# Genetic Analysis of Yield Components in the PI 438489B by 'Hamilton' Recombinant Inbred Line (RIL) Population of Soybean [*Glycine max* (L.) Merr.]

Ragin Bobby<sup>1</sup>, Richard Bazelle<sup>1</sup>, Willsheana Clark<sup>1</sup>, Stella K. Kantartzi<sup>2</sup>, Khalid Meksem<sup>2</sup>, Masum Akond<sup>1</sup> & My Abdelmajid Kassem<sup>1</sup>

<sup>1</sup> Plant Genomics and Biotechnology Lab, Department of Biological Sciences, Fayetteville State University, Fayetteville, NC, USA

<sup>2</sup> Department of Plant, Soil, and Agricultural Systems, Southern Illinois University, Carbondale, IL, USA

Correspondence: My Abdelmajid Kassem, Department of Biological Sciences, Fayetteville State University, Fayetteville, NC 28301, USA. Tel: 1-910- 672-1692. E-mail: mkassem@uncfsu.edu

Received: May 7, 2012Accepted: June 5, 2012Online Published: July 26, 2012doi:10.5539/jas.v4n9p98URL: http://dx.doi.org/10.5539/jas.v4n9p98

# Abstract

Yield is a multi-factorial trait determined by several genetic traits and highly correlated with important agronomic traits in many crops including soybean. [*Glycine max* (L.)]. Plant height, seed and pod numbers, and seed weight are all components of yield and polygenic in nature. The objective of this study was to identify quantitative trait loci (QTL) for days to germination, days to flowering, plant height, pod number, seed number, 100-seed weight, and total seed weight in soybean using the using the PI 438489B by 'Hamilton' recombinant inbred line (RIL) population (PIxH, n=50). A total of 18 QTL were found on 10 different chromosomes. Three QTL for days to germination (qDG001-qDG003) have been identified on chromosomes 5b, 6, and 13b. Two QTL (qDF001 and qDF002) have been identified on chromosomes 9 and 13b, respectively. On QTL for plant height (qPH001) have been identified on chromosome 6. Four QTL for pod number (qSN001 and qSN002) have been identified on chromosomes 5b and 11b, respectively. Two QTL for 100-seed weight (qSW001 to qSW005) have been identified on chromosomes 5a, 6, 8, 9, and 11c, respectively. Two QTL for total seed weight (qTSW001 and qTSW002) have been identified on chromosomes 5b and 17c, respectively. The QTL identified here may be introduced in breeding programs to develop soybean cultivars with high yield potential.

**Keywords:** soybean, QTL, days to germination, days to flowering, plant height, pod number, seed number, seed weight

# 1. Introduction

Soybean (*Glycine max* L. Merr.) is a valuable crop due to its protein and oil contents which are suitable for human and animal consumption. One of the most important factors for soybean breeding is high-yield potential. Yield is a multi-factorial trait determined by several genetic traits and highly correlated with important agronomic traits. Agronomic characters as days to flowering, plant height, maturity and 100-seed weight are high correlated, positively or negatively, with yield in soybean (Panthee et al., 2007; Li et al., 2008).

It has been reported that yield component traits such as days to germination, days to flowering, plant height, pod number, seed number, seed weight (100-seeds), and total seed weight are polygenic and governed by many quantitative trait loci (QTL) (Tanksley, 1993; Mansur et al., 1996). In soybean, during the past 2 decades, many studies genetically mapped QTL for days to germination (SoyBase, 2012), days to flowering (SoyBase, 2012), plant height (Mansur et al., 1993; SoyBase, 2012), seed weight (Mansur et al., 1993; Yuan et al., 2002; SoyBase, 2012), and many other traits (Orf et al., 1999; Yuan et al., 2002; Kassem et al., 2004 & 2006; Kassem et al., 2007a; Kassem et al., 2007b; SoyBase, 2012) using different mapping populations. A linkage map can be created with molecular marker analysis, and then after phenotyping, a QTL analysis can be used to detect positions and genetic distances of markers among chromosomes that are associated with the QTL.

The genetic analysis of desirable traits in breeding programs has been made substantially easier with the advent of molecular markers and marker-based linkage maps (Paterso et al., 1991). Most of the soybean genetic linkage

maps to date were constructed with RFLPs, AFLPs, RAPDs, and SSRs. Choi et al. (2007) developed the first soybean transcript map by mapping 1141 SNP markers (derived from 1141 expressed gene sequences) onto the previous version of the soybean genetic map (Song et al., 2004), which included 1015 PCR-based markers (SSRs). On the basis of gene-based SNPs mapped, SNP markers were positioned in many of the 5 and 10 cM gaps that existed in the previous map. This map will be very useful for the case study of the diversity of gene function associated with these transcripts, as it will offer researchers an opportunity to identify potential candidate genes for >1,150 QTLs that have been reported to date. Hyten et al. (2010) recently reported on the latest version of soybean integrated genetic linkage map (Consensus Map 4.0) by adding 2,651 new SNP markers into the previous genetic map developed by Choi et al. (2007). There are a total of 5,500 genetic markers in this new genetic linkage map. New technologies for assaying genotypes for SNP allele type are expected to make SNP markers the replacement marker system for the currently used SSR marker systems, relative to future soybean genetics and breeding studies (Hyten et al., 2008). A particularly important advantage of the Illumina-based SNP allele detection over the SSR marker allele detection is the elimination of the tedious gel-based marker allele visualization required for the latter. The SNP-based genetic linkage map (Kassem et al., 2012) was used for genetic analysis of yield components of a soybean RIL population in this study.

The objective of this study is to map QTL for days to germination (DG), days to flowering (DF), plant height (PH), pod number (PN), seed number (SN), seed weight (SW), and total seed weight (TSW) using the PI 438489B By 'Hamilton' recombinant inbred line (RIL) population of soybean.

#### 2. Materials and Methods

#### 2.1 Plant Material

The 'PI 438489B' By 'Hamilton' RIL population (PIxH, *n*=50) used in this study was developed at the University of Missouri Agronomy Research Center (Yue, Arelli, & Sleper, 2001) and was extensively studied to map QTL for SDS resistance (Kassem et al., 2012), SCN resistance (Yue et al., 2001).

#### 2.2 Growth Conditions and Traits Measurements

Four seeds of the parents (PI 438489B and Hamilton) and RILs were sown in potting soil in pots of 15x14 cm and kept in the greenhouse at  $25\pm1^{\circ}$ C under natural daylight for 3 weeks. After then, plants have been divided in 2 groups (Group I, row spacing of 25 cm and Group II, row spacing of 50 cm) and planted in a field in Saint Pauls, NC (Bladen County).

Days to germination have been recorded in the greenhouse and plant heights have been recorded in the field at the maturity of all RILs and parental lines (R8,  $\sim$ 120 days after planting – DAP) just before harvest. After maturity, the plants have been harvested and pod numbers, seed numbers, seed weight (100-seed weight), and total seed weight have been recorded in the lab.

# 2.3 DNA Isolation, SNP Genotyping, and Genetic Map Construction

DNA isolation was performed as described earlier (Kassem et al., 2012). Single nucleotide polymorphism (SNP) genotyping of 1,536 SNPs was performed by the Golden Gate Assay according to Illumina, Inc. protocol (Hyten et al., 2008). The SNP-based genetic linkage map (Kassem et al., 2012) was used to map QTL in this study.

# 3. Statistical Data Analysis

QTL Mapping was performed by CIM using WinQTL Cartographer as described earlier (Kassem et al., 2012). Briefly, 1,000 permutations were performed in order to establish LOD values to declare QTLs at a significant level of  $P \le 0.05$ . The Model 6 and its default settings have not been changed. Correlation coefficients between all the traits studied (days to germination, days to flowering, plant height, pod number, seed number, seed yield and total seed weight) have been calculated and results are reported in a Pearson correlation matrix. All statistical analyses have been performed on JMP 9.0 (SAS Institute Inc., Cary, NC, USA).

# 4. Results and Discussion

# 4.1 Correlation Coefficients

Correlation coefficients for each pairwise combination of studied traits (days to germination, days to flowering, plant height, pod number, seed number, seed yield and total seed weight) from the recombinant inbred lines are presented in Table 1. Days to germination were negatively correlated with all traits and all correlations were significant except with seed yield. On the other hand, total seed weight was very strongly correlated with plant height, pod and seed number and also seed yield. As it was expected seed number was significantly correlated with pod number although the coefficient was moderate. Seed number had a low correlation with days to flowering,

moderate with plant height and high with pod number. Finally, plant height was found moderate but significant correlated with days to flowering.

Table 1. Correlation coefficients for agronomic traits in a soybean recombinant inbred population from a cross between PI438489B and Hamilton

	Days to	Plant Height	Pod #	Seed #	Seed yield	Total seed
	Flowering	(cm)			(g)	weight (g)
Days to Germination	-0.767***	-0.637***	-0.405**	-0.430**	-0.232ns	-0.363*
Days to Flowering		0.479**	0.285ns	0.323*	0.304ns	0.258ns
Plant Height (cm)			0.686***	0.608***	0.307ns	0.578***
Pod #				0.957***	0.273ns	0.932***
Seed #					0.233ns	0.940***
Seed yield (g)						0.416**

\* Significant at P < 0.05 probability level

\*\* Significant at P < 0.01 probability level

\*\*\* Significant at P < 0.001 probability level

#### 4.2 Yield Components QTL

Table 2. QTL that control days to germination (DG), days to flowering (DF), plant height (PH), pod number (PN), seed number (SN), seed weight (SW), and total seed weight (TSW) in the soybean PI 438489B by 'Hamilton' recombinant inbred line (RIL) population

Trait	No.	Chr./L.G.	QTL	Marker/Interval	Position	LOD	$R^{2}(\%)$
Days to Germination	1	5b/A1	qDG001	SS107921684-SS107912607	10.4-19.6	2.6	0.70
	2	6/C2	qDG002	SS107912977-SS107917113	33.4-40.4	3.2	0.10
	3	13b/F	qDG003	SS107919342-SS107924542	38.7-43.5	3.8	0.14
Days to Flowering	4	9/k	qDF001	SS107912559-SS107917952	0.0-8.0	2.6	0.14
	5	13b/F	qDF002	SS107919342-SS107924542	38.7-44.3	3.0	0.18
Plant Height	6	6/C2	qPH001	SS107921568-SS107929749	27.0-34.0	5.0	0.26
Pod Number	7	2/D1b	qPN001	SS107913056-SS107913946	42.0-49.4	3.7	0.11
	8	6/C2	qPN002	SS107922673-SS107925078	40.5-43.0	3.2	0.16
	9	8/A2	qPN003	SS107913537-SS107913537	10.5-12.5	2.7	0.14
	10	8/A2	qPN004	SS107914369-SS107916169	24.5-37.0	2.8	0.14
Seed Number	11	5b/A1	qSN001	SS107921684-SS107916326	10.0-32.5	3.2	0.16
	12	11b/B1	qSN002	SS107920383-SS107912753	0.0-20.0	3.2	0.15
Seed Yield	13	5a/A1	qSY001	SS107919282-SS107930364	4.0-20.0	2.7	0.09
	14	6/C2	qSY002	ss107912977-SS107925078	38.5-43.6	3.0	0.10
	15	8/A2	qSY003	SS107916169-SS107915722	0.0-10.5	2.5	0.11
	16	9/K	qSY004	SS107913799-SS107918132	16.8-24.8	2.8	0.06
	17	11c/B1	qSY005	SS107913440-SS107912515	0.0-13.5	4.0	0.17
Total Seed Weight	18	5b/A1	qTSW001	SS107912607-SS107916326	19.6-33.8	5.2	0.27
-	19	17c/D2	qTSW002	SS107927901-SS107915314	6.1-11.5	2.6	0.12





Figure 1. Locations of SNP markers and the QTL underlying days to germination (DG), days to flowering (DF), plant height (PH), pod number (PN), seed number (SN), seed weight (SW), and total seed weight (TSW) in the soybean PI 438489B by 'Hamilton' recombinant inbred line (RIL) population

Interestingly, QTL for pod number and oil content have been mapped at the same location of qDG001 and QTL for stem strength, seed yield, and oil content have been mapped at the same location of qDG002 (Qi et al., 2011;

Zhang et al., 2010; Chen et al., 2011; Guzman et al., 2007; Orf et al., 1999). A QTL for seed abortion was also mapped at the same position of *qDG003* and *qDF002* (Tischner, Allphin, Chase, Orf, & Lark, 2003). Similarly, QTL for seed yield and protein content have been mapped near *qDF001* (Mansur, Lark, Kross, & Oliveira, 1993; Yuan et al., 2002). The plant height QTL *qPH001* was mapped at a region containing QTL for stem strength, seed yield and oil content (Chen et al., 2011; Guzman et al., 2007; Orf et al., 1999).

The pod number QTL, *qPN001* has been mapped at the same location of other QTL for leaf length, width and shape (SoyBase, 2012); *qPN002* has been mapped close to QTL for stem strength, seed yield and oil content (Chen et al., 2011; Guzman et al., 2007; Orf et al., 1999); *qPN003* has been mapped close to QTL for oil content (Qi et al., 2011); and *qPN004* has been mapped close to QTL for stem strength and pod number (Chen et al., 2011; Zhang et al., 2010).

Interestingly, the seed number QTL, *qSN001* was mapped to a region containing QTL for seed weight and oil content while *qSN002* was found close to QTL for seed weight and seed number (Orf et al., 1999; Teng et al., 2009). Moreover, the seed yield QTL, *qSY001* was located on the same region containing QTL for sucrose content and soybean cyst nematode (SCN) resistance (Vierling, Faghihi, Ferris, & Ferris, 1996; Maughan, Saghai Maroof, & Buss, 2000); *qSY002* was mapped to a region harboring QTL for stem strength, seed yield, and oil content (Chen et al., 2011; Guzman et al., 2007; Orf et al., 1999); *qSY003* was mapped 4.5 cM from a QTL for oil content (Qi et al., 2011); *qSY004* was located near QTL protein and oil content (Mansur, Lark, Kross, & Oliveira, 1993; Hyten et al., 2004); and *qSY005* is positioned near a cluster of QTLs for seed weight, seed number, and hypocotyl length (Lee, Park, Lee, Park, & Boerma, 2001; Orf et al., 1999). The QTL for total seed weight, *qTSW001* was located near QTL for plant phosphate content and plant weight (Lian, Xiaohui, Mantong, Xiaolong, & Hong, 2010; Qi et al., 2011; Zhang et al., 2010).

In this study, we have identified clusters of QTL for DG, SN, and TSW on Chromosome 5, PH, DG, PN, and SY on Chromosome 6, SY and PN on Chromosome 8, and DF and SY on Chromosome 9. In addition, many of these clusters contain QTL governing other important agronomic traits especially QTL for disease resistance such as sudden death syndrome resistance and soybean cyst nematode resistance (Vierling, Faghihi, Ferris, & Ferris, 1996; Maughan, Saghai Maroof, & Buss, 2000; Kassem et al., 2012) which can lead to developing high yielding soybean cultivars with super resistance to several diseases.

#### Acknowledgements

This work is supported by the Department of Defense (DoD) through the grant W911NF-11-1-0178 to MAK. We also thank Ms. Pam Ratcliff and the undergraduate crew members for taking care of the plants in the greenhouse and the field.

#### References

- Chen, H., Shan, Z., Sha, A., Wu, B., Yang, Z., Chen, S., Zhou, R., & Zhou, X. (2011). Quantitative trait loci analysis of stem strength and related traits in soybean. *Euphytica*, 179, 485-497. http://dx.Doi.org/10.1007/s10681-011-0382-5
- Choi, I. Y., Hyten, D. L., Matukumalli, L. K., Song, Q., Chaky, J. M., Quigley, C. V., ... Cregan, P. B. (2007). A soybean transcript map: Gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* 176, 685-696. http://dx.doi.org/10.1534/genetics.107.070821
- Guzman, P., Diers, B., Neece, D., St. Martin, S., LeRoy, A., Grau, C., Hughes, T., & Nelson, R. (2007). QTL Associated with Yield in Three Backcross-Derived Populations of Soybean. *Crop Sci*, 47, 111-122. http://dx.doi.org/10.2135/cropsci2006.01.0003
- Hyten, D. L., Pantalone, V. R., Sams, C. E., Saxton, A. M., Landau-Ellis, D., Stefaniak, T., & Schmidt, M. E. (2004). Seed quality QTL in a prominent soybean population. *Theor. Appl. Genet, 109*, 552-561.
- Hyten, D., Song, Q., Choi, I. Y., Yoon, M. S., Specht, J., Matukumalli, L., ... Cregan, P. (2008). High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *Theor. Appl. Genet*, 116, 945-952.
- Hyten, D. L., Choi, I. Y., Song, Q., Specht, J. E., Carter, T. E., Shoemaker, R. C., ... Cregan, P. B. (2010). A high density integrated genetic linkage map of soybean and the development of a 1536 Universal Soy Linkage Panel for quantitative trait locus mapping. *Crop Sci, 50,* 960-968. http://dx.doi.org/10.2135/cropsci2009.06.0360
- Kassem, M. A., Meksem, K., Kang, C. H., Njiti, V. N., Kilo, V., Wood, A. J., & Lightfoot, D. A. (2004). Loci underlying resistance to manganese toxicity mapped in a recombinant inbred line population of 'Essex' x 'Forrest'. *Plant and Soil, 260*, 197-204.

- Kassem, M. A., Shultz, J., Meksem, K., Cho, Y., Wood, A. J., Iqbal, M. J., & Lightfoot D. A. (2006). An updated 'Essex' by 'Forrest' linkage map and first composite map of QTL underlying six soybean traits. *Theor Appl Genet*, 113, 1015-1026. http://dx.doi.org/10.1007/s00122-006-0361-8.
- Kassem, M. A., Meksem, K., Wood, A. J., & Lightfoot, D. A. (2007a). A Microsatellite map developed from late maturity germplasm 'Essex' by 'Forrest' detects four QTL for soybean seed yield expected from early maturing germplasm. *Rev Biol & Biotech*, 6, 2-10.
- Kassem, M. A., Meksem, K. Wood, A. J., & Lightfoot, D. A. (2007b). Additional QTL for SDS and SCN resistances in a soybean recombinant inbred line population of Essex x Forrest. *Rev in Biol & Biotech, 6*, 11-19.
- Kassem, M. A., Ramos, L., Leandro, L., Mbofung, G., Hyten, D. L., Kantartzi, S. K., ... Meksem, K. (2012). The PI 438489B by Hamilton SNP-Based Genetic Linkage Map of Soybean [*Glycine max* (L.) Merr.] Identified Quantitative Trait Loci that Underlie Seedling SDS Resistance. J Plant Genome Sciences, 1, 17-32. http://dx.doi.org/10.5147/jpgs.2012.0053
- Lee, S. H., Park, K. Y., Lee, H. S., Park, E. H., & Boerma, H. R. (2001). Genetic mapping of QTLs conditioning soybean sprout yield and quality. *Theor. Appl. Genet, 103*, 702-709. http://dx.doi.org/10.1007/s001220100595
- Li, W., Zheng, D. H., Van, K., & Lee, S. H. (2008). QTL mapping for major agronomic traits across two years in soybean (*Glycine max* L. Merr.). *J. Crop Sci. Biotechnol*, *11*, 171-190. http://dx.doi.org/10.1093/aob/mcm149
- Lian, Q., Xiaohui, C., Mantong, M., Xiaolong, Y., & Hong, L. (2010). QTL analysis of root traits as related to phosphorus efficiency in soybean. *Ann. of Bot, 106*, 223-234. http://dx.doi.org/10.1093/aob/mcq097
- Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B., & Lark K. G. (1996). Genetic Mapping of Agronomic Traits Using Recombinant Inbred Lines of Soybean. *Crop Sci*, *36*, 1327-1336.
- Mansur, L. M., Lark, K. G., Kross, H., & Oliveira, A. (1993). Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (Glycine max L.). *Theor. Appl. Genet*, 86, 907-913. http://dx.doi.org/10.1007/BF00211040
- Maughan, P. J., Saghai Maroof, M. A., & Buss G. R. (2000). Identification of quantitative trait loci controlling sucrose content in soybean (*Glycine max*). Mol. Breeding, 6, 105-111. http://dx.doi.org/10.1023/A:1009628614988
- Orf, J. H., Chase, K., Jarvik, T., Mansur, L. M., Cregan, P. B., Adler, F. R., & Lark, K. G. (1999). Genetics of Soybean Agronomic Traits: I. Comparison of Three Related Recombinant Inbred Populations. *Crop Sci*, 39, 1642-1651.
- Panthee, D. R., Pantalone, V. R., Saxton, A. M., West, D. R., & Sams, C. E. (2007). Quantitative trait loci for agronomic traits in soybean. *Plant Breeding*, 126, 51-57. http://dx.doi.org/10.1111/j.1439-0523.2006.01305.x
- Paterson, A. H., Tanksley, S. D., & Sorrells, M. E. (1991). DNA markers in plant improvement. *Adv. Agron, 46,* 39-90. http://dx.doi.org/10.1016/S0065-2113(08)60578-7
- Qi, Z., Wu, Q., Han, X., Sun, Y., Du, X., Liu, C., Jiang, H., Hu, G., & Chen, Q. (2011). Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes. *Euphytica*, *179*, 499-514. http://dx.doi.org/10.1007/s10681-011-0386-1
- SoyBase. (2012). http://soybase.agron.iastate.edu/ Last accessed in June 2012.
- Song, Q. J., Marek, L. F., Shoemaker, R. C., Lark, K. G., Concibido, V. C., Delannay, X., Specht, J. E., & Cregan, P. B. (2004). A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet*, 109, 122-128.
- Tanksley, S. D. (1993). Mapping polygenes. Annu. Rev. Genet, 27, 205-233.
- Teng, W., Han, Y., Du, Y., Sun, D., Zhang, Z., Qiu, L., Sun, G., & Li, W. (2009). QTL analyses of seed weight during the development of soybean (Glycine max L. Merr.). *Heredity*, 102, 372-380. http://dx.doi.org/10.1038/hdy.2008.108
- Tischner, T., Allphin, L., Chase, K., Orf, J. H., & Lark, K. G. (2003). Genetics of Seed Abortion and Reproductive Traits in Soybean. *Crop Sci*, 43, 464-473. http://dx.doi.org/10.2135/cropsci2010.03.0123

- Vierling, R. A., Faghihi, J., Ferris, V. R., & Ferris, J. M. (1996). Association of RFLP markers with loci conferring broad-based resistance to the soybean cyst nematode (*Heterodera glycines*). *Theor. Appl. Genet*, 92, 83-86. http://dx.doi.org/10.1007/BF00222955
- Yuan, J., Njiti, V. N., Meksem, K., Iqbal, M. J., Triwitayakorn, K., Kassem, M. A., Davis, G. T., Schmidt, M. E., & Lightfoot, D. A. (2002). Quantitative trait loci in Two Soybean Recombinant Inbred Line Populations Segregating for Yield and Disease Resistance. *Crop Sci*, 42, 271-277.
- Yue, P., Arelli, P. R., & Sleper, D. A. (2001). Molecular characterization of resistance to *Heterodera glycines* in soybean 'PI 438489B'. *Theor Appl Genet, 102*, 921-928. http://dx.doi.org/10.1007/s001220000453
- Zhang, D., Cheng, H., Wang, H., Zhang, H., Liu, C., & Yu, D. (2010). Identification of genomic regions determining flower and pod numbers development in soybean (*Glycine max* L). J. Genet. and Genom, 37, 545-556. http://dx.doi.org/10.1016/S1673-8527(09)60074-6