2006

Soybean Cultivars Resulted from More Recombination Events Than Unselected Lines in the Same Population

T. R. Stefaniak
University of Kentucky, trstef1@pop.uky.edu

D. L. Hyten Jr.
University of Tennessee - Knoxville, david.hyten@unl.edu

V. R. Pantalone
University of Tennessee - Knoxville, vpantalo@utk.edu

A. Klarer
University of Kentucky

T. W. Pfeiffer
University of Kentucky

Published in *Crop Science* 46:1 (2006), pp. 43–51; doi:10.2135/cropsci2005.0016

Submitted January 7, 2005; published online December 2, 2005.

Soybean Cultivars Resulted from More Recombination Events Than Unselected Lines in the Same Population

T. R. Stefaniak,[1] D. L. Hyten,[3] V. R. Pantalone,[3] A. Klarer,[2]

and T. W. Pfeiffer[2]

1. T. R. Stefaniak, 322 Plant Sciences Bld., Dep. of Agronomy, University of Kentucky, Lexington, KY 40546-0312
2. A. Klarer and T. W. Pfeiffer, Department of Agronomy, University of Kentucky, Lexington, KY 40546
3. D. L. Hyten and V. R. Pantalone, Department of Plant Sciences and Landscape Systems, University of Tennessee, Knoxville, TN 37901

Corresponding author – T. R. Stefaniak, email trstef1@pop.uky.edu

Abstract

The selection of superior adapted cultivars has contributed to the doubling of soybean [*Glycine max* (L.) Merr.] yields in the USA since 1930. Genetic variation was required for this selection to be effective. The objective of this study was to evaluate the importance of homologous chromosome meiotic recombination in the creation of soybean cultivars. A set of 10 adapted high-yielding cultivars selected from the cross 'Williams' × 'Essex' was compared with a set of 156 random recombinant inbred lines (RILs) from the same population. Crossover events were identified using 143 simple sequence repeat (SSR) markers spanning all 20 soybean chromosomes. The recombination rates were standardized among chromosomes by dividing the realized crossovers by the potential crossovers. The standardized recombination rate for the entire genome was significantly greater for the 10 cultivars (0.34) than for the RILs (0.29). The cultivars had numerically higher standardized recombination rates for 17 of the 20 chromosomes, significantly higher on chromosomes defined by the molecular linkage groups C2, L, and M. The interaction of linkage groups

with the two sets of lines was nonsignificant for standardized recombination. Soybean breeding progress has been accomplished in part by creating and capitalizing on new within-chromosome allele combinations.

Abbreviations: MAS, marker assisted selection; QTL, quantitative trait loci; RIL, random inbred lines; SSR, simple sequence repeat.

The improvement of any species through breeding requires the creation and selection of a novel combination of alleles in progeny from the genetic variation contained within the parents. These new combinations could arise from either independent chromosome assortment or homologous chromosome recombination. Breeding progress can be made only through variation from independent assortment if the superior alleles are found on different chromosomes (unlinked).

Variation in recombination frequency exists in plant breeding populations (Pfeiffer and Vogt, 1990; Tulsieram et al., 1992; Fatmi et al., 1993). Similarly, inbred progeny receive a range of alleles from one or the other of the population's parents (Kiem et al., 1991). The inheritance of alleles in adapted soybean cultivars does not necessarily follow the ratios estimated by the coefficient of parentage for a particular cross (Kisha and Diers, 1997) because the coefficient of parentage is a probability that necessarily ignores the preferential selection by plant breeders of alleles that favorably affect a phenotype. Because the site of crossing-over is random, selection for polygenic traits will alter the number and position of crossovers found in the lines a breeder chooses to advance as opposed to those found in the entire population. Therefore, for quantitative traits, meiotic recombination has been, and will continue to be, a mechanism on which breeders must capitalize to establish novel superior linkage blocks in these regions.

Demarly (1979) introduced the term "linkat" for a collection of favorable alleles that are linked and tend to be inherited intact because of the competitive advantage they give the individuals that possess them. These linkats contribute to the preferential inheritance of sets of favorable accumulated alleles that cause the actual pattern of inheritance to deviate from that predicted by the coefficient of parentage (Kisha and Diers, 1997). Using restriction fragment length polymorphism (RFLP) markers, Lorenzen et al. (1996) identified two linkage groups in four soybean cultivars that have long chromosome sections inherited intact from the same parents despite the fact that these cultivars were developed in separate breeding programs. If a collection of genes fundamental to domestication were established in linkats long ago, it would be likely that this collection would be found intact in many modern cultivars despite a lack of a common pedigree. Similarly, new favorable linkage blocks may be created by selection following hybridization. For example, Lorenzen et al. (1996) identified five cultivars that were the result of a crossover in the same location with the same parental alleles selected on either side of the crossover.

Seed yield in soybean is a polygenic trait, which could be greatly affected by recombination. Whether these genes interact additively or epistatically is of great importance with respect to the efficacy of utilizing enhanced recombination in a breeding strategy. If yield genes act additively, then enhanced recombination that brings these alleles together will

be beneficial. If these genes interact epistatically, then recombination's value is uncertain until the coupling-repulsion status of the alleles is clarified (Hanson and Hayman, 1963).

Cregan et al. (1999) have combined data from mapping populations into a highly saturated linkage map of the soybean genome containing classical, SSR, and RFLP markers. This map may be used to screen parents for polymorphisms useful in evaluating crossing-over in the progeny of these parents. One can then track whether linkage blocks were inherited intact or broken in the progeny of a cross. The high degree of map saturation makes it likely that some markers are linked to yield quantitative trait loci (QTL).

The objectives of this study were to (i) evaluate the importance of homologous chromosome meiotic recombination in the creation of soybean cultivars by comparing standardized recombination rates between a random population and a set of adapted cultivars derived from the same cross and (ii) infer the relationship between genomic regions with high or low crossover rates and the location of previously identified QTL in the soybean genome.

Materials and Methods

Genetic Material

The genotypes investigated in this study fell into one of two selection sets. All genotypes were from the cross 'Williams' (Bernard and Lindahl, 1972) × 'Essex' (Smith and Camper, 1973). Ten released adapted cultivars comprised the high yield selection (cultivar) set. These genotypes, 'Pennyrile,' 'S4240,' 'RA452,' 'A4268,' 'RA481,' 'A3860,' 'A3127,' '9441,' '9471,' and 'Coker 393,' came from various breeding programs (Gabe, 1994) and are assumed to be the product of transgressive segregation. These cultivars were probably selected as F4 or F5 derived lines, depending on the breeding strategy utilized by the particular breeding program from which the cultivars were developed. Selection at this level of inbreeding was typical in U.S. soybean breeding programs during this time frame. Seeds of cultivars were obtained from the soybean germplasm collection or from the breeding program that developed the cultivar. An unselected set of 156 lines from this same cross was compared with the cultivar set. The random lines were created by single seed descent and were advanced to the $F_{6:8}$ generation (Hyten, 2002). These lines will be subsequently referred to as the random RIL set. The RIL population was previously used in a QTL study that involved agronomic trait testing in five environments (Hyten et al., 2004). Three other similarly sized populations of unselected RILs were utilized from the crosses 'Peking' × 'Hamilton' (Nickell et al., 1990), 'Pershing' (Anand and Shannon, 1985) × 'Hamilton,' and 'Peking' × 'Essex' to validate our analysis of the SSR markers on linkage group C2.

SSR Analysis

About 5 g of leaf material was collected from plants from each random line and each selected cultivar. This leaf material was stored frozen until desiccated in a lyophilizer. The DNA from the 10 cultivars was extracted following the CTAB procedure of Kiem et al. (1988). The DNA from the 156 RILs was extracted with the Qiagen DNAeasy mini prep kit (Qiagen, Valencia, CA). The DNA was quantified with a fluorometer and diluted to 10

ng/mL. A total of 10 mL of this DNA suspension was run on a 1% (w/v) agarose gel to check for quality and to verify concentration.

Williams and Essex were screened with 568 SSR markers to identify polymorphisms. The SSR primer sequences were obtained from the Soybase (1995) web site. From this screening, 277 markers were found to be polymorphic between Williams and Essex, and 143 of the 277 were used to obtain crossover data (fig. 1). These markers covered all 20 soybean linkage groups (Cregan et al., 1999), with a minimum of four and a mean of 7.15 markers per linkage group.

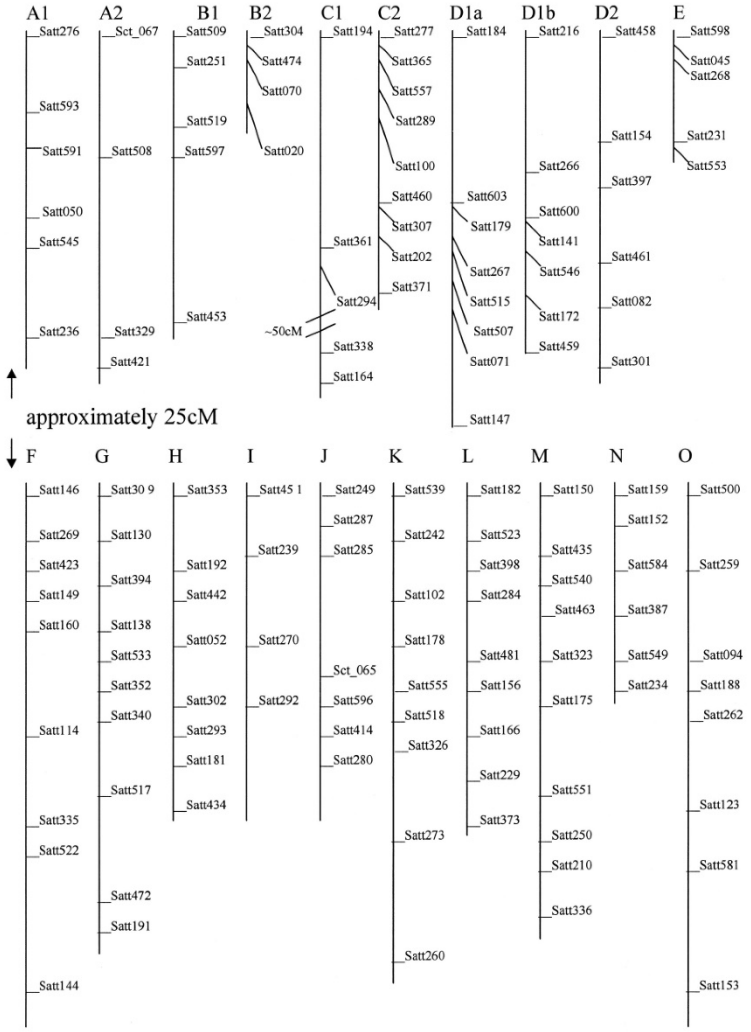

Figure 1. Schematic representation of the relative positions of the 143 SSR markers on soybean's 20 molecular linkage groups (Cregan et al. 1999) used to identify regions of crossing-over in the progeny of Williams × Essex.

The polymerase chain reaction (PCR) was conducted as defined on Soybase (1995). Amplified PCR fragments were separated by either metaphor agarose or polyacrylamide gel electrophoresis, depending on the size of the polymorphism between Williams and Essex. Polymorphisms greater than 10 base pairs were run for 4 h at constant 70 V on 3% (w/v) metaphor agarose gels stained with ethidium bromide. Polymorphisms that were less than 10 base pairs were run on 6% (w/v) nondenaturing polyacrylamide gels at 200 V for 4 h and stained with ethidium bromide.

Genetic Map

Marker order was initially defined on the basis of the public soybean composite map (Cregan et al., 1999). Because some of the markers used in this study are placed on the composite map with distances less than 5 cM between them, the RIL population was mapped by the Mapmaker program (Lander et al., 1987; Lincoln et al., 1992) to verify the map order. This was accomplished by first defining 20 linkage groups and then using the "assign" command to place markers onto soybean's 20 chromosomes. In this analysis, all markers were assigned to their defined linkage group except those found on linkage group A2, which Mapmaker designated as unlinked. Next, the order of the markers on the individual linkage groups was determined using the "compare" command. The four markers unassigned by the "assign" command were assumed to be on linkage group A2 and analyzed as such by Mapmaker and agreed with the published map for 14 marker order calculated linkage groups. The remaining six linkage groups (A1, A2, B2, C1, C2, and D1a) were ordered differently compared with the published map. On each linkage group, the order of two markers in close proximity (< 5.0 cM) of each other was inverted. For each of these linkage groups one of the two markers was dropped from the data set.

Recombination was detected by first listing the markers for each chromosome in the order found on the genetic map. Next, crossing over was counted by following each line's marker scores along the chromosome and noting where a line's score changes from one parent's allele to the other. All crossovers for each line and linkage group were then simply summed and standardized by putting this sum in the numerator of a fraction with the potential crossovers (number of markers per chromosome – 1) in the denominator. The analyzed variable was standardized crossovers. Crossovers were standardized to permit analysis among linkage groups in which different numbers of markers were available and to adjust individual lines in which the allele designation at a locus may have been unassigned.

Data Analysis

The expected 1:1 segregation ratio of the inheritance of parental alleles in the RILs was tested by the chi-square test. The Yates correction factor for a chi-square test with one degree of freedom was not used because the expected number in each class ($n = 78$) was relatively large (Bailey, 1961). Because of the large number of loci analyzed, deviations from this ratio at a single locus were considered significant at $p < 0.01$. Deviations from expected on an entire chromosome were considered significant at $p < 0.05$. Segregation ratios were not analyzed for the cultivars. The number of cultivars was small for a chi-square test, and

selection was expected to favor positive alleles negating the 1:1 segregation ratio expectation.

The data were analyzed for all 20 linkage groups together (sources of variation: selection sets, linkage groups, selection sets × linkage groups, all factors fixed) as well as all 20 linkage groups individually (source of variation: selection sets) by PROC GLM of SAS (SAS Institute, Cary, NC, release 8.1). Standardized crossovers were compared between the cultivars and all the 156 RILS. Ten sets of 10 RILs, drawn from the larger set of 156 RILs, were also created to compare standardized crossovers in equal sample sizes between the cultivars and the random lines using the analysis above. The locations of regions where comparatively few or many crossovers were detected in the RILs compared with the cultivars were related to QTL data from previous studies. These comparisons on individual linkage groups, in specific regions, did not include replication and hence were not analyzed statistically.

Results

Parental Allele Distribution

Chi-square significance was analyzed only for the full RIL set because of insufficient sample sizes of the other groups, and the following chi-square results pertain to the RIL set only. In this experiment, only 8% of the markers had significant chi-square values. Over all 143 SSR markers used in this experiment the chi-square value of 0.0004 was not significant and corroborates the expected 1:1 inheritance ratio of parental alleles in the RILs (table 1). Chi-square values were significant ($p < 0.05$) for the linkage groups B2, C2, and L. On linkage groups B2 and L, a greater number of Williams alleles accounted for the significant chi-square value, while the Essex alleles were in excess for C2. Of the 143 markers analyzed, 12 had significant ($p < 0.01$) chi-square values. Five of these were on linkage group C2 and all five exhibited a preponderance of Essex alleles. Of the remaining seven markers with significant chi-square values, five had a greater number of Williams alleles and each of the seven was on a different linkage group.

Table 1. χ^2 Values for 1:1 segregation of Williams and Essex alleles calculated for each SSR marker, individual linkage groups, and all 143 SSR markers in 156 $F_{6:8}$ random inbred lines.

MLG§ marker	χ^2	MLG§ marker	χ^2	MLG§ marker	χ^2	MLG§ marker	χ^2
A1	0.004	D1a	0.99	G	1.87	L	5.97*†
Satt276	6.14	Satt184	0.54	Satt309	5.77	Satt182	0.01
Satt593	1.47	Satt603	0.40	Satt130	0.08	Satt523	1.45
Satt591	1.8	Satt179	0.17	Satt394	0.00	Satt398	8.31**†
Satt050	0.42	Satt267	7.14**‡	Satt138	2.88	Satt284	0.18
Satt545	0.29	Satt515	2.31	Satt533	0.17	Satt481	2.19
Satt236	0.18	Satt507	0.44	Satt352	2.52	Satt156	1.57
A2	2.10	Satt071	1.09	Satt340	1.45	Satt166	0.01
Sct_067	2.35	Satt147	5.92	Satt517	0.16	Satt229	1.91
Satt508	2.00	D1b	2.61	Satt472	0.32	Satt373	0.67

Satt329	0.77	Satt216	0.01	Satt191	8.45**†	M	3.27
Satt421	0.78	Satt266	0.11	H	2.55	Satt150	2.61
B1	1.09	Satt600	1.49	Satt353	2.56	Satt435	11.22**†
Satt509	1.07	Satt141	1.12	Satt192	0.01	Satt540	0.06
Satt251	0.29	Satt546	0.13	Satt442	0.01	Satt463	0.01
Satt519	5.93	Satt172	2.52	Satt052	15.87**‡	Satt323	0.76
Satt597	0.00	Satt459	0.33	Satt302	0.00	Satt175	4.14
Satt453	0.14	D2	0.98	Satt293	4.68	Satt551	0.03
B2	7.03**†	Satt458	0.59	Satt181	0.52	Satt250	0.01
Satt304	5.25	Satt154	0.04	Satt434	0.15	Satt210	0.00
Satt474	1.46	Satt397	0.00	I	0.46	Satt336	2.70
Satt070	1.71	Satt461	0.01	Satt451	3.28	N	1.60
Satt020	0.71	Satt082	0.06	Satt239	0.48	Satt159	0.00
C1	0.57	Satt301	2.63	Satt270	0.95	Satt152	1.88
Satt194	3.23	E	0.01	Satt292	1.22	Satt584	0.03
Satt361	0.21	Satt598	0.01	J	2.72	Satt387	0.11
Satt294	0.01	Satt045	0.35	Satt249	1.74	Satt549	0.33
Satt338	0.06	Satt268	0.70	Satt287	0.03	Satt234	7.68**†
Satt164	0.32	Satt231	0.43	Satt285	0.02	O	0.16
C2	66.27**‡	Satt553	0.39	Sct_065	4.50	Satt500	0.08
Satt277	9.00**‡	F	1.68	Satt596	2.86	Satt259	2.79
Satt365	9.65**‡	Satt146	0.55	Satt414	31.27**†	Satt094	0.82
Satt557	6.38*‡	Satt269	1.00	Satt280	4.35	Satt188	0.17
Satt289	10.94**‡	Satt423	0.34	K	0.64	Satt262	1.58
Satt100	6.34*‡	Satt149	0.33	Satt539	2.14	Satt123	0.82
Satt460	12.97**‡	Satt160	0.54	Satt242	1.69	Satt581	0.45
Satt307	12.25**‡	Satt114	0.06	Satt102	0.44	Satt153	0.43
Satt202	5.32*‡	Satt335	0.65	Satt178	1.44		
Satt371	1.27	Satt522	6.13	Satt555	0.68	All markers	0.00
		Satt144	0.79	Satt518	1.51		
				Satt326	0.10		
				Satt273	2.13		
				Satt260	0.32		

* Significant at $\alpha = 0.05$. Deviations from expected were declared significant at $p < 0.01$ for an individual locus and at $p < 0.05$ for an entire linkage group. Individual loci on linkage group C2 which have significant deviation from 1:1 at $p < 0.05$ are indicated to highlight the unique segregation pattern on this chromosome.

** Significant at $\alpha = 0.01$. Deviations from expected were declared significant at $p < 0.01$ for an individual locus and at $p < 0.05$ for an entire linkage group. Individual loci on linkage group C2 which have significant deviation from 1:1 at $p < 0.05$ are indicated to highlight the unique segregation pattern on this chromosome.

† = Williams allele(s) present in excess.

‡ = Essex allele(s) present in excess.

§ MLG, molecular linkage group (Cregan et al., 1999).

Crossover Means for All 20 Linkage Groups

Standardized crossovers exhibited a wide range in the progeny of the Williams × Essex cross (fig. 2). Among lines within sets the ranges of mean standardized crossovers for all 20 linkage groups were 0.17 to 0.39 for the RILs and 0.29 to 0.43 for the cultivars. The selection set mean standardized crossover value averaged over all 20 linkage groups was significantly lower ($p < 0.05$) for the 156 RILs (0.29) than for the 10 cultivars (0.34).

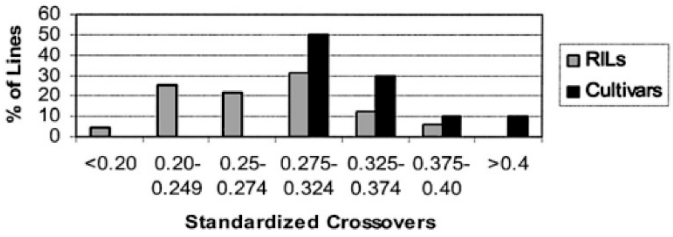

Figure 2. Frequency distribution of standardized crossovers (actual crossovers/potential crossovers where potential crossovers = number of marker loci per linkage group – 1) averaged across 20 linkage groups for 156 random inbred lines and 10 cultivars from the soybean population Williams × Essex.

Crossover Means for Individual Linkage Groups

For individual linkage groups, the cultivar set had significantly greater crossing-over than the RIL set on three of the 20 linkage groups, C2, L, and M (fig. 3) and numerically higher standardized crossover means in all but five of the 20 linkage groups A1, B1, C1, J, and O. There was no significant linkage group by selection set interaction in this analysis.

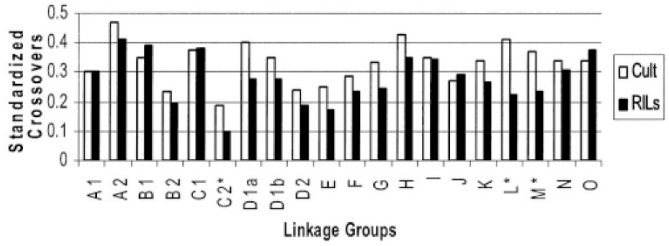

Figure 3. Standardized crossover (actual crossovers/potential crossovers) means for the cultivars and RILs (random lines) on the 20 soybean chromosomes [defined by the 20 molecular linkage groups (Cregan et al. 1999)] * Cultivars have significantly ($p < 0.05$) greater standardized crossovers than random lines.

Crossover Means for All Linkage Groups with Selection Sets of Equal Size

The disparity in sample sizes (10 cultivars vs. 156 RILs) between our selection sets might influence the results. To address this, we created 10 sets of 10 RILs drawn at random from the 156 lines. All 10 sets had a significantly lower 20 chromosome standardized crossover mean (range 0.25–0.29) than the cultivar set (0.34). On the basis of this result, all further discussion is with respect to the 156 RILs and 10 cultivar comparisons.

Discussion

Crossover Means for All 20 Linkage Groups

Where favorable interactions between genes are already established, selection for superior performance should favor the retention of favorable linkage groups through the reduction of recombination. Minimal recombination would tend to preserve linkage blocks, and selection for favorable linkage blocks would maintain epistatic interactions. The overall analysis of the recombination rates between the RILs and the cultivar sets on all 20 linkage groups suggests that a large number of favorable highly epistatic intrachomosomal allele interactions are not present in Williams or Essex. The higher level of recombination detected in the cultivars compared with the RILs suggests that an appreciable source of genetic variance exploited in the advancement of lines from this cross was additive.

Although creating enhanced recombination through intermating was proposed as a breeding strategy (Hanson, 1959a, 1959b), selection favoring maximal recombination has not been supported by previous soybean breeding experimentation. For example, in a well-documented study, short-term genetic gain was reduced in progeny lines when mating schemes were used that increased the opportunity for recombination (Piper and Fehr, 1987). That experiment, however, could not measure the resulting difference in recombination between the progeny of the different mating schemes. Therefore, it was unknown how much recombination was increased, much less in which regions this greater recombination occurred. The pool of progeny lines from which Piper and Fehr (1987) selected the individuals for further intermating was 300 individuals. Two thousand lines is a conservative estimate of the number of lines from which the 10 Williams × Essex progeny cultivars were selected. Perhaps greater sampling of the genetic variation (sampling at least 2000 lines vs. 300 lines) allowed identification of the true transgressive segregates in the various cultivar development programs. Alternatively, the Williams × Essex progeny population could be unique among soybean breeding populations in the manner in which transgressive segregates are created. We already know that this population is unique as a source of demonstrated transgressive segregates in that at least 12 cultivars (two were no longer available for this study) were selected from this cross (Gabe, 1994). The Williams × Essex cross is the only soybean cross we know of that produced transgressive segregates in a large enough number suitable for a comparison of this type.

Crossover Means for Individual Linkage Groups

A nine-marker 50-cM region of linkage group C2 was analyzed (fig. 4). Within a 15-cM region, Williams alleles at both markers Satt 460 and Satt 307 were seen at a frequency of 0.27 in the RILs compared with 0.60 in the cultivars (table 2). In relation to the RILs, a majority of the cultivars have an intact linkage group inherited from the Williams parent. This indicates the existence in Williams of a favorable linkage block in this region. This superiority may result from positive epistatic interactions within the linkage block. It is also possible that the genes composing this block interact additively in Williams and, therefore, are superior when inherited together. In this case it, is not clear that the preservation of a linkage block indicates that the gene interaction is epistatic or additive. However, the

lack of any crossovers in this region in the cultivars hints that an epistatic interaction may occur.

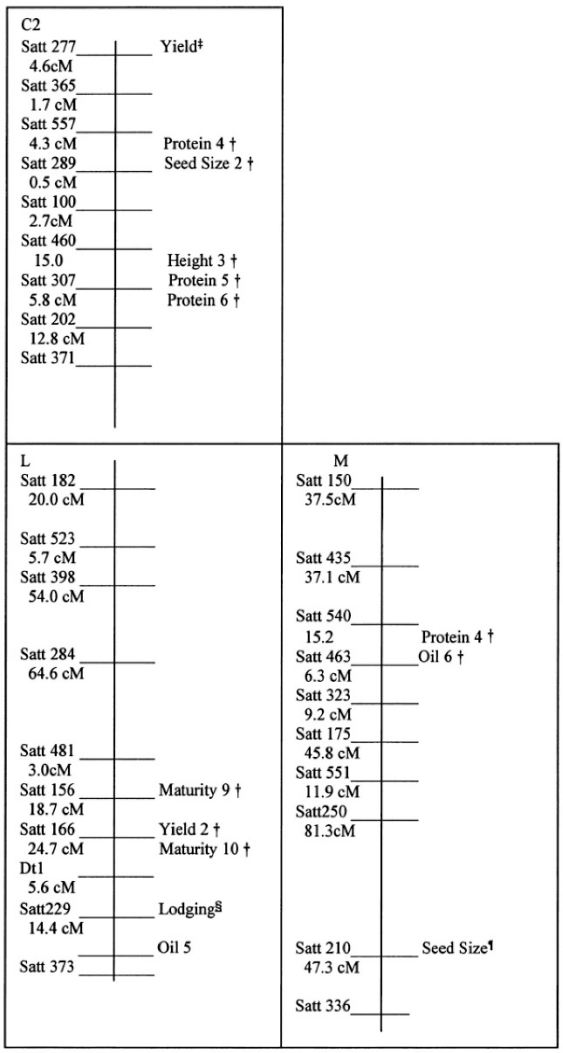

Figure 4. Schematic representation of the three linkage groups for which, cultivars had significantly greater recombination than the random lines. Molecular linkage groups and marker positions are as found on the published soybean genetic map (Cregan et al. 1999) and confirmed in this population. Distances between markers were calculated by Mapmaker (Lander et al. 1987; Lincoln et al. 1992) for the Williams × Essex population. QTL were previously reported. †QTL followed by numbers are so designated by Hyten 2002. ‡Yield QTL reported in Orf et al. (1999). §Lodging QTL reported in Lee et al. (1996b). ¶Seed size QTL reported in Csanadi et al. (2001).

Table 2. Numbers of individuals with parental alleles at particular linkage blocks on linkage group C2† with notable differences between cultivars and random lines.

		Random lines	Cultivars
Locus-loci interval	Alleles‡	RIL with designated alleles/total lines § (frequency)	Cultivars with designated alleles/total lines (frequency)
Satt 460-Satt 307			
	WW	36/135 (0.27)	6/10 (0.60)
	EE	70/135 (0.52)	4/10 (0.40)
	WE	11/135 (0.08)	0/10
	EW	18/135 (0.13)	0/10
Satt 289			
	W	40/118 (0.34)	8/10 (0.80)
	E	78/118 (0.66)	2/10 (0.20)

† Complete marker interval Satt277 Satt365 Satt557 Satt289 Satt100 Satt460 Satt307 Satt202 Satt371.
‡ W denotes Williams allele; E denotes Essex allele.
§ Number of lines which have marker data for all loci in the interval.

A QTL for seed size (SS2; Hyten, 2002) is tightly linked to Satt 289 on linkage group C2 (fig. 4). At this locus, 80% of the cultivars and only 34% of the RILs have the Williams allele (table 2), which conditions a seed size increase of 4.5 mg seed^{-1} (Hyten, 2002). Other authors have detected a positive correlation between seed size QTL and seed yield (Csanadi et al., 2001; Mansur et al., 1993; Mian et al., 1996). The QTL SS2 appears to have a selective advantage in this population for higher yield.

The significantly higher crossing-over seen in the cultivars compared with the 156 RILs on linkage group L was not uniformly distributed. In the interval from Satt 182 to Satt 156 (fig. 4), all of the cultivars had at least one crossover, with nine having more than one crossover as compared with a mean 0.61 crossovers in the RILs. In the interval between Satt 156 and Satt 373 (fig. 4) we detected less recombination in the cultivars. Only one of the cultivars had more than one crossover in this region, while 18% of the RILs had more than one.

The QTLs for early maturity (M9 and M10) were found on linkage group L in the interval Satt 156 to Satt166 (Hyten, 2002; fig. 4). Ten percent of the cultivars had only Essex alleles at the Satt166 and Satt 229 markers, compared with 0.33 in the RILs. A frequency of 0.80 of the cultivars had the region from Satt 166 to Satt 229 inherited intact from the Williams parent (table 3). Conversely, a frequency of only 0.37 of the RILs had this intact block from Williams. A substitution of an Essex allele at the M10 QTL was estimated to affect a 2.6-d shortening of days to maturity. Later maturity in soybean is often correlated with higher seed yield (Mansur et al., 1993), and selection for yield related to later maturity (selection of the Williams allele at M10) may have driven the accumulation of this 30-cM chromosome region from Williams. Hyten (2002) also linked the Satt 156 marker, which immediately precedes the Satt 166 marker, to the M9 maturity QTL. None of the cultivars had the block from Satt 156 to Satt 229 inherited intact from the Essex parent (table 3), while 0.23 of the RILs received this block intact from Essex.

Table 3. Numbers of individuals with parental alleles at particular linkage blocks on linkage group L† with notable differences between cultivars and random lines.

		Random lines	Cultivars
Locus-loci interval	Alleles‡	RIL with designated alleles/total lines § (frequency)	Cultivars with designated alleles/total lines (frequency)
Satt 166–229			
	WW	54/144 (0.37)	8/10 (0.80)
	EE	47/144 (0.33)	1/10
	WE	17/144 (0.12)	1/10 (0.10)
	EW	26/144 (0.18)	0/10 (0.10)
Satt 156–229			
	WWW	43/133 (0.32)	4/10 (0.40)
	EEE	31/133 (0.23)	0/10 (0.10)
	WEE	9/133 (0.07)	1/10 (0.10)
	EWW	8/133 (0.06)	4/10 (0.40)

† Complete marker interval Satt156, Satt166, Satt229.

‡ W denotes Williams allele; E denotes Essex allele.

§ Number of lines which have homozygous marker data for all loci in the interval.

Selection for indeterminate growth habit would also favor the accumulation of Williams alleles in this region of linkage group L. All of the cultivars are indeterminate, conditioned by the Dt_1 allele from Williams. The Dt_1 locus is between Satt 166 and Satt 229 (fig. 4). In fact, selection for indeterminate growth habit may have led to only one of the cultivars receiving the minor yield QTL (yield 2) allele from Essex, which conditioned a 198 kg ha^{-1} increase in yield (Hyten, 2002). The cultivar Pennyrile resulted from a double crossover in this region to obtain the Essex alleles at Satt 166 and Satt 229 as well as the Dt_1 allele from Williams.

Hyten et al. (2004) mapped an oil QTL (O6) to linkage group M (fig. 4). The QTL was mapped to a region between Satt 540 and Satt 463. In this region, a frequency of 0.7 of the cultivars had Essex alleles at these two markers. A substitution of an Essex allele here conditioned a 2.4 g kg^{-1} reduction in seed oil content. Only a frequency of 0.51 of the RILs had Essex alleles in these locations. A protein concentration increasing QTL (P4) was also mapped on linkage group M (Hyten et al., 2004) (fig. 4). In one environment the QTL was mapped to a location half way between Satt 540 and Satt 463. A frequency of 0.51 of the RILs and 0.70 of the cultivars has the Essex allele, which conditions the increased protein concentration QTL at Satt 463. It appears this region may have contributed to increasing protein and decreasing oil concentrations in the cultivars.

Other researchers have mapped a QTL (fig. 4) for seed size on linkage group M near the Satt 210 marker (Csanadi et al., 2001). Only a small difference in the percentages of RILs (53%) and the cultivars (60%) that received the seed size increasing Williams allele at this locus was seen in the Williams × Essex population. However, in the cultivars 66% of the Williams alleles at Satt 210 were associated with a crossover between Satt 551 and Satt 250, which is mapped within 6.3 cM of Satt 210 on the published map (Cregan et al., 1999). Only 19% of the RILs having the Williams allele at Satt 210 exhibit this same crossover. It appears

that the cultivars, through recombination, might have collected a set of alleles beneficial for increasing yield on the basis of seed size contributions to yield.

Eight percent of the SSR markers had segregation ratios that deviated significantly ($p < 0.01$) from the 1:1 expected ratio (15% at $p < 0.05$) in the RIL set. Significant chi-square values for deviation from expected segregation ratios of molecular markers have been reported in soybean (17% of the markers; Zhang et al., 2004), rice (43% of the markers, Thomson et al., 2003; 81% of the markers, Nguyen et al., 2003), oilseed rape (Pilet et al., 2001), maize (Jampatong et al., 2002), and wheat (Liu et al., 2001; Campbell et al., 2003). Two (C2 and L) of the three linkage groups that had significantly greater standardized crossovers in the cultivars did not fit the 1:1 expected Mendelian segregation ratio for the entire linkage group in the RILs at ($p < 0.05$). Genetic control of preferential distribution of one parent's alleles in the formation of gametes in a segregating population in the absence of selection is one possible explanation for these observations. Natural selection in favor of one parent's alleles in the gametophyte or embryo is another possible explanation. The ten cultivars contained predominately Williams alleles in the Satt 289-Satt 307 region of linkage group C2. In the RIL population all markers in this region exhibited an excess number of Essex alleles (table 1). Thus it is hard to argue that the excess alleles were strongly favored by natural selection.

Is this pattern of excess Essex alleles real or an artifact of the experimental analysis of the markers? The same chi-square analysis was repeated on data from three other RIL populations using the same markers on linkage group C2. In the RILs from the Peking × Hamilton population, four of the 10 markers had segregation ratios significantly ($p < 0.01$) skewed toward the Peking parent, and the entire linkage group was skewed toward the Peking alleles ($p < 0.01$). None of the loci on C2 in the Pershing × Hamilton or Peking by Essex populations had distorted segregation ratios. This does suggest that these markers are reliable to use in measuring recombination and that the distortion measured in the Williams × Essex population is real. Perhaps Peking and Essex each possess a segregation distortion allele that is nonfunctional in the cross of these two parents together but functional in crosses of either Peking or Essex with another parent lacking the allele. We cannot suggest a way to predict the presence of segregation distortion in soybean populations before making a cross.

Numerous authors have reported the detection of QTL and have suggested methods for using this information in marker-assisted selection (MAS) (Diers and Shoemaker 1992; Mansur et al., 1993, 1996; Lark et al., 1995; Lee et al., 1996a, 1996b; Brummer et al., 1997; Qui et al., 1999). The progeny selected as cultivars from the Williams × Essex population had more total crossing-over than did the RILs, and the cultivars had greater recombination in some chromosome regions where QTL have not been detected. Perhaps selection for greater recombination in predetermined locations without knowledge of specific QTL in these locations could also be a beneficial strategy for MAS. Additionally, selection for more or less recombination in regions where we have knowledge of the presence of favorable or unfavorable combinations of alleles is possible.

References

Anand, S.C., and J.G. Shannon. 1985. Registration of Pershing soybean. Crop Sci. 25:367.

Bailey, N.T.J. 1961. Introduction to the mathematical theory of genetic linkage. Oxford Univ. Press, London.

Bernard, R.L., and D.A. Lindahl. 1972. Registration of Williams soybean. Crop Sci. 12:716.

Brummer, E.C., G.L. Graef, J. Orf, J.R. Wilcox, and R.C. Shoemaker. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci. 37:370–378.

Campbell, B.T., P.S. Baenziger, K.S. Gill, K.M. Eskridge, H. Budak, M. Erayman, I. Dweikat, and Y. Yen. 2003. Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. Crop Sci. 43:1493–1505.

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. 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., and R.C. Shoemaker. 1992. Restriction fragment length polymorphism analysis of soybean fatty acid content. J. Am. Oil Chem. Soc. 69:1242–1244.

Demarly, Y. 1979. The concept of linkat. p. 257–265. In C. Zeven and D. M. Harten (ed.) Proc. Conf. Broadening Genetic Base of Crops, Wageningen, the Netherlands.

Fatmi, A., C. Poneleit, and T.W. Pfeiffer. 1993. Variability of recombination frequencies in the Iowa Stiff Stalk synthetic. Theor. Appl. Genet. 86:859–866.

Gabe, H. 1994. Indeterminate short season soybeans for the south: Some breeding and agronomic considerations. p. 152–130. In Proceedings of the Southern Soybean Conference, Memphis, Tennessee. 14–16 Feb 1994. American Soybean Association, St. Louis, MO.

Hanson, W. 1959a. Theoretical distribution of the initial linkage block lengths intact in the gametes of a population intermated for n generations. Genetics 44:839–845.

Hanson, W. 1959b. The breakup of initial linkage blocks under selected mating systems. Genetics 44:857–868.

Hanson, W., and B. Hayman. 1963. Linkage effects on additive genetic variance among homozygous lines arising from a cross between two homozygous parents. Genetics 48:755–766.

Hyten, D.L. 2002. QTL mapping and identification of G3E interactions of agronomic and seed quality traits in soybean. Thesis presented for the Master of Science Degree at the University of Tennessee.

Hyten, D.L., V.R. Pantalone, C.E. Sams, A.M. Saxton, D. Landau- Ellis, T.R. Stefaniak,, and M.E. Schmidt. 2004. Seed quality in a prominent soybean population. Theor Appl. Genet. 109:552–556.

Jampatong, C., M.D. McMullen, B.D. Barry, L.L. Darrah, P.F. Byrne, and H. Kross. 2002. Quantitative trait loci for the first- and second-generation European corn borer resistance derived from the maize inbred Mo47. Crop Sci. 42:584–593.

Kisha, T., and B.W. Diers. 1997. Allele contribution of parents to selected progeny from two-way crosses. Soybean Genet. Newsl. 24:190–193.

Kiem, P., T.C. Olson, and R.C. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. 15:150–152.

Kiem, P., W.O. Beavis, J.M. Schupp, B. Baltazar, L. Mansur, R.E. Freestone, M. Vahedian, and D.M. Webb. 1991. RFLP analysis of soybean breeding populations: I. Genetic structure differences due to inbreeding methods. Crop Sci. 34:55–61.

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 maps of experimental and natural populations. Genomics 1:174–181.

Lark, K.G., K. Chase, F. Adler, L. Mansur, and J. Orf. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. Proc. Natl. Acad. Sci. USA 92:4656–4660.

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, Jr., D.A. 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.

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, S., C.A. Griffey, and M.A. Saghai Maroof. 2001. Identification of molecular markers associated with adult plant resistance to powdery mildew in common wheat cultivar Massey. Crop Sci. 41:1268– 1275.

Lorenzen, L., S. Lin, and R.C. Shoemaker. 1996. Soybean pedigree analysis using map base molecular markers: Recombination during cultivar development. Theor. Appl. Genet. 93:1251–1260.

Mansur, L. M., Lark, K.G., Kross H., and A. Oliveira A. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean. Theor Appl Genet. 86:907–913.

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

Mian, M.A.R., M.A. Bailey, J.P. Tamulonis, E.R. Shipe, T.E. Carter, Jr., W.A. Parrott, D.A. Ashley, R.S. Hussey, and H.R. Boerma. 1996. Molecular markers associated with seed weight in two soybean populations. Theor Appl Genet. 93:1011–1016.

Nguyen, B.D., S.B. Darshan, B.C. Bui, T.V. Nguyen, L.N. Pham, and H.T. Nguyen. 2003. Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, Oryza rufipogon Griff., into indica rice (Oryza sativa L.). Theor. Appl. Genet. 106:583–593.

Nickell, C.D., D.J. Thomas, and P. Stephens. 1990. Registration of Hamilton soybean. Crop Sci. 30:1364.

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. Comparisons of three related recombinant inbred populations. Crop Sci. 39:1642–1651.

Pfeiffer, T., and S. Vogt. 1990. Variability for recombination frequency in the AP12 soybean population. Crop Sci. 30:545–549.

Pilet, M.L., G. Duplan, M. Archipiano, P. Barret, C. Baron, R. Horvais, X. Tanguy, M.O. Lucas, M. Renard, and R. Delmourme. 2001. Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. Crop Sci. 41:197–205.

Piper, T., and W. Fehr. 1987. Yield improvement in a soybean population by utilizing alternative strategies of recurrent selection. Crop Sci. 27:172–178.

Qui, B.X., P.R. Arrelli, and D.A. Sleper. 1999. RFLP markers associated with soybean cyst nematode resistance and seed composition in a Peking × Essex population. Theor. Appl. Genet. 98:356–364.

Smith, T.J., and H.M. Camper. 1973. Registration of 'Essex' soybean. Crop Sci. 13:495.

Soybase. 1995. A genome database for Glycine. http://soybase.com; verified August 30, 2005.

Thomson, M.J., T.H. Hai, A.M. McClung, X.-H. Lai, M.E. Hinga, K.B. Lobos, Y. Xu, C.P. Martinez, and S.R. McCouch. 2003. Mapping quantitative trait loci for yield, yield components, and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson. Theor. Appl. Genet. 107:479–493.

Tulsieram, L., W. Compton, R. Morris, M. Thomascompton, and K. Eskridge. 1992. Analysis of genetic-recombination in maize populations using molecular markers. Theor. Appl. Genet. 84:65–72.

Zhang, W.K., Y.J. Wang, G.Z. Luo, J.S. Zhang, C.Y. He, X.L. Wu, J.Y. Gai, and S.Y. Chen. 2004. QTL mapping of ten agronomic traits on the soybean genetic map and their association with EST markers. Theor. Appl. Genet. 108:1131–1139.