Primary Trisomics and SSR Markers as Tools to Associate Chromosomes with Linkage Groups in Soybean

P. B. Cregan,* K. P. Kollipara, S. J. Xu, R. J. Singh, S. E. Fogarty, and T. Hymowitz

ABSTRACT

Primary trisomics provide an excellent cytogenetic tool to associate genes and linkage groups with their respective chromosomes. A complete set of 20 primary trisomics (2x + 1 = 41) has been established in soybean [Glycine max (L.) Merr.]. A linkage map of soybean with 20 consensus linkage groups has recently been defined. Because simple sequence repeat (SSR) markers map to defined single positions in the soybean genome, the association of a SSR locus with a chromosome will provide an unambiguous association of a linkage group to a specific chromosome. The objective of this work was to demonstrate the use of SSR markers to associate linkage groups with chromosomes by means of primary trisomics. One population of F2 plants was developed from an F1 hybrid trisomic for chromosome 13 (Triplo 13) and a second F2 population was obtained from a F1 hybrid trisomic for chromosome 5 (Triplo 5). Polymorphic SSR markers from different consensus linkage groups were tested on a subset of 20 plants from each population to identify markers that appeared to show segregation that deviated from normal (1:2:1) disomic inheritance. Markers not associated with the specific chromosome segregated in a disomic (1:2:1) fashion. Markers identified in this manner were further examined in the complete population of F2 plants to identify those that demonstrated trisomic segregation (6:11:1). By this approach, Triplo 13 was associated with linkage group F and Triplo 5 with linkage group A1. This result was verified by the examination of seven SSR loci on linkage group F and eight loci from linkage group A1 with each showing trisomic segregation with the Triplo 13- and Triplo 5-derived F2 populations, respectively. These results demonstrate the first association of molecular linkage groups with chromosomes in soybean and indicate that SSR markers provide a tool to associate the remaining 18 trisomics with their respective linkage groups.

A primary trisomic individual carries the normal complement of chromosomes plus an additional copy of one chromosome. Primary trisomics provide a way to associate genes and linkage groups with their respective chromosomes. In diploid species including maize (Zea mays L.) (Rhoades and McClintock, 1935), tomato (Lycopersicon esculentum Mill.) (Rick and Barton, 1954), barley (Hordeum vulgare L.) (Tsuchiya, 1967), and rice (Oryza sativa L.) (Khush et al., 1984), a complete set of primary trisomic individuals (each chromosome is trisomic in one member of the set) has allowed the establishment of cytogenetic maps. The trisomic sets have been useful in positioning classical genes and molecular markers on specific chromosomes by the use of modified genetic ratios associated with trisomic inheritance or by dosage effects from the extra chromosome (Tsuchiya, 1983; Carlson, 1988; Khush et al., 1984; McClouch et al., 1988; and Young et al., 1987).

In soybean [Glycine max (L.) Merr.], the complete set of primary trisomics (2x + 1 = 41) has been difficult to establish because mitotic metaphase chromosomes are morphologically indistinguishable. Early attempts by soybean cytologists resulted in the characterization of five trisomics (Gwyn and Palmer, 1989; Gwyn et al., 1985; Palmer, 1976; Sadanaga and Grindeland, 1984; Skorupska et al., 1989). In a species with small chromosome size and few differences in the morphology of mitotic metaphase chromosomes, pachytene chromosome analysis was shown to be a useful tool in the identification of individual chromosomes (Singh and Hymowitz, 1991). This approach was used to construct the first cytological map of soybean on the basis of chromosome length and euchromatin and heterochromatin distribution. As a result of the work of Singh and Hymowitz (1991), Ahmad et al. (1992), and Ahmad and Hymowitz (1994), 13 of the possible 20 primary trisomics were identified by pachytene chromosome analysis. Recently, these same researchers completed the characterization of the complete set of soybean primary trisomics (Xu et al. 2000).

In soybean, the association of linkage groups with specific chromosomes via trisomic analysis is in its initial stages. Hedges and Palmer (1991) associated the isozyme marker Dia1 (diaphorase) with Tri D (trisomic D). Tri D was identified as Triplo 4 by pachytene analysis by Singh and Hymowitz (1991). Sadanaga and Grindeland (1984) associated the w1 (flower color) locus (classical linkage group 8) with the satellite chromosome using Tri S which is now referred to as Triplo 13 (Singh and Hymowitz, 1991). More recently, Xu et al. (2000) confirmed the presence of w1 on Triplo 13 and also placed

Abbreviations: cM, centimorgan; PCR, polymerase chain reaction; MLG, molecular linkage group; RFLP, restriction fragment length polymorphisms; SSR, simple sequence repeat.
v2 (variegated leaf) on chromosome 5. They also associated eu1 (urease null), ixl (lipoxgenase null), and p2 (puberulent) on chromosomes 5, 13, and 20, respectively.

DNA marker technology has progressed rapidly in soybean. In the past few years, much interest has focused on the development of polymerase chain reaction (PCR) based single locus DNA markers with multiple alleles. The highly polymorphic nature (i.e., multiallelism) of simple sequence repeat (SSR) or microsatellite DNA markers is quite clear as evidenced by initial work of Akkaya et al. (1992) and Morgante and Olivieri (1993). Subsequent reports (Rongwen et al., 1995; Maughan et al., 1995; Powell et al., 1996; Diwan and Cregan, 1997) have described highly polymorphic microsatellite loci with as many as 26 alleles. SSR markers are the latest addition to the molecular genetic map of soybean that has developed over the past 10 yr. The first such map, published by Keim et al. (1990), was based upon 150 restriction fragment length polymorphism (RFLP) loci and contained 26 linkage groups. Subsequent reports by Shoemaker and Olson (1993), Shoemaker and Specht (1995), Mansur et al. (1996), Keim et al. (1997), and Cregan et al. (1999) have added greatly to the number of available DNA markers and their assembly into linkage maps with a total length of between 2400 and 3000 centimorgans (cM). The most recent map (Cregan et al., 1999) was actually a compilation of three maps based upon the genotyping of plants in three mapping populations. A total of more than 1400 DNA markers was positioned in this work, of which more than 600 were SSR markers. This resulted in the assembly of 20 consensus linkage groups. In some instances, two or more linkage groups on one map were joined not on the basis of a significant statistical association between markers in the groups, but on the basis of markers in common with a single linkage group in another population. Nonetheless, the final distillation into 20 molecular linkage groups led to the assumption that these groups corresponded to the 20 soybean chromosomes.

In addition to the synthesis of a linkage map containing 20 molecular linkage groups, Cregan et al. (1999) reported the association of 18 of the 20 classical linkage groups (Palmer and Shoemaker, 1998) with a molecular linkage group based on in situ segregation or linkage reports in the literature. Only classical linkage group 6 was not associated with a molecular linkage group. Thus, while the molecular and classical linkage maps are now substantially integrated, no association of these maps with specific soybean chromosomes has been made. It is therefore the objective of this paper to demonstrate the use of SSR markers to associate consensus molecular linkage groups with their respective soybean chromosome.

**Materials and Methods**

**Plant Materials**

**Triplo 13**

A plant (UT95-135, Xu et al., 2000) identified as trisomic for chromosome 13 (Triplo 13) was used as the female parent in a cross with an experimental line that was homozygous for the ixl (lipoxgenase null) gene. Root-tip squashes were used to identify trisomic F1 plants. The F1 seeds were germinated in a greenhouse sand bench at the Univ. of Illinois, Urbana, IL, and actively growing root tips were collected at 7 to 10 d after seeding and chromosomes were counted as described by Xu et al. (2000). F2 seeds were collected from a single trisomic F1 plant and a chip was removed from each of the seeds to determine the presence (Lxl-) or absence (lxIXlX) of lipoxyn- genase as described by Hildebrand and Hymowitz (1981). A total of 63 F2 plants was grown in the greenhouse at Beltsville, MD, and DNA was isolated from each as described by Keim et al. (1988).

**Triplo 5**

A plant (UT95-109, Xu et al., 2000) identified as trisomic for chromosome 5 (Triplo 5) was used as the female parent in a cross with an experimental line that was homozygous for the mutant alleles n1 (Kunitz trypsin inhibitor null), eu1 (urease null), and sl (beta amylase null) (Singh et al., 1998). Root-tip squashes were used to identify trisomic F1 plants at the Univ. of Illinois as described above. F2 seeds were collected from a single trisomic F1 plant. A total of 56 F2 plants was grown in the greenhouse at Beltsville, MD. DNA was isolated from the leaf tissue as described above. Plants were allowed to mature and seeds were collected from each plant. A sample of F2 seeds from each of the 56 F2 plants was used to determine the presence (Eu1-) or absence (eu1 eu1) of seed urease as described by Kloth et al. (1987).

**Simple Sequence Repeat Marker Analysis**

DNA of the parents used to produce the Triplo 13 F1 from which the F2 plants arose was not available. Therefore, to determine which SSR markers were polymorphic in this cross, two SSR markers from each of the 20 linkage groups defined by Cregan et al. (1999) were initially used to assay six plants from each set of F2 plants using 32P-labeled PCR products separated on DNA sequencing gels as described by Cregan and Quigley (1997). Additional loci were used if necessary to identify at least one polymorphic locus per linkage group. In the case of the Triplo 5 population, the SSR allele present in the n1 eu1 sl experimental line was compared with that of the cultivar Clark to identify polymorphic loci. Clark was used because the Triplo 5 parent used to create the F1 population had been backcrossed into the cultivar Clark 63.

Once polymorphic markers were identified, a group of 20 F2 plants from each population was assayed to identify loci that deviated significantly from 1:2:1 (homozygous for the SSR allele contributed by the female parent: heterozygous for the SSR allele contributed by the male parent). It was assumed that the anticipated excess of genotypes in one homozygous class along with the deficiency in the other would allow the putative identification of SSR loci mapping to the trisomic chromosome. Following the identification of SSR markers showing a putative association with either Triplo 13 or Triplo 5, each of these markers was used to genotype the entire population of F2 plants derived from the appropriate Triplo cross. The data obtained from the genotyping of the entire F2 populations were analyzed by \( \chi^2 \) for goodness of fit to a 1:2:1 segregation ratio and to a 6:1:1 segregation ratio. The latter ratio would be expected with trisomic inheritance and 50% female transmission of the extra chromosome (Singh, 1993). When the \( \chi^2 \) analysis indicated significant deviation from a 1:2:1 ratio at a given SSR locus, other markers in the same linkage group were used to genotype the complete set of plants from the appropriate F2 population.
Table 1. Simple sequence repeat markers used in the initial analysis of 20 $F_2$ plants from populations derived from trisomic $F_1$ plants from crosses of Triplo 13 $\times$ Lx1 and Triplo 5 $\times$ ti eu sp1, segregation ratios obtained, and $\chi^2$ probability for goodness of fit to a 1:2:1 ratio.

<table>
<thead>
<tr>
<th>SSR marker and F1 linkage group</th>
<th>Segregation ratio</th>
<th>$\chi^2$ probability for goodness of fit to 1:2:1</th>
<th>Plant no. of $F_2$ plants in the under-represented homozygous recessive class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triplo 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOYGPATR C1</td>
<td>5:12:3</td>
<td>0.55</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>GMGLPS12 I</td>
<td>4:12:3</td>
<td>0.49</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>GMGLRUBP F</td>
<td>10:9:1</td>
<td>0.02</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>Sat_022 D2</td>
<td>5:11:3</td>
<td>0.64</td>
<td>3, 28, 33, 68</td>
</tr>
<tr>
<td>Satt022 N</td>
<td>7:9:5</td>
<td>0.67</td>
<td>50% female transmission of the trisomic chromosome.</td>
</tr>
<tr>
<td>Satt030 F</td>
<td>8:11:1</td>
<td>0.08</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>SOYLGBC O</td>
<td>6:6:5</td>
<td>0.45</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Triplo 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMABAB N</td>
<td>5:10:4</td>
<td>0.92</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>Sat_044 K</td>
<td>8:9:3</td>
<td>0.26</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>Sat_064 G</td>
<td>8:7:5</td>
<td>0.26</td>
<td>18, 44, 45, 48, 70</td>
</tr>
<tr>
<td>Satt002 D2</td>
<td>6:10:3</td>
<td>0.61</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Satt005 D1b</td>
<td>5:8:7</td>
<td>0.55</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Satt006 L</td>
<td>6:11:3</td>
<td>0.58</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Satt030 F</td>
<td>7:8:5</td>
<td>0.55</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Sat0713 A1</td>
<td>12:7:0</td>
<td>&lt;0.001</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Sat0706 L</td>
<td>4:11:4</td>
<td>0.79</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Sat0709 C2</td>
<td>5:12:3</td>
<td>0.55</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Satt0809 A2</td>
<td>7:9:2</td>
<td>0.25</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Sat1002 C2</td>
<td>3:14:3</td>
<td>0.20</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Set_001 J</td>
<td>4:9:6</td>
<td>0.79</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>Set_028 C2</td>
<td>5:14:1</td>
<td>0.09</td>
<td>3, 28, 33, 45, 68</td>
</tr>
<tr>
<td>SOYGPATR C1</td>
<td>5:10:3</td>
<td>0.72</td>
<td>3, 28, 33, 45, 68</td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSION

Triplo 13

The initial screening of 20 $F_2$ plants was undertaken with SSR markers from only six linkage groups. Early in the screening, a possible association was indicated between linkage group F and chromosome 13 by apparent deviations from a 1:2:1 ratio obtained from the analysis of segregation data with SSR markers GMRUBP and Satt030 (Table 1). The analysis of the complete population of 63 plants with Satt030 resulted in a segregation ratio of 24:34:5 and a highly significant $\chi^2$ value indicating deviation from a 1:2:1 ratio (Table 2). Six additional loci on linkage group F yielded similar deviations from 1:2:1 segregation. None of the seven loci deviated from the expectation of 6:11:1 trisomic inheritance (Table 2). The deviation of these loci from the 1:2:1 ratio plus the fit to 6:11:1 confirmed the association of chromosome 13 with linkage group F. The molecular marker phenotypes of 32 of the 63 plants at the Sat_090 locus are shown in Fig. 1.

The analysis of seed lipoxygenase indicated that $F_2$ plant numbers 1, 28, 33, 45, and 68 were lipoxygenase nulls. This segregation ratio of 58 Lx1::3 lhx1Lx1 deviates significantly ($<0.001$) from a 3:1 ratio that would be anticipated for normal disomic segregation. The 58:5 ratio was a good fit to a 17:1 ratio ($\chi^2 = 0.68, P = 0.41$) that would be expected with trisomic inheritance and 50% female transmission of the trisomic chromosome. This result verifies the location of the Lx1 locus on chromosome 13 as previously reported (Xu et al., 2000). At the Sat_090 locus, $F_2$ plant numbers 1, 3, 28, 33, and 68 were homozygous for the under-represented allele. Thus, in four of five instances the molecular phenotype at the Sat_090 locus corresponded with the lipoxygenase null phenotype. This suggests that the Lx1 locus is situated close to Sat_090 because relatively little recombination has occurred between the two loci. In contrast, only one of the five plants that was homozygous for the under-represented allele at the Satt030, Satt193, Satt343, and Satt569 loci were lipoxygenase nulls. These loci are located at the opposite end of linkage group F from the Sat_090 locus (Cregan et al., 1999).

Triplo 5

The 20 plants were screened with SSR markers from 13 of the 20 linkage groups and the segregation tested for goodness of fit to a 1:2:1 ratio. Marker Sct_028 yielded a segregation ratio of 5:14:1 with a $\chi^2$ value of 4.8 ($P = 0.09$) suggesting a possible deviation from 1:2:1 segregation and the association between chromosome 5 and linkage group C2. However, tests of two additional loci on linkage group C2 (Satt079 and Satt100) did not sup-
port this association (Table 1). It might be desirable to genotype a larger number of plants initially to avoid false positive associations such as occurred in the case of Sct_028. However, as demonstrated in this instance, false positives were quickly identified by analysis of the population with other loci in the same linkage group.

The initial genotyping of 20 plants with Satt073 gave skewed segregation and the $\chi^2$ value for goodness of fit to a 1:2:1 ratio indicated highly significant deviation (Table 1). The analysis of the complete population of 56 F2 plants with Satt073 resulted in a segregation ration of 24:29:3 and a highly significant $\chi^2$ for deviation from a 1:2:1 ratio (Table 2). Seven additional loci on linkage group A1 showed deviation from 1:2:1 segregation and good fit to 6:11:1 segregation (Table 2) and thus confirmed the association of chromosome 5 with linkage A1. The molecular marker phenotypes of 28 of the 56 plants at the Satt300 locus are shown in Fig. 1.

In addition to the molecular marker data, the F1 seed produced by each F2 plant was characterized for the presence of seed urease. F2 plant numbers 30, 38, and 60 were urease nulls. Thus, the ratio of urease normal:urease nulls was 53:3 which is a highly significant ($P = 0.0007$) deviation from the expected 3:1 segregation ratio for a single gene. In contrast, this segregation fit trisomic 17:1 segregation ($P = 0.68$). This result suggests that the urease gene is on linkage group A1. The plant number of each of the individual plants in the under-represented homozygous recessive class for each SSR locus is given in Table 2. In the case of Satt364, Satt300, Satt155, and Satt073, F2 plant numbers 30, 38, and 60 were in the under-represented homozygous recessive class. The exact correspondence of homozygosity at the marker loci and at the urease locus suggests that the urease gene is located relatively near these molecular marker loci. Thus, it is not surprising that Satt364, Satt300, Satt155, and Satt073 are mapped to an interval that spans only 6.5 cM on the University of Utah linkage group A1-U07 (Cregan et al., 1999). However, Satt471 also maps to this interval but at this locus, plants 30, 38, and 60 were urease nulls. Thus, the ratio of urease normal:urease nulls was 53:3 which is a highly significant ($P = 0.0007$) deviation from the expected 3:1 segregation ratio for a single gene. In contrast, this segregation fit trisomic 17:1 segregation ($P = 0.68$). This result suggests that the urease gene is on linkage group A1. The plant number of each of the individual plants in the under-represented homozygous recessive class for each SSR locus is given in Table 2. In the case of Satt364, Satt300, Satt155, and Satt073, F2 plant numbers 30, 38, and 60 were in the under-represented homozygous recessive class. The exact correspondence of homozygosity at the marker loci and at the urease locus suggests that the urease gene is located relatively near these molecular marker loci. Thus, it is not surprising that Satt364, Satt300, Satt155, and Satt073 are mapped to an interval that spans only 6.5 cM on the University of Utah linkage group A1-U07 (Cregan et al., 1999). However, Satt471 also maps to this interval but at this locus, plants 30, 38, and 60 were urease nulls. Thus, the ratio of urease normal:urease nulls was 53:3 which is a highly significant ($P = 0.0007$) deviation from the expected 3:1 segregation ratio for a single gene. In contrast, this segregation fit trisomic 17:1 segregation ($P = 0.68$). This result suggests that the urease gene is on linkage group A1. The plant number of each of the individual plants in the under-represented homozygous recessive class for each SSR locus is given in Table 2. In the case of Satt364, Satt300, Satt155, and Satt073, F2 plant numbers 30, 38, and 60 were in the under-represented homozygous recessive class. The exact correspondence of homozygosity at the marker loci and at the urease locus suggests that the urease gene is located relatively near these molecular marker loci. Thus, it is not surprising that Satt364, Satt300, Satt155, and Satt073 are mapped to an interval that spans only 6.5 cM on the University of Utah linkage group A1-U07 (Cregan et al., 1999). However, Satt471 also maps to this interval but at this locus, plants 30,
and linkage groups is the highly polymorphic nature of SSR loci. Thus, the likelihood that a given locus will be polymorphic in a particular population is greater with SSR markers than with any other marker type. The populations used in the current study were created by crossing experimental lines that were fairly similar in their genetic backgrounds. The genetic backgrounds of the two genotypes carrying Triplo 13 and Triplo 5, respectively, were related to the cultivar Clark. Likewise, the genotypes used as the donors of the urease null and the lipoygenase null were related to Clark. As a result, only about 15% of the SSR markers tested were polymorphic in the initial assay used to identify polymorphic SSR loci. A higher level of polymorphism would expedite the process of associating linkage groups with chromosomes. At this time, the 20 soybean primary trisomics are being backcrossed into the cultivar Clark 63 so that each can be examined in a common genetic background. In order to provide a high level of SSR length polymorphism, each of the primary trisomics are being crossed with \textit{Glycine soja} PI 407287. PI 407287 is a vigorous genotype that produces completely fertile hybrids with cultivated soybean (Cregan et al., 1989). Crosses of PI 407287 have been made to each of the 20 primary trisomics. Cytological studies to identify trisomic F1 plants are currently underway at the University of Illinois. In defining 20 consensus linkage groups, Cregan et al. (1999) assumed each to be associated with a separate chromosome. In the near future, it is anticipated that this assumption will be tested by the use of the newly developed complete set of soybean primary trisomics.

**ACKNOWLEDGMENTS**

The authors thank Ms. Sarah Hyatt and Mr. DaRel Barkdendale for their excellent technical assistance during the course of the work reported here. This research was supported in part by funds from United Soybean Board (USB Grants #8207 and 9222) to PBC. Research at the University of Illinois supported in part by the Illinois Agricultural Experiment Station, Illinois Soybean Program Operating Board grant 97-23-182-3HY and C-FAR grant 1.5-95411.

**REFERENCES**


