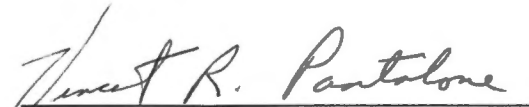


To the Graduate Council:

I am submitting herewith a thesis written by David Lee Hyten, Jr. entitled "QTL Mapping and Identification of GxE Interactions of Agronomic and Seed Quality Traits in Soybean." I have examined the final paper copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Plant and Soil Sciences.

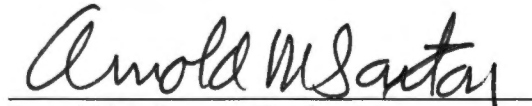


Vincent R. Pantalone, Major Professor

We have read this thesis
and recommend its acceptance:



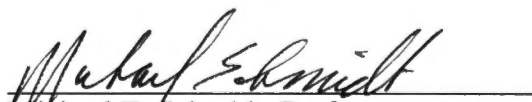
Carl E. Sams, Professor



Arnold M. Saxton, Professor



Perry B. Cregan, Research Geneticist



Michael E. Schmidt, Professor

Acceptance for the Council:



Vice Provost and Dean of
Graduate Studies

**QTL Mapping and Identification of GxE Interactions of Agronomic and
Seed Quality Traits in Soybean**

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

David Lee Hyten, Jr.
May 2002

AG-VET-MED.

Thesis
2002
.H98

DEDICATION

This thesis is dedicated to my parents, David Lee Hyten, Sr. and Barbara Joyce Hyten, for supporting and giving me motivation all of these years, to my two sisters and their families, Jill Stover and Kimberly Vaughan, they have always been great friends that have always been there to help me when I needed it, and to my fiancée, Aimee Sheer, for giving me unlimited love, support, and encouragement.

ACKNOWLEDGEMENTS

I wish to thank all those who helped me in completing my degree in Master of Science in Plant and Soil Science. I thank Dr. Vincent R. Pantalone for being a great major professor. He has not only been a professor to me but a good friend that has helped me throughout my project under his direction. I thank my other committee members, Dr. Perry B. Cregan, Dr. Carl E. Sams, Dr. Arnold M. Saxton, and Dr. Michael E. Schmidt for giving me great advice on my project along with being very encouraging. I thank the department of Plant Science and Landscape Systems. I thank all the people that have supported and helped me on this project, Dr. Roger Boerma, Meredith Crosby, Dr. Brian Diers, Lee Ellis and the workers at the Knoxville Plant Science Farm, James Klein and the workers at the Southern Illinois University Soybean Breeding Program, Debbie Landau-Ellis, Cindy Longmire, Beth Meyer, Gordon Percell and the workers at the West Tennessee Experiment Station, Dilip Panthee, Amanda Pierce, Jessie Smith, Dr. Madeleine Spencer, Yamika Stokes, Jonathan Swanks, Kevin Swanks, The Tennessee Soybean Promotion Board, Donna Thomas and the USDA Northern Regional Research Center at Peoria, IL, and The United Soybean Board, without their help this project would not have been possible.

I thank my family, fiancée, and friends, whose suggestions and encouragement made this work possible.

ABSTRACT

The detection of quantitative trait loci (QTL) in soybean [*Glycine max* (L) Merr] enables regions of chromosomes to be identified, which contain genes that regulate the expression of important agronomic and seed quality traits. The objective of this study was to identify QTL and genotype x environment (GxE) interactions contributing to agronomic and seed quality traits in soybean. A recombinant inbred line (RIL) population was created from two prominent ancestors (Essex and Williams) of currently available U.S. cultivars. One hundred simple sequence repeat (SSR) markers spaced throughout the genome were mapped in this population. Agronomic and seed quality traits were measured in six environments spanning two years. QTL were identified with GxE interaction for agronomic and seed quality traits, using composite interval mapping and multiple trait analysis. A total of 11 maturity, six height, seven lodging, two yield, six oil, 13 protein, nine seed size, four palmitate, seven stearate, four oleate, five linoleate, and eight linolenate QTL were detected in this population. The QTL found in this study may be of use in marker assisted selection to enable breeders to more finely tune their ability to make further genetic gains.

TABLE OF CONTENTS

Chapter	PAGE
1. Introduction.....	1
2. Analysis of Agronomic Traits in a Soybean Population to Discover QTL and Establish GxE Interaction.....	3
Introduction.....	3
Materials and Methods.....	6
Results and Discussion.....	8
Conclusion.....	23
3. Molecular Mapping and Identification of GxE Interactions of Soybean QTL for Seed Quality Traits.....	24
Introduction.....	24
Materials and Methods.....	26
Results and Discussion.....	29
Conclusion.....	39
4. Molecular Mapping and Identification of GxE Interactions of Soybean Fatty Acid Modifier QTL	40
Introduction.....	40
Materials and Methods.....	42
Results and Discussion.....	45
Conclusion.....	56
References.....	57
Vita.....	64

LIST OF TABLES

TABLE	PAGE
2.1. Mean, standard deviation, range, and parental means for soybean agronomic traits in an Essex x Williams F ₆ RIL population grown in six environments.....	10
2.2. Position, effects, and significance of environments associated with soybean maturity in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	11
2.3. Position, effects, and significance of environments associated with soybean height in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	16
2.4. Position, effects, and significance of environments associated with soybean lodging in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	20
2.5. Position, effects, and significance of environments associated with soybean yield in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	22
3.1. Mean, standard deviation, range, parental means, and heritability for soybean seed traits in an Essex x Williams F ₆ RIL population grown in six environments.....	30
3.2. Position, effects, and significance of environments associated with soybean seed oil concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	32
3.3. Position, effects, and significance of environments associated with soybean seed protein concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	33
3.4. Position, effects, and significance of environments associated with soybean seed size concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	38
4.1. Mean, standard deviation, range, parental means, and heritability for soybean seed fatty acid concentration in an Essex x Williams F ₆ RIL population grown in five environments.....	46

4.2. Position, effects, and significance of environments associated with soybean seed palmitate concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	48
4.3. Position, effects, and significance of environments associated with soybean seed stearate concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	49
4.4. Position, effects, and significance of environments associated with soybean seed oleate concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	50
4.5. Position, effects, and significance of environments associated with soybean seed linoleate concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	51
4.6. Position, effects, and significance of environments associated with soybean seed linolenate concentration in an Essex x Williams F ₆ RIL population, based on QTL Cartographer analysis.....	52

LIST OF FIGURES

FIGURE	PAGE
2.1. Maturity and Height LOD scores from six combined environments graphed along soybean linkage group C2.....	18
4.1. Stearate LOD scores from A) one environment (KES-C/00) B) combined over five environments, graphed along soybean linkage group C2.....	54

Chapter 1

Introduction

Soybean is grown worldwide with major production in the United States, Brazil, China, Argentina, and India (Liu, 1997). There have been continuous gains in soybean yield, increased disease resistance, and improvements in other traits of agronomic importance over the past several decades (Fehr 1987). The diverse uses of soybean and increased demand for soy products require improvements in traits that enhance agronomic production and seed quality.

Many of the agronomic and seed quality traits are quantitative traits controlled by multiple genes having small or large effects. Quantitative trait loci (QTL) are regions on the chromosome that control quantitative traits. Molecular markers such as simple sequence repeat (SSR) and restriction fragment length polymorphism markers have made it easier to map QTL in soybean. There have been many QTL discovered for yield, protein, oil, maturity, height, and other traits (SoyBase, 1995).

Methods for QTL mapping have improved and now include single marker analysis of variance, interval mapping (Lander and Bostein, 1989), and composite interval mapping (CIM) (Zeng, 1993; Zeng, 1994). CIM is a QTL mapping procedure that tests an interval on a genetic map for significant QTL while using the most important markers to control for genetic background. This gives a location of the QTL and an estimate of its effect in a population.

In inbreeding species such as soybean, recombinant inbred lines (RIL) are advantageous since only additive effects are measured and they allow the opportunity to

test over multiple environments. Multiple trait analysis (MTA) is a QTL mapping procedure that will combine multiple environments in one model, taking advantage of the correlated structure of the data to improve the estimate of QTL chromosomal location and test for genotype x environment (GxE) interaction (Jiang and Zeng, 1995).

Significant GxE interaction can cause the overall additive effect in multiple environments to be severely diminished. Knowledge of significant GxE interaction for QTL would be beneficial in developing selection strategies such as marker assisted selection.

This study attempts to locate important agronomic and seed quality QTL and assay their level of GxE interaction. We will map SSR markers in an F₆ RIL population created from the cross between 'Essex' (Smith and Camper, 1973) and 'Williams' (Bernard and Lindahl, 1972). We will identify the genomic location of QTL and their GxE interaction for agronomic and seed quality traits measured over six field environments. This population should be considered a model population for studying segregating QTL because the parents contributed to the genetic background of many northern and southern U.S. cultivars and elite breeding lines (Sneller, 1994; Gizlice et al., 1996). In addition, this study is mapping segregating QTL in a north by south cross, which should produce a large amount of transgressive segregation. Combined with other research this study may help to broaden the understanding of QTL governing agronomic and seed quality traits.

Chapter 2

Analysis of Agronomic Traits in a Soybean Population to Discover QTL and Establish GxE Interaction

This chapter is a lightly revised version of a paper by the same name that will be submitted to the journal *Euphytica* in 2002 by David L. Hyten, Vincent R. Pantalone, Debbie Landau-Ellis, Arnold M. Saxton, Michael E. Schmidt, and Carl E. Sams.

My use of “we” and “our” in this chapter refers to my co-authors and myself. My primary contributions to this paper include (1) organization, development, and management of all aspects of the project, (2) mapping all of the molecular markers, (3) collecting all the field data at the Knoxville environments, (4) analyzing all data, and (5) all of the writing except for some minor revisions.

Introduction

There have been continuous gains in soybean [*Glycine max* (L.) Merr.] yield, increased disease resistance, and improvements in other traits of agronomic importance over the past several decades (Fehr, 1987). Unfortunately, breeding for higher yielding lines has created a problem of low genetic diversity among elite lines (Sneller, 1994; Gizlice et al., 1996). This low genetic diversity may cause future genetic gains to become more difficult. A proposed solution is to utilize the 15,000+ accessions of *G. max* maintained at the USDA soybean germplasm collection to introduce new genetic diversity into elite lines, enabling genetic gains in yield (Brown-Guedira et al., 2000).

The introduction of beneficial alleles through hybridization with diverse exotic germplasm is appealing. A negative consequence will be that accessions may contain many agronomically undesirable alleles closely linked to desirable alleles, making selection of superior lines difficult and time consuming. New tools such as molecular marker assisted selection (MAS) and marker assisted backcrossing have been proposed to

help incorporate unadapted germplasm rapidly (Reyna and Sneller, 2001; Frisch et al., 1999; Hospital et al., 1992). It has been shown that both of these new tools are useful when a breeder is incorporating one or two genes into an elite line from an undesirable line (Li et al., 2001).

Yield, maturity, height, and lodging have all been shown to be quantitative traits under the control of quantitative trait loci (QTL). These traits are largely affected by multiple genes. It is likely that the QTL are genes with promoters that are up or down-regulated by environmental factors. Breeders typically deal with this problem by selecting lines that show significant improvement in an agronomic trait over multiple environments where the line is to be adapted (Fehr, 1987). This has allowed the deployment of productive cultivars adaptable over broad geographical areas. Following this reasoning, QTL studies have used recombinant inbred lines (RIL), since they are homozygous and thus identical genotypes can be grown over multiple environments (Orf et al., 1999). The average RIL trait value from different environments is compared to molecular marker data to discover and map a QTL.

Further knowledge can be gained from QTL research by examining the existence of significant genotype x environment (GxE) interaction. A gene may only be turned on in a specific environment, due to a specific cue that affects the quantitative trait. Multiple trait analysis (MTA) is a QTL mapping procedure that will combine the multiple environments in one model, taking advantage of the correlated structure of the data to determine a better estimate of the chromosomal location and tests whether there is a significant GxE interaction for the QTL (Jiang and Zeng, 1995). Significant GxE interaction can cause the overall additive effect in multiple environments to be severely

diminished. Knowledge of significant GxE interaction could be beneficial in developing selection strategies to break tight linkages utilizing MAS.

Multiple trait analysis provides a better estimate of genomic location and will help soybean researchers to confirm candidate genes that have been found to affect agronomic traits in other species. The first step in confirmation of these candidate genes will be to locate them in the same regions of QTL. Then candidate genes will be assayed for environmental sensitivity using multiple environments that were used to identify the QTL answering the biological question of GxE interaction.

This study attempts to locate important agronomic QTL and assay their level of GxE interaction. We will map simple sequence repeat (SSR) markers in an F₆ RIL population created from the cross between 'Essex' (Smith and Camper, 1973) and 'Williams' (Bernard and Lindahl, 1972). We will identify the genomic location of QTL and their GxE interaction for agronomic traits measured over six field environments. This population should be considered a model population for studying segregating QTL because the parents contributed to the genetic background of many northern and southern U.S. cultivars and elite breeding lines (Sneller, 1994; Gizlice et al., 1996). In addition, this study is mapping segregating QTL in a north by south cross, which should produce a large amount of transgressive segregation. Combined with other research this study may help to broaden the understanding of QTL governing agronomic traits so that future specific candidate genes may be verified, and strategies developed on how to break tight linkages to integrate new germplasm.

Materials and Methods

One hundred and fifty-six F_6 RIL were created at The University of Tennessee, Knoxville Experiment Station (KES), Knoxville, TN from the cross between the soybean cultivars Essex and Williams. Sufficient seed was available to plant 131 RIL as $F_{6:8}$ lines during the year 2000 in a randomized complete block design with three replications in each of three environments [KES-C1 field and KES-S field, and at the West Tennessee Experiment Station (WTES) at Jackson, TN]. Two 6.1 m rows per plot were planted at two environments and one 6.1 m row per plot were planted in one environment. In 2001, the 131 RIL as $F_{6:9}$ lines were divided into three maturity groups (early, medium, and late) with 49 early, 37 medium, and 47 late with two lines overlapping in the medium and late maturity groups. Within each maturity group, a randomized complete block design with three replications each at two environments (KES-V field and Harrisburg, IL) and two replications at the third environment (WTES) were planted in four 6.1 m row plots. At each location, the different maturity group replications were randomly assigned to a location in the field. Maturity was recorded as the number of days after planting when 95% of the pods had reached mature pod color (Fehr, 1987). Plant height was measured as the average row height from the base of the plant to the top at maturity. Lodging was scored from 1 (all plants erect) to 5 (all plants prostrate). Yield was measured from rows end-trimmed to 4.9 m at all locations with all rows harvested for the one and two-row plots and the center two rows harvested for the four-row plots.

DNA was isolated from all 156 RIL by taking a single leaf from five plants and combining them into one sample. Then Qiagen Plant Easy DNA Extraction Kit (Qiagen, Hilden, Germany) was used to obtain purified DNA.

Essex and Williams were tested with 568 SSR markers previously placed on the integrated soybean genetic linkage map (Cregan et al., 1999). The sequence information for the soybean SSR primers used is publicly available from the USDA Internet site <http://129.186.26.94/ssr.html>; verified February 28, 2002 and relates to the integrated molecular linkage map of soybean (Cregan et al., 1999). Polymorphic markers were chosen to achieve a genome wide coverage of all 20 molecular linkage groups previously identified (Cregan et al., 1999).

PCR reactions were performed in a 384-well MBS Hybaid thermocycler (Hybaid Franklin, MA). The 10 μ l PCR mix consisted of 40 ng template DNA, 1x PCR Buffer, 0.2 mM dNTP mixture (Pharmacia, Piscataway, NJ), 1.0 μ M forward and reverse primer, and 0.5 units of KlenTaq (Ab Peptides, Inc, St. Louis, MO). The amplification conditions consisted of 94°C for 5 min followed by 35 cycles of 94°C denaturation for 25 s, 48°C annealing for 30 s, and 72°C extension for 30 s, ending with 1 cycle of 72°C for 5 min. PCR products were separated using 6% non-denaturing polyacrylamide gel electrophoresis with ethidium bromide staining for visualization. The RIL were scored (1 = Essex allele, 2 = heterozygote, and 3 = Williams allele) based on DNA band visualization. Because we utilized an F₆-derived RIL population, heterozygote scores were dropped from the data set to allow only additive effects to be detected.

One hundred and fifty-six lines were mapped with 100 SSR markers. MAPMAKER/EXP (Lander et al., 1987; Lincoln et al., 1992) was used to estimate genetic distances between SSR markers. A minimum likelihood of odds (LOD) ≥ 3.0 and a maximum distance ≤ 50 centimorgan (cM) were used to test linkages among markers.

QTL analysis was performed using the phenotypic traits and SSR RIL scores on 131 RIL planted in the field environments. A one-way ANOVA (SAS version 8.0) was initially used to identify candidate quantitative trait loci (results not reported). Composite interval mapping and multiple trait analysis were performed using QTL Cartographer (Basten et al., 1994; Basten et al., 2001). CIM was used to identify and map significant QTL. Multiple trait analysis was used to confirm and to refine the position of QTL found in CIM using the joint LOD score or the LOD from the GxE interaction test option of MTA. One thousand permutation tests were performed on each trait and respective environment to establish empirical LOD thresholds at the 5% alpha level (Churchill and Doerge, 1994). Heritability estimates were computed as

$$h^2 = \sigma^2_g / [\sigma^2_g + (\sigma^2_{ge}/e) + (\sigma^2/re)]$$

where h^2 = heritability; σ^2_g = genotypic variance; σ^2_{ge} = genotype x environment variance; σ^2 = error variance; r = number of replications, and e = number of environments (Nyquist, 1991).

QTL from this study were compared to similar QTL previously discovered (SoyBase, 1995). A QTL was considered new if either no similar QTL had been previously been associated with the linkage group or if the QTL was at least 20 cM distant from a previously reported QTL for that trait, based on the integrated soybean genetic linkage map (Cregan et al., 1999).

Results and Discussion

A total of 278 SSR markers were found to be polymorphic between these two parents. With the limited number of 100 markers mapped in this population, the integrated soybean genetic linkage map was used to help align the 23 genetic linkage

groups formed from 89 linked SSR markers and two phenotypic markers (flower color and pubescence) with the 20 known consensus linkage groups (Cregan et al., 1999). The Essex x Williams genetic linkage map covered about 869.4 cM, which is about 32% of the currently known recombination distance of 2748 cM, averaged from the integrated soybean genetic linkage map (Cregan et al., 1999). Continued placement of markers in the future should help to assure more coverage of the genome so that additional QTL can be detected.

A large difference in maturity was detected between the parental lines (Table 2.1). Likewise, there was a wide range of maturity among the progeny. This was expected since Essex is an early maturity group (MG) V cultivar, adapted to southern USA and Williams is a late MG III cultivar, adapted to mid-west USA (SoyBase, 1995).

There were a total of 11 intervals that were possible maturity QTL (denoted M1 to M11) detected in this population on six different linkage groups (LG) (Table 2.2). Maturity QTL M6 located on LG D1b has not been previously reported and may be a newly discovered QTL. M6 was the only QTL for maturity that did not have a significant GxE interaction. Maturity QTL M7 was not detected at the LOD 2.5 threshold at any of the environments but when the data was combined, significance was achieved along with a significant GxE interaction. Maturity QTL M8 located on LG J has been previously reported as linked to Satt287 to have an R^2 of 8% (Orf et al., 1999), which supports the 2-5% R^2 found at the different environments in our study. The maturity QTL M6-8 were all minor QTL that were either detected in one location or a few locations.

Table 2.1. Mean, standard deviation, range, parental means, and heritability for soybean agronomic traits in an Essex x Williams F₆ RIL population grown in six environments

Trait	Mean	Standard Deviation	Range	Essex	Williams	h ²
Yield (kg ha ⁻¹)	2323	261	1680-2879	2672	2499	0.76
Maturity (days)	123	8.0	107-140	129	117	0.98
Height (cm)	84	15.8	33-117	74	87	0.97
Lodging ^a	2.0	0.5	1.0-3.1	1.7	1.7	0.81

^a Lodging was scored from 1 (all plants erect) to 5 (all plants prostrate)

Table 2.2. Position, effects, and significance of environments associated with soybean maturity in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment ^d	R ² (%)	LOD	Effect ^e
M1	C2	T-Satt365	3.2	3.18	Combined	78	48.4	8.2
				3.18	WTES/00	73	28.2	8.3
				3.18	KES-S/00	77	47.5	9.2
				3.18	KES-C/00	61	27.6	6.7
				3.18	KES-V/01	64	36.4	8.4
				3.18	Harrisburg/01	67	35.4	6.0
				3.18	WTES/01	59	32.5	7.6
				3.18	GxE	-	48.1	2.9
M2	C2	Satt557-Satt289	3.3	1.34	KES-S/00	77	47.5	9.2
				1.34	WTES/01	59	28.1	7.7
				1.96	GxE	-	45.0	2.8
M3	C2	Satt100-Satt460	2.7	2.04	KES-S/00	76	21.2	9.3
				2.04	GxE	-	23.5	2.2
M4	C2	Satt460-Satt307	14.2	4.07	Combined	77	26.8	8.2
				4.07	WTES/00	73	21.9	9.5
				4.07	KES-S/00	76	27.2	9.2
				6.07	KES-C/00	66	10.9	6.7
				6.07	KES-V/01	68	19.1	8.5
				4.07	Harrisburg/01	65	18.4	5.8
				4.07	WTES/01	59	16.8	7.5
				4.07	GxE	-	35.5	2.2
M6	D1b	Satt141-Satt600	2.6	1.99	Harrisburg/01	4	3.3	1.4
				2.58	MTA	-	5.3	0.9
M7	F	Satt144-Satt522	21.9	12.01	Combined	3	2.5	-1.5
				14.01	GxE	-	3.1	-0.3
M8	J	Satt287-Satt285	15.1	0.01	Combined	2	2.6	1.2
				0.01	KES-S/00	2	2.8	1.4
				0.01	Harrisburg/01	5	4.3	1.5
				0.01	GxE	-	6.6	1.1

^a M1 to M11 denote the eleven maturity QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d KES-C/00 Knoxville Experiment Station (C1-field) at Knoxville, TN in 2000

WTES/00 West Tennessee Experiment Station at Jackson, TN in 2000

KES-S/00 Knoxville Experiment Station (S-field) at Knoxville, TN in 2000

KES-V/01 Knoxville Experiment Station (V-field) at Knoxville, TN in 2001

Harrisburg/01 Harrisburg, IL in 2001

WTES/01 West Tennessee Experiment Station at Jackson, TN in 2001

Combined data from three locations averaged

GxE Genetic vs. environmental interaction

MTA Multiple trait analysis

^e Average change (days) resulting from Essex allele substitution for the QTL

Table 2.2. Continued.

QTL	LG	Interval	Interval length (cM)	QTL position (cM)	Environment	R ² (%)	LOD	Effect
M9	L	Satt156-Satt166	16.6	13.97	Combined	6	6.1	-2.1
				15.97	WTES/00	7	6.8	-2.8
				11.97	KES-S/00	4	3.8	-2.0
				1.97	Harrisburg/01	5	4.4	-1.6
				1.97	WTES/01	6	4.2	-2.2
				13.97	GxE	-	3.9	-1.7
M10	L	Satt166-Satt229	29.5	8.05	Combined	9	6.6	-2.6
				10.05	KES-C/00	8	3.2	-2.2
				10.05	GxE	-	12.0	-1.4
M11	M	Satt540-463	18.4	0.01	Combined	6	8.0	2.2
				0.01	WTES/00	7	7.5	2.7
				2.01	KES-S/00	8	7.8	2.6
				12.01	KES-C/00	7	3.9	2.1
				0.01	KES-V/01	4	4.1	1.8
				2.01	Harrisburg/01	4	3.2	1.4
				0.01	GxE	-	4.5	1.7

The QTL located on linkage groups C2, L, and M were detected across multiple environments, and were found in other studies, suggesting that these may be QTL with underlying genes that regulate maturity. Although all of these QTL were detected in multiple environments, they all had significant GxE interaction. This is evident by the range of the additive effects in the multiple environments (Table 2.2). Linkage group C2 has been known to contain the *E1* (Bernard, 1971) and the newly discovered *E7* maturity genes (Cober et al., 2001). Cober et al. (2001) placed the *E1* gene 1.3 cM distal to the *T* (pubescence) gene and the *E7* gene 4.4 cM proximal of the *T* gene suggesting a gene order of *E1-T-E7*. Cober et al. (2001) noted the possibility of an alternative gene order of *T-E1-E7* but their study did not have molecular data available to establish definitive map order. Based on our study, maturity QTL M1 (detected in all environments) is the locus with the largest effect on this chromosome and should correspond to the reported distance between the *T* gene and the *E1* gene. Maturity QTL M2-4 would then become candidates for the *E7* gene. Maturity QTL M4 was detected in all environments, which makes it a good candidate for the *E7* gene. Our study supports the map order of *T-E1-E7*.

A window size of 10 cM was used in CIM and MTA. Any marker in this window is not used as a cofactor when an interval is being tested. This 10 cM window is used so that closely linked QTL are detected, but it increases the chance of a Type I error. For this reason, QTL tightly linked in this study, such as M2 and M3 could be maturity QTL or ghost QTL created from the artifact effect of a QTL within the 10 cM gap such as M1. If M2 and M3 are QTL, then map order differences could be due to the multiple maturity QTL closely linked to each other.

Several molecular studies have placed maturity QTL linked to Satt365, Satt489, Satt557, Satt291, Satt079K365_1, K474_1, K474_2, and A397_1 (Orf et al., 1999; Mansur et al., 1996; Mansur et al., 1993; Kiem et al., 1990) on LG C2 where maturity QTL M1-4 are found, suggesting that multiple *E* genes may reside in these intervals. All previously mentioned studies showed that the variance explained (R^2) ranged 13-23%.

The *E3* gene has been previously reported to be located on LG L (Cregan et al., 1999). In our population there were two maturity QTL (M9 and M10) found on LG L. It is unknown if either of these QTL may be the *E3* gene due to a low R^2 of 4-9%. Many studies have reported R^2 as low as 3-9% in this area linked to Satt166, Sat_099, Dt1, and O109_1, but there have been reports of larger R^2 ranging 10-29% from QTL linked to Satt006, Satt527, Satt373, G173_1, and A489_1 (Orf et al., 1999; Mansur et al., 1996; Specht et al., 2001; Lee et al., 1996a). It is doubtful that M9 or M10 is the *E3* gene since the six environments tested in this study all showed these two QTL to be minor QTL. Likely, *E3* is fixed in the Essex by Williams population and M13 and M14 are minor segregating genes that affect maturity, linked to a fixed *E3* gene.

This lends a solution for future research in distinguishing if minor QTL are tightly linked to major QTL for a specific trait. The first step is to create a population by either selecting parents or by selecting RIL that are fixed for the major QTL but are segregating for closely linked markers and minor QTL. With the major gene fixed, it would be possible to fine map the minor QTL to discover the true genetic contribution from the minor QTL that modulate expression of the trait.

Linkage group M contains maturity QTL M11 that is detected over six environments and has been reported by multiple studies to be linked to Satt150, Satt567,

Sat_003, Satt150, A584_3, R079_1, and A060_2, (Orf et al., 1999; Mansur et al., 1996; Specht et al., 2001; Lark et al., 1994; Mansur et al., 1993). The M11 QTL found in Essex x Williams RIL is possibly one of the many that other researchers have found in this region of LG M. The R^2 for the QTL at different environments in this study ranged from 4-8% while in other studies some of the QTL found in this region linked to R079_1 were as low as 6.8% (Mansur et al., 1996). Another study, with a different population found the same marker R079_1 linked to a major QTL that has an R^2 of 45% (Lark et al., 1994). This suggests that there is a minor and a major gene affecting maturity closely linked in this region of LG M.

The two parental lines showed some separation in height but most variability in the progeny is due to transgressive segregation (Table 2.1). Height was significantly correlated to maturity ($r=0.48$, $p<.0001$). There were a total of six QTL for height (all with significant GxE interaction) found in this population on linkage groups C2, F, and L (Table 2.3). The height QTL on these linkage groups were detected across environments and found in multiple studies, suggesting that these QTL may be present in different populations.

Height QTL H1-3 located on LG C2 are all considered major QTL having an R^2 above 12% for each environment. The GxE interaction found in MTA is verified by the fluctuation of the variance CIM explained by these QTL in the different environments. For example, height QTL H1 has a lower R^2 of 13% at KES-S location while the same QTL has a high R^2 of 32% at the KES-V location. Height QTL H2-3 also fluctuated for the proportion of the variability they explained. The R^2 reported in other studies ranged from 14%-20% in this same region linked to Satt079, Satt277, Satt489, K418_4, and

Table 2.3. Position, effects, and significance of environments for QTL associated with soybean height in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
H1	C2	Satt277-T	3.2	0.01	Combined	24	13.7	8.3
				0.01	WTES/00	18	9.0	8.0
				2.01	KES-S/00	13	6.8	6.9
				0.01	KES-V/01	32	14.7	11.4
				0.01	Harrisburg/01	16	7.4	7.0
				0.01	WTES/01	20	8.7	8.3
				2.01	GxE	-	15.5	1.9
H2	C2	T-Satt365	3.2	3.18	Combined	26	15.1	9.3
				3.18	WTES/00	22	10.6	9.3
				3.18	KES-S/00	14	7.8	7.7
				2.01	KES-C/00	24	10.2	5.7
				3.18	KES-V/01	36	17.0	13.0
				3.18	Harrisburg/01	16	6.9	7.4
				3.18	WTES/01	21	8.9	9.2
				3.18	GxE	-	18.4	2.2
H3	C2	Satt557-Satt289	3.3	1.96	KES-S/00	12	6.2	6.8
				1.96	KES-V/01	34	12.8	12.5
				1.96	Harrisburg/01	12	4.6	6.4
				1.96	WTES/01	17	6.2	8.1
		Satt100-Satt460	2.7	0.04	GxE	-	3.7	1.4
		Satt460-Satt307	14.2	2.07	WTES/00	12	4.7	6.7
				2.07	KES-C/00	12	3.7	3.8
H4	F	Satt114-Satt335	24.8	0.01	Combined	6	3.6	-3.9
				0.01	KES-V/01	5	3.0	-4.5
				0.01	WTES/01	8	3.5	-5.0
				0.01	GxE	-	4.1	-0.3
H5	L	Satt156-Satt166	16.6	15.97	WTES/00	46	18.9	-12.9
				13.97	KES-S/00	48	17.5	-14.2
				13.97	KES-C/00	31	10.3	-6.4
				15.97	KES-V/01	35	15.6	-12.3
				15.97	Harrisburg/01	35	13.1	-9.9
				13.97	WTES/01	34	11.5	-10.3
				11.97	GxE	-	17.0	-3.4
H6	L	Satt166-Satt229	29.5	4.05	Combined	55	24.3	-13.1
				4.05	WTES/00	59	21.6	-14.8
				6.05	KES-S/00	66	25.1	-16.4
				6.05	KES-C/00	52	15.5	-8.0
				6.05	KES-V/01	49	19.1	-15.0
				4.05	Harrisburg/01	41	14.0	-10.8
				4.05	WTES/01	36	11.6	-10.8
				6.05	GxE	-	27.6	-4.2

^a H1 to H6 denote the six height QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (cm) resulting from Essex allele substitution for the QTL

A397_1 on LG C2 (Mansur et al., 1993; Mansur et al., 1996; Lark et al., 1995; Mian et al., 1998; Orf et al., 1999; Specht et al., 2001). This serves as an example of how information on GxE interaction could be lost when combining environments to search for QTL.

The height QTL H2 on LG C2 map to the same location as the major maturity QTL M3 on this same linkage group (Figure 2.1), which would explain the correlation found between the two traits. This suggests that this region of the chromosome, which contains the *E* genes for maturity, has a pleiotropic effect on both height and maturity by signaling the plant to go from vegetative growth to reproductive growth. This hypothesis would explain why common genes for the two traits were detected in this chromosomal region.

Height QTL H4, found on linkage group F, has also been confirmed by other researchers. This QTL has an R^2 of 6% with the locations combined, which is close to the variability that has been explained in other studies with markers A186_1, P157_1, K644_1, and SOYHSP176, which showed an R^2 with a range of 4-8% (Lee et al., 1996a; Mian et al., 1998). The identification in our study, and confirmation at the same region of this LG in other studies, we have identified H5 as a minor height QTL.

Linkage group L contains two major height QTL for this population. The two height QTL could be due to the closely linked *Dt1* gene that is responsible for determinacy (Bernard, 1972; Cregan et al., 1999). Williams is an indeterminate cultivar while Essex is a determinate cultivar. Unfortunately, determinacy was not recorded for this experiment, and was not mapped. If the *Dt1* gene is responsible for detecting two

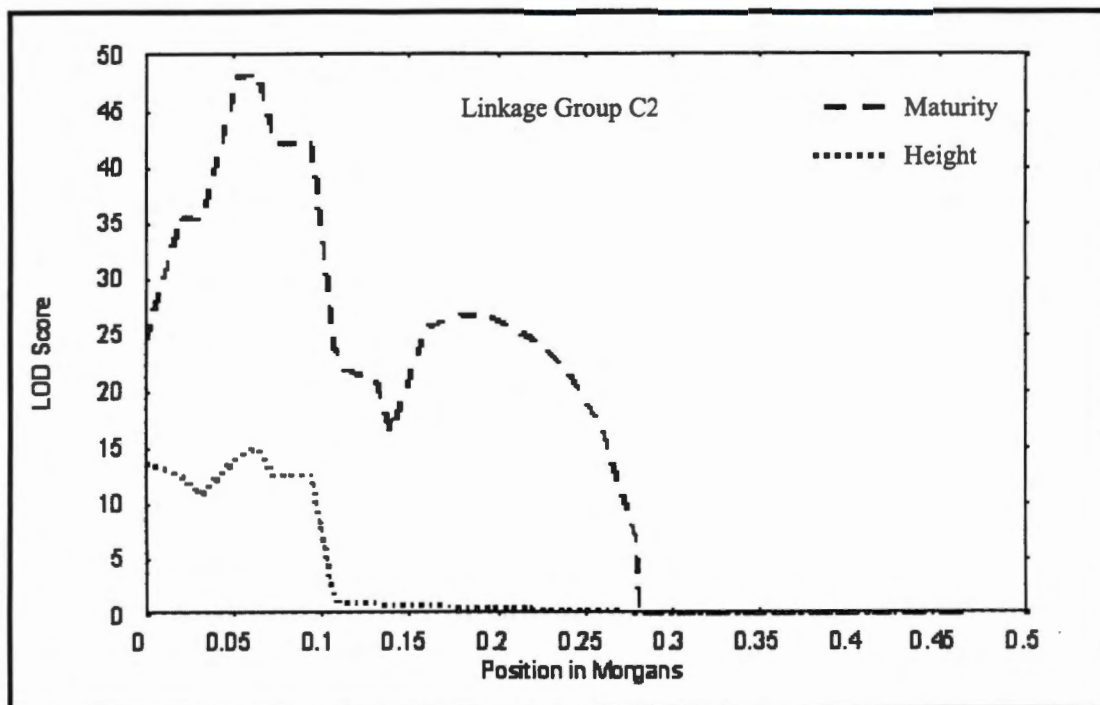


Figure 2.1. Maturity and Height LOD scores from six combined environments graphed along soybean linkage group C2

significant QTL, then the *DT1* gene effect on height is environmentally affected as evidenced by significant GxE interaction of H5.

There have been many studies that have found major height QTL in this same region linked to Satt166, Satt006, Satt527, A489_1, K385_1, G173_1, and Dt1 suggesting that there is a highly stable height QTL among populations (Mansur et al., 1996; Mansur et al., 1993; Lee et al., 1996b; Lark et al., 1995; Orf et al., 1999; Specht et al., 2001). In a study between PI97100 x Coker237 which is an indeterminate by a determinate population, there were major height QTL linked to K385_1 and Dt1 which contributed an R^2 of 67.7% and 64.1%, respectively (Lee et al., 1996b). However the GxE interaction that we identified in H5 has not been previously described.

There have also been QTL found in this chromosomal region that affect height when the two parents were *both* indeterminate (Specht et al., 2001; Orf et al., 1999; Lark et al., 1995). These other studies reported height QTL linked to A489_1, Dt1, and G173_1 having a lower R^2 (approximately 20-30%) indicating that even with the *DT1* gene fixed, there were other minor genes controlling height. This is an example of how genes that affect one trait can be clustered in the same chromosomal segment, and need to be mapped with the local major gene fixed in order for the true effect of the minor genes to be discovered.

The two parental lines are similar in lodging so the variability in the progeny is apparently due to transgressive segregation (Table 2.1). Lodging was significantly correlated to height ($r=0.62$ $p<0.0001$) and maturity ($r=0.39$ $p<0.0001$). There were a total of seven lodging QTL found in this population on LG C2, D1b, F, and L (Table 2.4).

Table 2.4. Position, effects, and significance of environments for QTL associated with soybean lodging in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
L1	C2	T-Satt365	3.2	3.18	Combined	12	6.4	0.2
				2.01	KES-C/00	14	4.5	0.1
				3.18	KES-V/01	14	6.1	0.4
				3.18	Harrisburg/01	18	8.2	0.5
				3.18	GxE	-	11.3	0.02
L2	C2	Satt460-Satt307	14.2	6.07	KES-C/00	20	3.5	-0.1
				8.07	MTA	-	5.2	-0.07
L3	D1b	Satt266-Satt141	36.3	0.01	Harrisburg/01	9	4.4	0.3
				0.01	GxE	-	5.7	-0.01
L4	F	W1-Satt160	2.4	2.05	Harrisburg/01	5	2.8	0.3
				2.05	GxE	-	3.7	-0.01
L5	L	Satt523-Satt182	18.0	15.99	Combined	11	5.1	-0.2
				15.99	KES-V/01	22	7.4	-0.4
				15.99	Harrisburg/01	13	5.8	-0.4
				15.99	WTES/01	15	5.1	-0.3
				11.99	GxE	-	8.3	0.01
L6	L	Satt481-Satt156	1.9	0.01	Combined	5	3.2	-0.1
				0.01	KES-V/01	6	2.8	-0.2
				0.01	GxE	-	3.0	0.01
L7	L	Satt166-Satt229	29.5	10.05	KES-S/00	20	5.4	-0.1
				4.05	GxE	-	4.7	-0.04

^a L1 to L7 denote the seven lodging QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change [1 (all plants erect) to 5 (all plants prostrate)] resulting from Essex allele substitution for the QTL

Most of the lodging QTL on these linkage groups exhibited significant GxE interaction; only lodging QTL L2 was stable across environments.

Lodging QTL L1-2 were detected on LG C2 and L5-7 were detected on LG L. This is probably due to pleiotropic effects from height and maturity QTL located at the same locations. This argument is supported by the fact that the effect of the Essex allele at these regions is consistent in virtually every interval, across environments, for all three traits. Several studies have found lodging QTL on LG C2 and LG L and those same studies reported height QTL in the same regions (Mansur et al., 1993; Mansur et al., 1996; Lee et al., 1996ab; Mian et al., 1998; Orf et al., 1999; Specht et al., 2001).

The lodging QTL L3 and L4 are minor QTL found on LG D1b and F respectively. They were significant in only one location and in MTA. Although LG F does have a height QTL, that significant interval is not linked to the intervals containing the lodging QTL. The two lodging QTL on LG D1b and F have not been previously reported, suggesting environmental sensitivity to our northernmost environment (Harrisburg). This suggests a need for verification before declaring them lodging QTL of value to breeders. There was severe lodging overall in the Harrisburg location, which may have led to an increased ability to detect QTL at that environment.

The two parental lines showed little separation in yield and most of the variability in the progeny was due to transgressive segregation (Table 2.1). Yield was negatively correlated to lodging ($r=-0.23$, $p<0.01$) in this RIL population. There were two yield QTL found in this population on LG A1 and L in specific environments each with significant GxE interaction (Table 2.5).

Table 2.5. Position, effects, and significance of environments associated with soybean yield in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
Y1	A1	Satt545-Satt236	24.5	23.99	Harrisburg/01	12	4.3	127
				19.99	GxE	-	3.6	21
Y2	L	Satt166-Satt229	29.5	2.05	WTES/01	13	4.4	198
				2.05	GxE	-	8.1	4

^a Y1 and Y2 denote the two yield QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (kg ha⁻¹) resulting from Essex allele substitution for the QTL

The yield QTL Y1 has not been previously reported. However it is not stable over environments, and could relate to a gene that benefits yield due to some environmental factor exerted in our northernmost environment. The yield QTL Y2 on linkage group L has been previously reported as a yield QTL linked to Satt561 and Dt1 with R^2 of 0.63% and 1.7%, respectively (Orf et al., 1999; Specht et al., 2001).

The high heritability for yield in this population (Table 2.1) and the lack of major QTL suggests that more areas of the genome need to be mapped in this population. With major QTL found for the other agronomic traits, there may be a highly heritable QTL that affects yield, which is unrelated to the traits mapped in this paper. Alternatively, yield is a complex trait, necessitating a collection of rare recombination events. Larger mapping populations may be required to detect yield QTL.

Conclusion

Several QTL have been mapped in this population for maturity, height, lodging, and yield. QTL for maturity, height, and lodging were found to be clustered on LG C2 and L. When a major gene on the linkage group is segregating among the RIL, detected effects of the minor QTL are influenced.

Utilization of composite interval mapping and multiple trait analysis demonstrated how most of the QTL discovered for the agronomic traits have significant GxE interaction. The estimates for the QTL detected in this study may help breeders outline appropriate MAS strategies. For example, they would be useful for knowing where to concentrate markers to detect when tight linkages have been broken between undesirable alleles and desirable alleles when integrating new or exotic germplasm for soybean improvement.

Chapter 3

Molecular Mapping and Identification of GxE Interactions of Soybean

QTL for Seed Quality Traits

This chapter is a lightly revised version of a paper by the same name that will be submitted to the journal *Theoretical and Applied Genetics* in 2002 by David L. Hyten, Vincent R. Pantalone, Carl E. Sams, Arnold M. Saxton, Debbie Landau-Ellis, and Michael E. Schmidt.

My use of “we” and “our” in this chapter refers to my co-authors and myself. My primary contributions to this paper include (1) organization, development, and management of all aspects of the project, (2) mapping all of the molecular markers, (3) collecting all the field data at the Knoxville environments, (4) analyzing all data, and (5) all of the writing except for some minor revisions.

Introduction

Soybean is grown worldwide with the United States, Brazil, China, Argentina, and India being the major producers (Liu, 1997). The major economic role of soybean production is for use in soybean meal, a primary component of livestock feed that provides protein. An important by-product of meal production is soybean oil, extracted from the crushed meal. Soybean oil is a vegetable oil used worldwide in cooking, frying fats and in spreadable margarines. Seed size is important in soybean products used for direct human consumption such as tofu, miso, and texturized soy vegetable protein.

An increase in seed protein concentration in commonly grown cultivars is needed to keep soybean competitive as a livestock feed. The typical negative correlation between protein and oil has caused difficulties in increasing the protein while keeping seed oil concentration at current levels. Protein, oil, and seed size are quantitative traits that are environmentally sensitive and are controlled by multiple genes having small or large effects. Recently there have been many studies that have used molecular markers to

map quantitative trait loci (QTL), the regions of the chromosome responsible for quantitative traits such as protein, oil, and seed size (Diers et al., 1992; Sebolt et al., 2000; Csanadi et al., 2001). The genetic effect that QTL have on quantitative traits could then be selected for using marker assisted selection (MAS), hence improving the trait in new cultivars. One approach of using MAS is to introduce alleles from exotic germplasm, such as *Glycine soja* Sieb, and Zucc., which will increase protein content in soybean (Sebolt et al., 2000).

The two ancestral cultivars 'Essex' (Smith and Camper, 1973) and 'Williams' (Bernard and Lindahl, 1972) have contributed to the genetic background of many northern and southern USA cultivars and elite breeding lines (Sneller, 1994; Gizlice et al., 1996). QTL discovered segregating between these two cultivars should be able to serve as a baseline for the effect QTL already present in commonly grown cultivars have on seed protein concentration. This would allow researchers to know which QTL from exotic germplasm would be beneficial to integrate into commonly grown soybeans.

If MAS introgression of QTL is going to be successful, then the true genetic effect of the exotic allele along with the alleles already present needs to be known. It has recently been noted that QTL discovered in one geographical region might not have the predicted effect in another geographical region (Reyna et al., 2001). This is likely due to the current mapping techniques used which map QTL using phenotypic data collected over multiple environments (usually four or less) and an overall mean is calculated to avoid specific environmental effects (Orf et al., 1999). Others have mapped QTL in specific environments and compared the stability of the QTL over the multiple environments to determine the usefulness of a QTL (Mian et al., 1996).

While it is appealing to look at multiple environments to determine environmental sensitivity or to average over environments to eliminate environmental sensitivity, the true genetic effect that a QTL contributes is overlooked and the usefulness of the QTL overestimated. A recent QTL mapping technique, multiple trait analysis (MTA) was developed that can take full advantage of phenotypic data collected on recombinant inbred lines (RIL) over multiple locations (Jiang and Zeng, 1995). MTA will analyze each environment in one model, utilizing the correlated structure of data to assign a better estimate of map location and to determine the genotype x environment (GxE) interaction. QTL mapped in a RIL population using MTA will have a better estimate of the additive effect along with the environmental sensitivity of the QTL. The objective of this study was to map protein, oil, and seed size QTL in multiple environments using composite interval mapping (CIM) along with MTA to establish representative additive effects and GxE interaction of the QTL.

Materials and Methods

One hundred and fifty-six F_6 RIL were created at The University of Tennessee, Knoxville Experiment Station (KES), Knoxville, TN from the cross between the soybean cultivars Essex and Williams. Sufficient seed were available to plant 131 RIL as $F_{6:8}$ lines during the year 2000 in a randomized complete block design with three replications in each of three environments [KES-C field and KES-S field, and the West Tennessee Experiment Station (WTES) at Jackson, TN]. Two 6.1 m rows per plot were planted at KES-S and WTES. One 6.1 m row per plot was planted at KES-C. In 2001, the 131 RIL as $F_{6:9}$ lines were divided into three maturity groups (early, medium, and late) with 49 early, 37 medium, and 47 late, with two lines overlapping in the medium and late

maturity groups. Within each maturity group, a randomized complete block design with three replications each at KES-V and Harrisburg, IL and two replications at WTES were planted in four 6.1 m row plots. At each environment, the different maturity group replications were randomly assigned to a location within the field.

Seed protein and oil concentration were determined at the USDA Northern Regional Research Center at Peoria, IL. Measurements were taken on a moisture-free basis using near-infrared reflectance spectroscopy on 21 to 25 g of whole beans from each plot. Seed size was determined by weighing 100 seeds from each plot.

DNA was isolated from all 156 RIL by taking a single leaf from five plants and combining them into one sample. Then Qiagen Plant Easy DNA Extraction Kit (Qiagen, Hilden, Germany) was used to obtain purified DNA.

Essex and Williams were tested with 568 simple sequence repeat (SSR) markers previously placed on the integrated soybean genetic linkage map (Cregan et al., 1999). The sequence information for the soybean SSR primers used is publicly available from the USDA Internet site <http://129.186.26.94/ssr.html>; verified February 28, 2002 and relates to the integrated molecular linkage map of soybean (Cregan et al., 1999). Polymorphic markers were chosen to achieve a genome wide coverage of all 20 molecular linkage groups previously identified (Cregan et al., 1999).

PCR reactions were performed in a 384-well MBS Hybaid thermocycler (Hybaid Franklin, MA). The 10 μ l PCR mix consisted of 40 ng template DNA, 1x PCR Buffer, 0.2 mM dNTP mixture (Pharmacia, Piscataway, NJ), 1.0 μ M forward and reverse primer, and 0.5 units of Klentaq (Ab Peptides, Inc, St. Louis, MO). The amplification conditions consisted of 94°C for 5 min followed by 35 cycles of 94°C denaturation for 25 s, 48°C

annealing for 30 s, and 72°C extension for 30 s, ending with 1 cycle of 72°C for 5 min. PCR products were separated using 6% non-denaturing polyacrylamide gel electrophoresis with ethidium bromide staining for visualization. The RIL were scored (1 = Essex allele, 2 = heterozygote, and 3 = Williams allele) based on DNA band visualization. Because we utilized an F₆-derived RIL population, heterozygote scores were dropped from the data set to allow only additive effects to be detected.

One hundred and fifty-six lines were mapped with 100 SSR markers. MAPMAKER/EXP (Lander et al., 1987; Lincoln et al., 1992) was used to estimate genetic distances between SSR markers. A minimum likelihood of odds (LOD) ≥ 3.0 and a maximum distance ≤ 50 centimorgan (cM) were used to test linkages among markers.

QTL analysis was performed using the phenotypic traits and SSR RIL scores on 131 RIL planted in the field environments. A one-way ANOVA (SAS version 8.0) was initially used to identify candidate quantitative trait loci (results not reported). Composite interval mapping and multiple trait analysis were performed using QTL Cartographer (Basten et al., 1994; Basten et al., 2001). CIM was used to identify and map significant QTL. Multiple trait analysis was used to confirm and to refine the position of QTL found in CIM using the joint LOD score or the LOD from the GxE interaction test option of MTA. One thousand permutation tests were performed on each trait and respective environment to establish empirical LOD thresholds at the 5% alpha level (Churchill and Doerge, 1994). Heritability estimates were computed as

$$h^2 = \sigma^2_g / [\sigma^2_g + (\sigma^2_{ge}/e) + (\sigma^2_{re})]$$

where h^2 = heritability; σ_g^2 = genotypic variance; σ_{ge}^2 = genotype x environment variance; σ^2 = error variance; r = number of replications, and e = number of environments (Nyquist, 1991).

QTL from this study were compared to similar QTL previously discovered (SoyBase, 1995). A QTL was considered new if either no similar QTL had been previously been associated with the linkage group or if the QTL was at least 20 cM distant from a previously reported QTL for that trait, based on the integrated soybean genetic linkage map (Cregan et al., 1999).

Results and Discussion

A total of 278 SSR markers were found to be polymorphic between these two parents. With the limited number of 100 markers mapped in this population, the integrated soybean genetic linkage map was used to help align the 23 genetic linkage groups formed from 89 linked SSR markers and two phenotypic markers (flower color and pubescence) with the 20 known consensus linkage groups (Cregan et al., 1999). The Essex x Williams genetic linkage map covered about 869.4 cM, which is about 32% of the currently known recombination distance of 2748 cM, averaged from the integrated soybean genetic linkage map (Cregan et al., 1999). Continued placement of markers in the future should help to assure more coverage of the genome so that additional QTL can be detected.

The summary of seed protein and oil concentration is presented in Table 3.1. Transgressive segregation was observed in this population, which was expected due to the parents' diversity with one parent from the northern USA germplasm pool and the other parent from the southern USA germplasm pool. There were a total of six oil QTL

Table 3.1. Mean, standard deviation, range, parental means, and heritability for soybean seed traits in an Essex x Williams F₆ RIL population grown in six environments

Seed Traits	Mean	Standard Deviation	Range	Essex	Williams	h ²
Seed size (mg seed ⁻¹)	131	15	99-166	120	154	0.95
Oil conc. (g kg ⁻¹)	203.5	5.1	189.8-215.2	199.6	209.1	0.84
Protein conc. (g kg ⁻¹)	422.7	9.7	399.7-450.9	429.5	424.7	0.91

and 13 protein QTL found in this population (Table 3.2 and Table 3.3, respectively). Oil QTL O2 and O6 were found in the same intervals as protein QTL P5 and P12 (respectively) with a negative pleiotropy between the two QTL, which verifies the negative correlation found between the two traits in this population ($r=-0.60$, $p<0.0001$). Only four of the protein QTL (P6 and P9-11) and two of the oil QTL (O2 and O3) did not have a significant GxE interaction. All protein and oil QTL with significant GxE interaction ($\text{LOD} \geq 3.0$) showed a decrease in the additive effect the QTL contributed in this population. The more environmentally sensitive the QTL, the less beneficial the QTL will be in different environments. Just as yield is tested in multiple environments to understand the yield potential of a soybean line, QTL for any trait has to be tested in multiple environments to discover the potential additive effect of that QTL.

Protein concentration QTL P1, located on linkage group (LG) C1, was only significant in one environment and showed a significant GxE interaction. The additive effect of 2.5 g kg^{-1} was diminished to 1.1 g kg^{-1} when mapped over multiple environments (Table 3.3). Lee et al. (1996) discovered this same QTL linked to RFLP marker *gc197_1* mapped over three locations (without the benefit of MTA) with an additive effect of 6 g kg^{-1} , which is a larger additive effect than in our population.

The protein QTL P2-5 on LG C2 were all significant for GxE interaction (Table 3.3). Protein QTL P2-4 showed a decrease in the additive effect due to the GxE interaction, hence they may not be protein QTL that would be effective to a breeder. The protein QTL P5 was significant across all environments but the additive effect varied across those environments. Even with the significant GxE interaction, this QTL may be

Table 3.2. Position, effects, and significance of environments for QTL associated with soybean seed oil concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment ^d	R ² (%)	LOD	Effect ^e
O1	C2	Satt277-T	3.2	2.01	WTES/00	8	3.1	-2.5
				2.01	KES-S/00	27	12.0	4.5
		T-Satt365	3.2	0.01	KES-C/00	14	6.2	3.4
				3.18	Harrisburg/01	17	6.5	-2.9
				0.01	WTES/01	17	7.1	-3.0
				3.18	GxE	-	23.6	-0.5
O2	C2	Satt460-Satt307	14.2	2.07	KES-S/00	19	6.6	3.7
				4.07	MTA	-	3.3	1.4
O3	H	Satt192-Satt442	5.0	2.01	Combined	9	3.3	-1.5
				2.01	MTA	-	4.3	-1.4
O4	L	Satt156-Satt166	16.6	13.97	Combined	7	2.6	-1.5
				11.97	KES-V/01	12	3.7	-2.3
				7.97	GxE	-	4.0	-0.5
O5	L	Satt166-Satt229	29.5	14.05	Combined	23	5.6	-2.6
				26.05	KES-S/00	10	3.0	-2.6
				8.05	KES-V/01	15	3.9	-2.5
				10.05	GxE	-	12.1	-1.4
		Satt229-Satt373	12.4	0.01	KES-C/00	8	3.5	-2.2
O6	M	Satt540-Satt463	18.4	0.01	WTES/00	8	3.6	-2.4
				4.01	Harrisburg/01	9	2.9	-1.7
				6.01	WTES/01	18	5.7	-2.7
				4.01	GxE	-	6.8	-1.7

^a O1 to O6 denote the six seed oil content QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d KES-C/00 Knoxville Experiment Station (C1-field) at Knoxville, TN in 2000

WTES/00 West Tennessee Experiment Station at Jackson, TN in 2000

KES-S/00 Knoxville Experiment Station (S-field) at Knoxville, TN in 2000

KES-V/01 Knoxville Experiment Station (V-field) at Knoxville, TN in 2001

Harrisburg/01 Harrisburg, IL in 2001

WTES/01 West Tennessee Experiment Station at Jackson, TN in 2001

Combined data from three locations averaged

GxE Genetic vs. environmental interaction

MTA Multiple trait analysis

^e Average change (g kg⁻¹) resulting from Essex allele substitution for the QTL

Table 3.3. Position, effects, and significance of environments for QTL associated with soybean seed protein concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
P1	C1	Satt338-Satt164	6.5	0.01	KES-V/01	4	3.7	2.5
				0.01	GxE	-	4.0	1.1
P2	C2	Satt277-T	3.2	0.01	Combined	8	6.1	-3.1
				2.01	KES-S/00	10	4.9	-4.7
				2.01	KES-C/00	15	7.8	-6.5
				2.01	KES-V/01	9	6.6	-3.8
				2.01	GxE	-	15.4	-1.6
P3	C2	T-Satt365	3.2	2.01	Combined	7	5.1	-3.3
				3.18	KES-S/00	13	6.5	-5.8
				3.18	KES-C/00	16	8.7	-7.4
				3.18	KES-V/01	9	6.4	-4.1
				3.18	GxE	-	20.1	-1.8
P4	C2	Satt557-Satt289	3.3	1.96	Combined	6	4.1	-3.2
				1.96	KES-V/01	7	5.1	-4.0
				1.96	GxE	-	16.2	-1.9
P5	C2	Satt460-Satt307	14.2	14.07	Combined	11	8.1	-4.5
				14.07	WTES/00	13	5.1	-4.2
				6.07	KES-S/00	32	10.8	-8.2
				6.07	KES-C/00	33	11.7	-8.6
				12.07	KES-V/01	25	12.4	-6.7
				8.07	Harrisburg/01	22	8.0	-5.6
				8.07	GxE	-	11.7	-4.2
P6	C2	Satt307-Satt202	6.5	2.07	KES-V/01	11	6.2	-4.9
				2.07	MTA	-	8.5	-4.3
P7	D1b	Satt172-Satt459	17.8	14.14	KES-V/01	6	3.5	2.9
				10.14	GxE	-	3.1	1.7
P8	F	Satt114-Satt335	24.8	24.01	Combined	4	3.0	2.0
				16.01	GxE	-	4.3	1.2
P9	F	Satt144-Satt522	21.9	4.01	Combined	5	2.8	2.3
				4.01	WTES/00	10	3.7	3.7
				0.01	KES-V/01	5	4.3	2.6
				6.01	Harrisburg/01	8	3.3	3.1
				0.01	MTA	-	6.7	2.6
P10	I	Satt292-Satt148	21.5	2.01	Combined	5	3.6	2.5
				14.01	MTA	-	3.4	1.8
P11	K	Satt273-Satt260	22.4	14.07	Combined	5	2.6	-2.3
				8.07	MTA	-	3.4	-1.8
P12	M	Satt540-Satt463	18.4	8.01	WTES/01	14	2.9	3.8
				12.01	GxE	-	4.2	0.5
		Satt463-Satt323	4.1	4.05	Combined	6	4.4	3.0
P13	N	Satt159-Satt584	13.7	4.01	Combined	10	4.5	3.6
				2.01	GxE	-	5.2	0.3

^a P1 to P13 denote the six seed protein content QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c Distance from the left marker of the interval

^d Average change (g kg⁻¹) resulting from Essex allele substitution for the QTL

beneficial due to the northern parent (Williams) having an additive effect of 4.2 g kg^{-1} over multiple environments.

The protein content of soybeans produced in northern USA is typically lower than that in southern USA (Wilson, 2002). It would be interesting to screen an array of current northern cultivars for the presence or absence of the Williams P5 allele. If this allele is typically absent from current northern USA cultivars, its broad-based incorporation could rapidly and systematically improve the USA domestic soybean protein production. If this allele is absent from southern cultivars, less dramatic gains could be realized because considerably less land area is devoted to southern soybean production. Protein QTL P6 also had a significant additive effect without a significant GxE interaction, but it may be a residual QTL resulting from the proximity and strong additive effect of P5.

Csandi et al. (2001) discovered a protein QTL close to P5 and P6 in two early maturing cultivars adapted to Central Europe, having an additive effect of 4.1 g kg^{-1} , which is virtually the same estimated additive effect of the Williams allele in our population. Protein QTL P5's additive effect could be misrepresented as contributing a much higher additive effect if it is not mapped in multiple environments covering a large geographical area. Four of the environments showed additive effects higher than 4.2 g kg^{-1} and if it was only mapped in those four environments, breeders would expect a larger response to increase protein from integrating that allele into southern USA germplasm. Nevertheless P5 appears to be a good candidate for further research and MAS.

Oil QTL O1 and O2 were also found on LG C2 (Table 3.2). Oil QTL O1 had a GxE interaction where the Essex allele contributed to a higher oil concentration in the

eastern environments and the same allele contributed to lower oil concentration in the western environments of our study. O2 did not have a significant GxE interaction. However it would lower oil concentration in the seed if the protein QTL P5 were used in MAS to raise protein.

Protein QTL P9 has been verified by Qiu et al. (1999), who found the same beneficial allele as the one in Essex, having an additive effect of 4 g kg^{-1} . In our study, this QTL did not have a significant GxE interaction, and additive effects ranged from 2.3 to 3.7 g kg^{-1} over environments. Multiple trait analysis predicts that P9 should contribute 2.6 g kg^{-1} over multiple environments, and may be useful to integrate into northern USA germplasm.

Protein QTL P11 may be the same QTL found by Csandi et al. (2001), which mapped a protein QTL linked to Satt196 with an additive effect of 3.2 g kg^{-1} . Protein QTL P11 was a beneficial northern allele from Williams that had an additive effect of 1.8 g kg^{-1} . P11 did not show significant GxE interaction. This is an allele that might be beneficial to southern breeders to integrate into southern USA germplasm.

Protein QTL P12 had a significant GxE interaction that greatly diminished the additive effect of the allele (Table 3.3). The diminished additive effect along with the presence of the opposing oil QTL O6 (Table 3.2) makes this protein QTL P12 of little use to breeders. There is a maturity QTL that has been found closely linked to protein QTL P12 inferring that there is pleiotropy due to the beneficial allele from the later maturing line (Essex). A protein QTL near this interval has been reported by Csandi et al. (2001) and Specht et al. (2001) to have an additive effect of 3.9 g kg^{-1} and 8.3 g kg^{-1} , which

would have agreed our study had we not employed MTA and detected a significant GxE interaction.

Protein QTL P13 is another reported QTL that had a significant GxE interaction, which diminished the additive effect. Lee et al. (1996) reported a protein QTL linked to A071_2 with an additive effect of 5 g kg⁻¹. This QTL is not a QTL useful to breeders due to the significant GxE interaction and resulting additive effect of only 0.3 g kg⁻¹ protein. Mapping in multiple environments and using multiple trait analysis is helpful in identifying which QTL are potentially beneficial for breeders to utilize for MAS.

Protein QTL P7, P8, and P10 have not been previously reported QTL. Multiple trait analysis increases the power to detect QTL, but increases the chance of declaring a QTL that is not actually significant (type I error) (Jiang et al., 1995). Hence, other researchers need to verify these newly detected protein QTL in other populations before they become candidate QTL for MAS.

Oil QTL O3, located on LG H, did not have a significant GxE interaction (Table 3.2). Brummer et al. (1997) also identified this QTL linked to A069_1, associated with oil, which was stable across environments over three years of research. This supports our study of using MTA to identify QTL that do not have a significant GxE interaction. Oil QTL O4 and O5 are located on LG L with significant GxE interaction. The additive effect was diminished enough for O4 that it would not be a useful QTL. Oil QTL O5 has an additive effect that may be useful to increase seed oil concentration. Our study shows that its additive effect is less than that reported by other researchers. Mansur et al. (1996) reported an oil QTL linked to Satt006 with an additive effect of 5.7 g kg⁻¹, which is larger than the 1.4 g kg⁻¹ the Williams allele contributes in our study. It should also be noted

that there have been maturity QTL found in this same interval that may have a pleiotropic effect on seed oil concentration (Table 2.2 and Table 3.2).

The summary of seed size is presented in Table 3.1. Transgressive segregation was observed in this population, which was expected due to the parents' diversity of one parent from northern germplasm and the other parent from southern germplasm. There were a total of nine seed size QTL (Table 3.4). All seed size QTL with significant GxE interaction ($\text{LOD} \geq 3.0$) showed a decrease in the additive effect that the QTL contributed in this population.

Seed size QTL SS1 located on LG A1 and SS2-3 located on LG C2 had significant GxE interactions that diminished the additive effect (Table 3.4). If these QTL had been mapped solely based on CIM, then they would be considered QTL worth pursuing in MAS, and limited genetic gain would be realized as a consequence. These QTL would not be useful for breeders to select, due to their environmental sensitivity.

Seed size QTL SS4, SS5, and SS9 also had significant GxE interactions with smaller effects than predicted by CIM (Table 3.4). Specht et al. (2001) reported a seed size QTL linked to Satt342 on LG F that has an additive effect of 2.5 mg seed^{-1} . This QTL corresponds to SS4 and would agree with the CIM estimate found in this population, while MTA gives an estimate of 1.6 mg seed^{-1} . Mian et al. (1996) estimated a seed size QTL with an additive effect of 7.0 mg seed^{-1} near SS5. While the estimate for SS5 with CIM in our population is not as high as 7.0 mg seed^{-1} , it is similar in value. The estimate from MTA shows that with environmental influence, this is not a major QTL but only a minor QTL allele from Essex contributing 1.8 mg seed^{-1} . Seed size QTL SS9 on LG L has been confirmed by many other studies linked to Dt1, K385_1, Satt229,

Table 3.4. Position, effects, and significance of environments for QTL associated with soybean seed size in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
SS1	A1	Satt050-Satt545	35.0	10.00	KES-C/00	9	3.0	4.3
				10.00	GxE	-	3.5	0.9
SS2	C2	T-Satt365	3.2	3.18	WTES/00	7	3.2	-4.0
				2.01	Harrisburg/01	15	9.2	-7.8
				2.01	GxE	-	18.4	-0.2
		Satt557-Satt289	3.3	3.28	Combined	8	5.1	-4.5
				3.28	WTES/00	7	3.5	-4.1
				3.28	KES-V/01	12	7.2	-8.5
SS3	C2	Satt460-Satt307	14.2	2.07	Combined	7	3.9	-4.1
				6.07	GxE	-	11.7	-0.1
SS4	F	Satt146-Satt269	9.1	6.01	KES-S/00	5	2.8	3.5
				8.01	GxE	-	3.2	1.6
SS5	F	Satt114-Satt335	24.8	12.01	Combined	7	3.1	3.9
				10.01	WTES/00	8	2.8	4.0
				8.01	KES-S/00	14	5.6	5.8
				0.01	KES-C/00	8	5.3	4.2
				20.01	KES-V/01	7	3.5	5.9
				0.01	GxE	-	3.2	1.8
SS6	G	Satt394-Satt340	5.8	0.01	KES-S/00	4	2.8	-3.3
				4.01	MTA	-	3.2	-1.6
SS7	G	Satt340-Satt352	7.2	6.05	KES-C/00	8	4.8	-4.1
				6.05	GxE	-	3.4	-1.9
SS8	G	Satt352-Satt138	8.5	4.05	KES-C/00	10	5.7	-4.7
				4.05	MTA	-	3.8	-2.6
SS9	L	Satt156-Satt166	16.6	13.97	Combined	31	14.7	-8.7
				13.97	WTES/00	22	7.5	-6.9
				13.97	KES-C/00	28	14.1	-7.7
				11.97	WTES/01	23	7.9	-6.4
				15.97	GxE	-	13.0	-5.0
		Satt166-Satt229	29.5	2.05	Combined	31	14.6	-8.6
				2.05	WTES/00	23	8.0	-6.9
				0.05	KES-S/00	35	17.1	-9.7
				6.05	KES-V/01	30	10.0	-12.4
				0.05	Harrisburg/01	34	14.8	-10.9
				4.05	WTES/01	22	7.6	-6.1

^a SS1 to SS9 denote the nine seed size QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (mg seed⁻¹) resulting from Essex allele substitution for the QTL

Sat_099, and Satt527 with an additive effect ranging from 3.0-13.0 mg seed⁻¹ (Maughan et al., 1996; Orf et al., 1999; Csandi et al., 2001; Mian et al., 1996). This confirms the variability found in our study with CIM due to the environment, while MTA gives an additive effect of 5.0 mg seed⁻¹. Despite the environmental influence, seed size QTL SS4, 5, and 9 have an effect on seed size that breeders could utilize.

Seed size QTL SS6-8 are located on LG G, which contains the only environmentally stable QTL (SS6 and SS8) for seed size found in this study (Table 3.4). Even though these two QTL were environmentally stable, MTA predicted a reduced additive effect when compared to CIM. Maughan et al. (1996) also reported a QTL near SS6-8 linked to A816_1 that had an additive effect of 9.0 mg seed⁻¹, which is much higher than that found in our study.

Conclusion

There were a total of 13 protein, six oil, and nine seed size QTL found in this F₆ RIL population created from crossing Essex and Williams. Most of the QTL found had significant GxE interaction that reduced their additive effect. The additive effect found from multiple trait analysis is an improvement from previous mapping techniques that do not test for GxE interaction. The QTL that were verified in this population could be used in marker assisted selection to improve seed quality traits. The predicted additive effects of the QTL should be useful to breeders when they are deciding which QTL to utilize for soybean improvement.

Chapter 4

Molecular Mapping and Identification of GxE Interactions of Soybean

Fatty Acid Modifier QTL

This chapter is a lightly revised version of a paper by the same name that will be submitted to the journal *Crop Science* in 2002 by David L. Hyten, Vincent R. Pantalone, Debbie Landau-Ellis, Arnold M. Saxton, Elizabeth J. Meyer, Michael E. Schmidt, and Carl E. Sams.

My use of “our” in this chapter refers to my co-authors and myself. My primary contributions to this paper include (1) organization, development, and management of all aspects of the project, (2) mapping all of the molecular markers, (3) collecting all the field data at the Knoxville environments, (4) analyzing all data, and (5) all of the writing except for some minor revisions.

Introduction

One of the main uses of soybean is for extracted vegetable oil in food products. Soybean oil is predominately composed of five fatty acids: palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2), and linolenate (18:3). There has been a concerted effort to breed soybeans that have fatty acid levels differing from common soybean cultivars. Consumers of soybean oil are concerned about the concentration of saturated fats such as palmitate. There are also concerns for levels of *trans*-fatty acids formed routinely during the hydrogenation process of soybean oil, since they may adversely affect human health (Willett, 1994; Beare-Rogers, 1995). Processors want an increase in oleate along with a reduction of linolenate to improve storage life of soybean oil.

Several genes have been discovered that alter fatty acid content considerably. Some are naturally occurring, and others were produced through mutation (Wilson et al., 2002; Erickson et al., 1988; Fehr et al., 1991; Rennie et al., 1988; Wilcox and Cavins, 1985). It has been suggested that there are modifying genes along with the major fatty

acid genes that affect the different fatty acids levels (Rebetzke et al., 1998; Lundeen et al., 1987; Graef et al., 1988). These modifier genes have small effects that cause the fatty acid profiles of populations to act as a quantitative trait.

Quantitative traits are controlled by multiple genes having small or large effects. Quantitative trait loci (QTL) are regions on the chromosome that control quantitative traits. Molecular markers such as simple sequence repeat (SSR) markers and restriction fragment length polymorphism (RFLP) markers have made it easier to map QTL in soybean. There have been many QTL discovered for yield, protein, oil, maturity, height, etc. (SoyBase, 1995). Methods for QTL mapping have also improved and now include single marker analysis of variance, interval mapping (Lander and Bostein, 1989), and composite interval mapping (CIM) (Zeng, 1993; Zeng, 1994). CIM is a QTL mapping procedure that tests an interval on a genetic map for significant QTL while using the most important markers to control for genetic background. This gives a location and an estimated effect of the QTL in a population.

In inbreeding species such as soybean, recombinant inbred lines (RIL) are advantageous since only additive effects are measured and they can be tested over multiple environments. Multiple trait analysis (MTA) is a QTL mapping procedure that combines multiple environments in one model, taking advantage of the correlated structure of the data to improve the estimate of QTL chromosomal location and test for GxE interaction (Jiang and Zeng, 1995). Significant genotype x environment (GxE) interaction can cause the overall additive effect in multiple environments to be severely diminished. Knowledge of significant GxE interaction for fatty acid QTL would be beneficial in developing selection strategies such as marker assisted selection (MAS).

Integration of new fatty acid modifier alleles into current elite cultivars and breeding lines to meet consumer and processor needs will be easier with the knowledge of existing fatty acid QTL and their environmental sensitivity. The objective of this study was to map soybean seed fatty acid composition modifier QTL and determine which modifier QTL have significant GxE interaction. We have created an F₆ derived RIL population from the cross of 'Essex' (Smith and Camper, 1973) by 'Williams' (Bernard and Lindahl, 1972). This cross has contributed to the genetic background of many northern and southern USA cultivars and elite breeding lines (Sneller, 1994; Gizlice et al., 1996). This population does not segregate for any known major fatty acid genes, hence only modifier QTL will be detectable for soybean fatty acid composition.

Materials and Methods

One hundred and fifty-six F₆ RIL were created at The University of Tennessee, Knoxville Experiment Station (KES), Knoxville, TN from the cross between the soybean cultivars Essex and Williams. Sufficient seed was available to plant 131 RIL as F_{6:8} lines during the year 2000 in a randomized complete block design with three replications in each of two environments [KES-C field and the West Tennessee Experiment Station (WTES) at Jackson, TN] with one 6.1 m row per plot at KES-C and two 6.1 m rows per plot at WTES. In 2001, the 131 RIL as F_{6:9} lines were divided into three maturity groups (early, medium, and late) with 49 early, 37 medium, and 47 late with two lines overlapping in the medium and late maturity groups. Within each maturity group, a randomized complete block design with three replications each at KES-V and Harrisburg, IL and two replications at WTES were planted in four 6.1 m row plots. At each

environment, the different maturity group replications were randomly assigned to a location within the field.

DNA was isolated from all 156 RIL by taking a single leaf from five plants and combining them into one sample. Then Qiagen Plant Easy DNA Extraction Kit (Qiagen, Hilden, Germany) was used to obtain purified DNA.

Essex and Williams were tested with 568 SSR markers previously placed on the integrated soybean genetic linkage map (Cregan et al., 1999). The sequence information for the soybean SSR primers used is publicly available from the USDA Internet site <http://129.186.26.94/ssr.html>; verified February 28, 2002 and relates to the integrated molecular linkage map of soybean (Cregan et al., 1999). Polymorphic markers were chosen to achieve a genome wide coverage of all 20 molecular linkage groups previously identified (Cregan et al., 1999).

PCR reactions were performed in a 384-well MBS Hybaid thermocycler (Hybaid Franklin, MA). The 10 µl PCR mix consisted of 40 ng template DNA, 1x PCR Buffer, 0.2 mM dNTP mixture (Pharmacia, Piscataway, NJ), 1.0 µM forward and reverse primer, and 0.5 units of Klentaq (Ab Peptides, Inc, St. Louis, MO). The amplification conditions consisted of 94°C for 5 min followed by 35 cycles of 94°C denaturation for 25 s, 48°C annealing for 30 s, and 72°C extension for 30 s, ending with 1 cycle of 72°C for 5 min. PCR products were separated using 6% non-denaturing polyacrylamide gel electrophoresis with ethidium bromide staining for visualization. The RIL were scored (1 = Essex allele, 2 = heterozygote, and 3 = Williams allele) based on DNA band

visualization. Because we utilized an F₆-derived RIL population, heterozygote scores were dropped from the data set to allow only additive effects to be detected.

Seed fatty acid composition was analyzed using gas chromatography of the methyl esters. Preparation of seeds consisted of three seeds crushed and soaked overnight in 1.5 ml Chloroform:Hexane:Methanol (8:5:2 v/v/v). One hundred microliters of the extract was transferred to a 1.5 ml autosampler vial and 75 µl of methylating reagent [sodium methoxyde/methanol:petroleum ether:ethyl ether (1:4:2 v/v/v) was added. The samples were then run on a Hewlett-Packard model HP 6890 gas chromatograph (Palo Alto, CA) equipped with a model 7673 auto-sampler, 7673 flame ionization detector, and an immobilized 30 x 0.53 mm inner diameter, AllTech AT-Silar capillary column with 0.5µm fused stationary phase. Operating conditions were: carrier, He (20ml/min); 20:1 (v/v) split injection; injection temperature, 250°C; detector temperature, 275°C, and column temperature, 240°C.

One hundred and fifty-six lines were mapped with 100 SSR markers. MAPMAKER/EXP (Lander et al., 1987; Lincoln et al., 1992) was used to estimate genetic distances between SSR markers. A minimum likelihood of odds (LOD) ≥ 3.0 and a maximum distance ≤ 50 centimorgan (cM) were used to test linkages among markers.

QTL analysis was performed using the phenotypic traits and SSR RIL scores on 131 RIL planted in the field environments. A one-way ANOVA (SAS version 8.0) was initially used to identify candidate QTL (only results from Satt070 reported). Composite interval mapping and MTA were performed using QTL Cartographer (Basten et al., 1994; Basten et al., 2001). Composite interval mapping was used to identify and map significant QTL. Multiple trait analysis was used to confirm and to refine the position of

QTL found in CIM using the joint LOD score or the LOD from the GxE interaction test option of MTA. One thousand permutation tests were performed on each trait and respective environment to establish empirical LOD thresholds at the 5% alpha level (Churchill and Doerge, 1994). Heritability estimates were computed as

$$h^2 = \sigma_g^2 / [\sigma_g^2 + (\sigma_{ge}^2/e) + (\sigma^2/re)]$$

where h^2 = heritability; σ_g^2 = genotypic variance; σ_{ge}^2 = genotype x environment variance; σ^2 = error variance; r = number of replications, and e = number of environments (Nyquist 1991).

Results and Discussion

A total of 278 SSR markers were found to be polymorphic between these two parents. With the limited number of 100 markers mapped in this population, the integrated soybean genetic linkage map was used to help align the 23 genetic linkage groups formed from 89 linked SSR markers and two phenotypic markers (flower color and pubescence) with the 20 known consensus linkage groups (Cregan et al., 1999). The Essex x Williams genetic linkage map covered about 869.4 cM, which is about 32% of the currently known recombination distance of 2748 cM, averaged from the integrated soybean genetic linkage map (Cregan et al., 1999). Continued placement of markers in the future should help to assure more coverage of the genome so that additional QTL can be detected.

As expected, the population means fall within the normal range for fatty acid composition that is typical of most currently grown soybean cultivars (Table 4.1). Nevertheless, the diversity of the two parents did lead to a range within the RIL population, which enabled modifier fatty acid QTL to be detectable. Palmitate, stearate,

Table 4.1. Mean, standard deviation, range, heritability, and parental means for soybean seed fatty acid concentration in an F₆-derived RIL population of Essex x Williams soybean grown in five environments

Fatty Acid	Mean	Standard Deviation	Range	Essex	Williams	h ²
Palmitate (g kg ⁻¹)	110	3.6	102-119	114	106	0.56
Stearate (g kg ⁻¹)	41	3.3	33-52	40	39	0.74
Oleate (g kg ⁻¹)	237	29.9	182-340	215	253	0.71
Linoleate (g kg ⁻¹)	536	24.7	453-575	549	534	0.70
Linolenate (g kg ⁻¹)	77	6.6	58-90	82	69	0.43

linoleate, and linolenate all have small ranges for the RIL population allowing only modifier QTL with small additive effects to be detected and mapped. The range for oleate was larger, which could be due to a few major modifier genes or to several minor modifier genes of the fatty acid pathway, segregating in this population.

Four palmitate QTL were detected on linkage group (LG) D2, G, K, and L (Table 4.2). Multiple trait analysis revealed no significant GxE interaction, but did confirm the significance of each QTL. Palmitate was the only fatty acid that did not show a significant GxE interaction. Other studies have found palmitate content is not as affected by environment as the other fatty acids (Primomo et al., 2002).

A major and a minor QTL for reduced palmitate content was recently mapped from the low palmitic line 'N87-2122-4' (Li et al., 2002). The major and minor modifying QTL mapped in N87-2122-4 are different than the modifier QTL found in our population. The modifier genes found in our study are already present in currently grown cultivars and the desired modifier allele (for either reduced or increased palmitate) could be retained when trying to introgress new QTL via MAS in order to achieve further genetic gains.

Linkage group C2 and L contained common QTL for stearate, oleate, linoleate, and linolenate that had significant GxE interaction (Table 4.3-4.6). The additive effect from the GxE interaction by MTA was usually lower than the additive effect predicted by CIM. This reduction due to GxE may be the reason that QTL selected for by MAS do not exhibit the additive effect that was predicted by other QTL mapping procedures. Despite the precision reductions due to GxE interaction, modifier QTL may still be

Table 4.2. Position, effects, and significance of environments for QTL associated with soybean seed palmitate concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment ^d	R ² (%)	LOD	Effect ^e
FP1	D2	Satt458-Satt154	25.2	4.01	Combined	13	3.4	1.3
				8.01	KES-C/00	13	2.8	1.6
				0.01	Harrisburg/01	10	3.7	1.3
				0.01	WTES/01	7	2.8	1.2
				4.01	MTA	-	5.9	1.4
FP2	G	Satt340-Satt352	7.2	0.05	Harrisburg/01	6	2.6	-1.1
				0.05	MTA	-	3.9	-0.8
FP3	K	Satt273-Satt260	22.4	20.07	Combined	11	3.4	1.2
				12.07	KES-C/00	13	3.2	1.6
				22.07	KES-V/01	8	2.9	1.4
				20.07	Harrisburg/01	13	4.5	1.4
				18.07	MTA	-	6.8	1.3
FP4	L	Satt166-Satt229	29.5	8.05	Combined	21	4.9	1.7
				12.05	KES-V/01	18	3.6	2.0
				4.05	Harrisburg/01	17	5.3	1.7
				0.05	WTES/01	10	4.0	1.6
				4.05	MTA	-	7.0	1.8

^a FP1 to FP4 denote the four palmitate QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d KES-C1/00 Knoxville Experiment Station (C1-field) at Knoxville, TN in 2000

WTES/00 West Tennessee Experiment Station at Jackson, TN in 2000

KES-V/01 Knoxville Experiment Station (V-field) at Knoxville, TN in 2001

Harrisburg/01 Harrisburg, IL in 2001

WTES/01 West Tennessee Experiment Station at Jackson, TN in 2001

Combined data from five locations averaged

GxE Genotype vs. environmental interaction

MTA Multiple trait analysis

^e Average change (g kg⁻¹) resulting from Essex allele substitution

Table 4.3. Position, effects, and significance of environments for QTL associated with soybean seed stearate concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
FS1	C2	Satt277-T	3.2	2.01	KES-C/00	25	7.7	-2.4
				2.01	GxE	-	11.2	-0.1
FS2	C2	T-Satt365	3.2	3.18	KES-C/00	28	8.9	-2.6
				2.01	GxE	-	10.9	-0.1
FS3	C2	Satt460-Satt307	14.2	10.07	Combined	14	4.9	-1.3
				10.07	WTES/00	20	7.0	-1.7
				4.07	KES-C/00	23	5.8	-2.3
				6.07	KES-V/01	12	3.1	-1.2
				14.07	WTES/01	9	3.4	-1.1
				6.07	GxE	-	10.5	-1.0
FS4	K	Satt518-Satt273	17.8	14.05	WTES/01	9	3.0	-1.0
				10.05	GxE	-	3.1	-0.4
FS5	K	Satt273-Satt260	22.4	2.07	WTES/01	8	2.9	-1.0
				6.07	MTA	-	4.1	-0.7
FS6	L	Satt156-Satt166	16.6	3.97	Combined	13	4.6	-1.2
				5.97	WTES/00	12	4.7	-1.3
				15.97	GxE	-	6.1	-0.5
FS7	L	Satt166-Satt229	29.5	0.05	Harrisburg/01	11	4.2	-14
				8.05	GxE	-	6.3	0.1

^a FS1 to FS7 denote the seven stearate QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change resulting from Essex allele substitution for the QTL except for GxE

Table 4.4. Position, effects, and significance of environments for QTL associated with soybean seed oleate concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
FO1	C2	T-Satt365	3.2	0.01	Combined	5	4.0	-7.5
				3.18	Harrisburg/01	17	9.9	-17.4
				3.18	GxE	-	14.5	-7.6
FO2	F	Satt144-Satt522	21.9	0.01	Combined	6	4.2	7.4
				0.01	Harrisburg/01	12	7.8	13.7
				4.01	GxE	-	6.3	6.7
FO3	L	Satt523-Satt182	18	5.99	KES-V/01	8	3.1	-12.4
				9.99	GxE	-	3.9	-3.9
FO4	L	Satt166-Satt229	29.5	0.05	Combined	35	19.9	-19.0
				0.05	WTES/00	26	11.1	-16.2
				0.05	KES-C/00	15	7.0	-12.4
				0.05	WTES/01	35	14.6	-26.9
				2.05	Harrisburg/01	19	8.5	-17.5
				2.05	WTES/01	26	6.6	-21.7
				0.05	GxE	-	8.7	-11.6

^a FO1 to FO4 denote the four oleate QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (g kg⁻¹) resulting from Essex allele substitution for the QTL

Table 4.5. Position, effects, and significance of environments for QTL associated with soybean seed linoleate concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
FL1	C2	Satt277-T	3.2	2.01	KES-C/00	14	6.4	11.3
				2.01	GxE	-	10.6	4.8
FL2	C2	T-Satt365	3.2	0.01	Combined	5	3.8	5.9
				3.18	KES-C/00	14	6.7	12.0
				3.18	Harrisburg/01	10	6.5	10.4
				3.18	GxE	-	12.5	6.3
FL3	F	Satt144-Satt522	21.9	0.01	Combined	7	5.6	-7.0
				0.01	KES-C/00	6	3.1	-7.3
				0.01	Harrisburg/01	13	8.8	-11.5
				0.01	GxE	-	5.0	-5.2
FL4	L	Satt523-Satt182	18.0	11.99	KES-V/01	8	3.3	9.8
				9.99	GxE	-	3.4	2.6
FL5	L	Satt156-Satt166	16.6	13.97	WTES/00	20	10.2	12.7
				13.97	KES-C/00	9	3.7	8.6
				11.97	Harrisburg/01	21	9.0	14.5
				11.97	WTES/01	20	4.7	16.9
				15.97	GxE	-	7.6	9.1

^a FL1 to FL5 denote the five linoleate QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (g kg⁻¹) resulting from Essex allele substitution for the QTL

Table 4.6. Position, effects, and significance of environments for QTL associated with soybean seed linolenate concentration in an Essex x Williams F₆ RIL population, based on QTL Cartographer analysis

QTL ^a	LG ^b	Interval	Interval length (cM)	QTL position ^c (cM)	Environment	R ² (%)	LOD	Effect ^d
FN1	A1	Satt050-Satt545	35.0	0.0	KES-C/00	9	5.7	-1.9
				1.99	GxE	-	3.7	-1.2
FN2	C2	T-Satt365	3.2	0.01	Combined	4	2.9	1.4
				2.01	Harrisburg/01	20	8.4	5.5
				2.01	GxE	-	10.1	0.4
FN3	F	Satt146-Satt269	9.1	0.01	Combined	7	5.7	-1.9
				0.01	KES-C/00	4	2.6	-1.2
				4.01	KES-V/01	9	3.1	-3.3
				0.01	MTA	-	6.1	-1.1
FN4	I	Satt292-Satt148	21.5	0.01	KES-C/00	4	3.0	1.2
				6.01	GxE	-	3.1	0.7
FN5	K	Satt273-Satt260	22.4	20.07	KES-C/00	9	5.2	1.7
				22.07	MTA	-	4.7	1.0
FN6	L	Satt156-Satt166	16.6	13.97	Combined	18	10.4	3.0
				15.97	WTES/00	23	10.5	3.0
				15.97	KES-C/00	19	10.5	2.5
				11.97	KES-V/01	14	4.5	4.2
				3.97	WTES/01	12	3.3	3.7
				15.97	MTA	-	10.7	2.2
FN7	L	Satt166-Satt229	29.5	2.05	WTES/00	13	5.7	2.5
				4.05	Harrisburg/01	11	3.9	3.8
				10.05	GxE	-	3.3	1.8
FN8	N	Satt584-Satt387	27.5	12.01	MTA	-	2.6	-0.5
		Satt387-Satt549	18.9	1.98	Combined	4	2.9	-1.5

^a FN1 to FN8 denote the eight linolenate QTL detected

^b Linkage group QTL is located based on the integrated soybean genetic linkage map

^c The distance from the left marker of the interval

^d Average change (g kg⁻¹) resulting from Essex allele substitution for the QTL

selectable that have a predictable effect in the environment in which the QTL were mapped.

Figure 4.1 is an example of how MTA can be used to determine if peaks from CIM could be ghost peaks that have no real effect on a trait and only exist due to close linkage to a major gene. The first two peaks (FS1 and FS2) were confirmed by MTA to have a GxE interaction with a small additive effect (Table 4.3) while the third peak was not significant in MTA. Using MTA the only QTL in this region that had any significant effect on stearate was FS3.

Although FS3 had significant effects, Rebetzke et al. (2001) observed that maturity could affect fatty acid levels. In our study, the population varied from maturity group (MG) III to MG V, and maturity genes were mapped on LG C2 and L in the same area of the fatty acid QTL (Table 2.2). The stearate and oleate QTL on C2 and L all had negative additive effects resulting from an Essex allele substitution. The linoleate and linolenate QTL had a positive additive effect resulting from an Essex allele substitution. An increase in temperature will reduce linoleate and linolenate and increase oleate and stearate acid (Howell and Collins, 1957; Wolf et al., 1982; Dornbos and Mullen, 1992; Primomo et al., 2002). The fatty acid QTL mapped at LG C2 and L could be due to a pleiotropic effect with the maturity QTL on the same linkage group, if during the period of seed fill, the early maturing lines were exposed to higher temperatures than the later maturing lines. Nevertheless, our study is important in describing modifier QTL that may be verified by other researchers, in populations where maturity effects are not present.

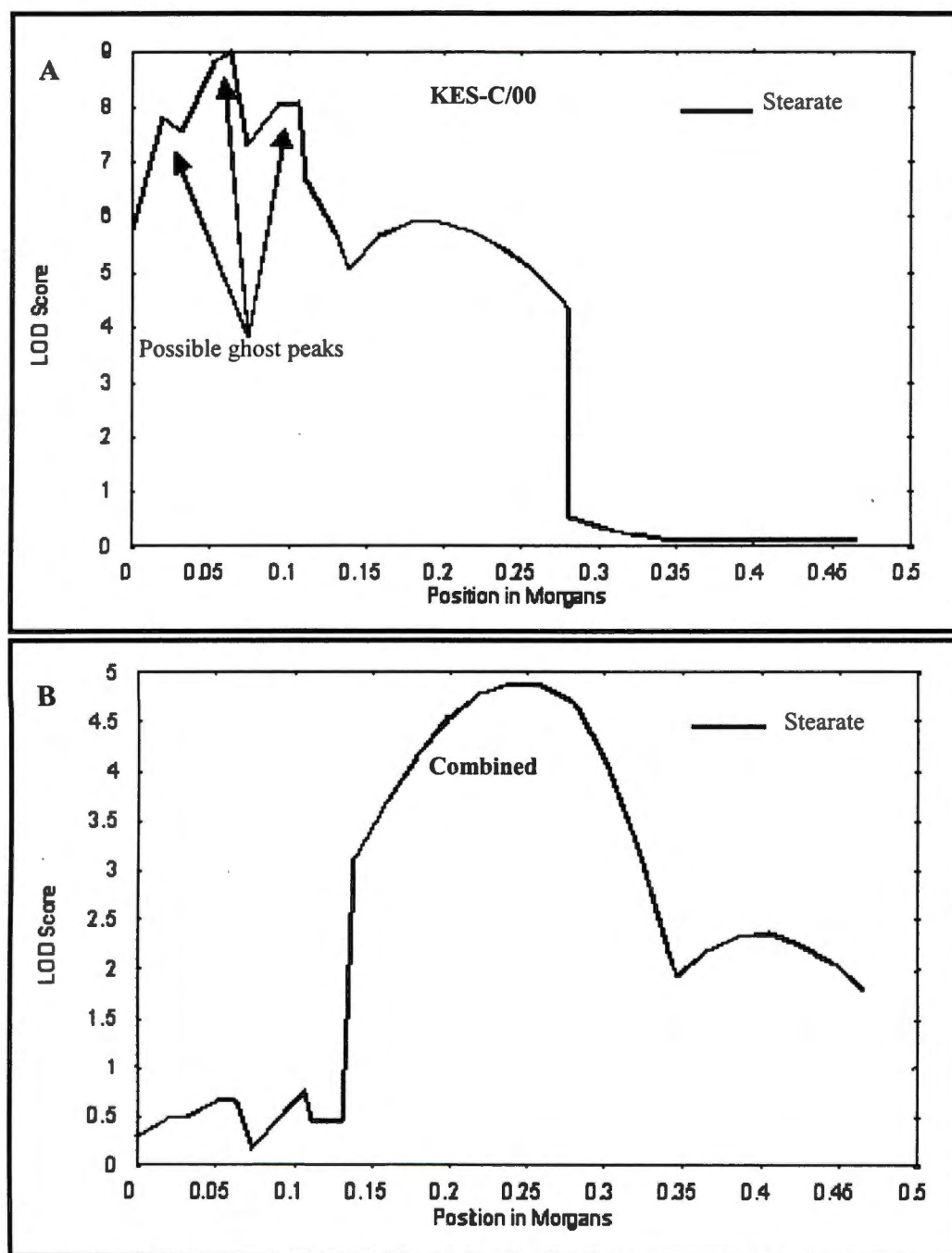


Figure 4.1. Stearate LOD scores from A) one environment (KES-C/00) B) combined over five environments, graphed along soybean linkage group C2

SSR marker Satt070 on LG B2 was unlinked to any of the other markers mapped in our population and was significantly associated with stearate ($p < .0001$, $R^2 = 47\%$). This modifier QTL apparently influences a major gene (*Fas*) in the same region, which governs stearic acid concentration. Spencer (personal communication) recently detected Satt070 at 1.9 cM from the *Fas* locus having an R^2 of 67% in a population, which contained a natural mutation governing a two-fold increase in stearic acid (Pantalone et al., 2002; Wilson et al., 2002). In our study the Essex allele of Satt070 had an additive effect of 4.5 g kg^{-1} . Stearate also has two QTL on LG K that were significant in one location and in MTA. Stearate QTL FS4 had a significant GxE effect while FS5 did not have significant GxE but was confirmed in MTA. The second QTL may have some benefit in lowering stearate but it needs to be confirmed in another population.

Oleate, linoleate, and linolenate QTL were located on LG F (Table 4.4-4.6). The oleate QTL FO2 and linoleate QTL FL3 mapped in approximately the same position using MTA. This is likely a QTL with a pleiotropic effect on both fatty acids that also has a significant GxE interaction, in which the Essex allele will increase oleate and decrease linoleate. Mapmaker analysis placed FN3, also located on LG F (Table 4.6) in a separate linkage group from FO2 and FL3 eliminating the possibility of pleiotropy.

All modifier QTL for oleate and linoleate had significant GxE interaction (Table 4.4 and 4.5), while linolenate showed four QTL that were not significant for GxE interaction (Table 4.6). Linolenate QTL FN3, FN5, FN6, and FN8 were QTL with significant additive effects across multiple environments. Diers et al. (1992) reported a linolenate QTL linked to A023_1, which could be the same QTL as FN6. These QTL

could be used along with other modifier genes in populations to reduce the amount of linolenate in new commercially grown cultivars.

Conclusion

Several modifier QTL for fatty acid composition of soybean seed were mapped in the RIL population created from an Essex by Williams cross. These QTL should be present in many of the cultivars grown in northern and southern USA. Multiple trait analysis allowed us to not only confirm QTL found by composite interval mapping, but also allowed us to label QTL with significant genotype by environment interaction. This knowledge will aid in the use of MAS in modifying soybean fatty acid composition, by predicting the genetic effect a selected QTL will have to achieve further genetic gains.

References

References

- Basten, C.J., B.S. Weir, and Z.-B. Zeng, 1994. Zmap-a QTL cartographer. In *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production: Computing Strategies and Software*, edited by C. Smith, J.S. Gavora, B. Benkel, J. Chesnais, W. Fairfull, J.P. Gibson, B.W. Kennedy and E.B. Burnside. 22:65-66. Published by the Organizing Committee, 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada.
- Basten, C.J., B.S. Weir, and Z.-B. Zeng. 2001. QTL Cartographer. Ver. 1.15. Dep. Statistics. North Carolina State University. Raleigh, NC.
- Beare-Rogers, J. 1995. Food fats and fatty acids in human nutrition. P. 1-7. In R. Przybylski and B.E. McDonald (ed.) *Development and processing vegetable oils for human nutrition*. AOCS, Champaign, IL.
- Bernard, R. L. 1971. Two genes for time of flowering and maturity in soybeans. *Crop Sci.* 11:242-244.
- Bernard, R.L. 1972a. Two genes affecting stem termination in soybeans. *Crop Sci.* 12:235-239.
- Bernard, R.L., D.A. Lindahl. 1972b. Registration of 'Williams' Soybean. *Crop Sci.* 12:716.
- Brown-Guedira, G.L., J.A. Thompson, R.L. Nelson, and M.L. Warburton. 2000. Evaluation of Genetic Diversity of Soybean Introductions and North American Ancestors Using RAPD and SSR Markers. *Crop Sci.* 40:815-823.
- Burton, J. W. 1987. Quantitative genetics: results relevant to soybean breeding. p. 211-247. In J. R. Wilcox (ed.) *Soybeans: Improvement, Production, and Uses*, 2nd ed. Agron. Monogr. 16. ASA, CSSA, SSSA, Madison, WI.
- Churchill, G., and R. Doerge, 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971.
- Cober, E.R., and H.D. Vodeng. 2001. A New Soybean Maturity and Photoperiod-Sensitivity Locus Linked to *E1* and *T*. *Crop Sci.* 41:698-701.
- Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. Van Toai, D.G. Lohnaes, T. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39:1464-1490.
- Csanadi, G., J. Vollmann, G. Stift, and T. Lelley. 2001. Seed quality QTLs identified in a molecular map of early maturing soybean. *Theor. Appl. Genet.* 103:912-919.

- Diers, B.W., P. Keim, W.R. Fehr, and R.C. Shoemaker. 1992. RFLP analysis of soybean seed protein and oil content. *Theor. Appl. Genet.* 83:608-612.
- Diers, B.W., and R. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. *J. Am. Oil Chem. Soc.* 69:1242-1244.
- Dornbos, D.L., and R.E. Mullen. 1992. Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. *J. Am. Oil Chem. Soc.* 69:228-231.
- Erickson, E.A., J.R. Wilcox, and J.F. Cavins. 1988. Inheritance of altered palmitic acid percentages in two soybean mutants. *J. Hered.* 79:465-468.
- Fehr, W.R. 1987. Quantitative genetics: results relevant to soybean breeding. p. 249-293. *In* J. R. Wilcox (ed.) *Soybeans: Improvement, Production, and Uses*, 2nd ed. Agron. Monogr. 16. ASA, CSSA, SSSA, Madison, WI.
- Fehr, W.R., G.A. Welke, E.G. Hammond, D.N. Duvick, and S.R. Cianzio. 1991. Inheritance of reduced palmitic acid content in seed oil of soybean. *Crop Sci.* 31:88-89.
- Frisch, M., M. Bohn, A.E. Melchinger. 1999. Minimum sample size and optimal positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. *Crop Sci.* 39:967-975.
- Gizlice, Z., T.E. Carter, Jr., T.M. Gerig, and J.W. Burton. 1996. Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. *Crop Sci.* 36:753-765.
- Graef, G.L., W.R. Fehr, L.A. Miller, E.G. Hammond, and S.R. Cianzio. 1988. Inheritance of fatty acid composition in a soybean mutant with low linolenic acid. *Crop Sci.* 28:55-58.
- Hospital, F., C. Chevalet, and P. Mulsant. 1992. Using markers in gene introgression breeding programs. *Genetics* 132:1119-1210.
- Howell, R.W., and R.F. Collins. 1957. Factors affecting linolenic and linoleic acid content of soybean oil. *Agron. J.* 49:593-597.
- Jiang, C. and Z-B. Zeng. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140:1111-1127.

- Keim P., B.W. Diers, T.C. Olson, and R.C. Shoemaker. 1990. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735-742.
- Lander, E.S. and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using rflp linkage maps. *Genetics* 121:185-199.
- Lander E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln, and L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage map of experimental and natural populations. *Genomics* 1:174-181.
- Lark, K.G., J.M. Weisemann, B.F. Matthews, R. Palmer, K. Chase, and T. Macalma. 1993. A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir'. *Theor. Appl. Genet.* 86:901-906.
- Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter, Jr., D.A. Ashley, R.S. Hussey, W.A. Parrott, and H.R. Boerma. 1996a. Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Sci.* 36:728-735.
- Lee, S.H., M.A. Bailey, M.A.R. Mian, T.E. Carter, E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996. RFLP loci associated with soybean seed protein and oil content across populations and locations. *Theor. Appl. Genet.* 93:649-657.
- Lee, S.H., M.A. Bailey, M.A.R. Mian, E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, and H.R. Boerma. 1996b. Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. *Theor. Appl. Genet.* 92:516-523.
- Li, Z., L. Jakkula, R.S. Hussey, J.P. Tamulonis, H.R. Boerma. 2001. SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. *Theor. Appl. Genet.* 103:1167-1173.
- Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma. 2002. Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. *Crop Sci.* 42:373-378.
- Lincoln, S.E., M.J. Daly, and E. Lander. 1992. Constructing genetic maps with MAPMAKER/EXP 3.0. 3rd edn, Whitehead Institute Technical Report.
- Liu, KeShun. 1997. *Soybeans: Chemistry, Technology and Utilization*. Pp. 279 International Thomson Publishing, New York.

- Lundeen, P.O., W.R. Fehr, E.G. Hammond, and S.R. Cianzio. 1987. Association of alleles for high stearic acid with agronomic characters of soybean. *Crop Sci.* 27:1102-1105.
- Mansur, L.M., K.G. Lark, H. Kross, and A. Oliveira. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean [*Glycine max* (L.) Merr.]. *Theor. Appl. Genet.* 86:907-913.
- 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.
- Matthews, B.F., T.E. Devine, J.M. Weisemann, H.S. Beard, K.S. Lewers, M.H. MacDonald, Y-B. Park, R. Maiti, J-J. Lin, J. Kuo, M.J. Pedroni, P.B. Cregan, and J.A. Saunders. 2001. Incorporation of Sequenced cDNA and Genomic Markers into the Soybean Genetic Map. *Crop Sci.* 41:516-521.
- Maughan, P.J., M.A. Saghai-Marooof, and G.R. Buss. 1996. Molecular-marker analysis of seed-weight: genomic locations, gene action, and evidence for orthologous evolution among three legume species. *Theor. Appl. Genet.* 93:574-579.
- Mian, M.A.R., D.A. Ashley, W.K. Vencill, and H.R. Boerma. 1998. QTLs conditioning early growth in a soybean population segregating for growth habit. *Theor. appl. Genet.* 97:1210-1216.
- Mian, A.R., M.A. Bailey, J.P. Tamulonis, E.R. Shipe, T.E. Carter Jr., W.A. Parrott, D.A. Ashley, R.S. Hussey, H.R. Boerma. 1996. Molecular markers associated with seed weight in two soybean populations. *Theor. Appl. Genet.* 93:1011-1016.
- Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10(3): 235-322.
- 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.
- Pantalone, V.R., R.F. Wilson, W.P. Novitzky, and J.W. Burton. 2002. Genetic regulation of elevated stearic acid concentration in soybean oil. *J. Am. Oil Chem. Soc.* (in press)
- Primomo, V.S., D.E. Falk, G.R. Ablett, J.W. Tanner, and I. Rajcan. 2002. Genotype x environment interactions, stability, and agronomic performance of soybean with altered fatty acid profiles. *Crop Sci.* 42:37-44.

- Qui, B.X., P.R. Arelli, D.A. Sleper. 1999. RFLP markers associated with soybean cyst nematode resistance and seed composition in a 'Peking' x 'Essex' population. *Theor. Appl. Genet.* 98:356-364.
- Reyna, N., and C.H. Sneller. 2001. Evaluation of Marker-Assisted Introgression of Yield QTL Alleles into Adapted Soybean. *Crop Sci.* 41:1317-1321.
- Rebetzke, G.J., J.W. Burton, T.E. Carter, Jr., and R.F. Wilson. 1998. Genetic variation for modifiers controlling reduced saturated fatty acid content in soybean. *Crop Sci.* 38:303-308.
- Rebetzke, G.J., V.R. Pantalone, J.W. Burton, T.E. Carter, Jr., and R.F. Wilson. 2001. Genetic background and environment influence palmitate content of soybean seed oil. *Crop Sci.* 41:1731-1736.
- Rennie, B.D., J. Zilka, M.M. Cramer, and W.D. Beversdorf. 1988. Genetic analysis of low linolenic acid levels in the soybean line PI 361088B. *Crop Sci.* 28:655-657.
- Reyna, N., and C.H. Sneller. 2001. Evaluation of Marker-Assisted Introgression of Yield QTL Alleles into Adapted Soybean. *Crop Sci.* 41:1317-1321.
- SAS Institute. 1989. SAS/STAT User's Guide, Version 6, 4th edn, Vol 1. SAS Institute, Inc., Cary, North Carolina.
- Sebolt, A.M., R.C. Shoemaker, and B.W. Diers. 2000. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. *Crop Sci.* 40:1438-1444.
- Smith, T.J., and H.M. Camper. 1973. Registration of 'Essex' soybean. *Crop Sci.* 13:495.
- Sneller, C.H. 1994. Pedigree analysis of elite soybean lines. *Crop Sci.* 34:1515-1522.
- SoyBase. 1995. A genome database for *Glycine*. <http://genome.cornell.edu/cgi-bin/WebAce/webace?db=soybase>; verified March 1, 2002.
- Specht, J.E., K. Chase, M. Macrander, B.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.
- Spencer, M.M., Plant Science and Landscape Design Department, The University of Tennessee, Knoxville, TN 37996-4500, USA.
- Wilcox, J.R., and J.F. Cavins. 1985. Inheritance of low linolenic acid content of the seed oil of a mutant in *Glycine max*. *Theor. Appl. Genet.* 71:74-78.

- Willett, W.C. 1994. Diet and health: What should we eat? *Science* 264:532-537.
- Wilson, R.F., J.W. Burton, V.R. Pantalone, R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean. p. 95-113. *In* T.M. Kuo and H.W. Gardner (ed.) *Lipid Biotechnology*. First Edition: Marcel Dekker, Inc., New York.
- Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black. 1982. Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars. *J. Am. Oil Chem. Soc.* 59:230-232.
- Zeng, Z-B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl. Acad. Sci. USA* 90:10972-10976.
- Zeng, Z-B. 1994. Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468.

Vita

David Lee Hyten, Jr. was born in Belleville, IL on December 8, 1976. He was raised in McLeansboro, IL and went to grade school and junior high school at McLeansboro Junior High School. He graduated from McLeansboro High School in 1995. From there, he went to Southern Illinois University at Carbondale and received a B.A. in microbiology with a minor in chemistry in 1999. He also worked as a Researcher I managing the marker assisted selection lab at Southern Illinois University Soybean Breeding Program from 1998-2000.

David enrolled at the University of Tennessee and received an M.S. in Plant and Soil Science in 2002. David's future plans are to go to the University of Maryland and pursue a doctorate in Natural Resources Sciences.

3784 2850 46
88-87-02 NFB