Theor Appl Genet (2001) 103:912–919

G. Csanádi · J. Vollmann · G. Stift · T. Lelley

Seed quality QTLs identified in a molecular map of early maturing soybean

Received: 9 October 2000 / Accepted: 26 February 2001

Abstract This study identified QTLs influencing seed quality characters in a cross of two early maturing soybean (*Glycine max* [L.] Merr.) cultivars (Ma.Belle and Proto) adapted to the short growing seasons of Central Europe. A molecular linkage map was constructed by using 113 SSR, 6 RAPD and 1 RFLP markers segregating in 82 individuals of an F_2 population. The map consists of 23 linkage groups and corresponds well to previously published soybean maps. Using phenotypic data of the F_2-derived lines grown in five environments, four markers for protein content, three for oil content and eight for seed weight were identified. Four from fifteen seed quality QTL-regions identified in the present study were also found by other authors. Markers associated with seed weight QTLs were consistent across all environments and proved to have effects large enough to be useful in a marker-assisted breeding program, whereas protein and oil QTLs showed environmental interactions.

Keywords *Glycine max* (L.) · SSR · RAPD · RFLP · Marker-assisted selection

Introduction

Soybean [*Glycine max* (L.) Merr., 2n=2*x*=40] is one of the World's major crops grown for seed protein and oil content. In Asia it is used mainly for its protein content in human consumption. In North America, it is used mostly for animal feed and for oil production (Smith and Huyser 1987). In Europe it is almost exclusively grown for its protein content to feed livestock. However, soybean varieties of early maturity groups adapted to short-season growing conditions of Central Europe have a lower average protein content than available on the world market (Vollmann et al. 2000). Therefore, to improve the competitive value of Austrian grown soybean varieties, increasing the seed protein content of adapted germplasm is highly desirable.

Recent developments in the use of molecular markers can make plant breeding more efficient (Paterson et al. 1991). Marker-aided selection (Lande and Thompson 1990) can serve as a tool to substantially increase the efficiency of selecting appropriate genotypes. The first molecular genetic maps of soybean were constructed using different kinds of parents and segregating populations. Shoemaker and Olson (1993) used an interspecific cross between *Glycine max* and *Glycine soja*, Lark et al. (1993) used recombinant inbred lines from a cross between Minsoy and Noir, and Muehlbauer et al. (1991) used near-isogenic lines (NILs) between the varieties Clark and Harosoy. Meanwhile, a large number of reports on the molecular mapping of characters such as seed protein and oil content (Diers et al. 1992; Mansur et al. 1993, 1996; Lark et al. 1994; Lee et al. 1996; Brummer et al. 1997; Qiu et al. 1999) shows the great interest for using molecular markers in soybean breeding.

In this study, molecular markers associated with seed quality characters were identified in a segregating soybean population of the early maturity gene pool. As a first step a molecular-marker map was established and compared with the already published maps. Microsatellite markers, frequently referred to as simple sequence repeats (SSRs), were used in constructing the core map. Such a map can later be saturated by other markers such as RFLPs, RAPDs and AFLPs. In a second step, markers associated with seed protein content, oil content and seed weight were identified by means of a single-marker approach and interval mapping.

Communicated by J.W. Snape

G. Csanádi · G. Stift · T. Lelley (✉)
Department of Plant Biotechnology,
Institute for Agrobiotechnology, Konrad Lorenz Strasse 20,
A-3430 Tulln, Austria
e-mail: lelley@ifa-tulln.ac.at
Tel.: +43 2272/66280 204

J. Vollmann
Department of Plant Breeding,
University of Agricultural Sciences Vienna,
Gregor Mendel Strasse 33, A-1180 Vienna, Austria

Present address:
G. Csanádi, Department of Biotechnology, University of Szeged,
Temesvari krt. 62, H-6726 Szeged, Hungary

Table 1 Description of the two parental genotypes

Genotype	Pedigree	Maturity Group	Origin
Proto	M70-504/M69-42	0	University of Minnesota
Ma. Belle	Unknown	00	Agriculture Canada

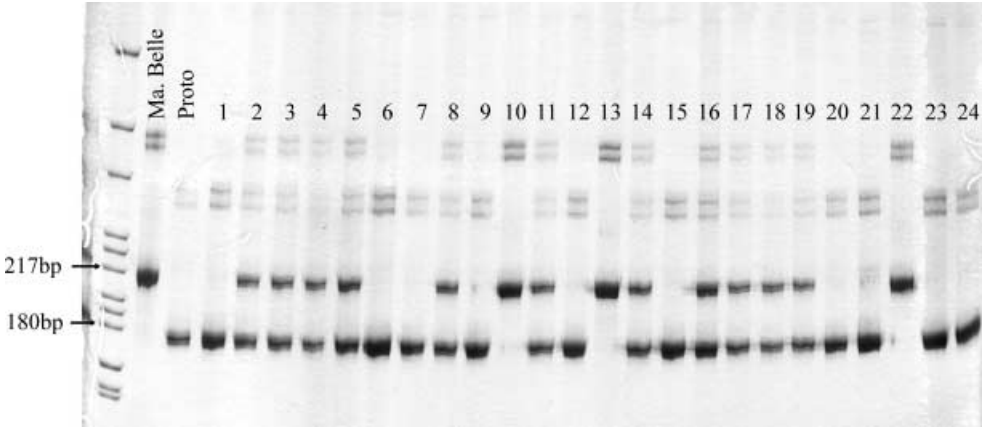

Fig. 1 Example of fragment separation for Ma.Belle, Proto and 24 F_2-plants in a 12% polyacrylamide gel for SSR marker Satt316 with size marker pBR322/*Msp*I

Materials and methods

Plant material

The two parent genotypes, Ma.Belle and Proto (Table 1), originate from a collection of early maturing soybean accessions evaluated at the University of Agricultural Sciences Vienna. These two genotypes were selected because of their difference in seed protein content (Proto 462 g kg^{-1} Ma.Belle 390 g kg^{-1}) and their large genetic distance (Doldi et al. 1997). Ma.Belle and Proto were crossed, and the F_1 plants were self-pollinated resulting in an F_2 population of 82 individuals which were used for the mapping experiment.

Phenotyping

For phenotypic evaluations of seed protein content, oil content and seed weight, F_2-derived lines were grown in five environments in the east of Austria, i.e. Raasdorf and Vienna in 1997, and Raasdorf, Vienna and Pama in 1998. In each environment, genotypes were planted in single-row plots of 2-m length using generalized lattice designs in two replications. Seed protein and oil content were determined in g kg^{-1} on a dry matter basis using near-infrared reflectance spectroscopy (NIRS) as described in Vollmann et al. (1996). For each environment, raw data were evaluated statistically using the PLABSTAT software package (Utz 1995). For subsequent genetic analyses, lattice-adjusted mean values were used.

DNA analysis

DNA was isolated from freeze-dried leaf tissue. Total DNA was prepared according to Keim et al. (1988a) with slight modifications due to the high level of polysaccharides in field-grown plants. Chloroform/isoamylalcohol was applied twice and, after precipitation, DNA was washed in 15 mM NH_4Ac dissolved in 80% ethanol. For RFLP detection 100 probes (BioGenetic Services, Brookings, South Dakota, USA) were used, originating from a genomic DNA library constructed by Keim et al. (1988b). They were pre-selected for unique or low-copy number fragments using *Pst*I enzyme digestion (Keim et al. 1988b). For radioactive hybridisation the buffer (C+G) of Church and Gilbert (1984), under the conditions described by Kiss et al. (1993) for alpha-P^{32} dCTP nucleotide, was used. For RAPD analysis primer Kits 02 and 03 were purchased from Advanced Biotechnologies Ltd., UK. The 15-µl reaction mixture contained 40 ng of genomic DNA, 10 pmol of random primer, 1× reaction buffer (containing 1.5 mM $MgCl_2$), 0.6 U of Taq DNA Polymerase (Genecraft, Münster, G) and 7 nmol of dNTPs. The amplification products were loaded either on a 1.5% agarose gel and stained with ethidium bromide, or on a 10% native acrylamide gel (Amersham-Pharmacia-Biotech, Hoefer Se 600, Uppsala, S), with running conditions for 1.30 h at 85 W constant at 15°C, and then silver-stained (Stift, unpublished). For microsatellite (SSR) amplification 561 soybean primer pairs were used. They were developed at the USDA-ARS, Soybean and Alfalfa Research Laboratory, Beltsville, Md. 20705, USA, (Cregan et al. 1999). The PCR reaction and cycling were performed following Akkaya et al. (1995) and Cregan et al. (1999). The amplification products were loaded on a 12% polyacrylamide gel (Amersham-Pharmacia-Hoefer, Hoefer Se 600, Uppsala, S), run for 2 h at 85 W constant and 15°C, followed by silver staining (Fig. 1). After staining, gels were dried using the model Drygel Sr, Slabgel Dryer SE 1160 (Amersham-Pharmacia-Biotech, Uppsala, S).

Statistical analysis

MAPMAKER/EXP 3.0 (Lander et al. 1987; Lincoln et al. 1992a) was used to construct the genetic linkage map. We used the two-point commands “group”, “map” and “lod table”, and the multipoint analysis “compare” and “pairwise” commands, for building the linkage groups. New markers were added by using the “try” multipoint analysis command and the results were checked by the “ripple” command. The error detection ratio was set at 5%. The Haldane mapping function was used with a default LOD score of 3.0 and a 50 cM maximal distance. Dashed lines along the bars of the linkage groups mark the regions where genetic distance between markers exceeded 50 cM (see Fig. 2). The linkage groups were designated LG em, with “em” referring to “early maturity.” The corresponding linkage group names of the USDA/Iowa State University map (Cregan et al. 1999) are also indicated in Fig. 2.

Single marker-QTL associations were identified by one-way ANOVA (SAS/STAT software, SAS Institute 1988) using phenotypic data from single environments and from all the environments

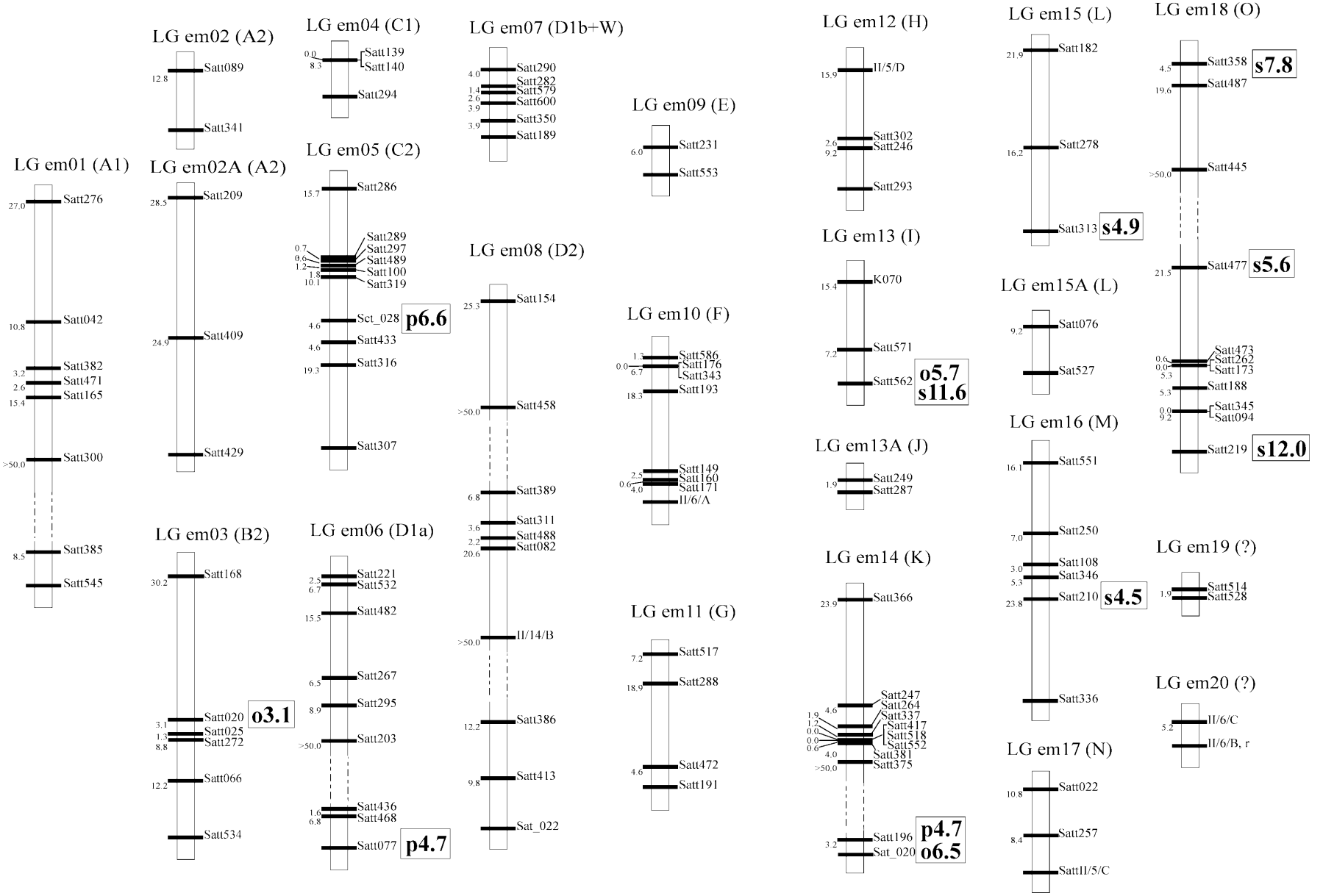
LG em01 (A1)
27.0 Satt276
10.8 Satt042
3.2 Satt382
2.6 Satt471
15.4 Satt165
>50.0 Satt300
8.5 Satt385
Satt545
LG em02 (A2)
12.8 Satt089
Satt341
LG em02A (A2)
28.5 Satt209
24.9 Satt409
Satt429
LG em03 (B2)
30.2 Satt168
3.1 Satt020 o3.1
1.3 Satt025
8.8 Satt272
12.2 Satt066
Satt534
LG em04 (C1)
0.0 Satt139
8.3 Satt140
Satt294
LG em05 (C2)
15.7 Satt286
0.7 Satt289
0.6 Satt297
1.2 Satt489
1.8 Satt100
10.1 Satt319
4.6 Sct_028 p6.6
4.6 Satt433
19.3 Satt316
Satt307
LG em06 (D1a)
2.5 Satt221
6.7 Satt532
15.5 Satt482
6.5 Satt267
8.9 Satt295
>50.0 Satt203
1.6 Satt436
6.8 Satt468
Satt077 p4.7
LG em07 (D1b+W)
4.0 Satt290
1.4 Satt282
2.6 Satt579
3.9 Satt600
3.9 Satt350
Satt189
LG em08 (D2)
25.3 Satt154
>50.0 Satt458
6.8 Satt389
3.6 Satt311
2.2 Satt488
20.6 Satt082
>50.0 II/14/B
12.2 Satt386
9.8 Satt413
Sat_022
LG em09 (E)
6.0 Satt231
Satt553
LG em10 (F)
1.3 Satt586
0.0 Satt176
6.7 Satt343
18.3 Satt193
2.5 Satt149
0.6 Satt160
4.0 Satt171
II/6/A
LG em11 (G)
7.2 Satt517
18.9 Satt288
4.6 Satt472
Satt191
LG em12 (H)
15.9 II/5/D
2.6 Satt302
9.2 Satt246
Satt293
LG em13 (I)
15.4 K070
7.2 Satt571
Satt562 o5.7 s11.6
LG em13A (J)
1.9 Satt249
Satt287
LG em14 (K)
23.9 Satt366
1.9 Satt247
4.6 Satt264
1.2 Satt337
0.0 Satt417
0.0 Satt518
0.6 Satt552
4.0 Satt381
>50.0 Satt375
3.2 Satt196 p4.7
Sat_020 o6.5
LG em15 (L)
21.9 Satt182
16.2 Satt278
Satt313 s4.9
LG em15A (L)
9.2 Satt076
Sat527
LG em16 (M)
16.1 Satt551
7.0 Satt250
3.0 Satt108
5.3 Satt346
23.8 Satt210 s4.5
Satt336
LG em17 (N)
10.8 Satt022
8.4 Satt257
SattII/5/C
LG em18 (O)
4.5 Satt358 s7.8
19.6 Satt487
>50.0 Satt445
21.5 Satt477 s5.6
0.6 Satt473
0.0 Satt262
5.3 Satt173
5.3 Satt188
0.0 Satt345
9.2 Satt094
Satt219 s12.0
LG em19 (?)
1.9 Satt514
Satt528
LG em20 (?)
5.2 II/6/C
II/6/B, r

Table 2 Summary of phenotypic data

Character	Ma.Belle	Proto	Progeny mean	LSD (0.05)	h^2 (%)
Protein content (g kg^{-1})	402	423	417	15	58.4
Oil content (g kg^{-1})	199	185	189	13	59.2
1000 seed weight (g)	215	191	208	17	83.0

combined. Associations between markers and phenotypic characters were evaluated by *F*-tests. QTLs were only considered as significant at a level of $P<0.01$ in the combined analysis across environments. In the combined analysis across all environments, in order to remove environmental effects, data from each environment were standardized. This allowed an accurate estimation of the amount of variation explained by each marker locus (R^2).

For simple interval mapping using MAPMAKER/QTL 1.1 (Paterson et al. 1988; Lincoln et al. 1992b), the significance level was set at LOD 2.0 (default) and each 2-cM interval (default) was checked for the presence of QTLs. The LOD score peak was used to estimate the most-likely QTL position on the linkage map. QTL regions were also verified across environments by composite interval mapping using the multiple regression-based software program PLABQTL at a minimum LOD score level of 2.5 (Utz and Melchinger 1996). Markers of a linkage group were selected as cofactors for analysis using the cov SELECT command.

Results

The map and its characteristics

The genetic map based on the segregating population of the two early ripening genotypes Ma.Belle and Proto consists of 120 different molecular markers (Fig. 2). The core map is made up almost exclusively of SSR markers.

From all the marker types used in this study, SSRs displayed the highest polymorphism ratio, 160 out of 546 (i.e. 29.3%). From these 160, 113 were mapped on the F_2 population, and 36 were unlinked. From the remaining 11, six (Satt023, Satt088, Satt326, Satt334, Satt410 and Satt425) showed only the band of one of the two parents in the F_2. In five cases (Satt291, Satt309, Satt324, Satt506 and Satt563) the gels of the F_2 population could not be evaluated.

Out of 100 RFLP markers, only 15 (15%) were polymorphic and only one (K070) could be integrated into the map. We also tested five RAPD primers using polyacrylamide for fragment separation and obtained 20 polymorphic bands. The location of six RAPD loci (II/5/C, II/5/D, II/6/A, II/6/B, II/6/C and II/14/B) was determined. Apart from the six SSR primers with only one band in the F_2, as mentioned above, no segregation distortion of SSR and RFLP markers occurred. The dominantly inherited RAPD showed a 1:3.5 recessive/dominant segregation. The complete map spans 757.5 cM (6.3 cM/marker density) and falls into 23 linkage groups.

◀ **Fig. 2** Linkage map of the early maturing soybean F_2-population (n=82) of the cross Ma.Belle×Proto consisting of 23 linkage groups integrating 113 SSRs, 6 RAPDs and 1 RFLP marker. The designation of linkage groups from the USDA/Iowa State University integrated linkage map is given in parenthesis. Effects of particular QTL regions on seed protein content (*p*), oil content (*o*) and seed weight (*s*) are indicated on the map, followed by R^2-values from analysis of variance results combined across all environments

QTLs for seed protein and oil content, and for seed-weight

The summary of seed protein data is presented in Table 2. Using a one-way ANOVA, 15 markers (14 SSRs and 1 RAPD) were identified showing significant ($P<0.01$) association with seed protein content in at least one environment on LG em05 (Satt319, Sct_028, Satt433, Satt316, Satt307), LG em06 (Satt468, Satt077), LG em12 (II/5/D, Satt246, Satt293), LG em14 (Satt196, Sat_020), LG em11 (Satt472), LG em18 (Satt477) and Satt567 (unlinked). From these 15 markers, only four (Satt77, Sct_028, Satt567, Satt196) have shown a significant association with protein content across all environments (Table 3). Jointly, they explain about 23% of the overall variation.

Using the phenotypic data of the different environments separately and simple interval mapping, three main intervals were detected for seed protein content on LG em04, LG em05 and LG em13. The interval identified on LG em05 (Satt319-Satt307) also covers all the markers identified by ANOVA. Combining protein data from all environments and using composite interval mapping verified one QTL region close to marker Sct_028 on LG em05 at a LOD-value of 2.7 (Table 3).

The summary statistics of the seed oil content of parents and progeny are listed in Table 2. One-way ANOVA detected SSR markers significantly ($P<0.01$) linked to seed oil content in at least one environment and to each other in linkage groups LG em03 (Satt272 and Satt020), LG em14 (Satt196, Sat_020), LG em18 (Satt358, Satt487, Satt445, Satt473, Satt262, Satt173, Satt188, Satt345, Satt094, Satt219), LG em13 (Satt562), and LG em04 (Satt139). From all these markers, only Satt020, Satt196 and Satt562, proved to be significant across environments (Table 3), explaining about 15% of the variation.

Simple interval mapping using single environments only, confirmed the presence of a QTL in LG em14 (Satt196, Sat_020) and identified two additional seed oil QTL-intervals in LG em05 (Satt319, Sct_028, Satt433, Satt316, Satt307) and in LG em13 (Satt571, Satt562). The average LOD of these QTLs is 2.78 (the threshold is 2.0). However, none of these regions proved to be significant in a combined analysis across all environments using composite interval mapping.

Table 3 Soybean microsatellite markers and marker-associated effects on seed protein and oil content and on 1000-seed weight in F_2-progeny of the cross Ma.Belle×Proto (analysis of variance results combined across five environments). * Linkage group designations of the USDA / Iowa State University map in parenthesis. ** LOD-score from composite interval mapping across all environments

Marker	Linkage-group	Sig. level (*F*-test)	R^2 (%)	LOD-score**	Allelic means		
					Ma.Belle	Heterozygous	Proto
					Protein content (g kg^{-1})		
Satt077	6 (D1a)*	0.0009	4.7	–	417.4	419.1	413.0
Sct_028	5 (C2)	0.0069	6.6	2.7	414.6	417.0	418.7
Satt567	Unlinked (M)	0.0020	7.1	–	417.0	418.4	413.1
Satt196	14 (K)	0.0009	4.7	–	416.9	418.8	413.7
					Oil content (g kg^{-1})		
Satt020	3 (B2)	0.0013	3.1	–	206.9	209.4	204.9
Satt196	14 (K)	0.0004	6.5	–	188.5	186.1	191.7
Satt562	13 (I)	0.0040	5.7	–	184.6	189.1	190.1
					1000-seed weight (g)		
Satt210	16 (M)	<0.0001	4.5	–	202.6	208.3	213.1
Satt219	18 (O)	<0.0001	12.0	4.3	214.0	207.5	196.2
Satt229	Unlinked (L)	<0.0001	4.8	–	212.8	207.8	200.2
Satt306	Unlinked (M)	<0.0001	6.8	–	212.7	203.8	191.3
Satt313	15 (L)	<0.0001	4.9	–	203.7	212.9	211.7
Satt358	18 (O)	<0.0001	7.8	–	215.8	205.9	202.4
Satt477	18 (O)	<0.0001	5.6	2.6	209.3	210.7	199.8
Satt562	13 (I)	<0.0001	11.6	–	218.0	208.9	200.4

In the investigated population, the phenotypic data of seed protein and oil content were negatively correlated (r=–0.77, P<0.0001). The negative correlation between the two characters was also reflected in their association with molecular markers. In the QTL region of marker Satt196 on LG em14, for instance, the allele from Ma.Belle codes for high protein and low oil content, whereas the allele from Proto is associated with low protein and high oil content (Table 3).

Similarly, in the QTL-region of marker Sct_028 on LG em05, the Proto allele codes for high, and the Ma.Belle allele for low protein content (Table 3), whereas for oil content the respective marker alleles reveal opposite effects. The data on oil content were, however, not significant across all environments and therefore were not included in Table 3. In the region of marker Sct_028 on LG em05, the LOD-score plots can not separate the two QTLs for seed protein and oil content, respectively, indicating a close linkage of the two loci (Fig. 3).

A large number of markers displayed association with seed weight at a high significance level (P<0.0001) on LG em10 (Satt343, Satt193, Satt149, Satt160), LG em16 (Satt250, Satt108, Satt346, Satt210, Satt336), LG em18 (Satt358, Satt487, Satt477, Satt473, Satt262, Satt173, Satt188, Satt345, Satt094, Satt219, Satt243), LG em06 (Satt482), LG em08 (Satt413), LG em11 (Satt288), LG em13 (Satt562), and LG em15 (Satt313); two more markers (Satt229 and Satt306) were unlinked. Out of this large number of markers, eight showed very strong association with seed weight across all environments, explaining more than 55% of the total variation observed (Table 3).

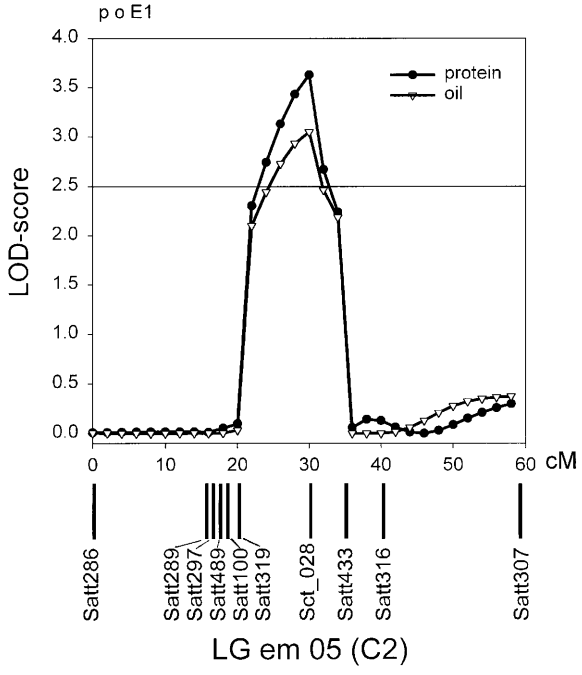

Fig. 3 LOD score plots from composite interval mapping of QTL regions for both seed protein and oil content identified near the position of marker Sct_028 on LG em 05 (data from the Raasdorf 1997 trial)

Simple interval mapping confirmed the presence of a QTL on LG em10 between Satt149 and Satt343, on LG em13 between K070 and Satt562, and on LG em18 between Satt173 and Satt219. Combining seed weight data from all environments and using composite interval

mapping, two significant QTL regions could be verified close to marker Satt219 (interval Satt094-Satt219; LOD=4.3) and near Satt445 (interval Satt445-Satt477; LOD=2.6) on LG em18 (Table 3).

Discussion

The main goal of the present study was to identify QTLs for seed quality characteristics in a segregating soybean population produced by the crossing of two early maturing genotypes. A molecular genetic map was constructed using 1 RFLP, 6 RAPD and 113 SSR markers. Due to the narrow range of adaptation of the early maturing varieties, we expected a low level of polymorphism. In our mapping population, 160 out of the 546 tested SSR primer pairs were polymorphic (29.3%), which is comparable to data reported by Arahana et al. (1999) who detected 16–30% polymorphism depending on the population. The six markers which were found to be monomorphic in the F_2 population, indicate heterogeneity of the two parents. The DNA used in testing for parental polymorphism did not originate from the same plants which were used for crossing.

By comparison, from the 100 RFLP markers tested, only 15% were polymorphic in our population. These results underline the usefulness of SSR markers for soybean. Another major advantage of this set of markers is that they all reportedly map to a single locus (Cregan et al. 1999) which is supported by the present study (Fig. 1). This makes SSRs ideal for integrating genetic linkage maps constructed by different laboratories using different mapping populations. Integrated maps may contain information from both classical and molecular genetic maps (Shoemaker and Specht 1995; Cregan et al. 1999).

Mapping 113 SSR loci to 23 linkage groups did not reveal any major discrepancy from the maps published by other groups (USDA/Iowa State University, University of Utah, University of Nebraska; Cregan et al. 1999). Some linkage groups are not represented in our map (e.g. LG B1 or LG P, USDA/Iowa) and we identified two LGs (LG em19, LG em20) which do not correspond to any of the LGs in the other maps. The two SSR markers making up LG em19 are well integrated into LG D2 of the USDA/Iowa map which corresponds to our LG em08. They should occur in the region between the markers Satt389 and Satt082 (Fig. 2). Such discrepancies could probably be eliminated by further saturation of our core map with markers such as RAPDs, AFLPs or ISSRs.

Based on data from single environments, 45 markers linked to seed quality QTLs were found. These are localized on ten main linkage groups. The low density of the map did not allow efficient use of interval mapping. When one-way ANOVA was used across all environments, the number of useful markers was reduced to 15 (Table 3). Thereby, the selection of markers was made based on the highest level of significance over environments.

In comparison with previously published data, four out of our 15 seed quality character QTLs (one on LG em05 and em13, and two on em14) were localized to regions in which other authors also reported similar QTLs. For example, in linkage group LG em05, marker Sct_028 is a protein marker of intermediate strength (P=0.0069). Brummer et al. (1997) found RFLP marker A538 to be associated with a seed protein content QTL in a 3-year experiment at P=0.01. This marker, in the University of Nebraska map (Cregan et al. 1999; LG C2), is only 1.3 cM away from Sct_028.

In linkage group LG em14, marker Satt196 was strongly associated with both protein and oil content (P=0.0009 and P=0.0004, respectively). Brummer et al. (1997) found RFLP marker K387 associated with oil content at the 3-year average level of P=0.002. K387 is 6.9 cM from Satt196 on the USDA/Iowa map (Cregan et al. 1999; LG K).

In linkage group LG em15, marker Satt313 was very strongly associated with seed weight (P<0.0001). In the same linkage group (L in the USDA/Iowa map), Mian et al. (1996) also found a very strong marker for seed weight. In their map, this marker is 10 cM from the determinate growth habit locus (*Dt1*). *Dt1* is not integrated into the USDA/Iowa linkage map. In the University of Utah linkage map, Satt313 is more than 50 cM apart from *Dt1*. In the same linkage group, Orf et al. (1999) found marker Satt527 to be associated with seed weight at LOD=3.1. Satt527 is approximately 20 cM apart from *Dt1* on the University of Utah map and therefore could be linked to the same seed weight QTL as found by Mian et al. (1996). Satt527 is approximately 25 cM further from Satt313 in the USDA/Iowa map. In LG em15 we found only three polymorphic markers. Satt527 was arranged in LG em15A, obviously because of the long distance between Satt313 and Satt076. Most probably our marker Satt313 is linked to a different seed weight QTL than the ones described by the two other authors and therefore is specific to our population.

In LG em 13, Satt562 was associated with QTLs influencing oil content and seed weight (Table 3). This marker is located in a region of chromosome I of the USDA/Iowa map, in which QTLs for protein content, oil content, seed weight and other agronomic characters have recently been reported by Sebolt et al. (2000), who introgressed QTL alleles from the wild species *G. soja* into different *G. max* backgrounds.

In linkage group LG em18, Satt219, Satt358 and Satt477 were found to be very strongly linked to seed weight (P<0.0001). In none of the previously mentioned publications was a seed weight QTL described in this linkage group (LG O of the USDA/Iowa map). Therefore, these QTLs appear to be specific for early maturing soybean varieties.

The negative correlation between protein and oil content has repeatedly been verified by molecular markers, i.e. QTL alleles coding for high protein content were also associated with low oil content and *vice versa* (Lark et al. 1994; Lee et al. 1996; Sebolt et al. 2000). A close

linkage between QTLs influencing protein and oil content is also suggested in this study by the mapping of LG em05 (Fig. 3). A much higher map density and a larger plant population would be necessary to resolve whether those protein and oil QTLs are controlled by the same gene or by different genes closely linked in repulsion phase.

Environmental effects on QTLs are illustrated in the case of the association of Sct_028 and protein content. This specific association was highly significant in two (P<0.0001), significant in one (P<0.1) and not significant in the remaining two environments. The calculated average significance level over the five environments was P=0.0069 (Table 3). At the same time, the combined analysis for Sct_028 showed a significant QTL by environment interaction (P<0.001). Brummer et al. (1997) published similar results for the same QTL. Comparing these protein data with those found for seed weight, which showed a stable and high significance across all environments, illustrates QTLs with a different degree of environmental interaction. The results also reflect the different heritability values obtained for these two traits (Table 2). Therefore, at this point, while in our population seed weight could easily be selected for using molecular markers, the efficient use of markers to select for protein or oil content would require additional mapping efforts.

Acknowledgements This research was financially supported by the Austrian Science Fund (FWF, Project No. P10743-ÖBI and P10663-ÖBI). We are especially grateful to Dr. Perry B. Cregan for providing all the SSR primers used. The experiments carried out in this work comply with the current laws of Austria.

References

Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci 35:1439–1445

Arahana VS, Graef GL, Specht JE, Steadman JR (1999) Identification of QTLs for *Sclerotina* stem rot resistance in soybean using SSR markers. Plant and Animal Genome Conf VII., Jan 17–21, San Diego, California, Posterabstract P241

Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker R (1997) Mapping QTLs for seed protein and oil content in eight soybean populations. Crop Sci 37:370–378

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Diers BW, Fehr W, Keim P, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. Theor Appl Genet 83:608–612

Doldi ML, Vollmann J, Lelley T (1997) Genetic diversity in soybean as determined by RAPD and microsatellite analysis. Plant Breed 116:331–335

Keim P, Olson TC, Shoemaker RC (1988a) A rapid protocol for isolating soybean DNA. Soybean Genet Newslett 15:150–152

Keim P, Groose R, Shoemaker RC (1988b) Construction of a random recombinant DNA library that is primarily single copy sequences. Soybean Genet Newslett 15:148–150

Kiss GB, Csanádi G, Kálmán K, Kaló P, Ökrész L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol Gen Genet 238:129–137

Lande R, Thompson R (1990) Efficiency of marker-aided selection in the improvement of quantitative traits. Genetics 124: 743–756

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lark KG, Weisemann JM, Matthews BF, Palmer R, Chase K, Macalma T (1993) A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir 1.' Theor Appl Genet 86:901–906

Lark KG, Orf J, Mansur LM (1994) Epistatic expression of quantitative trait loci (QTLs) in soybean [*Glycine max* (L.) Merr.] determined by QTL association with RFLP alleles. Theor Appl Genet 88:486–489

Lee SH, Bailey MA, Mian MAR, Carter TE, Shipe ER, Ashley DA, Parrott WA, Hussey RS, Boerma HR (1996) RFLP loci associated with soybean seed protein and oil content across populations and locations. Theor Appl Genet 93:649–657

Lincoln S, Daly M, Lander E (1992a) Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, 3rd edn

Lincoln S, Daly M, Lander E (1992b) Mapping genes controlling quantitative traits with MAPMAKER/QTL 1.1. Whitehead Institute Technical report. 2nd edn

Mansur LM, Lark KG, Kross H, Oliveira A (1993) Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet 86:907–913

Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci 36:1327–1336

Mian MAR, Bailey MA, Tamulonis JP, Shipe ER, Carter TE, Parrott WA, Ashley DA, Hussey RS, Boerma HR (1996) Molecular markers associated with seed weight in two soybean populations. Theor Appl Genet 93:1011–1016

Muehlbauer GJ, Staswick PE, Specht JE, Graef GL, Shoemaker RC, Keim P (1991) RFLP mapping using near-isogenic lines in the soybean (*Glycine max* L. Merr.). Theor Appl Genet 81:189–198

Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, Lark KG (1999) Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. Crop Sci 39:1642–1651

Paterson A, Lander E, Lincoln S, Hewitt J, Peterson S, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors using a complete RFLP linkage map. Nature 335:721–726

Paterson AH, Tanksley SD, Sorrells ME (1991) DNA markers in plant improvement. Adv Agron 46:39–89

Qiu BX, Arelli PR, Sleper DA (1999) RFLP markers associated with soybean cyst nematode resistance and seed composition in a "Peking"×"Essex" population. Theor Appl Genet 98:356–364

SAS Institute (1988) SAS/STAT user's guide. Release 6.03 ed. SAS Inst., Inc., Cary, NC

Sebolt AM, Shoemaker RC Diers BW (2000) Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. Crop Sci 40:1438–1444

Shoemaker RC, Olson TC (1993) Molecular linkage map of soybean (*Glycine max* L. Merr.). In: O'Brien SJ (ed) Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press, New York, pp 6131–6138

Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. Crop Sci 36: 436–446

Smith KJ, Huyser W (1987) World distribution and significance of soybean. In: Wilcox JR (ed) Soybeans: improvement, production and uses, 2nd edn. Agronomy 16:23–48

Utz HF (1995) PLABSTAT – plant breeding statistical program. Version 2M. Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, Universität Hohenheim, Stuttgart, Germany

Utz HF Melchinger AE (1996) PLABQTL – a computer program to map QTL. Version 1.1. Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, Universität Hohenheim, Stuttgart, Germany

Vollmann J, El Hadad T, Gretzmacher R, Ruckenbauer P (1996) Seed protein content of soybean as affected by spatial variation in field experiments. Plant Breed 115:501–507

Vollmann J, Fritz CN, Wagentristl H, Ruckenbauer P (2000) Environmental and genetic variation of soybean seed protein content under Central European growing conditions. J Sci Food Agric 80:1300–1306