Environmental and genetic variation of soybean seed protein content under Central European growing conditions

Johann Vollmann, Christina N Fritz, Helmut Wagentristl and Peter Ruckenbauer

Plant Breeding Department, University of Agricultural Sciences Vienna, Gregor Mendel Strasse 33, A-1180 Vienna, Austria
Experimental Station Gross Enzersdorf, Schlosshofer Strasse 31, A-2301 Gross Enzersdorf, Austria

Abstract: Seed protein content is important for both feed and food utilisation of soybean. In soybeans grown in Central Europe, considerable variation in protein content was due to seasonal influences, as demonstrated in different experiments from a breeding programme. In soybean genotypes of early maturity groups, average to high protein content (range 399–476 g kg\(^{-1}\)) was found in years with high air temperature and moderate rates of rainfall during the seed-filling period, whereas seed protein content was drastically reduced (range 265–347 g kg\(^{-1}\)) in seasons of insufficient nitrogen fixation or higher amounts of precipitation during seed filling. In a set of 60 genotypes, protein content was increased both by late nitrogen fertilisation before the onset of seed filling and by inoculation of seed with nitrogen-fixing rhizobia. Despite the high degree of environmental modification, genetic variation of seed protein content was considerable, and genotype × environment interaction was of low magnitude. Therefore selection of early maturing soybean genotypes with improved seed protein content appears to be feasible and is only limited by the moderately negative correlation between protein content and seed yield.

Keywords: soybean; seed protein content; environmental variation; nitrogen fixation; genetic improvement

INTRODUCTION

Soybean (Glycine max (L) Merr), which contains about 200 g oil and 400 g protein kg\(^{-1}\) seed dry matter, is the major oilseed and protein crop in many regions worldwide, providing approximately 60% of the world supply of vegetable protein. In Europe, more than 1.2 × 10\(^{6}\) ha of soybean was grown in 1998, mainly in Italy, France, Russia, Romania, Yugoslavia, Croatia and Austria. In Central European countries such as Austria the interest in soybean is due to its high seed protein content, whereas vegetable oil is preferably produced by cultivating oilseed rape, sunflower or other species. Consequently, the soybean oil is preferably produced by cultivating oilseed rape, sunflower or other species. Consequently, the soybean crop processed in Austria is mainly applied in livestock feeding, eg as full-fat soybean for pig fattening, whereas a smaller proportion is used in the food industry or in tofu making.

For both feed and food utilisation a high and stable seed protein content is desirable. However, in northern regions of soybean production over a number of seasons. Similarly, protein content was reported to be low in the northern locations of north-east China and in northern sites of Europe, where large seasonal variations were observed in protein content. Apart from other findings, it was recently demonstrated that low root zone temperatures reduce nitrogen fixation, which might explain the low protein contents commonly found in northern areas of soybean cultivation. Nevertheless, several soybean cultivars of early maturity groups have been developed with considerably improved seed protein content. Moreover, agronomic research has been initiated in many soybean production areas focusing on the management of seed quality characters.

Recent information on the seed protein content of Central European-grown soybean, which might be of interest to both agronomists and processors, is not available. For this reason, the magnitude of environmental and genetic sources of variation in seed protein content is being studied in a number of performance trials, nitrogen supply experiments and protein screenings within a soybean breeding programme carried out in the soybean growing area of Austria.

Correspondence to: Johann Vollmann, Plant Breeding Department, University of Agricultural Sciences Vienna, Gregor Mendel Strasse 33, A-1180 Vienna, Austria

Contract/grant sponsor: Austrian Science Foundation; contract/grant number: P10663-OBI

(Received 26 August 1999; revised version received 4 January 2000; accepted 14 February 2000)
MATERIALS AND METHODS

Genetic materials
Most of the soybean cultivars and breeding lines investigated for agronomic characters and seed protein content in this study are early maturing genotypes of soybean maturity group 00, which are commonly used in breeding programmes in Central Europe. Moreover, a few genotypes of maturity groups 000 and 0 respectively were included in particular experiments as standards. Breeding lines were usually derived from biparental crosses between well-adapted genotypes of early maturity (‘conventional crosses’). In crosses intended for development of populations with improved protein content, high-protein genotypes such as cv Proto⁹ or germplasm line BARC-6¹⁴ were utilised as protein donor parents. In performance trials and in the protein ranking experiment, cultivars and advanced breeding lines (F₃-derived lines in F₇ or later generations) were investigated, whereas in a screening experiment, F₂-derived lines in F₃ and F₄ generations were employed, which had only been preselected for early maturity in F₂.

Performance trials
Performance trials were carried out at Gross Enzersdorf (Lower Austria) from 1993 to 1998. Each experiment was planted as a 10 × 5 generalised lattice design with 50 entries in two replications. A set of eight standard genotypes was included in each of the trials, whereas breeding lines were usually tested for two to three seasons and then replaced by other lines as a result of selection. Individual plots were 5.5 × 1.25 m² in size, with four rows grown per plot. Soybean seeds were inoculated with Nodular-G (Serbios Inc, Badia Polesine, Italy), a commercial preparation of Bradyrhizobium japonicum (Kirchner) Jordan, for promoting nodulation and symbiotic nitrogen fixation. In addition, mineral fertilisers were applied at rates of 30 kg ha⁻¹ N, 70 kg ha⁻¹ P₂O₅ and 140 kg ha⁻¹ K₂O prior to sowing. Field trials were planted during the last week of April in each year at sowing rates of 80 seeds m⁻². Plots were end-trimmed prior to harvesting, and trials were harvested during the first week of October in each year.

Protein ranking experiment
The ranking of 60 soybean genotypes by seed protein content under varying nitrogen supply conditions was evaluated in 1996 and 1997 in Gross Enzersdorf in a single-row plot experiment (2.5 m row length, 70 cm row distance, sowing rate of 100 seeds per row, sowing dates 29 April 1996 and 1 May 1997) using a split-plot design with two replications. Three different nitrogen supply treatments, i.e. ‘control’ (no nitrogen fertilisation), ‘inoculation’ (seed inoculation with B japonicum at sowing time) and ‘nitrogen at RI’ (50 kg ha⁻¹ nitrogen applied at the beginning of the flowering stage)¹⁵ respectively, were set as the whole-plot factor, whereas the 60 genotypes were assigned to split plots. Moreover, genotypes were arranged in a 10 × 6 generalised lattice design in order to control spatial field variations within the relatively large whole plots.

Screening of breeding lines for seed protein content
In a screening experiment for protein content, 96 F₃-derived breeding lines and four parent cultivars were evaluated in single-row plots (2.5 m row length, 70 cm row distance, sowing rate of 100 seeds per row) in two replications in lattice designs at three Austrian locations, Gross Enzersdorf (Lower Austria), Pama (Burgenland) and Gleisdorf (Styria), in 1998. Trials were planted on 29 April, 30 April and 7 May 1998 at Gross Enzersdorf, Pama and Gleisdorf respectively. Breeding lines from both conventional crosses (M₆X-89/Apache, Dom/Ceresia) and high-protein crosses (M₆X-89/Proto, Dom/BARC-6) were included in this experiment. Yielding ability of individual breeding lines was estimated using a visual score of pod set at the time of full maturity (scale: ‘1’=lowest number of pods per plant, ‘3’=highest number of pods per plant; scale in steps of 0.5 score units).

Data collection and statistical analysis
After harvest, a 25 g sample of dry seeds from each experimental unit was finely ground using a Cyclotec 1093 sample mill (Foss Tecator, Höganas, Sweden). Thereafter, seed protein and oil contents were determined by near-infrared reflectance spectroscopy (NIRS) using an InfraAlyzer 450 spectrophotometer and IDAS spectroscopic data analysis software (Bran & Luebbe, Norderstedt, Germany). In each year, 20–25 seed samples were used as reference samples in order to expand and validate the spectroscopic calibration equations for accurate protein and oil content calculation. Seed protein and oil contents were expressed in g kg⁻¹ on a dry matter basis, whereas grain yield and seed size (weight of 1000 seeds in g) were given on an 80 mg g⁻¹ seed moisture basis.

Each individual field trial was analysed statistically using the appropriate lattice design ANOVA model and the PLABSTAT software program.¹⁶ Lattice-adjusted plot values and genotype mean values were then used for data analysis across environments and for illustrating variation. For visualising field trends of protein content in a contour plot, covariates from neighbour analysis were calculated as described by Vollmann et al.¹⁷

RESULTS
A considerable variation in soybean seed protein content across different growing seasons is presented in Fig 1a. In 1994 a protein content of more than 450 g kg⁻¹ was found in some genotypes, whereas in 1996 a protein content lower than 300 g kg⁻¹ was determined for a number of trial entries. Despite the high degree of variation across seasons, the ranking of particular genotypes for protein content was rather constant in all seasons except 1996 (Fig 1b). In Table 1, yield level, protein and oil contents, the amount of...
rainfall and the average air temperature during the developmental stages are given for the different growing seasons. In 1993 and 1994, low rates of precipitation and high temperatures clearly favoured protein instead of oil synthesis, whereas the extremely low protein content in 1996 was probably due to insufficient nitrogen availability, as nearly no nodule formation could be observed in that particular season. Moreover, in 1996 the lowest temperature during the seed-filling period was observed, which also might have affected seed protein content. In 1995, 1996 and 1998 the formation of an enhanced oil content was promoted by high amounts of rainfall during the seed-filling period (R3 to R8 stages). A moderately negative correlation between seed protein content and grain yield was observed in nearly all seasons (Table 2), whereas the correlation between protein content and time to maturity was positive in high-protein seasons and negative in low-protein seasons. The relationship between protein and oil content was consistently negative for all growing seasons (Table 2). In the exceptionally low-protein season 1996 the early maturing cv Ultra was higher in protein content than high-protein cvs Proto and Apache (Fig 1b). In this season of low nitrogen availability, high amounts of rainfall occurred late in August and in September, when cv Ultra had already reached full maturity (data not shown). In genotypes of later maturity, oil synthesis was enhanced by late water availability, which reduced the protein fraction of the seed as a consequence. This view is also supported by the finding of a clearly negative correlation between protein content and time to maturity in 1996 (Table 2).

As with seasonal influences, variations in soybean protein content were also found between different geographic locations, as summarised in Fig 2 for a breeding line screening experiment carried out in 1998. Moreover, an investigation of the on-site seed-filling period was observed, which also might have affected seed protein content. In 1995, 1996 and 1998 the formation of an enhanced oil content was promoted by high amounts of rainfall during the seed-filling period (R3 to R8 stages). A moderately negative correlation between seed protein content and grain yield was observed in nearly all seasons (Table 2), whereas the correlation between protein content and time to maturity was positive in high-protein seasons and negative in low-protein seasons. The relationship between protein and oil content was consistently negative for all growing seasons (Table 2). In the exceptionally low-protein season 1996 the early maturing cv Ultra was higher in protein content than high-protein cvs Proto and Apache (Fig 1b). In this season of low nitrogen availability, high amounts of rainfall occurred late in August and in September, when cv Ultra had already reached full maturity (data not shown). In genotypes of later maturity, oil synthesis was enhanced by late water availability, which reduced the protein fraction of the seed as a consequence. This view is also supported by the finding of a clearly negative correlation between protein content and time to maturity in 1996 (Table 2).

As with seasonal influences, variations in soybean protein content were also found between different geographic locations, as summarised in Fig 2 for a breeding line screening experiment carried out in 1998. Moreover, an investigation of the on-site

Table 1. Yield level and seed protein and oil contents in soybean performance trials and amount of rainfall and average temperature during developmental stages of soybean in different growing seasons (1993–1998)

<table>
<thead>
<tr>
<th>Year</th>
<th>Yield level a</th>
<th>Protein content b</th>
<th>Oil content c</th>
<th>Amount of rainfall (mm)</th>
<th>Average air temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V1 to R2 stages b</td>
<td>R3 to R8 stages c</td>
<td>During growing period</td>
<td>V1 to R2 stages b</td>
<td>R3 to R8 stages c</td>
</tr>
<tr>
<td>1993</td>
<td>2643</td>
<td>423.7</td>
<td>196.7</td>
<td>175</td>
<td>115</td>
</tr>
<tr>
<td>1994</td>
<td>1826</td>
<td>433.7</td>
<td>174.4</td>
<td>120</td>
<td>72</td>
</tr>
<tr>
<td>1995</td>
<td>2993</td>
<td>332.0</td>
<td>228.6</td>
<td>219</td>
<td>171</td>
</tr>
<tr>
<td>1996</td>
<td>2493</td>
<td>301.8</td>
<td>224.0</td>
<td>215</td>
<td>154</td>
</tr>
<tr>
<td>1997</td>
<td>2595</td>
<td>357.8</td>
<td>211.3</td>
<td>313</td>
<td>62</td>
</tr>
<tr>
<td>1998</td>
<td>2323</td>
<td>360.4</td>
<td>248.4</td>
<td>210</td>
<td>196</td>
</tr>
<tr>
<td>1991–1997</td>
<td></td>
<td></td>
<td></td>
<td>297</td>
<td></td>
</tr>
</tbody>
</table>

* Trial’s total mean yield (kgha⁻¹), protein content (g·kg⁻¹) and oil content (g·kg⁻¹) respectively.  
*Soybean developmental stages from germination to full flowering according to Fehr and Caviness.  
*Soybean developmental stages from initial pod to full maturity (= seed-filling period).
variation of seed protein content by utilising covariates from neighbour analysis revealed clear patterns of spatial dependence of protein content within particular fields, as shown in Fig 3; deviations from the average protein content were between $-12$ and $+14 \, g \, kg^{-1}$ for a relatively small field section $32 \times 32 \, m^2$ in size.

The influence of both nitrogen supply and genetic make-up on seed protein content is demonstrated in the protein ranking experiment summarised in Fig 4: an additional supply of $50 \, kg \, ha^{-1}$ nitrogen at the flowering stage resulted in an average increase of $25 \, g \, kg^{-1}$ protein, whereas protein content was only increased by $5 \, g \, kg^{-1}$ through symbiotic nitrogen fixation, as compared to an untreated control variant. Although statistically significant ($LSD_{0.05} = 3.7 \, g \, kg^{-1}$ for comparisons of treatment means), the small increase in protein content between the untreated control and the inoculation treatment must not be considered as the sole contribution of symbiotic nitrogen fixation to protein content, because nodulation was also observed in plants of both the untreated control and the nitrogen fertilisation variants, which is due to the presence of nitrogen-fixing bacteria from soybeans grown in that field during earlier years. The ranking of genotypes by protein content appeared to be rather similar for the different nitrogen regimes applied. Moreover, genotypes were ranked similarly across years (Spearman rank correlation coefficient $r_s = 0.78$), which was also confirmed by an analysis of variance, in which genotype $\times$ treatment interactions were of a much lower magnitude than was the effect of genotype on protein content (Table 3). For this reason, the low degree of genotype $\times$ year interaction appears to be mainly a non-crossover type of interaction, and protein content data of genotypes were combined over years (Fig 4). Apart from seed protein content, seed weight and oil content were also influenced by the nitrogen regime,
and the influence of the year was predominant for all the seed characters investigated (Table 3).

In a screening experiment of breeding lines (see Fig 2 for variation of protein content in three locations), seed protein content of genotypes derived from high-protein crosses was about 12 g kg\(^{-1}\) higher than that of progeny from conventional crosses, whereas pod set score was slightly lower in the high-protein crosses (Fig 5). The correlation between pod set score and protein content was slightly lower in the high-protein crosses of progeny from conventional crosses, whereas pod set yield level should be possible. This moderately negative correlation suggests that selection of breeding lines for variation of protein content in three locations (Fig 2) have been published for soybean production for tofu making. Variations of seed protein content within a particular field location consequently in tofu characteristics for soybeans grown in different locations; they identified favourable locations to the low level of rainfall. Similarly, in experiments on stress during the seed-filling period, both drought and high-air-temperature conditions enhanced protein content by approximately 30–50 g kg\(^{-1}\), whereas oil content and grain yield were reduced.\(^{18}\) At constant levels of water supply, both protein and oil contents were increased at high temperature, and fatty acid composition was most affected by temperature, whereas amino acid composition was stable.\(^{19}\) In other investigations dealing with water stress and irrigation of early maturing soybeans, the importance of timing the irrigation was emphasised: protein content was highest for irrigation after the flowering stage, whereas protein content was lower with continuous water supply.\(^{12,20}\) This might be due to higher yield levels or higher oil content at optimum rates of water supply, which would dilute a given amount of protein. Consequently, high rates of rainfall during the seed-filling periods in 1995, 1996 and 1998 resulted in increased oil contents (Table 1). In 1996, seed protein content was lowest as a result of nitrogen deficiency due to insufficient nodulation. In this particular season, soil temperature was below 15 °C during the V0 to V4 stages of growth. It has been demonstrated that nodulation is drastically reduced and the onset of nitrogen fixation is delayed at low root zone temperature,\(^{1}\) thus decreasing grain and protein yields.\(^{21}\) For this reason, Pazdernik et al\(^{22}\) suggested selecting soybean genotypes for early nodulation characteristics to overcome nitrogen deficiency in the initial vegetative developmental stages.

Reports on variation of soybean seed protein content across locations (Fig 2) have been published frequently. In a recent investigation, Bhardwaj et al\(^{23}\) reported differences in seed protein content and subsequently in tofu characteristics for soybeans grown in different locations; they identified favourable locations for soybean production for tofu making. Variations of seed protein content within a particular field location (Fig 3) were found of a similar magnitude as described earlier.\(^{17}\) Field variations in protein content have been shown to be closely correlated to variations in grain yield, oil content and seed size.\(^{24}\) These spatial variations are due to heterogeneity in various soil

**DISCUSSION**

In the present investigation, seed protein content was highest for soybean crops grown under moderately dry conditions and high temperature during the seed-filling period (Table 1). In 1993 a high yield level was maintained apart from the high protein content, whereas in 1994 the grain yield was reduced owing to the low level of rainfall. Similarly, in experiments on stress during the seed-filling period, both drought and high-air-temperature conditions enhanced protein content by approximately 30–50 g kg\(^{-1}\), whereas oil content and grain yield were reduced.\(^{18}\) At constant levels of water supply, both protein and oil contents were increased at high temperature, and fatty acid composition was most affected by temperature, whereas amino acid composition was stable.\(^{19}\) In other investigations dealing with water stress and irrigation of early maturing soybeans, the importance of timing the irrigation was emphasised: protein content was highest for irrigation after the flowering stage, whereas protein content was lower with continuous water supply.\(^{12,20}\) This might be due to higher yield levels or higher oil content at optimum rates of water supply, which would dilute a given amount of protein. Consequently, high rates of rainfall during the seed-filling periods in 1995, 1996 and 1998 resulted in increased oil contents (Table 1). In 1996, seed protein content was lowest as a result of nitrogen deficiency due to insufficient nodulation. In this particular season, soil temperature was below 15 °C during the V0 to V4 stages of growth. It has been demonstrated that nodulation is drastically reduced and the onset of nitrogen fixation is delayed at low root zone temperature,\(^{1}\) thus decreasing grain and protein yields.\(^{21}\) For this reason, Pazdernik et al\(^{22}\) suggested selecting soybean genotypes for early nodulation characteristics to overcome nitrogen deficiency in the initial vegetative developmental stages.

Reports on variation of soybean seed protein content across locations (Fig 2) have been published frequently. In a recent investigation, Bhardwaj et al\(^{23}\) reported differences in seed protein content and subsequently in tofu characteristics for soybeans grown in different locations; they identified favourable locations for soybean production for tofu making. Variations of seed protein content within a particular field location (Fig 3) were found of a similar magnitude as described earlier.\(^{17}\) Field variations in protein content have been shown to be closely correlated to variations in grain yield, oil content and seed size.\(^{24}\) These spatial variations are due to heterogeneity in various soil

---

**Table 3.** Degrees of freedom (DF) and mean squares from the analysis of variance for seed quality parameters in 60 soybean genotypes and three nitrogen treatments of the protein ranking experiment.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Df</th>
<th>Seed weight</th>
<th>Protein content</th>
<th>Oil content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replication within year</td>
<td>2</td>
<td>747.9</td>
<td>1912.0*</td>
<td>724.6</td>
</tr>
<tr>
<td>Year</td>
<td>1</td>
<td>265397.8***</td>
<td>347723.6***</td>
<td>304033.1***</td>
</tr>
<tr>
<td>Treatment</td>
<td>2</td>
<td>9298.5***</td>
<td>42657.5***</td>
<td>8803.1***</td>
</tr>
<tr>
<td>Year × treatment</td>
<td>2</td>
<td>5216.4**</td>
<td>27828.8***</td>
<td>4581.9**</td>
</tr>
<tr>
<td>Error (a)</td>
<td>4</td>
<td>113.8</td>
<td>233.4</td>
<td>154.9*</td>
</tr>
<tr>
<td>Genotype</td>
<td>59</td>
<td>3804.0***</td>
<td>3333.3***</td>
<td>1321.8***</td>
</tr>
<tr>
<td>Genotype × year</td>
<td>59</td>
<td>473.7***</td>
<td>431.3***</td>
<td>153.1***</td>
</tr>
<tr>
<td>Genotype × treatment</td>
<td>118</td>
<td>127.1*</td>
<td>244.9***</td>
<td>107.6***</td>
</tr>
<tr>
<td>Genotype × year × treatment</td>
<td>118</td>
<td>154.1**</td>
<td>262.2***</td>
<td>121.7***</td>
</tr>
<tr>
<td>Error (b)</td>
<td>354</td>
<td>98.7</td>
<td>172.4</td>
<td>63.8</td>
</tr>
</tbody>
</table>

*\(, **, ***\) Significant (F-test) at the 0.05, 0.01 and 0.001 levels respectively.

---

![Figure 5.](image.png)

**Figure 5.** Scatter plot of pod set scores as an estimate of yielding ability (visual score from '1' = low to '3' = high number of pods per plant) and seed protein content for 96 breeding lines and four standard cultivars grown in a screening experiment at the three locations Gross Enzersdorf, Pama and Gleisdorf in 1998 (data are mean values over the three locations).
parameters, eg soil water content and nitrogen uptake. Within-field variations, which were also described in wheat for grain yield and protein content, are of major concern in plant breeding experiments, as the response to selection might be reduced.

The availability of nitrogen during the seed-filling period is another important prerequisite of a high seed protein content. Late nitrogen fertilisation during the seed-filling period was reported to increase nitrogen uptake in pods and seed protein content at some of a number of locations. However, Paek et al demonstrated that the increase in seed protein content after late nitrogen application was associated with poorer protein quality, because the low-sulphur beta-subunit of 7S protein was more strongly expressed than other fractions of seed storage protein. Nitrogen application at an earlier stage of development (ie flowering) did not affect seed protein content, but increased grain yield in some studies (eg Ref 30). In the present investigation (protein ranking experiment), nitrogen fertilisation at the flowering stage was superior to both the control and rhizobium inoculation in increasing seed protein content in a set of genotypes (Fig 4). Moreover, the low magnitude of interaction between genotype and nitrogen treatment in seed quality characters (Table 3) suggests that different genotypes would react similarly to those management variants, and that selection of genotypes would have given a similar result under each of the nitrogen treatments. The small but significant increase in protein content after inoculation as compared to the control (Fig 4) may be due to higher nitrogen fixation efficiency of the *Bradyrhizobium japonicum* variant used than of bacterial strains indigenous in the field. However, as nitrogen fixation is a highly complex phenomenon, this may be due to insufficient nitrogen fixation in cool seasons or to high rates of precipitation and low temperature during the seed-filling period. In soybean genotypes of early maturity groups, genetic variation in protein is significant under different nitrogen regimes and could be used in selection programmes for a stepwise improvement of seed protein content.

**ACKNOWLEDGEMENTS**

This contribution was supported by a research grant provided by the Austrian Science Foundation (FWF Research Project P10663-OBI). Thanks are due to Mrs J Winkler of Saatzucht Glesdorf for providing experimental facilities and to Mrs A Damboeck and R Tumpold of Vienna for technical assistance.

**REFERENCES**

18. Dornbos Jr DL and Mullen RE, Soybean seed protein and oil