ELSEVIER

Plant Science 152 (2000) 181–189

PLANT SCIENCE

www.elsevier.com/locate/plantsci

Genetics of seed quality in soybean analysed by capillary gel electrophoresis

Sophie Gerber *, Françoise Fabre, Claude Planchon

INP/ENSAT, Laboratoire de Biotechnologie et d'Amélioration des Plantes, BP 107, Auzeville Tolosane, 31326 Castanet-Tolosan Cedex, France

Received 3 June 1999; received in revised form 9 November 1999; accepted 9 November 1999

Abstract

Soybean recombinant inbred lines resulting from a cross between two lines of different genetic origins, each with a high protein content, were analyzed for their seed protein composition and for different quantitative traits involved in the establishment of seed quality, including N_2 fixation. Capillary gel electrophoresis (CGE) was used to separate and quantify the seed protein components according to their molecular weights. The relative peak surfaces were compared to the quantitative traits. Several significant associations were found between seed or symbiotic fixation traits and peak surfaces. Two peaks are inversely associated with seed yield and seed protein content. Another peak is related to symbiotic dinitrogen fixation parameters. The proteins constituting these peaks could be interesting candidates to further study seed quality establishment.

Keywords: Soybean; Seed quality; Seed protein; Capillary gel electrophoresis

1. Introduction

Soybean is primarily produced for its proteins (it is the first source of vegetable proteins in the world) and oil, for animal feeds and human foods. The main goals of soybean breeding programs are to increase the content and the quality of the protein and oil [1]. One of the main drawback regarding these goals are the negative correlations between protein content on one hand, oil content and yield on the other hand. Nonetheless, improving simultaneously both traits, or at least maintaining one when the other is improved, seems possible [2]. Soybean plants can acquire a substantial portion of their nitrogen requirements from N_2 fixation, and seed yield is dependent on the amount of nitrogen in the seed originating from symbiotic fixation [3]. The nitrogen source supplied influences the protein composition and the protein concentration of soybean seeds [4]. Increasing N_2 fixation efficiency could be a means to achieve a simultaneous improvement in yield and protein content [5].

The objective of the present work was to try to reach a better understanding of the genetics of seed quality in soybean. The study focused on the main physiological traits of the plants, including N_2 fixation, which is related to the final objective, i.e. seed quality. Focusing on seed protein composition, capillary gel electrophoresis (CGE) of the total seed proteins was used to describe it more precisely. Apart from studies to evaluate polymorphism of some specific proteins [6] CGE of proteins, to our knowledge, has rarely been used for genetic analysis. This technique allows the separation of the protein components according to their molecular weight. It offers the rare opportunity to quantify each of these components. By this way, we have access to an aspect of the regulation processes, which could better explain the quantitative trait variation observed between individuals:

* Present address. INRA, Laboratoire de génétique et amélioration des arbres forestiers, BP 45, 33611 Gazinet Cedex; Tel.: +33-5-5797-9081; fax: +33-5-5797-9088.

E-mail address: gerber@pierroton.inra.fr (S. Gerber)

0168-9452/00/$ - see front matter
PII: S0168-9452(99)00233-2

polymorphism in gene regulation is thought to be an important basis for phenotypic changes [7]. Moreover, the study of proteins, which collaborate to the building of the phenotype, may help to understand it better, even if they are still anonymous. They may be valuable tools to dissect quantitative traits, compared to anonymous DNA markers with no phenotypic meaning. Such an example can already be found in soybean where primers of known gene sequences revealed polymorphism associated with quantitative traits [8]. Relationship between protein quantity variation and agronomical traits was demonstrated in some species; maize [9] and pine [10]. In the present work, the recombinant inbred lines resulting from the cross between two high-protein content soybean cultivars of different genetic origins were analyzed.

2. Materials and methods

2.1. Plant material

The two parental lines were selected after a diallel analysis that had shown the opportunity for yield and protein content to be improved by their cross [11]. The 111 recombinant inbred lines (F6 generation) derived by single seed descent from the cross between Provar (female, maturity group II, American origin) and X514-95 (maturity group O-I, French origin) were evaluated in the field for their earliness and seed traits. The F7 generation was evaluated in the growth chamber for their N_2 fixation ability.

2.2. Field experiment conditions and measurements

The experiment was carried out at the INRA (Institut National de la Recherche Agronomique) station in Toulouse, France (43°36′ N, 1°27′ E) on a fine mixed Udic Haplustalf soil (pH 6.4, organic matter 2% and total N 1.5 g kg^{-1}). The F6 plants were grown in 1.5 m rows spaced 50 cm apart with 5 cm intervals within the row. The soil was inoculated with the G49 strain of *Bradyrhizobium japonicum* (Lipha, Lyon, France). The seeds were sown on May 6, 1996. Soil moisture was maintained by irrigation.

All traits were characterized at maturity, including yield components: earliness: total cycle duration to maturity expressed as 'heat units' [12]; seed yield: weight of seeds at 0% relative humidity harvested from the whole row (g row^{-1}); seed number per row; seed size: weight of 1000 seeds (g).

Seed protein content (%) was evaluated by the Kjeldahl method (Tecator Kjeltec, Auto 1030, Hänagäs, Sweden) and seed protein yield (g row^{-1}) was also determined [13].

The lipoxygenase 1 and lipoxygenase 2 + 3 activities were measured on seeds by polarographic method using a Hansatech DWI Oxygraph (Norfolk, UK) as described by Zougari et al [14]. Total lipoxygenase activity was calculated as the sum of lipoxygenase 1 and 2 + 3 activities. One lipoxygenase activity unit is defined as the uptake of one micromole oxygen per min and results are expressed in units per gram of seed (U g $seed^{-1}$). Only 35 of the 111 recombinant inbred lines were analyzed for lipoxygenase activity.

2.3. Growth chamber experiment conditions and measurements

The study was conducted under controlled conditions: 14 h a day photoperiod, about 400 μE m^{-2} s^{-1} irradiance, 25/20°C light/dark temperatures and 70% relative humidity.

The F7 plants were grown in pots (of diameter 18 cm and 16 cm height) filled with sand (0.5/5 mm grain size). Seeds were germinated in vermiculite for 3 days at 25°C and seedlings (one per pot) were planted on sand. Two inoculations were made with G49 strain of *B. japonicum* (Lipha) on seeds in germination (10^6 bacteria per seed), and by watering the plants with an inoculum suspension (10^6 bacteria per liter) after planting out. Plants were fertilized twice a day by soaking the pots during 30 min in the IRAT solution (Institut de Recherche en Agronomie Tropicale, France) supplied in magnesium [15] and containing 3.2 mM $Ca(NO_3)_2$.

Dinitrogen fixation was estimated by measuring the acetylene reduction activity (ARA) using an in situ method [16]. The pot was tightly sealed and acetylene was injected into the soil around the root system; the acetylene volume amounted to 10% of total porosity of the mixture in pot. After a 7 min incubation, samples were removed to determine ethylene concentration by gas chromatography (Delsi Model DI 200, Paris, France). The measure-

ments were carried out at reproductive stages R5 and R6 [17] and the results are expressed in μmol C_2H_4 h^{-1} $plant^{-1}$. The integrated ARA (mmol C_2H_4 $plant^{-1}$) was calculated according to Patterson and Larue [18] through the period between the two stages. Nodule number and dry weight were determined per plant at R6 stage after the last ARA measurement. Specific activities (ARAs or integrated ARAs) were also determinated and expressed in nmol C_2H_4 h^{-1} mg $nodule^{-1}$ and in μmol C_2H_4 mg $nodule^{-1}$.

Eventually, 17 traits corresponding to two or three replicates per genotype were available on the 111 recombinant inbred lines or on part of them.

2.4. Capillary electrophoresis

2.4.1. Protein extraction

Protein analysis was made on F6 seeds harvested during the field experiment described above. Seed coats were removed from five mature seeds randomly sampled for each genotype. They were ground in a ball mill during 1.5 min. Proteins were extracted from the fine powder obtained using a modified method by Anderson et al [19]. The powder was crushed in a microfuge tube with 20 μl mg^{-1} of extraction buffer. The extraction mixture contained 3 M urea, 2% ampholytes (pH 3–10), 5% dithiothreitol and 4% FSN 100 (a fluorocarbon-based detergent supplied by duPont). Proteins were allowed to extract for 1 h at room temperature. The tube was vortexed briefly and then centrifuged for 5 min at $10\,000 \times g$. The clear part of the extract was removed with a micropipette and frozen at -80°C until it was analyzed. The extracts contained ≈ 12.5 μg of protein per μl.

The two parental lines were sampled eight times. Each of the 111 recombinant inbred lines was sampled two to three times.

2.4.2. Electrophoresis conditions

Each protein extract was analyzed using the eCAP SDS 14-200 kit supplied with the P/ACE™ Instrument by Beckman.

The following ingredients were mixed in a microfuge tube: 5 μl protein extract, 25 μl sample buffer and 2.5 μl orange G reference marker (both supplied in the kit), 1.25 μl 2-mercaptoethanol, 21.25 μl deionized water. The mix was shortly vortexed and boiled for 10 min at 98°C. 45 μl of the protein sample mix was placed in a 400 μl vial and put on a top of a spring into a 4 ml vial, all these elements are provided by the manufacturer.

CGE itself was performed in a 27 cm coated capillary filled with SDS 14-200 Gel Buffer, according to the instructions given by the manufacturer (eCAP SDS 14-200 kit-P/ACE™. Capillary electrophoresis size separation of SDS proteins, Beckman, 1994). The vial filled with gel was not sonicated as recommended but placed in vacuum during the 10 min to remove air bubbles. It was replaced every 100 analysis. The UV detector was set at 214 nm, the first rinse, with HCl, lasted 1 min, the second, with Gel Buffer, lasted 3 min, and the sample injection time lasted 30 s. The separation was performed at 8.1 kV and 20°C and lasted 16 min. Control samples were run, but the extraction mix contained water instead of protein sample.

Overall, 253 separations were available and integrated for individual and total peak surfaces using the software provided by the manufacturer (System Gold®).

2.4.3. Corrected correlation

Linear correlation between individual peak surfaces and the different traits were measured. To take the residual variance into account, a corrected correlation was computed according to the following model.

The peak surface Y_{ik} of the kth repetition of the ith genotype can be written:

$$Y_{ik} = g_{Y,i} + e_{Y,ik}$$

where $g_{Y,i}$ stands for the effect of the genotype on the peak suface and $e_{Y,ik}$ for the residual.

Similarly, the trait X_{ik} of the kth repetition of the ith genotype can be written:

$$X_{ik} = g_{X,i} + e_{X,ik}$$

where $g_{X,i}$ stands for the effect of the genotype on the trait and $e_{X,ik}$ for the residual.

We are interested in computing the correlation:

$$\mathrm{cor}^2(X,Y) = \frac{\mathrm{cov}^2(G_X,G_Y)}{\mathrm{var}(G_X)\mathrm{var}(G_Y)}$$

which can be estimated by the expression:

$$\mathrm{cor}^2(X,Y) = \frac{\mathrm{cov}^2(X\cdot,Y\cdot)}{(\mathrm{var}(X\cdot)-\mathrm{var}(E_{X,}\cdot))(\mathrm{var}(Y\cdot)-\mathrm{var}(E_{Y,}\cdot))} \quad (1)$$

where $X\cdot$, $Y\cdot$, $E_{X,}\cdot$ and $E_{Y,}\cdot$ stand for the means over repetitions.

With two repetitions:

$$\mathrm{var}(E_{Y,}\cdot) = \frac{\sigma_e^2}{2} \quad \text{where} \quad \hat{\sigma}_e^2 = \frac{\sum_{i=1}^{n} (Y_{i1} - Y_{i2})^2}{2n} \qquad (2)$$

If $K_{(i)}$ repetitions are available per genotype i, then:

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^{n} \sum_{k=1}^{K_{(i)}} (Y_{ik} - Y_i\cdot)^2}{\sum_{i=1}^{n} (K_{(i)} - 1)} \qquad (3)$$

Similar Eq. (2) and Eq. (3) are written for X_i.

2.4.4. False positive

When a large number of statistical tests are carried out, Type I errors are likely to occur. Let T be the total number of tests computed, and α the significance level. The expected number of test corresponding to Type I errors follows a binomial law with parameters T and α. Consequently, the probability $P(K)$ of finding, by chance only, at least K tests significant at the α level among the T tests can be written:

$$P(K) = 1 - \sum_{i=0}^{K-1} \binom{T}{i} \alpha^i (1-\alpha)^{T-i} \qquad (4)$$

3. Results and discussion

After CGE, each sample gave a profile similar to the one presented in Fig. 1. Around 18 peaks were resolved, each described by its migration time in minutes, corresponding to protein products with increasing molecular weight. The surface of each peak was integrated. According to migration time of reference markers with known molecular weight, approximate molecular weights of each peak are calculated by the software and given with the migration time in Fig. 1. The last peak of the profile is likely to have an overestimated molecular weight according to its extreme position. Peak 'a' in Fig. 1 (appearing after 3 min of migration) is an artifact associated with the sample loading. The first highest peak, identified by 'o' in Fig. 1 (appearing after about 6.7 min of migration), coincides with the orange G reference marker added to each sample. As peak 1 was sometimes present in the control sample, we discarded it. Peaks 2, 3, 4, 5 and 6 were clearly resolved throughout all electrophoresis, their surfaces were recorded for each genotype and each repetition. The individual peak surfaces were divided by the total peak surface computed by the software minus the surface of peaks 'a' and 'o'. The mean of the relative peak surfaces was calculated over repetitions for each line.

Peak 2 was present in the X514-95 line but absent in the Provar line and exhibited Mendelian segregation (presence/absence) among the recombinant inbred lines. The others peaks were present in both parental lines. It can be noted that the relative surfaces of peaks 1, 3 and 6 were significantly different between Provar and W514-95 (data not shown). The relative surfaces of all the peaks exhibited an approximately normal distribution (data not shown), indicating that no major gene could explain the segregation observed among the recombinant inbred lines.

The correlations between the relative surfaces of the 5 peaks and the 17 quantitative traits measured on the recombinant inbred lines were computed. For peak 2, a student test was used to compare the mean of the lines having the peak, to the mean of those lacking the peak. No significant association was observed at the 1% level between this peak relative surface and any trait.

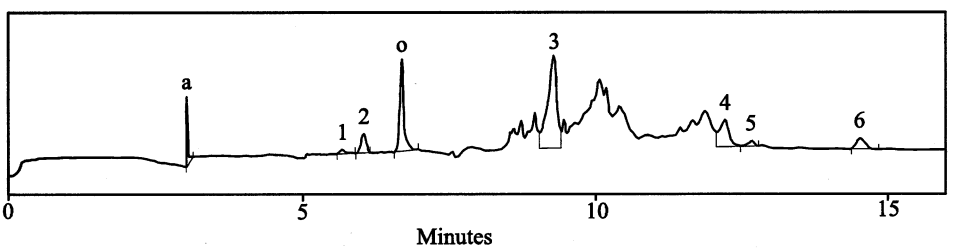

Fig. 1. Profile by CGE of total proteins from soybean seed. Approximate migration time (min) and molecular weights (kDa): Peak 'a': 3.04, 0.30, Peak 1: 5.68, 1.89, Peak 2: 6.05, 2.44, Peak 'o': 6.71, 3.86, Peak 3: 9.31, 23.21, Peak 4: 12.23, 175.55, Peak 5: 12.69, 240.72, Peak 6: 14.54, 864.90.

Table 1
Significant correlations between peaks surfaces, seed traits and N_2 fixation

Traits	Correlation	Significance level	95% Confidence interval	Corrected correlation	
	Peak 3				
Seed yield (g row^{-1})	−0.24	0.0066	−0.40	−0.051	−0.40
Protein content (%)	0.24	0.0055	0.06	0.41	0.41
	Peak 4				
Seed size (g)	0.22	0.0100	0.04	0.39	0.74
Earliness (HU)	0.24	0.0059	0.05	0.41	0.80
ARA at R6 stage (μmol C_2H_4 h^{-1} plant^{-1})	0.25	0.0044	0.06	0.42	0.83
Nodules dry weight (g plant^{-1})	0.25	0.0042	0.07	0.42	0.83
Lipoxygenases 2+3 (U g seed^{-1})	0.50	0.0011	0.20	0.71	0.50
	Peak 6				
Seed yield (g row^{-1})	0.35	1E-04	0.17	0.50	0.41
Seed number	0.35	0.0001	0.17	0.50	0.41
Earliness (HU)	0.26	0.0032	0.08	0.42	0.31
Protein content (%)	−0.41	3E-06	−0.56	−0.25	−0.49
Seed protein yield (g row^{-1})	0.27	0.0024	0.08	0.43	0.32
	Peak 3				
Peak 6	−0.27	0.0025	−0.43	−0.08	−0.54

The linear correlation between the 4 other peaks and the 17 traits, their significance level, their 99% confidence interval, and the corrected correlation (Eq. (1)) were computed. A total of 68 tests (4 peaks × 17 traits) was performed among which 12 tests were significant at the 1% level. No significant association was observed between peak 5 and any trait. The relative surfaces of peak 3 and 6 were also significantly correlated. The summary of significant correlations involving peak surfaces is given in Table 1. The correlations between all traits involved in the previous correlations and their significance level is given in Table 2. The plots of the peak relative surfaces versus the traits with which they are significantly correlated are given in Fig. 2. The probability $P(12)$ of finding, by chance only, at least 12 tests significant at the 1% level among the 68 tests is equal to 5×10^{-12} (Eq. (4)). This result clearly shows that there are more significant correlations between peak surfaces and quantitative traits than what would have been obtained by chance only. Consequently, some of these correlations must be relevant. The corrected correlations, which eliminate the residual variance due to experimental variation, have always a greater value than the simple correlations, indicating that they are underestimates of the true correlations.

As can be seen in Tables 1 and 2, peaks 3 and 6, which surfaces are negatively correlated, are both in relation to total seed protein content and to seed yield. These two peaks have an opposite relation to the traits. The increase in protein content corresponds to an increase in peak 3, whereas it corresponds to a decrease in peak 6. Inversely, the increase in seed yield is related to an increase in peak 6 and a decrease in peak 3 in about the same proportion. Moreover, peak 6 is significantly associated to the different parameters involved in the expression of seed yield as earliness and seed number (Table 2). Peaks 3 and 6 may correspond to protein participating in a different way to the establishment of seed protein content. Seed protein composition and seed yield are the major goals of the breeding programs in soybean. The global seed protein amount and the content in some storage proteins are known to be affected by the nitrogen nutrition, that is nitrate versus N_2 fixation [4]. Field methods can also have an impact [20]. Seed protein amount has been subjected to several quantitative trait loci (QTL) approaches, thanks to molecular markers. Few QTL, with large effect, have been found in different studies, these QTL being often specific to the population under investigation, the genetic background playing a role in QTL expression [21–24]. It can be

noted that, in several linkage study, loci affecting seed traits were repeatedly localized on linkage group F of the soybean genome [25–27].

Peak 4 seems to have a different implication in the functioning of the plant. Its relative surface is significantly and positively correlated to several traits related to symbiotic dinitrogen fixation, as ARA value and nodule dry weight analyzed at R6 stage two parameters closely associated (Table 2). The efficiency of symbiotic fixation is often associated with the total development cycle length [18] as shown in Table 2, which could explain the relationship between the surface of peak 4 and earliness. The lipoxygenase (2 + 3) content is also related to peak 4 (Table 1). This parameter and those involved in symbiotic fixation are significantly correlated (Table 2), which could be due to the symbiotic origin of the nitrogen transferred into seeds [28]. The 1, 2 and 3 lipoxygenases are enzymes representing 2% of the total seed proteins, so they could be one of the components of our CGE profile. They are responsible for the undesirable flavours associated with soybean products. They are active during germination for lipid mobilization, and thus for energy supply [29].

The corrected correlations concerning peak 4 have high values (Table 1). Peak 4 may be an interesting candidate to evaluate the symbiotic fixation ability more precisely. Moreover, dinitrogen fixation was advocated as the best parameter for a simultaneous improvement of yield and protein content [5], two traits that are negatively correlated [30]: peak 4 could give the opportunity to break this negative relationship.

This first insight into the quantitative soybean seed protein composition, thanks to CGE, suggests that some of the seed components could be interesting to study further. The nature of the components seen on the profile of the CGE could be determined by their amino-acid sequences, and compared to known proteins in databases. The relationship of the peak surfaces with the traits of interest should be confirmed in other crosses before this kind of investigation is performed. However, since quantity variations are considered, the regulation mechanisms of the protein amounts could be the key elements to be dealt with.

Molecular markers could be used in conjunction with a CGE approach. CGE in itself can provide markers, like peak 2, which was alternatively

Table 2
Phenotypic correlations between seed traits and N_2 fixation

	Seed number	Seed size	Earliness	Protein content
Seed yield (g row^{-1})	0.93**	0.26**	0.44**	−0.19*
Seed number		−0.09	0.42**	−0.22*
Seed size (g)			0.13	0.07
Earliness (HU)				−0.14
Protein content (%)				
Seed protein yield (g row^{-1})				
ARA at R6 stage (mmol C_2H_4 h^{-1} plant^{-1})				
Nodules dry weight (g plant^{-1})				

	Seed protein yield	ARA at R6 stage	Nodules dry weight	Lipoxygenases 2+3
Seed yield (g row^{-1})	0.97**	−0.01	−0.04	−0.37*
Seed number	0.90**	−0.03	−0.05	−0.44**
Seed size (g)	0.27**	0.03	0.005	0.18
Earliness (HU)	0.43**	0.32**	0.34**	0.12
Protein content (%)	0.02*	0.01	−0.02	0.20
Seed protein yield (g row^{-1})		−0.01	−0.03	−0.34*
ARA at R6 stage (mmol C_2H_4 h^{-1} plant^{-1})			0.65**	0.36*
Nodules dry weight (g plant^{-1})				0.29*

* significant at 5%.
** significant at 1%.

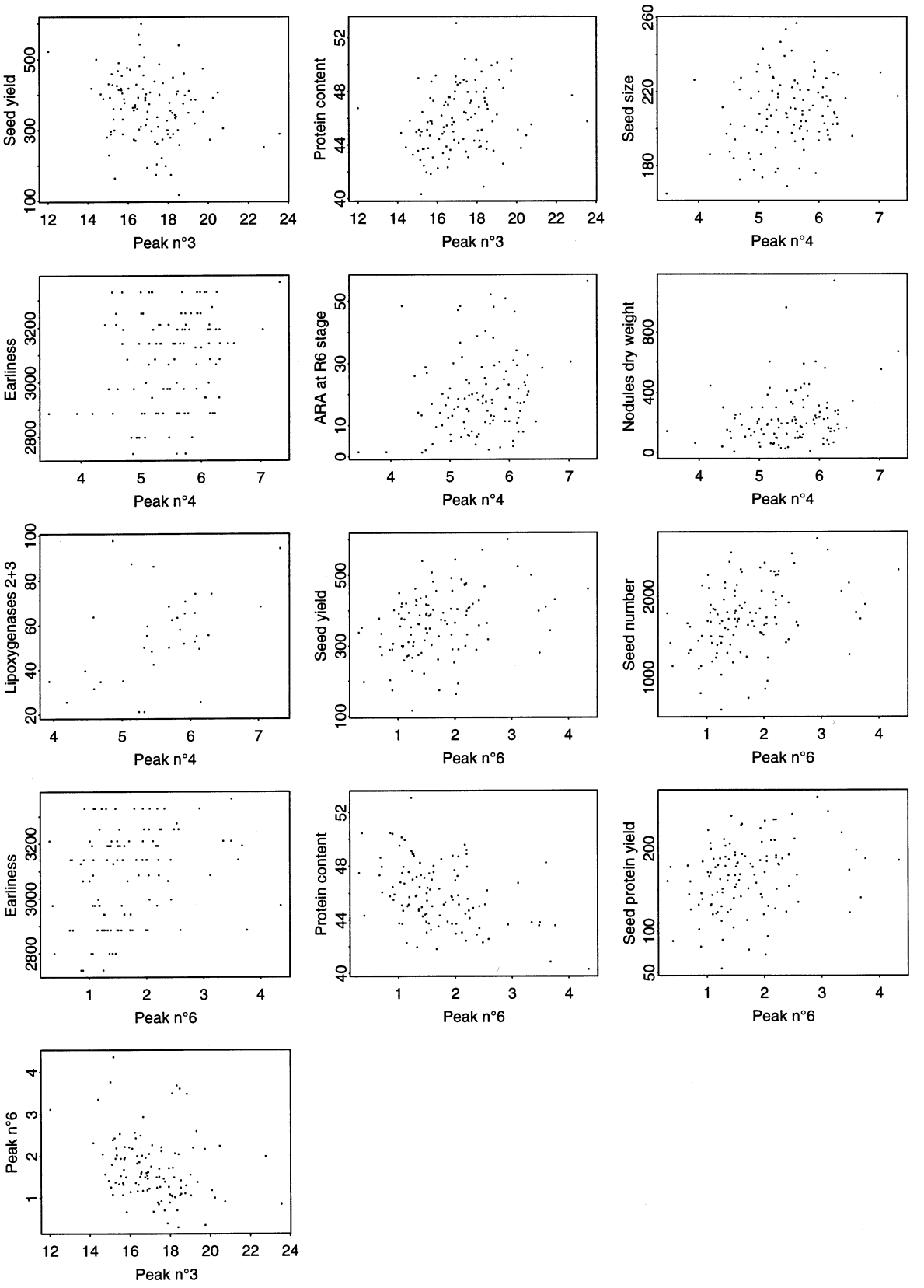

Fig. 2. Plots of the peak relative surfaces against the traits.

present/absent and followed a Mendelian segregation. Such a 'null allele' was observed in the subunit of a storage protein in soybean seed with classical electrophoresis [28]. Moreover, CGE gives an insight into the quantitative composition of the total seed proteins, and each of the peak relative surfaces can be used as a trait in a classical QTL approach. In the present study, we had some random amplified polymorphic DNA (RAPD) markers segregating among the recombinant inbred lines. No significant association (at the 1% level) between the data provided by CGE and these markers were detected, but there were too few markers compared to the large genome size of soybean. However, this approach could be easily fulfilled with the numerous molecular markers now available in this species [31–33]. A combined study of the proteins composing the CGE peaks, their regulation and a classical molecular markers/QTL analysis would allow a better description of the genetics of seed quality.

Acknowledgements

We are very grateful to Daniel Kahn for his hospitality in his laboratory, at INRA, Toulouse. To Philippe Rousseau for his support with the CGE analysis, to Sylviane Baurès for her technical assistance, to Frédéric Danjon for his help and to Bruno Goffinet for his suggestions.

References

[1] J.W. Burton, Soyabean (*Glycine max* (L) Merr), Field Crop. Res. 53 (1997) 171–186.
[2] R.A. Scott, K.D. Kephart, Selection for yield, protein, and oil in soybean crosses between adapted and introduced parents, Field Crop. Res. 49 (1997) 177–185.
[3] D.M. Ronis, D.J. Sammons, W.J. Kenworthy, J.J. Meisinger, Heritability of total and fixed N content of seed in two soybean populations, Crop Sci. 25 (1985) 1–4.
[4] N.C. Paek, J. Imsande, R.C. Shoemaker, R. Shibles, Nutritional control of soybean seed storage protein, Crop Sci. 37 (1997) 498–503.
[5] C. Planchon, N. Burias, M. Berger, Nitrogenase activity, protein content and yield in soybean, Eur. J. Agron. 1 (1992) 195–200.
[6] I. Recio, M. Ramos, L. Amigo, Study of the polymorphism of ovine α_{s1}- and α_{s2}-caseins by capillary electrophoresis, J. Dairy Res. 64 (1997) 525–534.
[7] C. Damerval, A. Maurice, J.M. Josse, D. de Vienne, Quantitative trait loci underlying gene product variation: a novel perspective for analysing regulation of genome expression, Genetics 137 (1994) 289–301.
[8] K. Zhang, K.M. Soliman, V.T. Sapra, Polymerase chain reaction detect polymorphisms and trait association in soybean, Biol. Plant. 40 (1998) 43–53.
[9] A. Leonardi, C. Damerval, Y. Hebert, A. Gallais, D. de Vienne, Association of protein amount polymorphism (PAP) among maize lines with performances of their hybrids, Theor. Appl. Genet. 82 (1991) 552–560.
[10] S. Gerber, M. Lascoux, A. Kremer, Relation between protein markers and quantitative traits in maritime pine (*Pinus pinaster* Ait.), Silvae Genet. 46 (1997) 286–291.
[11] S. Suc, Amélioration génétique de la teneur en protéines du soja (*Glycine max* L. Merrill), PhD thesis, INP-ENSA, Toulouse, France, 1983.
[12] R. Ecochard, Y. Ravelomanantsoa, Genetic correlations derived from full-sib relationships in soybean (*Glycine max* (L.) Merrill), Theor. Appl. Genet. 63 (1982) 9–15.
[13] F. Fabre, Amélioration génétique de la qualité de la graine et fixation symbiotique de l'azote chez le soja (*Glycine max* L.), PhD thesis, INP/ENSA, Toulouse, France, 1998.
[14] A. Zougari, S. Guy, C. Planchon, Genotypic lipoxygenase variation in soybean seeds and response to nitrogen nutrition, Plant Breed. 114 (1995) 313–316.
[15] N. Burias, C. Planchon, Increasing soybean productivity through selection for nitrogen fixation, Agron. J. 82 (1990) 1031–1034.
[16] J. Balandreau, Y. Dommergues, Mesure in situ de l'activité nitrogénasique, CR. Acad. Sci. 273 (1971) 2020–2022.
[17] W.R. Fehr, C.E. Caviness, Stages of soybean development, Coop. Ext. Serv., Iowa State Univercity, Special Report, 80, 1977.
[18] T.G. Patterson, T.A. Larue, Nitrogen fixation by soybeans: seasonal ans cultivar effects, and comparison of estimates, Crop Sci. 23 (1983) 488–492.
[19] N.G. Anderson, S.L. Tollaksen, F.H. Pascoe, L. Anderson, Two-dimensional electrophoretic analysis of wheat seed proteins, Crop Sci. 25 (1985) 667–674.
[20] D.L. Pazdernik, L.L. Hardman, J.H. Orf, F. Clotaire, Comparison of field methods for selection of protein and oil content in soybean, Can. J. Plant Sci. 76 (1996) 721–725.
[21] B.W. Diers, P. Keim, W.R. Fehr, R.C. Shoemaker, RFLP analysis of soybean seed protein and oil content, Theor. Appl. Genet. 83 (1992) 608–612.
[22] L.M. Mansur, J.H. Orf, K. Chase, T. Jarvik, P.B. Cregan, K.G. Lark, Genetic mapping of agronomic traits using recombinant inbred lines of soybean, Crop Sci. 36 (1996) 1327–1336.
[23] S.H. Lee, M.A. Bailey, M.A.R. Mian, T.E. Carter, E.R. Shipe, D.A. Ashley, W.A. Parrott, R.S. Hussey, H.R. Boerma, RFLP loci associated with soybean seed protein and oil content across populations and locations, Theor. Appl. Genet. 93 (1996) 649–657.
[24] E.C. Brummer, G.L. Graef, J. Orf, J.R. Wilcox, R.C. Shoemaker, Mapping QTL for seed protein and oil content in eight soybean populations, Crop Sci. 37 (1997) 370–378.

[25] B.W. Diers, V. Beilinson, N.C. Nielsen, R.C. Shoemaker, Genetic mapping of the Gy4 and Gy5 glycinin genes in soybean and the analysis of a variant of Gy4, Theor. Appl. Genet. 89 (1994) 297–304.
[26] M.A.R. Mian, M.A. Bailey, J.P. Tamulonis, E.R. Shipe, T.E. Carter, W.A. Parrott, D.A. Ashley, R.S. Hussey, H.R. Boerma, Molecular markers associated with seed weight in two soybean populations, Theor. Appl. Genet. 93 (1996) 1011–1016.
[27] Z. Chen, R.C. Shoemaker, Four genes affecting seed traits in soybeans map to linkage group, F. Journal of Heredity 89 (1998) 211–215.
[28] F.R. Warembourg, M.P. Fernandez, Distribution and remobilization of symbiotically fixed nitrogen in soybean (*Glycine max*), Physiol. Plant. 65 (1985) 281–286.
[29] I. Feussner, C. Wasternack, H. Kindl, H. Kühn, Lipoxygenase-catalysed oxygenation of storage lipids is implicated in lipid mobilization during germination, Proc. Natl. Acad. Sci. USA 92 (1995) 11849–11853.
[30] V.K. Wehrman, W.R. Fehr, S.R. Cianzo, J.F. Cavins, Transfer of high seed protein to high-yielding soybean cultivars, Crop Sci. 27 (1987) 927–931.
[31] N. Diwan, P.B. Cregan, Automated sizing of fluorescent-labeled Simple Sequence Repeat (SSR) markers to assay genetic variation in soybean, Theor. Appl. Genet. 95 (1997) 723–733.
[32] P. Keim, J.M. Schupp, S.E. Travis, K. Clayton, T. Zhu, L. Shi, A. Ferreira, D.M. Webb, A high-density soybean genetic map based on AFLP markers, Crop Sci. 37 (1997) 537–543.
[33] G. Wang, R. Mahalinga, H.T. Knap, C-A) and (G-A) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, *Glycine max* (L.) Merr, Theor. Appl. Genet. 96 (1998) 1086–1096.