG07-3925	10	115.6	2.0	82.3	3.75
	LG07-3926	10	115.5	2.0	80.0	3.75
	LG07-3927	10	115.0	2.2	80.9	3.79
	LG07-3932	10	115.1	2.3	80.9	3.87
Within allele class variation (P value)			<0.0001	0.0012	<0.0001	0.0002

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 12: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes for the Satt477 confirmation population B.

Population B Satt477†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 391583	LG07-3817	10	107.4	1.4	69.2	3.18
	LG07-3818	10	106.3	1.5	68.4	3.11
	LG07-3823	9	111.9	1.6	76.4	3.46
	LG07-3825	10	112.0	1.7	76.3	3.56
	LG07-3837	10	111.6	1.7	78.6	3.61
	LG07-3838	10	112.9	1.7	79.4	3.72
	LG07-3839	10	112.0	1.7	78.1	3.70
	LG07-3840	10	112.9	1.6	79.4	3.65
	LG07-3841	10	112.5	1.7	81.8	3.69
	LG07-3842	10	112.2	1.6	77.8	3.62
	LG07-3843	10	112.9	1.9	79.0	3.61
	LG07-3857	10	112.9	1.7	76.7	3.62
	LG07-3859	10	111.3	1.5	73.1	3.56
	LG07-3860	10	112.0	1.5	74.6	3.37
	LG07-3861	10	113.1	1.6	76.3	3.37
	LG07-3863	10	111.9	1.7	77.0	3.51
	LG07-3864	10	112.9	1.7	78.7	3.64
	LG07-3865	10	112.2	1.8	77.9	3.64
	LG07-3866	10	110.5	1.7	73.7	3.43
	LG07-3867	10	110.9	1.6	72.6	3.45
Kenwood	LG07-3819	10	107.0	1.5	69.1	3.09
	LG07-3820	10	106.8	1.4	69.9	3.16
	LG07-3821	10	106.9	1.4	68.4	3.22
	LG07-3822	10	107.7	1.5	68.6	3.26
	LG07-3826	10	106.6	1.4	66.7	3.13
	LG07-3827	10	107.0	1.4	67.0	3.20
	LG07-3836	10	107.1	1.3	70.6	3.23
	LG07-3844	10	113.2	1.6	77.8	3.66
	LG07-3845	10	112.2	1.7	79.2	3.57
	LG07-3846	10	113.2	1.7	79.0	3.75
	LG07-3847	10	111.7	1.7	79.4	3.57
	LG07-3848	10	113.6	1.6	81.3	3.74
	LG07-3849	10	112.7	1.7	79.3	3.54
	LG07-3862	10	106.8	1.5	69.7	3.19
	LG07-3868	10	113.3	1.8	80.2	3.59
	LG07-3869	10	112.1	1.5	76.7	3.47
	LG07-3870	10	111.7	1.7	73.6	3.41
	LG07-3871	10	111.8	1.6	77.2	3.59
	LG07-3872	10	112.0	1.7	76.1	3.47
	LG07-3873	10	112.4	1.4	77.4	3.40
	LG07-3874	10	113.4	1.6	77.0	3.49
Within allele class variation (P value)			<0.0001	<0.0001	<0.0001	<0.0001

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 13: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes for the Satt557 confirmation population A.

Population A Satt557†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 391583	LG07-3881	10	115.4	2.1	80.9	3.85
	LG07-3890	10	115.2	2.3	79.4	3.62
	LG07-3896	10	116.7	2.3	80.2	3.92
	LG07-3897	10	116.4	2.3	84.1	3.85
	LG07-3898	10	116.5	2.2	82.9	3.88
	LG07-3899	9	115.1	2.1	82.0	3.74
	LG07-3900	10	115.7	2.2	86.0	3.80
	LG07-3901	10	116.8	2.3	82.8	3.87
	LG07-3906	9	115.7	2.0	80.9	3.90
	LG07-3912	10	116.0	2.1	83.0	3.86
	LG07-3914	10	115.9	2.2	81.0	3.86
	LG07-3922	9	116.0	2.2	82.4	3.73
	LG07-3924	10	115.6	2.2	79.2	3.62
	LG07-3925	10	115.6	2.0	82.3	3.75
	LG07-3928	10	116.1	2.2	83.3	3.87
	LG07-3929	10	116.0	2.3	81.4	3.77
	LG07-3930	10	115.9	2.3	81.8	3.81
	LG07-3931	9	115.6	2.0	81.1	3.84
	LG07-3932	10	115.1	2.3	80.9	3.87
Kenwood	LG07-3875	10	107.8	1.7	73.7	3.62
	LG07-3876	9	109.6	1.9	72.8	3.58
	LG07-3877	10	108.8	2.0	74.2	3.61
	LG07-3878	10	108.9	1.8	73.2	3.67
	LG07-3893	10	109.7	1.9	73.9	3.56
	LG07-3894	10	109.9	1.9	72.0	3.68
	LG07-3895	10	108.8	2.0	77.8	3.85
	LG07-3902	9	108.8	1.7	72.9	3.62
	LG07-3905	10	113.7	1.8	79.6	3.76
	LG07-3911	10	114.1	2.0	81.1	3.73
	LG07-3913	10	109.6	2.1	76.6	3.78
	LG07-3915	10	109.5	1.9	73.2	3.53
	LG07-3916	10	110.4	2.0	75.7	3.62
	LG07-3917	10	109.5	1.8	71.7	3.49
	LG07-3918	10	111.0	2.2	87.6	3.87
	LG07-3919	9	108.3	1.7	74.8	3.76
	LG07-3920	10	109.4	1.9	74.1	3.64
	LG07-3921	10	108.9	1.7	74.3	3.57
	LG07-3927	10	115.0	2.2	80.9	3.79
Within allele class variation (P value)			<0.0001	0.0133	<0.0001	0.0009

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 14: The experimental line means across environments for each of the agronomic traits and the within allele class variation P Value observed for the two homozygous allele classes for the Satt557 confirmation population B.

Population B Satt557†	Entry	N‡	Mat (days)‡	Ldg (1-5)‡	Plant Height (cm)	Seed Yield (Mg/ha)
PI 391583	LG07-3837	10	111.6	1.7	78.6	3.61
	LG07-3838	10	112.9	1.7	79.4	3.72
	LG07-3839	10	112.0	1.7	78.1	3.70
	LG07-3840	10	112.9	1.6	79.4	3.65
	LG07-3841	10	112.5	1.7	81.8	3.69
	LG07-3842	10	112.2	1.6	77.8	3.62
	LG07-3843	10	112.9	1.9	79.0	3.61
	LG07-3844	10	113.2	1.6	77.8	3.66
	LG07-3845	10	112.2	1.7	79.2	3.57
	LG07-3846	10	113.2	1.7	79.0	3.75
	LG07-3847	10	111.7	1.7	79.4	3.57
	LG07-3849	10	112.7	1.7	79.3	3.54
	LG07-3850	10	112.8	1.9	78.1	3.64
	LG07-3851	10	113.6	1.7	79.6	3.64
	LG07-3852	10	112.6	1.6	78.1	3.62
	LG07-3853	10	113.4	1.7	80.7	3.57
	LG07-3854	9	115.3	1.7	82.0	3.53
	LG07-3855	10	113.7	1.8	83.1	3.65
	LG07-3856	10	113.6	1.9	80.3	3.71
	LG07-3864	10	112.9	1.7	78.7	3.64
Kenwood	LG07-3817	10	107.4	1.4	69.2	3.18
	LG07-3818	10	106.3	1.5	68.4	3.11
	LG07-3819	10	107.0	1.5	69.1	3.09
	LG07-3820	10	106.8	1.4	69.9	3.16
	LG07-3821	10	106.9	1.4	68.4	3.22
	LG07-3822	10	107.7	1.5	68.6	3.26
	LG07-3824	10	107.0	1.4	65.2	3.13
	LG07-3826	10	106.6	1.4	66.7	3.13
	LG07-3827	10	107.0	1.4	67.0	3.20
	LG07-3828	10	106.5	1.4	65.4	2.96
	LG07-3829	10	106.7	1.5	67.8	3.20
	LG07-3830	8	106.9	1.5	65.6	3.13
	LG07-3831	10	106.4	1.4	67.7	3.15
	LG07-3832	10	107.0	1.4	68.6	3.14
	LG07-3833	10	106.5	1.5	68.3	3.19
	LG07-3834	10	107.2	1.4	68.2	3.14
	LG07-3835	10	107.9	1.5	68.1	3.14
	LG07-3836	10	107.1	1.3	70.6	3.23
	LG07-3862	10	106.8	1.5	69.7	3.19
	LG07-3866	10	110.5	1.7	73.7	3.43
	LG07-3873	10	112.4	1.4	77.4	3.40
Within allele class variation (P value)			<0.0001	0.0722	<0.0001	0.0456

† The SSR marker linked to the putative QTL for seed yield tested in the population followed by the donor parent for the homozygous allele classes.
‡ N = the number of observations, Mat = plant maturity (Days after May 31), and Ldg = lodging score (1 = plant erect, 5 = prostrate).

Table 15: The population means for Urbana 2007 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Urbana 2007						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt225‡	PI 68508	18	99.2	3.2	87.0	2.76
	BSR 101	20	98.6	3.0	82.7	2.81
P Value			0.0513	>0.2	0.122	>0.2
Satt363	PI Parent	14	102.1	3.0	86.6	2.66
	BSR 101	18	101.6	3.2	85.6	2.59
P Value			0.0934	>0.2	>0.2	>0.2
Satt142	PI Parent	10	108.0	2.5	100.6	3.04
	BSR 101	18	108.8	2.6	102.6	3.01
P Value			0.1843	>0.2	>0.2	>0.2
Satt544	FC 04007B	17	110.6	3.4	103.1	2.99
	BSR 101	13	111.9	3.4	102.6	2.89
P Value			0.0222	>0.2	>0.2	0.0704

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 16: The population means for Arthur 2008 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Arthur 2008						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt142‡	PI Parent	20	114.3	3.5	84.6	2.91
	BSR 101	36	114.8	3.5	86.9	3.01
P Value			>0.2	>0.2	0.0518	>0.2
Satt544	FC 04007B	35	114.8	3.8	91.7	3.44
	BSR 101	26	115.7	3.9	94.9	3.44
P Value			0.1214	0.1263	0.1015	>0.2

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 17: The population means for Bellflower 2008 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Bellflower 2008						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt168‡	PI 68508	29	115.9	2.9	73.2	2.38
	BSR 101	36	115.8	3.0	71.8	2.39
P Value			>0.2	>0.2	>0.2	>0.2
Satt22	PI 68508	36	116.9	2.2	72.1	2.27
	BSR 101	40	117.2	2.2	68.2	2.35
P Value			>0.2	>0.2	0.0507	>0.2
Satt358	FC 04007B	38	119.4	2.4	72.3	2.58
	BSR 101	36	119.8	2.4	69.9	2.55
P Value			>0.2	>0.2	0.0638	>0.2
Satt363	PI Parent	25	122.1	3.1	66.6	2.44
	BSR 101	32	121.1	2.9	67.4	2.27
P Value			0.0431	>0.2	>0.2	0.0645
Satt405	PI Parent	41	122.0	3.3	88.8	2.96
	Kenwood	37	120.6	3.3	86.4	2.94
P Value			0.0019	>0.2	0.0943	>0.2
Population A Satt477	PI 391583	34	120.6	2.5	73.8	3.45
	Kenwood	47	121.1	2.4	74.4	3.44
P Value			0.0912	>0.2	>0.2	>0.2
Population A Satt557	PI 391583	38	122.0	2.3	73.4	3.49
	Kenwood	38	120.4	2.4	73.0	3.40
P Value			<0.0001	>0.2	>0.2	0.1561
Population B Satt477	PI 391583	40	119.9	2.4	68.5	3.13
	Kenwood	42	120.0	2.2	65.7	3.05
P Value			>0.2	0.0282	0.0969	>0.2
Population B Satt557	PI 391583	40	120.3	2.5	69.8	3.24
	Kenwood	41	119.4	2.0	63.0	2.88
P Value			0.0016	<0.0001	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 18: The population means for DeKalb 2008 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

DeKalb 2008						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt168‡	PI 68508	30	112.5	3.6	82.2	2.84
	BSR 101	35	112.7	3.6	79.2	2.77
P Value			>0.2	>0.2	0.0535	>0.2
Satt225	PI 68508	36	115.6	3.3	82.1	2.82
	BSR 101	40	115.2	3.2	81.1	2.98
P Value			>0.2	0.1321	>0.2	0.022
Satt358	FC 04007B	38	114.7	3.6	84.2	2.90
	BSR 101	36	114.6	3.5	82.9	2.91
P Value			>0.2	0.1409	0.1739	>0.2
Satt363	PI Parent	26	116.5	3.0	80.2	2.84
	BSR 101	36	115.8	2.7	78.1	2.79
P Value			0.0543	0.004	0.1293	>0.2

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 19: The population means for Pontiac 2008 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Pontiac 2008						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt405‡	PI Parent	40	99.4	2.3	101.3	3.15
	Kenwood	34	97.9	2.2	100.4	3.23
P Value			0.0005	0.1703	>0.2	0.1029
Population A Satt477	PI 391583	35	99.5	1.7	86.3	3.29
	Kenwood	48	100.5	1.7	88.2	3.39
P Value			>0.2	>0.2	>0.2	0.1237
Population A Satt557	PI 391583	38	104.5	1.8	90.6	3.45
	Kenwood	38	94.6	1.5	82.0	3.30
P Value			<0.0001	<0.0001	<0.0001	0.0245
Population B Satt477	PI 391583	40	94.6	1.0	82.7	3.29
	Kenwood	42	93.3	1.0	80.4	3.22
P Value			0.1338	>0.2	>0.2	>0.2
Population B Satt557	PI 391583	39	96.4	1.1	86.6	3.41
	Kenwood	41	90.5	1.0	73.3	3.01
P Value			<0.0001	0.0148	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 20: The population means for Urbana 2008 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Urbana 2008						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt142‡	PI Parent	20	120.2	4.4	80.0	2.91
	BSR 101	36	120.6	4.2	80.3	2.92
P Value			>0.2	>0.2	>0.2	>0.2
Satt544	FC 04007B	36	116.5	5.7	92.0	2.70
	BSR 101	28	116.7	5.9	100.1	2.78
P Value			>0.2	>0.2	0.0248	>0.2

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 21: The population means for Arthur 2009 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Arthur 2009						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt142‡	PI Parent	20	111.8	3.7	92.8	3.62
	BSR 101	36	111.8	3.8	93.6	3.76
P Value			>0.2	>0.2	>0.2	0.1189
Satt544	FC 04007B	35	111.4	4.3	102.2	3.04
	BSR 101	27	112.3	4.3	108.9	3.07
P Value			0.1455	>0.2	0.0022	>0.2

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 22: The population means for Bellflower 2009 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Bellflower 2009						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt168‡	PI 68508	30	115.8	2.9	80.0	3.51
	BSR 101	36	116.1	2.8	79.2	3.55
P Value			>0.2	0.1754	>0.2	>0.2
Satt225	PI 68508	34	120.8	3.2	80.2	3.58
	BSR 101	40	120.7	3.2	78.1	3.53
P Value			>0.2	0.0678	0.1439	>0.2
Satt358	FC 04007B	38	118.4	3.2	78.8	4.03
	BSR 101	36	118.6	3.2	77.3	4.06
P Value			>0.2	>0.2	0.0916	>0.2
Satt363	PI Parent	28	120.3	3.3	82.6	3.67
	BSR 101	36	120.0	3.2	82.9	3.60
P Value			>0.2	>0.2	>0.2	0.1067
Satt405	PI Parent	42	120.7	3.4	97.8	4.11
	Kenwood	38	120.3	3.2	95.9	4.12
P Value			0.0242	0.0051	0.0772	>0.2
Population A Satt477	PI 391583	35	117.7	2.6	82.5	4.42
	Kenwood	48	118.3	2.7	85.8	4.38
P Value			>0.2	0.1726	0.0303	>0.2
Population A Satt557	PI 391583	37	120.7	2.8	88.1	4.40
	Kenwood	37	113.3	2.5	82.7	4.28
P Value			<0.0001	0.0021	0.0003	0.0171
Population B Satt477	PI 391583	40	117.0	2.0	82.1	3.78
	Kenwood	42	114.3	1.9	80.3	3.67
P Value			0.0353	0.0562	>0.2	0.0735
Population B Satt557	PI 391583	40	117.6	2.1	85.4	3.88
	Kenwood	42	109.7	1.8	74.3	3.45
P Value			<0.0001	0.0008	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 23: The population means for Hume 2009 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Hume 2009						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt142‡	PI Parent	20	125.5	2.2	78.6	3.02
	BSR 101	36	124.6	2.1	78.6	3.07
P Value			>0.2	>0.2	>0.2	>0.2
Satt544	FC 04007B	36	124.4	3.5	78.3	2.98
	BSR 101	28	125.2	3.5	82.4	2.87
P Value			0.0239	>0.2	0.0063	0.091

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 24: The population means for Pontiac 2009 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Pontiac 2009						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt168‡	PI 68508	30	108.0	2.7	72.6	3.37
	BSR 101	36	108.0	2.7	72.1	3.41
P Value			-	>0.2	>0.2	>0.2
Satt225	PI 68508	36	113.4	3.3	77.7	3.69
	BSR 101	40	113.1	3.2	75.2	3.69
P Value			>0.2	>0.2	0.0087	>0.2
Satt358	FC 04007B	38	116.0	3.4	77.7	4.18
	BSR 101	36	116.0	3.4	77.3	4.15
P Value			-	0.1597	>0.2	>0.2
Satt363	PI Parent	27	114.7	3.2	78.8	4.04
	BSR 101	36	114.0	2.9	79.9	4.06
P Value			0.0507	0.0162	>0.2	>0.2
Satt405‡	PI Parent	39	116.0	3.5	94.3	4.29
	Kenwood	36	116.0	3.5	92.3	4.29
P Value			-	-	0.0588	>0.2
Population A Satt477	PI 391583	36	112.9	2.0	77.6	4.04
	Kenwood	46	113.4	2.2	79.5	4.06
P Value			>0.2	>0.2	0.1226	>0.2
Population A Satt557	PI 391583	35	115.3	2.6	81.7	4.02
	Kenwood	38	110.6	1.6	75.8	4.01
P Value			<0.0001	<0.0001	<0.0001	>0.2
Population B Satt477	PI 391583	40	112.2	1.4	77.1	4.20
	Kenwood	42	112.0	1.3	76.5	4.07
P Value			>0.2	0.1422	>0.2	0.0664
Population B Satt557	PI 391583	40	113.8	1.5	80.1	4.26
	Kenwood	42	109.2	1.1	71.4	3.83
P Value			<0.0001	<0.0001	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.

Table 25: The population means fir Urbana 2009 for each of the agronomic traits and the significant difference between the two homozygous allele classes for each confirmation population.

Urbana 2009						
Population	**Donor Parent**	**N†**	**Mat (days)†**	**Ldg (1-5)†**	**Plant Height (cm)**	**Seed Yield (Mg/ha)**
Satt168‡	PI 68508	30	108.0	2.1	69.0	2.85
	BSR 101	36	108.1	2.0	69.1	2.85
P Value			>0.2	0.1747	>0.2	>0.2
Satt225	PI 68508	35	115.5	2.3	74.6	3.06
	BSR 101	40	114.7	2.1	72.9	3.02
P Value			>0.2	0.0258	>0.2	>0.2
Satt358	FC 04007B	38	115.9	1.8	71.1	3.40
	BSR 101	36	116.1	1.7	69.8	3.40
P Value			>0.2	0.1742	>0.2	>0.2
Satt363	PI Parent	28	116.4	2.0	75.7	3.10
	BSR 101	36	116.2	1.9	73.4	3.07
P Value			>0.2	>0.2	0.038	>0.2
Satt142	PI Parent	20	115.6	3.2	78.4	3.09
	BSR 101	36	115.4	3.3	78.7	3.16
P Value			>0.2	>0.2	>0.2	>0.2
Satt544	FC 04007B	36	113.6	3.8	76.0	3.14
	BSR 101	28	114.5	3.8	77.6	3.20
P Value			0.141	>0.2	>0.2	>0.2
Satt405	PI Parent	42	116.5	2.3	92.6	3.77
	Kenwood	38	116.2	2.2	89.0	3.76
P Value			0.0142	>0.2	0.0006	>0.2
Population A Satt477	PI 391583	36	115.1	1.4	66.4	3.51
	Kenwood	46	115.2	1.4	68.9	3.58
P Value			>0.2	>0.2	>0.2	>0.2
Population A Satt557	PI 391583	38	116.9	1.4	71.6	3.74
	Kenwood	36	111.9	1.5	63.8	3.36
P Value			<0.0001	>0.2	<0.0001	<0.0001
Population B Satt477	PI 391583	39	114.5	1.4	66.8	3.21
	Kenwood	42	112.6	1.4	65.2	3.06
P Value			0.0825	>0.2	0.1742	0.0635
Population B Satt557	PI 391583	40	115.9	1.4	70.7	3.37
	Kenwood	42	107.9	1.3	59.1	2.71
P Value			<0.0001	0.0038	<0.0001	<0.0001

† N = the number of observations, Mat = plant maturity (Days after May 31), Ldg = lodging score (1 = plant erect, 5 = prostrate).
‡ The SSR marker linked to the putative QTL that was tested.