Genetic analysis of soybean hard seededness with molecular markers*

P. Keim 1, **, B. W. Diers 1 and R. C. Shoemaker 1, 2
1 Departments of Agronomy and Genetics, Iowa State University, Ames, IA 50011, USA
2 United States Department of Agriculture-Agricultural Research Service, Agronomy Hall, Ames, IA 50011, USA

Received June 23, 1989; Accepted November 2, 1989
Communicated by K. Tsunewaki

Summary. Hard seededness in soybean [Glycine max (L.) Merr.] is a quantitative trait that affects the germination rate, viability, and quality of stored seeds. We have used 72 restriction fragment length polymorphisms (RFLPs) to identify genomic regions containing quantitative trait loci (QTL) affecting hard seededness in a segregating population from a G. max by a Glycine soja (Sieb. & Zucc.) cross. Five independent RFLP markers were found to be associated with variation in the hard-seeded trait. These markers and the epistatic interactions between them explain 71% of the variation for hard seededness. A genomic region associated with the i locus accounted for 32% of the variation in this segregating population. This study illustrates one approach to physiological genetic studies in plants.

Key words: Soybean – RFLP – Quantitative trait loci – Germination – Hard seededness

Introduction

Imbibition is one of the first steps in breaking dormancy and initiating germination of a soybean seed. Dormancy and viability can be maintained for long periods in hard-seeded soybean accessions because their seed coats are impermeable to water (Rolston 1978). Some legume seeds have remained viable for more than 100 years (Rolston 1978). This is important ecologically in wild populations and economically in cultivated legumes. Typically, wild soybean accessions are hard seeded, whereas cultivated varieties are not. In the southern United States and in the tropics, hard seededness would be beneficial by contributing to the quality of stored soybeans and to their viability (Potts et al. 1978). Some breeding programs have recognized the importance of hard seededness and have introgressed this trait into adapted cultivars (Kilen and Hartwig 1978).

The physiological basis for seed-coat impermeability is not fully understood (Rolston 1978, Ballard 1973), but morphological (Egley and Paul 1981; Yaklich et al. 1984), enzymatic (Egley et al. 1983), and phenolic differences (Marbach and Mayer 1974, 1975) have been implicated. Even though as few as three genes may control variation in soybean hard seededness in some populations (Kilen and Hartwig 1978), the relationship between the physiology and the genetics of hard seededness has not yet been explored.

The use of genetics to understand physiological phenomena has been invaluable in bacteria and fungi (Gots and Benson 1974; Henry et al. 1984), and this success has fostered many attempts to develop comparable systems in plants (Terzaghi et al. 1985). The problems associated with these types of studies in plants have been enormous, and plant genetics has played a relatively minor role in physiological studies when compared with other systems. Arabidopsis seems the best candidate for integrating genetic and physiological studies, and notable success has been achieved (Bowman et al. 1989). However, this success is dependent upon the unique biology of Arabidopsis and may not be generally applicable to studies in other plants. Novel genetic strategies will be required for wide application of physiological-genetic studies in soybean.

Restriction fragment length polymorphisms (RFLPs) as genetic markers have facilitated the genetic study of traits conditioned by multiple genes (Burr et al. 1989).
Marker loci genetically linked to quantitative trait loci (QTLs) can be used to first identify genomic regions that contain important genes, and then to facilitate the genetic manipulation of the trait. We have developed such markers for soybean and used them to genetically characterize seed-coat hardness. Our goals were to identify genes conditioning this trait and to characterize their genetic action. This study illustrates a possible genetic strategy for studying physiological phenomena.

Materials and methods

A population containing great diversity in the hard-seeded trait was chosen for our studies. This was an F2 population (provided by Dr. W. R. Fehr) generated by crossing a Glycine max breeding line (AS1-356022 from Iowa State University) and a wild G. soja accession (PI 468916). Each of 60 F2 individuals was grown near Ames/IA and harvested separately. Thirty F3 seeds from each F2 individual were scarified and grown in a winter nursery (Isabela, Puerto Rico). The F4 seeds from each F2-derived line were bulk harvested and then used in evaluating the seed-hardness trait.

As a measure of seed-coat hardness (permeability), seed germination was evaluated in the absence of scarification. F4 seeds from each F2-derived line were placed on wet germination paper in the dark at room temperature for 7 days. The proportion of seeds that germinated was then determined. As a control, all ungerminated seeds were scarified and placed on wet germination paper for an additional 7 days. More than 98% of the ungerminated seeds germinated with scarification. This is consistent with high seed viability and suggests that the seed-coat permeability is a major barrier to germination. The germination assay was replicated twice.

RFLP genotypes of the 60 F2 plants were determined. DNA was isolated from leaves (Keim et al. 1988) and digested with one of five restriction endonucleases (EcoRI, EcoRV, HindIII, TaqI, or DraI). Digested DNA was then separated by agarose gel electrophoresis (Maniatis et al. 1982) and transferred to nylon membrane (Biotrace RP, Gelman) by using a VacuBlot apparatus (Applied BioNetics). Recombinant DNA clones were isolated from a PstI library (Keim and Shoemaker 1988) and screened for their ability to detect RFLPs in this population (P. Keim, B. W. Diers, R. C. Shoemaker unpublished results). Recombinant DNA was radioactively labelled with 32P (Boehringer-Mannheim random primer kit no. 1004760) and then used for molecular hybridization (Aputya et al. 1988) with DNA from F2 individuals. F2 genotypes for five isozymes (malate dehydrogenase, malic enzyme, isocitrate dehydrogenase, acid phosphatase, and diaphorase; Griffin 1986) and three morphological loci (ph, i, and dull seed coat; Palmer and Kilen 1987) were determined by screening F3 progeny.

RFLP genotypes established in the F3 plants were compared with germination scores in the F4 progeny of each plant. The families were sorted into genotypic classes (homozygous G. max, heterozygous, homozygous G. soja) for each of the 72 marker loci. F-tests (GLM, Statistical Analysis Systems, Cary/NC) were used to determine if significant differences existed for germination rates among the classes. Loci significant in this analysis were combined into a multivariate linear regression model to determine their combined effects. The significant markers were also tested for epistatic interactions by two-way analysis of variance. The significant interactions and main effects were combined in a multivariate regression model (GLM, SAS) to predict the total variation explained with markers. Genetic linkage among markers was determined by the maximum likelihood method and the computer program MAPMAKER (Lander et al. 1987). In this study, 72 one-way ANOVAs were performed on the same set of data. By using a significance level of p<0.05, there is a great risk of concluding that a marker is linked to a QTL when it is not, whereas choosing a lower level creates the possibility of missing important loci. This dilemma has been resolved by some researchers via a two-step design (Nienhuis et al. 1987). First, significant relationships are observed and then tested in a second set of progeny from the same population. This study constitutes only the observation stage, and markers detecting variation near the p<0.05 level will require further confirmation.

Results

In this study, the hard-seed trait was tested on ca. 150 seeds from each of 60 F2-derived lines. The germination response of the progeny lines was continuous from the extremes of 9% to nearly 100% germination (Fig. 1). The lack of discrete classes is consistent with a polygenically determined trait. Even so, some of the progeny exhibit the extreme parental phenotypes (AS1-356022 ca. 100%; PI 468916 ca. 1% germination). This also has been observed on other soybean populations (Kilen and Hartwig 1978), which suggested that a relatively small number of genes (ca. three) was involved in seed-coat hardness. The mean for this segregating population was 53% germination, which is not significantly different from the midparent value.

Populations can be partitioned into genotypic classes defined by genetic markers to identify quantitative trait loci (Edwards et al. 1987). In our study, seed-coat hardness data were sorted according to marker classes, and F-tests were used to determine if significant differences among genotypic classes existed. Seventy-two markers were used, 7 of which described significant effects at the p<0.05 level (Table 1). One explanation is that a QTL for hard seedness is located on the same chromosome as...