Theor Appl Genet (2008) 117:143–147
DOI 10.1007/s00122-008-0759-6

ORIGINAL PAPER

Soybean molecular linkage group B1 corresponds to classical linkage group 16 based on map location of the lf_2 gene

Thomas M. Seversike · Jeffery D. Ray · Jeffry L. Shultz · Larry C. Purcell

Received: 5 October 2007 / Accepted: 25 March 2008 / Published online: 8 April 2008
© Springer-Verlag 2008

Abstract The seven-leaflet character of soybean [*Glycine max* L. (Merr.)] is a single recessive trait conditioned by the lf_2 gene. The lf_2 gene is located on linkage group (LG) 16 of the classical soybean genetic map, but it has not been placed on the molecular map. The objective of this research was to identify the location of the lf_2 gene on the soybean molecular map using simple sequence repeat (SSR) markers. A backcross breeding method was used to create three- and seven-leaflet near-isogenic lines in genetic backgrounds of 'Traill', 'MN1401', and 'MN1801'. Eight mapping populations were derived from eight single heterozygous Lf_2 lf_2 plants. A total of 482 SSR markers that covered approximately every 10–20 cM of all soybean molecular LG were used to screen the mapping populations for polymorphisms. For the 115 SSRs that were identified as polymorphic, possible linkage between the lf_2 gene and the polymorphic SSR markers was determined. One SSR marker from the LG B1, Sat_272, was linked (LOD > 4.0) to the lf_2 gene in the Traill and MN1401 derived populations, with map distances ranging from 2.8 to 11.2 cM. Two additional markers (a SSR, Sat_270 and a SNP, A588c) located on LG B1 were also polymorphic and were identified as linked to the lf_2 gene in one of the populations. This research was successful in mapping the lf_2 gene to LG B1 of the soybean molecular map and therefore, provides evidence that molecular LG B1 corresponds to classical LG 16.

Abbreviations

LG linkage group
RFLP restriction fragment length polymorphism
SNP single nucleotide polymorphism
SSR simple sequence repeat

Communicated by I. Rajcan.

T. M. Seversike
Department of Crop Science,
North Carolina State University, Williams Hall 4411,
Box 7620, Raleigh, NC 27695-7620, USA

J. D. Ray
USDA-ARS, 141 Experiment Station Road,
JWDSRC, Stoneville, MS 38776-0345, USA

J. L. Shultz
Louisiana Tech University, 120 Carson Taylor Hall,
Ruston, LA 71272, USA

L. C. Purcell (✉)
Department of Crop, Soil, and Environmental Sciences,
University of Arkansas, 1366 W. Altheimer Drive,
Fayetteville, AR 72704, USA
e-mail: lpurcell@uark.edu

Introduction

The seven-leaflet trait in soybean was discovered by Fehr (1972) from a spontaneous mutation in the cultivar Hawkeye. The recessive lf_2 allele conditions seven-leaflet expression and the dominant Lf_2 allele conditions three-leaflet expression (Fehr 1972). Fehr (1972) reported that although the Lf_2 allele was completely dominant to the lf_2 allele in his three- leaflet cross × seven-leaflet cross, it was not completely dominant in all genetic backgrounds.

In addition to the lf_2 gene, there is another gene that controls leaflet number in soybean. The Lf_1 gene conditions expression of the five-leaflet trait in soybean. It is partially dominant to the lf_1 allele, which conditions expression of the three-leaflet trait (Fehr 1972). The Lf_1 gene is located at map position 126.7 on linkage group (LG) A2 of the soybean molecular map (Song et al. 2004), but it has not been linked to the classical genetic map. Conversely, the lf_2 gene

was found to be linked to the pubescence type (Pd_2) gene (dense pubescence) on LG 16 of the classical map (Devine 2003; Palmer et al. 2004), but neither the lf_2 nor Pd_2 loci have been placed on the molecular map.

The objective of this research was to map the gene conferring the seven-leaflet trait on the soybean molecular map. Segregating populations of near isolines for lf_2 were created in three genetic backgrounds, totaling 430 individuals. The populations were phenotyped, genotyped and evaluated for linkage with simple sequence repeat (SSR) markers, and more closely mapped with single nucleotide polymorphisms (SNPs).

Materials and methods

Population development

Plant introduction (PI) 548232 possesses the single recessive gene (lf_2) conferring the seven-leaflet trait, and it was mated with three-leaflet (Lf_2) cultivars Traill (Helms and Nelson 1998), MN1401 (Orf and Denny 2000a), and MN1801 (Orf and Denny 2000b). Detailed information on breeding methodology and phenotyping was described by Seversike et al. (2008).

The F_1 generation was allowed to self-pollinate, and homozygous recessive plants (lf_2lf_2) were backcrossed with the recurrent parents Traill, MN1401, and MN1801. Backcross populations were used to reduce extraneous phenotypic variation and reduce the potential number of polymorphic molecular markers that were unassociated with lf_2. All of the progenies were advanced to the BC_1F_2 generation, and Traill- and MN1801-derived plants were backcrossed two additional times and advanced to the BC_3F_2 generation. Two heterozygous BC_3F_2 plants of MN1801 and four BC_3F_2 plants of Traill were allowed to self pollinate and were harvested individually. $BC_3F_{2:3}$ seeds were sown in the field in 2005 and served as our mapping populations for MN1801 and Traill (Table 1).

Heterozygous (Lf_2lf_2) plants of MN1401 at the BC_1F_2 and BC_1F_3 generations were selected and allowed to self pollinate. Seeds from two BC_1F_4 heterozygous plants were individually harvested and sown in separate rows in the field in 2005. These $BC_1F_{4:5}$ segregating plants served as our mapping populations for MN1401 (Table 1).

Individual plants were phenotyped as homozygous recessive (seven leaflets), heterozygous (three leaflets with one or more leaves having four leaflets), and homozygous dominant (three leaflets only). Seeds from plants putatively classified as either homozygous dominant or heterozygous were harvested and progeny tested in the greenhouse. No attempt was made to phenotype plants for Pd_2, which is linked to lf_2 on the classical map (Devine 2003; Palmer et al. 2004).

Table 1 The number of homozygous dominant (Lf_2 Lf_2), heterozygous (Lf_2 lf_2), and homozygous recessive (lf_2 lf_2) individuals (phenotype) in each population

Population	Phenotype			Total	χ^2	Probability
	Lf_2 Lf_2 [a]	Lf_2 lf_2 [a]	lf_2 lf_2			
MN1401-1 $BC_1F_{4:5}$	10	34	13	57	2.4	0.30
MN1401-2 $BC_1F_{4:5}$	10	38	23	71	5.1	0.08
Traill-1 $BC_3F_{2:3}$	4	35	12	51	9.6	0.01
Traill-2 $BC_3F_{2:3}$	9	20	5	34	2.0	0.37
Traill-3 $BC_3F_{2:3}$	12	21	13	46	0.4	0.82
Traill-4 $BC_3F_{2:3}$	7	18	10	35	0.5	0.76
MN1801-1 $BC_3F_{2:3}$	10	29	14	29	1.1	0.58
MN1801-2 $BC_3F_{2:3}$	13	65	29	107	9.7	0.01

Chi-square test was based upon expected phenotypic ratios for segregation at a single locus (i.e., 1:2:1)

[a] Phenotypes of putative Lf_2 Lf_2 and Lf_2 lf_2 individuals were confirmed in a progeny test

Molecular markers

Simple sequence repeat (SSR) markers were used to screen for polymorphisms between the DNA from homozygous three- and seven-leaflet near-isogenic pairs. Leaf tissue samples were taken from lf_2 and Lf_2 genotypes from the populations developed in each of the backgrounds (Traill, MN1401, and MN1801) in the $BC_3F_{2:3}$, $BC_1F_{4:5}$, and $BC_3F_{2:3}$ generations, respectively (Table 1). The samples were freeze-dried and pulverized to a fine powder in a ball mill (Garcia Manufacturing, Visalia, CA). DNA was extracted from the samples using a CTAB (cetyltrimethylammonium bromide) buffer (Saghai-Maroof et al. 1984) with a phenol/chloroform purification step employing a heavy phase lock gel (5 Prime Inc, Gaithersburg, MD).

There were 482 SSR markers chosen for screening, and they covered approximately every 10 to 20 cM of all 20 soybean LGs (Song et al. 2004). Each SSR primer was fluorescently labeled with either FAM or HEX (Integrated DNA Technologies, Skokie, IL). The SSR primers were used to amplify the DNA from near-isogenic pairs using the polymerase chain reaction (PCR). Amplification conditions were 95°C for a 2-min activation step followed by 35 cycles of 94°C for 30 s, 46°C for 30 s, and 72°C for 30 s, with a final step of 72°C for 5 min. Sigma JumpStart Taq and buffers were used in the PCR according to manufacturer's instructions (Sigma-Aldrich, St. Louis, MO). Amplicons were separated and measured using an ABI 3730 XL sequencer (Applied Biosystems, Foster City, CA) in the USDA-ARS Midsouth Genomics Laboratory (Stoneville, MS), and amplicon length was determined using Gene Mapper 3.7 software (Applied Biosystems, Foster City, CA). When polymorphisms were identified, the SSR

primer was considered a possible candidate marker associated with the lf_2 gene and linkage analysis was performed.

The SSR marker Sat_272 located on LG B1 was determined to be putatively linked to the lf_2 gene, and all other reported SSR markers in that region (Song et al. 2004) were tested, but no polymorphisms were found. Subsequently, eight SNPs previously reported in that region of LG B1 (Choi et al. 2007) were also evaluated. An allele specific primer extension (ASPE) method was used for SNP detection (Lee et al. 2004) on a Luminex 100 (Luminex, Austin, TX) using "per region" settings and Miraibio CT/GT software. Median fluorescent intensity (MFI) was used to identify specific alleles, with the lowest MFI ratio used to score a homozygous allele at 80%.

In the region of LG B1, where Sat_272 was putatively found to be linked to the lf_2 gene, there were also seven previously mapped restriction fragment length polymorphism (RFLP) markers. To make these easier to evaluate, the RFLPs were converted into sequence-tagged site (STS) primers (Table 2). Each of these primer pairs was used to amplify genomic DNA from the parents. The PCR product was electrophoresed and all monomorphic amplicons were excised, cloned and sequenced. ASPE primers were designed from a total of seven of these sequences with the Primer 3 program (Rozen and Skaletsky 2000). Settings were for an 18–27 base pair oligonucleotide length, $T_m = 53–57°C$, and GC% of 20–80%. These were then evaluated as SNPs on the Luminex 100 as described above.

Linkage and statistical analysis

The phenotypic and molecular marker data were analyzed in SAS (V8, SAS Institute, Cary, NC) using a Chi-square 3×3 contingency table and an F-test to determine putative linkage to the lf_2 gene ($P < 0.001$). Linkage was confirmed and molecular distances calculated using JoinMap 4.0 software (Van Ooijen, 2006). JoinMap 4.0 was employed because it has the appropriate algorithms to analyze the different combinations of backcrosses and self-pollinated generations used in this study (e.g. $BC_3F_{2:3}$ and $BC_1F_{4:5}$). All linkages had a minimum LOD score of 5.0, except for Sat_247 in the Mn1401-2 population which had a LOD score of 3.75. However, comparison with the soybean consensus map supports the linkage of Sat_247 (Fig. 1).

Results

Near-isogenic lines for Traill, MN1401, and MN1801 were successfully created and used in field experiments evaluating the seven-leaflet trait (Seversike et al. 2008). With the exception of the Traill-1 and MN1801-2 populations, phenotypic segregation of all populations for the seven-leaflet trait fit the 1:2:1 ratio expected for a simple Mendelian trait (Table 1). For the Traill-1 and MN1801-2 populations there were fewer homozygous dominant individuals observed than expected and more heterozygous individuals observed than expected. In the SSR screening, 115 out of 482 SSR markers were found to be polymorphic between three- and seven-leaflet near-isogenic plants. Five SSR markers were polymorphic across all three backgrounds (Traill, MN1401, and MN1801), but subsequent analysis indicated that none of these markers was linked to the seven-leaflet trait.

Of the 115 polymorphic markers identified, only one, Sat_272, was putatively linked to the lf_2 gene in the two MN1401 populations and the four Traill populations. Sat_272 was not polymorphic in the MN1801 isoline pair. Sat_272 is located approximately 14.3 cM from one end of soybean LG B1 (Song et al. 2004), and the lf_2 locus mapped distal to this marker. There was a highly significant ($P < 0.0001$) association between Sat_272 and the lf_2 trait in the MN1401 and Traill populations, and linkage was

Table 2 List of PCR primers and probes designed to convert previously reported RFLPs to SNPs

SNP	LG B1 (cM)[a]	Forward/Reverse Primers (5′–3′)	ASPE Probes (5′–3′)	Source (GenBank #)
A333a	8.1	AGTTGTGGATTTCGGACTGT GAACCGGAGAAATCTGAGAA	ATTGTAAGTTCATTT CTTTTCTTT**A/G**	AQ842018
A333b	8.1	GAATGCTGCACTATCTGCAC GAGAAACACACATGCCAAAA	CTGGGTGGGAAAAC ATT**T/C**	AQ841817
A588c[b]	20.3	ACCAAGAAACTGCATCATCA TGGCTAGAGCATCAACACAT	TGAAGTGATCAGAC AAAACTTAC**T/A**	AQ841768
K011a	17.6	CCGGAATACCCAGATAACAG GGCTGATGAGAGAAGCAGTT	CAGCGTATTGAAGA ACATT**TA/CG**	AQ842190
A702a	29.0	GCAGCACTAGAGAGTGCAAA TATCCAGGGCATCAGAATTT	CTAGAGAGTGCAAA CTTTGT**GTA/ATG**	BH173805
A702b	29.0	AAGGAGCAATTGTTGAGAGG CAACCAAGTCAGTACGTCCA	AGTGTAAGTTATTCT ATTGATAGTGG**A/T**	AQ841991
A702c	29.0	TGCTAGTGCAGCTCTTTCTG TCATGGACGTACTGACTTGG	ATTTTCAGGCCTAAT **GCAA/TAC**	AQ841800

[a] Based on the current integrated map (Choi et al. 2007)

[b] Polymorphic in Mn1401-1 and Mn1401-2 mapping populations (see Fig. 1)

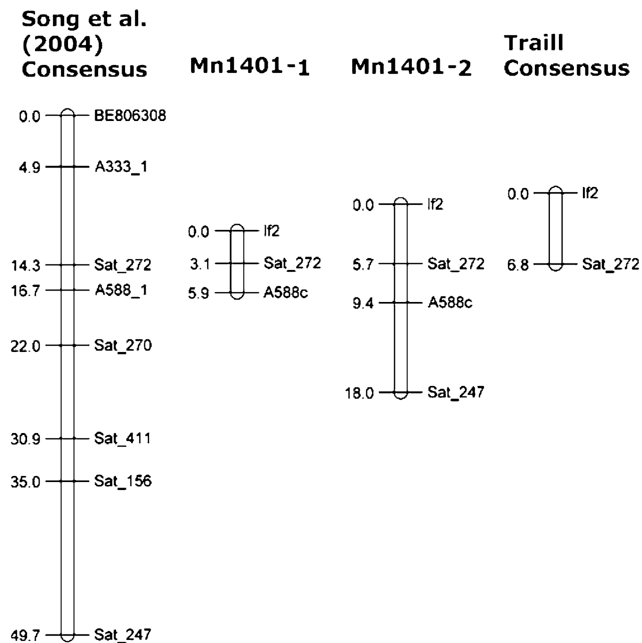

Fig. 1 Diagrammatic representation of the linkage position of the *lf*$_2$ gene. On the *left* is a portion of the published soybean consensus map for LG B1 (Song et al. 2004) for reference. The linkage groups for Mn1401-1 and Mn1401-2 show all the markers in the region of LG B1 that were polymorphic. The "Traill Consensus" linkage group is a composite of four Traill mapping populations (see Table 1)

confirmed with JoinMap 4 software. The distance between Sat_272 and the *lf*$_2$ gene was estimated to be approximately 3.1 cM (LOD 8.46) in the MN1401-1 population and 5.7 cM (LOD 11.0) in the MN1401-2 population (Fig. 1). For the four Traill populations, the map distances between Sat_272 and the *lf*$_2$ gene ranged from 2.8 to 11.2 cM (LOD 4.4 to 9.4). The variability in map distance among populations within the same cultivar background (MN1401 or Traill) may be due to differences in population size (Table 1) and the small amount of heterozygosity remaining in the backcross populations.

With the likely location of the *lf*$_2$ gene identified, other SSR markers within the region were screened. Additional linked markers at the same location would serve as strong confirmation of the location of the *lf*$_2$ gene. Although all SSRs within 50 cM of Sat_272 of LG B1 (Song et al. 2004) were screened, only one additional polymorphic SSR (Sat_247) was identified. Sat_247 was polymorphic in the MN1401-2 population but was not polymorphic in the MN1401-1, Traill or MN1801 populations. In the MN1401-2 population, linkage analysis indicated that Sat_247 was linked to the *lf*$_2$ gene at a distance of approximately 18.0 cM, and placed Sat_272 between Sat_247 and the *lf*$_2$ gene (Fig. 1).

Although the putative linkage of an additional SSR in the same region of LG B1 confirms the linkage with the *lf*$_2$ gene, the fact that Sat_247 was polymorphic in only one population led us to search for additional markers. A soybean SNP map (Choi et al. 2007) identified eight SNPs in the same region of interest on LG B1. These eight SNPs were tested for polymorphisms, but none of these SNPs were polymorphic in any of the mapping populations.

Early mapping efforts in soybean placed a number of RFLPs in the region of interest on LG B1 (Song et al. 2004). We converted these RFLPs to sequence tagged sites and evaluated them as SNPs in our mapping populations (Table 2). One of these SNPs (A588c) was polymorphic in both of the MN1401 populations. Linkage analysis showed A588c as putatively linked to the *lf*$_2$ gene at distances of 5.9 (MN1401-1) and 9.4 (MN1401-2) cM, and placed SSR Sat_272 between A588c and the *lf*$_2$ gene. Additionally, in the MN1401-2 population, which was polymorphic for Sat_247, analysis indicated that A588c was between Sat_247 and Sat_272 (Fig. 1). Alignment of the linked markers in this study with the soybean consensus map (Fig. 1) indicates that the *lf*$_2$ gene is likely located near the end of LG B1. No polymorphic markers were found between *lf*$_2$ locus and the end of the LG.

Discussion

Figure 1 shows the results of the linkage analysis for the polymorphic markers and the *lf*$_2$ gene in the two MN1401 populations and a consensus map of the Traill populations along with the published consensus map of the region of interest on LG B1 (Song et al. 2004). All published markers in this region were evaluated for polymorphisms in the mapping populations. However, only three polymorphic markers were identified (Sat_272, A588c, and Sat_247). Sat_272 was also polymorphic in both MN1401 populations and all four Traill populations, and in all of these populations, Sat_272 was putatively linked to the *lf*$_2$ gene in between 2.8 and 11.2 cM. A RFLP that was converted to a SNP (A588c) was poplymorphic in the two MN1401 populations and was putatively linked to the *lf*$_2$ gene between 5.9 and 9.4 cM. Lastly, in the MN1401-2 population, Sat_247 was polymorphic and was putatively linked to the *lf*$_2$ gene at a distance of about 18.0 cM. All three markers mapped to the same side of the *lf*$_2$ gene. Nevertheless, these multiple markers in multiple populations provide strong evidence of the genomic location of the *lf*$_2$ gene as near the end of LG B1 (Fig. 1). Based on our analysis, we place the *lf*$_2$ gene in between 3.1 and 11.5 cM on LG B1 of the soybean molecular map (Song et al. 2004).

Since the *lf*$_2$ gene is linked to the dense *Pd*$_2$ on the classical map (Devine 2003; Palmer et al. 2004), our research also indicates the genomic location of *Pd*$_2$ on LG B1. Additionally, our research provides strong evidence that molecular

LG B1 corresponds to classical LG 16, thereby, allowing the integration of a portion of the classical soybean map with the molecular map. These represent the only classically mapped traits thus far known to be on molecular LG B1.

Acknowledgments The authors thank Angelie Davis, Andy King, and Marilynn Davies for their excellent technical assistance.

References

Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, Hwang EY, Yi SI, Young ND, Shoemaker RC, van Tassell CP, Specht JE, Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics 176:685–696

Devine TE (2003) The Pd_2 and Lf_2 loci define soybean linkage group 16. Crop Sci 43:2028–2030

Fehr WR (1972) Genetic-control of leaflet number in soybeans. Crop Sci 12:221–224

Helms TC, Nelson BD (1998) Registration of 'Traill' soybean. Crop Sci 38:549

Lee SH, Walker DR, Cregan PB, Boerma HR (2004) Comparison of four flow cytometric SNP detection assays and their use in plant improvement. Theor Appl Genet 110:167–174

Orf JH, Denny RL (2000a) Registration of 'MN1401'. Crop Sci 40:1825

Orf JH, Denny RL (2000b) Registration of 'MN1801' soybean. Crop Sci 40:1826

Palmer RG, Pfeiffer TW, Buss GR, Kilen TC (2004) Qualitative genetics. In: Boerma HR, Specht JE (eds) Soybeans: improvement, production, and uses, third edn. ASA CSSA SSSA, Madison, pp 137–233

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley: Medelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci U S A 81:8014–8018

Seversike TM, Purcell LC, Gbur E, Chen P, Scott R (2008) Radiation interception and yield responses to increased leaflet number in early-maturing soybean genotypes. Crop Sci (accepted for publication)

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Van Ooijen JW (2006) JoinMap 4®, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen